



KUNGL
TEKNISKA
HÖGSKOLAN



Using the WWW as infrastructure for collaborative production of documents

Henry Rodríguez

TRITA-NA-0117

**Licentiate Dissertation
Royal Institute of Technology
Department of Numerical Analysis and Computer Science**

Akademisk avhandling som med tillstånd av Kungl. Tekniska Högskolan framlägges till offentlig granskning för teknisk licentiatexamen tisdag den 7 juni 2001, kl. 10.00 i seminarierum E2, Lindstedtsvägen 3, Kungl. Tekniska Högskolan, Stockholm.

ISBN 91-7283-128-6, TRITA-NA-0117, ISSN 0348-2952, IPLab-185

ISRN KTH/NA/R-01/17-SE

© Henry Rodríguez, juni 2001

KTH, Stockholm 2000



KUNGL
TEKNISKA
HÖGSKOLAN



Using the WWW as infrastructure for collaborative production of documents

Henry Rodríguez

TRITA-NA-0117

**Licentiate Dissertation
Royal Institute of Technology
Department of Numerical Analysis and Computer Science**

Abstract

Collaborative writing is a common practice in workplaces and academic environments. The possibility to exchange documents through the Internet has paved the way for collaboration among co-authors. To date, the tools available for collaborative writing have largely been limited to email and commenting support. However, communication, co-ordination and sharing information are factors of vital importance to succeed in a collaborative writing task. The World Wide Web provides a network infrastructure that can support those needs of co-authors.

This thesis presents two web-based prototypes, the Domain Help System and Col^laboracio, aiming to support document and information sharing, communication, and co-ordination. The main research issue has been how small or middle size groups work collaboratively to discuss, annotate and revise a document on the WWW. Special emphasis has been put on the need to support a dialogue among collaborators. A number of case studies have been carried out, and the tools have been developed using the results of observation on the users' interaction with the systems. The studies indicate that the WWW can successfully be used for sharing documents and discuss their content. Also, asynchronous text-based communication was shown to be an effective medium for collaborative writing. The development of the prototypes has given general experiences on the development of collaborative web-based tools. These experiences suggest that such tools should be designed in collaboration with the users. A method for the design of such tools emerges from these experiences.

Acknowledgements

There will be many names that will not appear on this page but that does not mean that I am not grateful to those persons, it is just a matter of space. The origin of this thesis is the DHS, a CID project in 1996. So I would like to thank CID for all the facilities they gave me to start this project. Thanks to Nils-Erik Gustafsson who led the work in the DHS and who taught me the importance of “keeping it simple”. I would like to thank to all my colleagues in the research group at the Interaction and Presentation Laboratory who from time to time have assisted me. Specially to Teresa Cerratto and Hee-Cheol (Ezra) Kim for interesting discussions on my work.

Special thanks to Sandra Brunsberg who kindly accepted to use the DHS in her course and for the many ideas we discussed during the experiments. I also want to recognize the work of my assistant supervisors Ann Lantz and Olle Bälter who patiently and iteratively have read and commented on my work. On the other hand I have found in both of them a person to talk to, not only about research.

The most special thanks to my supervisor, Kerstin Severinson-Eklundh from whom I have learned so much and who has shown me the right way and even shaped bright fully many of my rough ideas.

To my friends Frank Santiago who proofread some of my writing and specially to Brian Sherwood, who put a lot of effort into proofreading my writing unselfishly, thanks a lot.

Finally I would like to thank to my friend Vicente Cabrera who in long conversations use to rise me up when I was feeling down and for his friendship. My last words are just to say sorry for the long time I have been away because of my work from my beloved guitar.

I INTRODUCTION

1 Background.....	6
2 The focus of this thesis	7
3 Outline of the thesis	8

II PREVIOUS RESEARCH

1 The writing process.....	10
2 The influence of the medium on the writing process.....	11
3 The use of annotations	12
4 Collaborative writing systems and theories	14
4.1 Software that supports collaborative writing	15

III THE DOMAIN HELP SYSTEM: WEB ANNOTATION AND DISCUSSION AROUND DOCUMENTS

1 Introduction	19
2 Main characteristics of the DHS.....	20
2.1 Layout of DHS.....	20
2.2 How the DHS works	21
2.3 The DHS as a discussion/annotation tool	22
2.4 Awareness of new comments	23
2.5 Logging function.....	24
2.6 Summary of the features of DHS.....	24
3 General topics in the DHS case studies	26
3.1 How users identified themselves in the domains	26
3.2 What is considered a relevant comment.....	28
3.3 Restrictions to enter into the domains.....	28
3.4 Members of the groups and participants in the domains	28
4 Case Studies.....	29
4.1 Pilot study: the R-group.....	29
4.1.1 Method.....	29
4.1.2 Results	30
4.1.3 Discussion and conclusion	31
4.2 Study one: the DHS as a discussion tool	32
4.2.1 Method.....	33
4.2.2 Results from the interaction.....	34
4.2.3 Discussion and conclusion	38
4.3 Case study three: the DHS as an annotation system (A-samples)	39
4.3.1 Method.....	41
4.3.2 Results and Discussion	43
4.3.3 Student perceptions of the DHS	52
4.3.4 Discussion and conclusions	55

IV COL•LABORACIÓ: A WEB-BASED COLLABORATIVE WRITING TOOL - SUPPORTING DIALOGUE

1 Introduction	57
2 The Author-Document-Author Workspace.....	58
3 Description of the system Col•laboració	59
3.1 Description of the document-development module.....	59
3.1.1 The index frame in Col•laboració.....	60
3.1.2 Commands of the system.....	61
3.1.3 Logging function	62
4 The author management module (AMM).....	62
4.1 Information given in the AMM.....	62
4.1.1 Commands that can be made from the AMM	63
4.1.2 The use of email in the AMM.....	64
5 Case Studies of Col•laboració	65
5.1 Case study 1: Poster A	67
5.1.1 Method.....	67
5.1.2 Results	67
5.2 Case Study 2: Report Alpha.....	68
5.2.1 Method.....	68
5.2.2 Results	68
5.3 Case Study 3: Report Beta.....	69
5.3.1 Method.....	69
5.3.2 Results	70
5.4 Case Study 4: Poster B.....	70
5.4.1 Method.....	70
5.4.2 Results	70
5.5 Case Study 5: Masters' thesis proposal	71
5.5.1 Method.....	71
5.5.2 Results	72
5.6 Case Study 6: Flyer.....	72
5.6.1 Method.....	72
5.6.2 Results	73
5.7 Case Study 7: Workshop	73
5.7.1 Method.....	73
5.7.2 Results	74
5.8 Case Study 8: Short paper.....	75
5.8.1 Method.....	75
5.8.2 Results	75
6 Discussion.....	76
6.1 The comments in the case studies.....	77
6.1.1 Discussion about the text.....	78
6.1.2 Trigger of actions.....	78
6.1.3 The dialogue among co-authors	79
6.1.4 Co-authors' activities in the dialogue	81
6.1.5 A window to the outside world.....	82
6.2 Problems with text-based communication within the system Col•laboració.....	82
6.2.1 Carriage return “noise”.....	82
6.2.2 Expired Validity of Comments	83

6.2.3 Probably too much “Talking”84

V DESIGNING FROM INSIDE: THE COMMENT SPACE OF THE ATP DOCUMENT AS A DESIGN TOOL

1 Unobtrusive user participation86

2 Supporting communication.....86

 2.1 Communication among users.....86

 2.2 Communication among designers and users.....87

 2.3 Communication among designers.....87

 2.4 Problem identification and error repair by mutual participation.....88

 2.5 A designer’s logbook88

 2.6 Visual prototyping.....90

 2.7 Social issues.....90

 2.8 A brief quantitative analysis of the users’ participation91

3 Conclusions92

VI CONCLUSIONS

1 About the methods used.....93

2 The discussion experience94

3 The annotations experience95

4 The collaborative document production experience96

5 Final words98

REFERENCES..... 99

I INTRODUCTION

1 Background

The Internet has paved the way for collaboration by letting users easily share objects and access to information, e.g. documents. Nowadays, the WWW has progressed beyond the capacity to display static information so that interactivity among members in a group can be supported. The possibility to communicate at a distance has increased and novel ways of collaboration have ensued.

The studies presented in this thesis are concerned with the achievement of collaborative tasks mediated by computers by small or middle size (2-15 people) groups. I have developed a set of computer programs that these groups have used to discuss and develop shared documents. The thesis is a contribution to the research field Computer Supported Collaborative Work (CSCW). Researchers from different fields have used diverse approaches in the development of this area where communication, collaboration, and coordination are the main issues. There are many arenas where CSCW has encountered practical applications (e.g. decision making, teaching, writing) and it continues to spread because participants' goals and tasks in a cooperation process are of abundant diversity (Shneiderman, 1998).

Writing is an activity performed daily in an office as well as academic environments. At the end of the seventies the computer as a new technology to the writing process was introduced. Technology solves some problems but always brings new ones. For example, a document can be saved in different formats and the user might not understand the difference between them; writers use different, often incompatible, software making it difficult to share their work; writing tools contain a great number of features that the user might never use or learn how to use; attention must be given to formatting.

More importantly, several studies have shown that the writing process is affected by the use of computer systems. For example, the amount of initial planning during writing is considerably less using word processors than pen-and-paper (e.g. Haas 1989). Also, writers generally revise more and pay more attention to the formatting aspects of the text. For an in-depth review of the impact of technology on the writing process, see Haas (1996).

At first glance writing may appear to be an individual activity, but this is by no means the case. Much writing is carried out in collaboration, directly or indirectly, together with others. Approximately 85% of all business and academic documents are written by several authors according to Baecker, Glass, Mitchell & Posner (1994). However, few word processors are designed to support collaborative writing.

Writing for publication on the World Wide Web is becoming increasingly frequent and tools that facilitate the production of Web documents are now very common. In fact, most word processors include the feature "Save as HTML", which gives a format supported by the Web. Several Web-based systems have been adapted or developed to support

collaborative writing. For example, Ceilidh (Huges, Jake & Okelderry, 1998) is a Web-based conferencing tool that has been used in a collaborative writing environment and as a distance learning tool. WebDAV (Whitehead 1999) is a network protocol for remote collaborative writing on the Web. Users are now beginning to explore the Web as a tool to support the collaborative writing process.

A Web-based collaborative writing tool inherits many of the intrinsic characteristics of the WWW. For example, it can make use of HTML formatting features and it is also possible to use hypermedia resources and to link to other pages on the WWW, sharing “external” information with co-authors¹.

2 The focus of this thesis

A considerable amount of research was done on computer supported collaborative writing before the Web (e.g. Haake & Wilson, 1992; Baecker et al., 1992, 1993; Neuwirth et al. 1992, 1994; Beck & Bellotti, 1993; Sharples et al., 1993; Mitchel, 1996; Miles, 1993; Kraut, 1992). Nevertheless, the WWW as an infrastructure for collaborative writing is a new technology that has been introduced to the writing process. It might be expected that the use of the WWW in general would affect writing, probably as much as previous technology has done. Studies of the use of the WWW in the writing process could give key insights into new Web-based collaborative tools.

In an interview study with collaborating writers, Kim and Severinson-Eklundh (1998) found that co-authors need a better network infrastructure, that co-writing is usually performed asynchronously and in small groups, and that one co-author usually maintains the document during the whole writing process. They also found that mostly no dedicated tools for collaborative writing were used by the co-authors. These premises and the rapid expansion of the WWW as a network are the points of departure of this thesis.

This thesis deals with the following issues:

How do small and middle size groups collaboratively discuss, annotate and communicate around Web-documents?

How is asynchronous text-based communication used in a shared space by co-authors?

Finally, this thesis deals with the design of Web-based collaborative tools in the light of these issues.

Generally speaking, the writing process can be seen as three intertwined sub processes (for details see the next chapter).

1. Prewriting in which, for example, co-authors have to plan and negotiate the activity, discuss the content of the text and strategies to follow.

1. The term co-authors refers both roles: the writer and the reviewer. Note that these roles can be played by the same individual at different times during the writing process.

2. Translation of the ideas into text.
3. Reviewing the text in which, for example, co-authors read the text in order to find problems or errors, and make annotations about it.

These processes must be coordinated in the writing group, which can follow several different models.

We set up our studies in such a way that the users:

- were exclusively involved in the discussion of the content of a Web document, or
- were exclusively involved in the annotation of a Web document, or
- had to interact among them to produce a document.

Breaking down these processes, we could collect interesting data about the use of the WWW as infrastructure for collaborative writing.

At the time when this research started (1996), the WWW was a “hot” technology. We decided to develop our Web-based tools, so that we could shape them according to our research intentions. As a result, we have developed two systems. The first system, the DHS (Domain Help System)¹ was used to support the process of discussion and annotation of Web documents. The second system, Col•laboració² was used to support the collaborative writing process in which co-authors discussed, negotiated, and co-ordinated their work. The design of these systems has a common characteristic: it is intended to support the need for communication that co-authors might have while working asynchronously and separately during the writing process. We have focused mainly on designing and studying this communication in the form of a text-based dialogue.

DHS and Col•laboració are still prototypes and we do not claim that they are a perfect solution for Web-based collaborative writing. Many enhancements have been suggested by users and designers. These enhancements have been taken into account when new releases of the prototype were made. One important factor was that the use of both systems could be observed by me. This gave me the opportunity to follow the interaction with the users closely and control the data collected. Moreover, developing our own Web-based tools gave experience of design of Web-based collaborative tools. These experiences are presented in chapter V of this thesis.

3 Outline of the thesis

Chapter II presents a review of previous research in areas related to the thesis. This includes the writing process in general, how the medium influences the writing process and some research about the use of annotations. Theories about collaborative writing and existing software to support it are briefly presented.

1. At the very beginning this project aimed to design a help system. For administrative reasons the name is kept.

2. Col•laboració is the catalan for collaboration

Chapter III presents the DHS system. A pilot study is also presented as well as two case studies in which the DHS was used involving discussion and annotation of Web documents.

Chapter IV presents the Col•laboració system. The chapter also reports on the case studies in which Col•laboració has been used. The last case study is discussed in detail.

Chapter V describes experience from the design point of view in relation to the Web-based tools developed, in particular how the users contributed to the design of the systems.

Chapter VI present conclusions drawn from the empirical studies using DHS and Col•laboració, and some design implications for collaborative tools. Limitations of this work are also discussed, and a brief description of future work is given.

II PREVIOUS RESEARCH

This chapter presents previous research related to this thesis. We focus mainly on collaborative writing, with special attention to the communication needs co-authors might have during the collaborative writing task. To broaden our understanding of the problems involved, it is appropriate to explore several issues such as the writing process, the influence of the medium on this process, the use of annotations, collaborative writing theories and systems, and the WWW as infrastructure for the writing process.

1 The writing process

Writing is a complex process. It is a learned craft that requires a lot of practice, and no formula can guarantee a good document. Writing is also a form of communication; it is used for the transfer of information, ideas and concepts from one individual to another, or within a group.

Rohman (1965) divided writing into three stages: prewriting, drafting, and revision in which the product was the core of writing (plan-write-edit). Janet Emig in 1971 (cited by Roundy & Mair, 1983) was one of the first that saw writing as processes (not as stages) focusing on *how* writers compose rather than *what* they compose. According to Emig research on writing before 1968 was not based on any theories. In 1980, Donald Murray's "Writing as Process: how writers find their own meaning" (cited by Caudery, 1995) argues that the writer is constantly learning from the writing. Murray placed emphasis on the importance of a series of drafts in the writing process as the writer gradually discovered through writing what it was that s/he wanted to say."

Hayes and Flower (1980) developed a model of the writing process as a cognitive process of problem-solving. Their model (see figure 1) has three basic processes: 1) planning, which includes generating ideas, organizing, and goal setting; 2) translation of plans into text; 3) reviewing, which includes reading and editing. Processes in this model do not appear in a linear way but randomly, that is, a subprocess could start at any time and even be part of any other subprocess. The writer could shift from one subprocess to another in any phase of writing. This model allows for studying how these three processes are coordinated in the writing process and how they affect the final product. Focusing on problem solving, the writing process is seen from a new perspective. A set of methods for probing these processes came out, for example verbal protocol analysis. Moreover, empirical findings from studies on problem solving had been used to shed light on a number of problem areas in current writing theory. In their studies of writing as problem-solving, Flower and Hayes focused their attention on writers' strategies. They discovered that expert writers, in comparison with novice writers, constantly redefine their audience and assignment while composing. Expert writers also consider in their goals how they wish to affect the audience. That is, expert writers develop more elaborate plans, and continue to develop and modify them throughout the whole writing process. As a result, expert writers modify the content more during the writing and the revision processes.

Flower and Hayes' findings have outlined various techniques that these good writers use to alleviate the problem of composing.

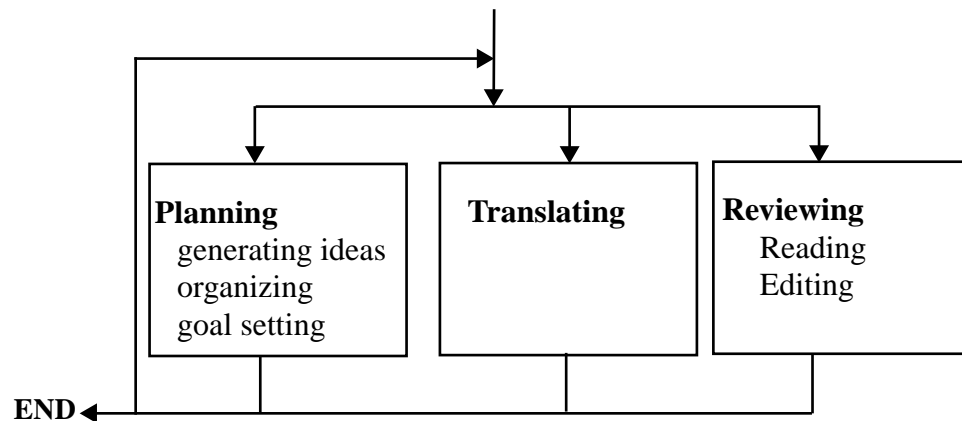


FIGURE 1. A simplified version of Hayes and Flower's three-phase model of the writing process (taken from Sharples et al. pg. 191)

Nystrand (1989) argues that writing is a social interactive process between readers and writers. In other words, a dialogue between the writer and the reader is performed to converge in a meaning of the text. The writer is trying to make the reader understand his/her intentions in order to reach a mutual understanding in a social context. The text itself plays a secondary role.

2 The influence of the medium on the writing process

“The medium is the message”, perhaps Marshall McLuhan’s most often quoted phrase, is relevant today because technology has been extensively used to create a new and diverse medium of communication. Technology has also affected writing tremendously; e.g. paper, moveable type, the typewriter, Linotype machine, and the computer. Christina Haas (1989) made an experimental study to assess writing products with: 1) personal computer, 2) workstation, and 3) pen-and-paper. The workstation had a bigger screen compared to the personal computer and different word processors were used in every condition. Results from these studies indicate that writers wrote at the same rate in all three conditions but they wrote longer on the workstation than with pen-and-paper and, therefore, produced longer texts. The quality of the text was significantly better on the workstation and with pen-and-paper than the personal computer. Haas’ second study assessed the writing process. The outcome of the experiment showed that more total planning and more initial planning was observed in the pen-and-paper condition. In the personal computer condition, users were attending more to the medium, rereading more and paid more conscious attention when revising. However, Severinson-Eklundh (1990) found that during revision on computer, more attention is paid to low-level error (e.g. misspelling, wording). In general, Haas found that “the pen-and-paper condition was consistently different than the computer conditions” (p. 163).

Writers using only word processors engaged in significantly less initial planning, conceptual planning, and total planning than when they used pen and paper. This

phenomenon was related to the difficulty experienced writers reported in getting “a sense of the text” of their texts and recalling them when using word processors (Haas, 1989; Severinson-Eklundh, 1992).

Kellogg (1994) affirms that word processors address just the typist aspect of the writing process and that idea processors should address the thinking aspect of the writing. He suggests that “idea processors may succeed in amplifying writing performance” (p. 184). He argues that outlining before translation of the ideas into text is tightly related to higher quality text compared to those who plan during the writing of the text itself. Kellogg (1989) points out that tools such as idea organisers may reduce the cognitive load of the user by focusing attention on the planning process while temporarily ignoring the translating and reviewing processes.

3 The use of annotations

During the writing process, from its genesis to the final product, annotations are almost certain to be made. Annotations can be created to fulfil different functions (e.g. planning, revising) and be directed to particular audiences (e.g. the next reader, a co-author, oneself). They also have distinct uses (e.g. to serve as a reminder, to command a change). Annotation can be used as an external store for knowledge or ideas that the writer has not yet given a coherent linear structure. For example, during the planning stage, the writer will probably make a note of some books, articles, or situations he or she would like to refer to in the text. Annotations could be used to advise on reading material, or schedule a meeting with the author. Also annotations could be done while reading; writers could record their own reactions, and interpretation of the text. In fact, annotating might be the most common activity when reviewing.

Neuwirth et al. (1987) presented the Notes program that keeps a link between each note and the specific region in the source text from which it came. They claim that computer-based note cards “would free the writers (1) to paraphrase because they would be able to recover the quotation and (2) to record their own elaboration, reactions, inferences, etc., because they could easily recover the context for them”. Computer-based note cards would also help writers create alternative organizational frameworks more easily. Moreover, the computer can hold and display a great amount of information which is significantly limited in paper.

Ovsiannikov, Arbib, and McNeill (1999) pointed out that digital annotation systems are in an underdeveloped stage and suggested an annotation technology that covers the architectural, functional, and user-interface aspects for annotation systems. They presented an empirical study of annotation on paper and demonstrated that electronic annotations are able to perform hypertext-oriented action, such as linking and sharing; they can be synchronized in real time; serve as the basis for conversations; searched; and have capabilities specific to the media type such as soundclip, reply, and so on.

The most common types of annotation interfaces are the split-screen, interlinear, and aligned interface. See figure 2.

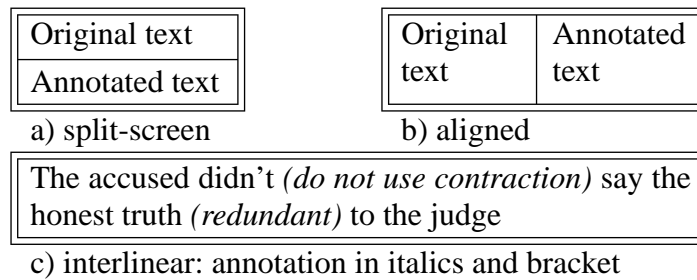


FIGURE 2. Three annotation interfaces

In the split-screen interface the user's screen is divided into two horizontal adjacent windows and in the aligned interfaces the screen is divided into two vertical adjacent windows. The original text is shown in one of the windows (usually the upper or the left one respectively) and the annotated text in the other. The interlinear interface annotations are differentiated from the original text by using format features (bold-face, italics, parenthesis, capitalisation, etc.) Wojahn, Neuwirth, & Bullock (1998) found that the production of annotation made by reviewers is affected by the annotation interface. They compared the three above mentioned annotation interfaces. It was shown that the time of the task (reviewing a document written by someone else) was not affected by the annotation interface. They found that those reviewers who used the aligned, and interlinear condition, communicated significantly more problems than those who used the split-screen condition.

Neuwirth et al. (1990) proposed a set of requirements for annotation interfaces during a co-authoring task:

- There is a minimum of motion required to make an annotation
- The primary text is easily distinguishable from the annotation text
- The annotations are visible “at a glance” while reading the primary text
- The relationship between the primary text and the annotation is easy to see
- Different annotators are readily distinguishable

Marshall's (1997) study focused on the form and function of annotations made on textbooks by students and discussed some issues and implications for the design of annotation tools for a digital library settings. She found that the annotations could be found in the text (also called in-situ), or in the margin (also called off-side). Annotations can be telegraphic, namely, using a personal opaque coding, or they can be explicit; usually textual. Finally, annotations can be removable or not. It was found that “annotation form arises in part from the characteristics of the material themselves, the imprints and the implements used to write them” (p. 134). For example, it turned out that students that use highlighters write fewer marginal notes than students who use pens. Finally, she suggested support for digital annotations: 1) in situ annotation (e.g. interlinear annotation), 2) annotations should be distinguishable from the source (c.f. Neuwirth et al., 1990), 3) non-interpretive marking (e.g. underline, highlight of text) as they are very much

used, 4) fluidity of form because annotations on paper were very rich in form, 5) informal coding; several annotators developed a personal system of annotation in which symbols and pen colour meant something to the reader, 6) public and private annotations, and 7) annotation should interrupt reading as little as possible.

It has also been found (Neuwirth et al., 1994) that voice annotations for reviewing documents have benefits compared to text-based annotations in the reviewing process of a document. Kraut et al. (1992) found that those who used voice annotations were more likely to comment on global problems of voice and structure of the text than those who were communicating via text. Moreover, voice annotation compared to text-based annotations were judged to be more useful; especially for structural and semantic problems, and more complete, justified and socially oriented. This implies that it is important to have different kind of modalities for this process. However, Kraut et al. note that voice annotation can be tedious to listen to and difficult to process.

4 Collaborative writing systems and theories

Collaborative writing is a very old and often-performed practice.

. Rimmershaw (1992) defines collaborative writing as follows:

“... any piece of writing, published or unpublished, ascribed or anonymous, to which more than one person has contributed, whether or not they grasped a pen, tapped a keyboard, or shuffled a mouse.”

Sharples (1993) describes four areas that should be studied to give a deeper understanding of collaborative writing and in this way consolidate and enforce that basis for the design of CSCW systems for collaborative writing. These areas are:

- Task issues: Includes strategies used by writers for partitioning and coordinating the work, and the interleaving of tasks: a) sequential - one co-author passes the document to another co-author after he/she has worked on the document; b) reciprocal - all co-authors work together to create the document; c) parallel - the writing is divided into sub-tasks and each co-author is responsible for one of these, for example: all co-authors work in parallel in different sections of the document..
- Group issues: Definition of roles, substitutability and interdependence in the group: a) Substitutability and interdependence between groups members; b) roles; c) management of conflicts; d) sub-groups.
- Communication issues: Describes the context in which communication takes place. The effects of media on communication and its structure: a) identifying the purpose of a communicated representation; b) communication in context; c) deindividualization and media effects; d) structured communication.
- External representation issues: Include types of representation used by writers, constraints and effects of media representation and management: a) what to represent;

b) constraints; c) communication of representation; d) effects of media on representation; e) version management.

Neuwirth et al. (1990) points out that members who are writing in collaboration should "understand plans, changes, and comments made by others"; namely, the communication process should be well supported. They also claims that definition of social roles and communication among the members about plans, constraints, and comments is crucial.

Galegher and Kraut (1992) compared groups of co-authors who used either computer-mediated or face-to-face communications and found that the groups that communicated by computer had more difficulties doing their work, but performed as well as those who met face-to-face. They also found that if the group had the means to work together easily, a substantial portion of the work was done jointly and a lot of time was given to communication so that members of the group could coordinate their ideas. On the other hand, groups lacking the means to work integratively worked more independently and invested less time in communicating about the task. Neuwirth et al.(1994) state that participants in collaboration can improve their chance of success by managing with whom, when, what, and how they communicate. Also, knowing what not to communicate can be just as important as communicating. They claim that "to provide computer support for interaction among collaborative writers requires understanding their patters of social interaction".

Kraut et al. (1988) affirmed that it is not easy for collaborators who are using computer-based tools to carry out their activities due to the incompatibility among programs and the computing environments. Thus, in a computer-supported collaborative writing task almost the whole responsibility of formatting text, for instance, is dropped upon one of the collaborators. As a result, the person who is responsible for the formatting takes total control of the manuscript. So the collaboration takes a hierarchical-centralized structure which, very likely, will make the outcome represent the preferences of the one who was in charge, more than those of the group as a whole.

Co-authors are likely to choose a rich communication medium (e.g. voice, face-to-face) when they are planning and revising. Face-to-face communication was found to be very helpful for coordination of planning and construction of long documents. However, co-authors avoided face-to-face communication when they were drafting the manuscript according to Kraut et al. (1992).

4.1 Software that supports collaborative writing

Some writing systems are explicitly intended to support collaborative efforts. Jones (1990) maintained that no single tool could as yet satisfy the majority of an author's needs.

Quilt (Fish et al. 1988) is a hypertext editing system that provides co-authoring services. It allows users to work on the same document but typically on different sections and at different times. The users are identified as either co-authors or commenters. It allows

annotations, computer conferencing, and notification facilities to support communication and information-sharing among collaborators on a document. Only certain privileged users, i.e. the co-authors, can modify the document.

NoteCards (Halasz, Moran & Trigg, 1987) is a hypertext system developed at Xerox PARC. NoteCards provides the user with a "semantic network" of electronic notecards interconnected by typed links. The system provides tools to organize, manage, and display the structure of the network. It allows multiple users to open and read the same notecard but only one of the users may modify its content at any given time. It is said that this limitation leads to drafts of the document being passed back and forth between the authors.

PREP (Neuwirth, Kaufer, Chandhok and Morris, 1990) is an editor that allows communication, planning, and organized annotation. The system defines chunks, which roughly correspond to ideas. A chunk can contain text, grids, trees or images. Chunks are stored in a database that is shared by collaborators. It focuses on providing a usable, visual representation of the information, pinpointing the differences between different versions of the document, and coupling among users. It provides an align-based interface where related information is linked across columns.

GROVE (Ellis, Gibbs, and Rein, 1991) is a synchronous distributed application where multiple users can simultaneously edit private, shared, and public views of a document. When users enter (or re-enter) the system, they receive an up-to-date document unless they choose to retrieve a previously-stored version. GROVE uses voice coordination and users themselves establish a social protocol. There is no locking function for editing or any turn-taking mechanism. The authors of GROVE reported that despite collations were surprisingly infrequent while using the system, the need to have a locking function could be relevant.

SEPIA (Haake & Wilson, 1992) is a hypertext authoring system that supports both communication and co-operation, but is limited to a single application. Co-authors can work in synchronous/asynchronous mode. In the independent mode (that is while working asynchronously) users may work on their own tasks without interfering with each other. In the loosely-coupled mode (working synchronously) users may share certain public information while working on their own tasks. This system aims not to affect an author's ability to write an individual document. It uses four activity-space browsers dedicated to the task of content generation and structuring, planning, arguing, and writing the final hyperdocument under a rhetorical perspective. Multiple browsers on the screen displaying different information are supported and it uses colour to indicate the state of an object. SEPIA allows co-authors to have verbal exchanges and each co-author sees the same browser in terms of size and content. Also, it has a telepointer device that is used for pointing to items visible in the shared view. A locking facility of the objects in the shared space is provided.

One of the first applications using the web infrastructure to support collaboration work was the **BSCW** (Basic Support for Collaborative Work) system (Appelt & Busbach, 1996). It is a web-based groupware application in which users can create and structure

repositories into which they can import documents which are then made available to all users who currently have access to the repository. Access rights in BSCW can be set on a per-object basis to control the operations available to different users. Locking and access control are clearly related. Horstmann and Bentley (1997) have built a number of applications which augment a standard web browser. These applications provide a much richer interface compared with pure HTML, e.g. drag-and-drop to the web browser to upload multiple documents is possible via the currently displayed BSCW's workspace folder.

Alliance (Romero & Decouchant, 1997) exploits special knowledge about the internal hierarchical structure of its documents to provide concurrence control on document subtrees. It uses the WWW as an infrastructure to accomplish distributed document management, asynchronous group awareness, and communication among co-authors. In Alliance, distributed document management is based on the full replication of documents on the site of each author participating in the writing task. It defines roles that are related to the document parts (also called fragments in the system). No synchronous work is supported within Alliance. It can handle temporary disconnection from the network without disrupting the cooperative editing.

Knister & Prakash (1990) developed a toolkit that can be used to convert single-user editors into group editors using minor modifications to the code of the editor. This toolkit is called **DistEdit**. Group editors are system oriented to support synchronous, distributed work. The DistEdit toolkit concentrated on providing more closely coupled real time interaction.

The resulting editor allows several people to jointly edit a shared document in a distributed environment. Participants can alternate between working independently and taking exclusive global control of the shared space while other users observe the changes as they occur. It does not provide structural information of the document. It maintains a copy of the state of the editor for each user.

WebDAV (Whitehead and Goland 1999) is a network protocol for remote collaborative writing on the web. The main idea behind WebDAV is to add collaborative authoring capability to the tools already used by the writers e.g. word processors. This is done by focusing on the network protocol. WebDAV extends basic HTTP with capabilities for overwriting prevention, properties, and namespace management. Using WebDAV co-authors are not forced to use the same tools, for example compatible word processors. It prevents lost updates caused by locking.

Ceilidh (Huges, Jake and Okelberry, 1998) is an asynchronous web-based group discussion tool that has been used for collaborative writing. Ceilidh takes advantage of two important characteristics of the WWW: its ease-of-use and its ability to create hyperlinks to other pages. It also has the ability to include graphics and even multimedia outcomes. Ceilidh automatically converts text, including carriage returns, URLs and email addresses in the body of the message into HTML. It uses threading. The content of each comment is presented one by one and a navigation bar through the comments is available.

Cadiz, Gupta and Grudin (2000) recently presented a case study of annotation created using the system **Microsoft office 2000**. The system is one of the first commercial products to support web annotations. It uses email for notification of new annotations. They observed that most annotations for a document occurred near the time of the first annotation, however, many documents continued to be annotated for several months. This shows the importance of annotations. The system anchors annotations by computing a unique signature for each paragraph. For this reason the annotation system can fail to match an annotation to the correct location when the text is edited. When this happens, the annotation is “orphaned.”

The **Anchored Conversations** system presented by Churchill et al. (2000) was presented at CHI 2000. It provides a synchronous text chat window that can be anchored to a specific point within a document, moved around like a post it note, and searched via a database.

III THE DOMAIN HELP SYSTEM: WEB ANNOTATION AND DISCUSSION AROUND DOCUMENTS

1 Introduction

The Domain Help System (DHS) project started in 1996. The original idea was to develop a new help system that could present information according to an iceberg model of information. The idea of this approach was that initially only a minimum of information is presented, namely that information which an experienced user could request for, and consecutively more information is available on demand. Like an iceberg the system shows only tips of information, but there is an entire mountain of information to be accessed as we descend deeper. An important requirement for this project was that a user should be able to easily add knowledge items to the system while he/she was interacting with it. We therefore decided to have a simple representation of knowledge items in the form of written comments or annotations.

However, the DHS is not at present a help system. The process of designing a new system, aiming to support a specific task, might be full of unexpected results. The original goals can easily swerve, and in this process the task is also redefined. Furthermore, the artifact (in this case the system) adapts itself to the requirements that the new task imposes and creates new possibilities that, in turn, may modify the task. Carroll (1991) defined this process as the task-artifact cycle. The design of the Domain Help System (DHS) bears witness of the task-artifact cycle through iterations of user participation.

The result of that project was a Web-based tool that, generally speaking, lets users select a Web document from a list and attach a comment to it. These comments are available immediately to all the users. The comments are presented in chronological order, which allows for the development of a dialogue. Joining a) the possibility to make an annotation or comment on a Web document, and b) the possibility to create a dialogue among those using the tool, we decided to use the DHS as a discussion or forum tool and simultaneously as an annotation tool for document production.

This chapter presents the current DHS user's interface and the case studies in which the DHS has been used. The DHS has been modified after each user study we carried out in terms of the interface and usability, and these changes will be reported on here. For the analysis of the case studies I used the comments submitted by the users while using the system, and also, I performed interviews and surveys. Additionally some of the case studies included log files that will be used in the analysis.

The use of the DHS can be separated into three contexts. The first context was in a research department and it was a pilot study (it will be called R-group case study). The purpose was to investigate how the DHS could be used and how users interacted with it. The R-group case study disclosed the potential of the system. The second and third context were in an educational framework.

In the second context the DHS was used as a forum tool in a CSCW course in which participants had to discuss the content of documents. This will be called the D-sample case studies and involved four versions of the course over a period of three years. The documents were written by the course participants and dealt with a topic given as an assignment to them by the teacher. In this context, the DHS was used to support asynchronous discussion for a middle size group (10-15) around a document.

In the third context, the DHS was used in a writing course given to Ph.D. students at KTH. Students were supposed to review or correct the documents that other class roommates wrote to fulfil an assignment. This will be called the A-sample case studies and involved data collected in four versions of the course over a period of four years. This case study led us to draw conclusions about the communication process between the writer and the reviewer about changes in the text. In contrast to the D-sample case studies, the comments made were related to the form of the document rather than their content.

Below some general aspects of the DHS are described. Later, we will describe the case studies and their results, followed by discussion.

2 Main characteristics of the DHS

The DHS is a Web-based system that lets users share Web documents and make comments on them. The comments are also shared by the users. A domain, in the DHS project, is a site on the Web that divides the Web browser into four frames which display: 1) a hypertext link list, 2) the content of the most recently activated link in the form of a document, 3) the comments users have made so far on that document, and 4) a button that pop-ups a comment-input window. The documents held by the system are included in the domain by the domain's editor (in all the studies, I was the editor). Comments could be related to the content, the formatting, the design of the document, or even be a response to a previous comment. A domain is said to be closed-from-inside because once a document or comment is "inside" the domain, users cannot delete or modify it. Users can just retrieve or add new comments. Only the domain's editor can update the files in the domain. In Rodríguez (1999) it is motivated the creation of a Domain Help System and is described the most important parts of its development pointing out the versions or prototype's characteristic and how differs a version of the DHS with the previous one.

2.1 Layout of DHS

When the user enters the system, the Web browser window is divided into four frames as follows (see figure 1):

- The left frame displays the hypertext link list (index frame).
- The right frame that is divided into three frames (information-frame).
 - The top-right frame displays the content of the most recently selected link (content-frame).

- The middle-right frame displays the comments made so far to the document that is presented in the content frame (comment-frame).
- The bottom-right frame displays a button that pops-up a window from which a new comment could be written and submitted (command-frame)

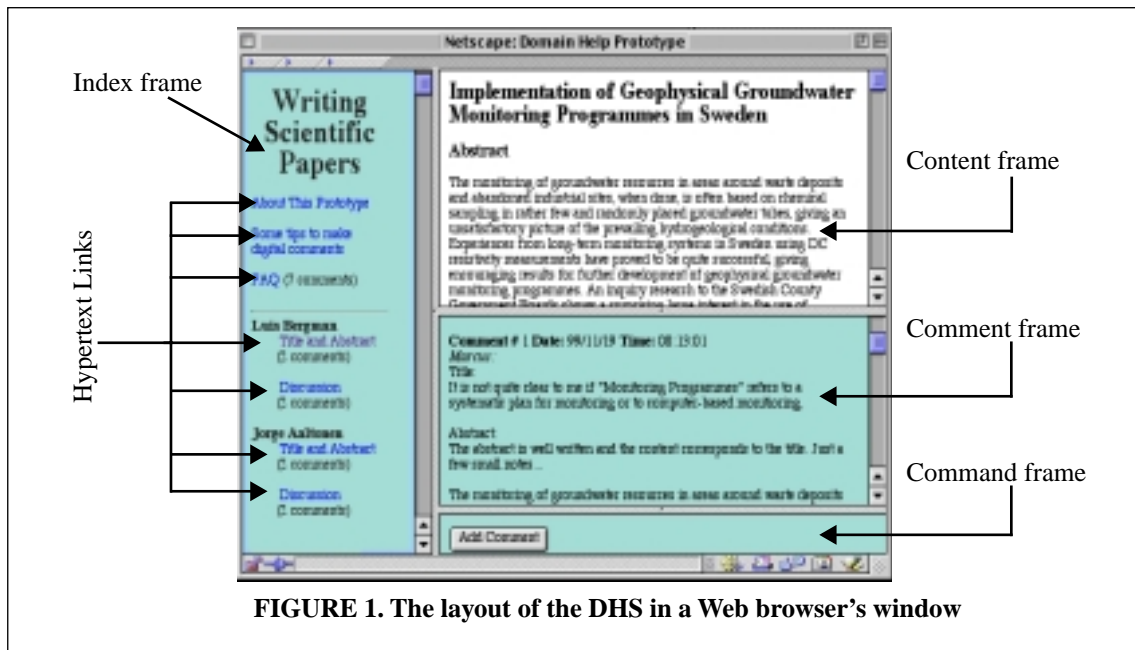


FIGURE 1. The layout of the DHS in a Web browser's window

2.2 How the DHS works

The domain's editor creates a domain and members of the group are asked to send their contributions in the form of documents. The domain's editor receives these documents, makes a HTML version of them if needed, and places them in the domain. The DHS automatically creates a hypertext link list including all the HTML files in the domain. Additionally, the domain's editor includes one document that explains how to use the system and how to deal with known problems of the system. This document is associated with link in the index frame of the domain with the text "About this prototype". In the following. This document will be referred to as the ATP document. To provide users with an easy and quick access to the ATP document the system reserved the first place in the index-frame to its link. The communication consisting of comments under the ATP document is discussed thoroughly in chapter V..

In detail, this is what happens when the user enters a domain.

- In the index-frame the hypertext link list of the documents that the domain contents is displayed.
- In the content-frame the ATP document is displayed.
- In the comment-frame the comments made so far on the ATP document are displayed. Comments in the domain are presented in a chronological order and a simple navigation bar is presented after each comment "First-Next-Previous-Latest". When all

the comments are loaded, the system scrolls automatically until the last comment is visible.

After entering the system, users can select any document from the index-frame. When the user selects a document (clicking on its link), the content-frame and the comment-frame are updated immediately. The content-frame displays the content of the just selected document and the comment-frame displays the comments made so far on it.

In the first versions of the DHS, to add a comment on the current document, (the one that is displayed in the content-frame) the user has to click on the “Add comment” button. This will open a separate window (see figure 2) in which users can write their name or nickname and the text of their comment. This window identifies the document to which it is related by adding the same hypertext link text as a heading in it.

Figure 2 shows the Add comment window for the ATP document. When the comment is submitted the Add comment windows is closed automatically, and the comment is appended immediately to the comment-frame; the just added comment is presented at first. After this, all members of the group have access to the comment whenever they select the corresponding link. If a document has no comments, a message saying so is presented to the user.

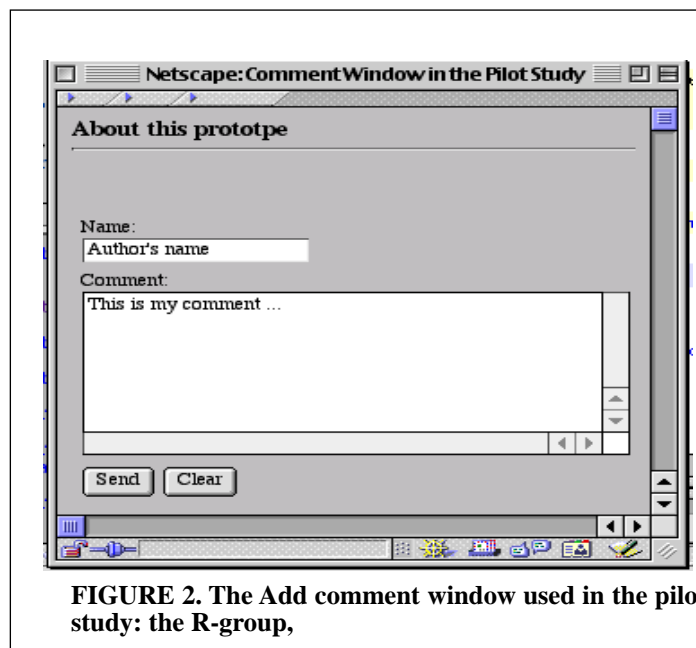


FIGURE 2. The Add comment window used in the pilot study: the R-group,

2.3 The DHS as a discussion/annotation tool

The DHS interface has been changed iteratively in response to the observations of the users', the interviews, and survey. One point worth noting here is that the changes made to the Add comment window have been significant for the system as a whole. The first significant change was to make this window bigger and divide it into two frames, where the left frames contains the same document that was in the content-frame when the add comment button was pressed (see figure 3)

Another important change was due to the use of the system in the annotation context first. In the first A-sample case studies, users frequently needed to copy-and-paste part of the content of the document they were supposed to review. To support this action a new button was included that pastes all the text that is found in the content-frame into the text area in which the comment is going to be written. In this way those users that want to quote the original text are supported by the system. Figure 3 shows the Add comment window that was used in the D-sample and the A-sample.

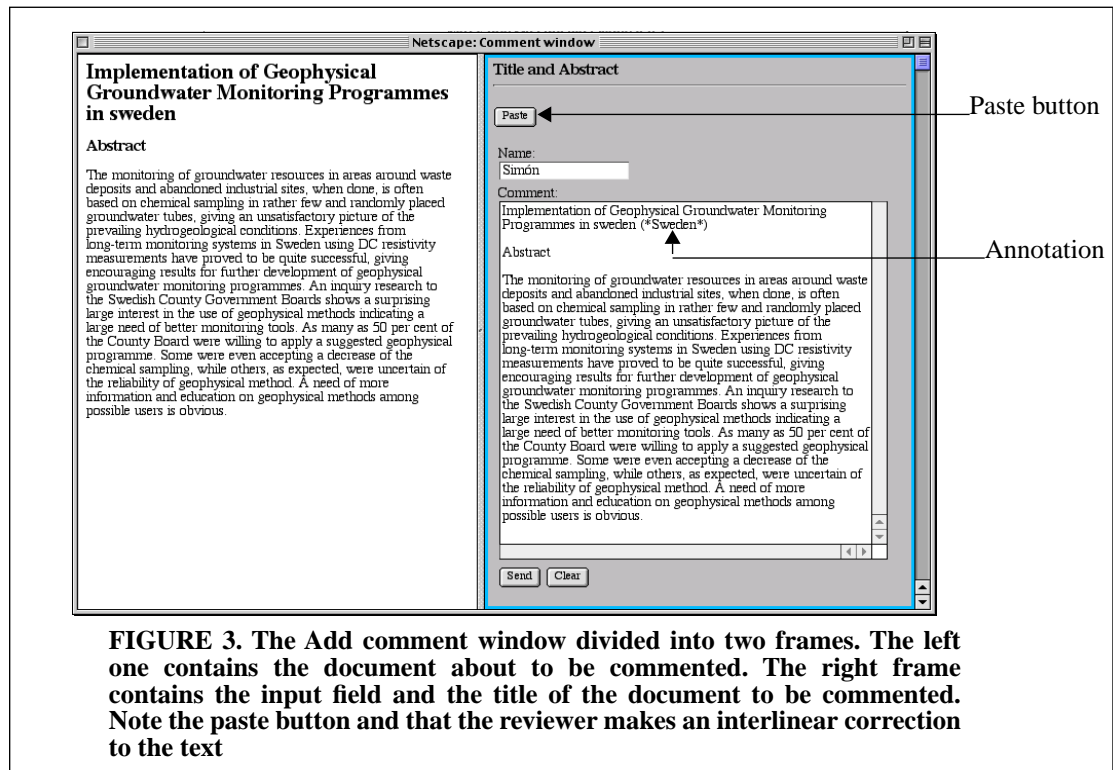


FIGURE 3. The Add comment window divided into two frames. The left one contains the input document and the title of the document to be commented. The right frame contains the input field and the title of the document to be commented. Note the paste button and that the reviewer makes an interlinear correction to the text

2.4 Awareness of new comments

One of the main problems that the first versions of DHS confronted was the lack of group awareness about changes. Consider this situation: user A is the author of document-A, say, user B makes a comment on document-A and on document-C. The only way for user A to know about this was to enter the system and select document-A and document-C. If user A enters the system it is very likely that he will select document-A and will, hopefully, notice that there is a new comment. It might not be the same for document-C, as user A is not its author. It might be frustrating for a user to spend his/her time visiting the domain several times and after browsing all the documents in it, find that everything was the same, especially in relation to the documents that he/she is particularly interested in. As a result of this frustration, the user might not come back to check the domain any more. To alleviate this problem two features were implemented. First: Email messages are sent automatically to the author(s) of the document when a comment is made to it. The comment is attached to the email and it is labelled in such a way that the receiver could recognize that the message is sent by the DHS system.

The second solution was to write a comment counter tag beside the links in the index-frames to indicate how many comments had been made when the session started, for example: Document-A (3 comments). Users might recognize that the comment counter tag of a document has changed. It also helps explorative browsing of the domains, since users will at a glance know whether a document in the domain has got comments and how many comments there are so far. The user might be interested to “see” the most (in number of comments) commented document’s content and its comment, or, conversely, see why a document has not been commented on at all.

2.5 Logging function

The DHS has a logging function in order to provide information about the users’ actions in the system, which documents they visited, how long time they spent on each document, how long time it took for a comment to be written, how long time the user spent in one session, etc. Basically, every time the user enters the system, selects a link from the index-frame, or when a comment is added, a record is appended in the log file as an action. An action here is defined as any change made on any of the users’ frames content in the DHS-site. That is, mouse movements, scrolling, text selection, key stroke are not considered actions. However, to what extent this reveals their real actions is unclear, as users could click on a document and at the same time receive a phone call, for example. This situation is not trailed by any logging function. Despite this uncertainty, I considered that this data might be valuable for some aspects of my studies. The log file registers the time of the action, the user’s IP number, which document was selected and the addition of a comment.

2.6 Summary of the features of DHS

To recapitulate the features that the current version of DHS offers as a result of user studies: First, there are two ways to make users aware about new comments in the domain. The first one, very individual oriented, is by sending an email to the writer of the document that has being commented, that is, **email notification** that a new comment has been added. The second, members oriented, a **comment counter tag** is attached to every document link in the index-frame indicating how many comments that document has.

The other important change was the layout of the Add comment window. The initial mode was a **single layout** (see figure 2), later a **combined layout** (see figure 3), was developed in which the input areas were beside the content of the document to be commented on.

Also, the **paste-text function** added is an important feature as it supports the annotation context in which DHS was also used. Finally, the **log file function** was introduced to trace users’ actions in the DHS.

In table 1 all the case studies in which DHS was used are presented. Additionally, it describes which of these features were used in the case studies. I have named the samples in the case studies in the following way. For the discussion context the letter D will be used as a prefix, and the letter A for the annotation context, in both cases the year in which the course was given is appended to the letter.

Table 1. Case studies in which the DHS was used

		Pilot Study	D-Sample					A-Sample					
Case study	R-group	D1998	D1999	D2000	D2001	D-sample total	A1997	A1998	A1999	A2000	A-sample total	Total ^a	
Number of participants	16	26	11	18	13	68	15	13	4	16	48	132	
Number of Contributions	17	21	11	18	13	63	18	16	30	98	162	242	
Contribution with comments	12	21	10	17	13	61	16	14	12	72	114	187	
Duration (days)	97	78	15	38	15	146	55	49	66	70	240	483	
Number of comments	77	130	33	61	103	327	65	70	22	132	289	693	
Number of words in comments	3789	12532	3751	4869	9868	31020	16719	9127	2264	17973	46183	80992	
Number of comments on ATP	48	25	1	4	13	43	0	0	0	2	2	93	
Number of words in ATP	2981	1916	57	189	2062	4224	0	0	0	62	62	7267	
Use mandatory	No	Yes	Yes	Yes	Yes	-	No	No	No	No	-	-	
Relevant Features	Log file	No	No	Yes	Yes	Yes	-	No	No	Yes	Yes	-	-
	Email awareness	No	No	Yes	Yes	Yes	-	No	Yes	Yes	Yes	-	-
	Comment counter tag	No	No	No	Yes	Yes	-	No	No	No	Yes	-	-
	Combined layout	No	No	No	Yes	Yes	-	No	No	Yes	Yes	-	-
	paste-text function	No	No	No	No	Yes	-	No	No	No	Yes	-	-

a. Total = R-group + D-sample total + A-sample total

3 General topics in the DHS case studies

There are some common aspects in the three contexts (in the research department, in the discussion context, and in the annotation context) in which the DHS was used. The way users identified themselves while using the system, which comments were considered relevant in order to be included in the analysis, the privacy policy we took into account for the domains, and which users were those whom we labelled as participants are the topics described in this section.

Additionally, there are some common characteristics for all the domains:

- Participants had to work in collaboration in real cases.
- Participants knew each other personally and there was a common physical location where they met regularly face-to-face (daily, once a week, or once a fortnight.).
- The groups of participants were not especially created to use the prototype, but were formed for other reasons, i.e. they were members of a class course, or they were in the same project, or they were working in the same department.
- Group members had access to different ways of communicating with each other, for example, using email, phone, face-to-face.
- The only way to include a document into the domain was through the domain's editor. Group members sent via email the document they wanted to include in the domain and the domain's editor made an HTML version of the document if it was needed.

3.1 How users identified themselves in the domains

When sending a comment, users could write their names or any nickname they wanted (the submission of a comment could be anonymous or attributed to the user). No automatic authentication feature is provided by the system; users have to type their names when sending a message.

Table 2. Number of users that always used the same nickname or several nicknames while sending a comment in the DHS

Case	Unique nickname (%)	Two or more nicknames(%)	Participants that submitted at least 2 comments
R-group	5 (63%)	3 (37%)	8/16
D1998	16 (70%)	7 (30%)	23/26
D1999	6 (75%)	2 (25%)	8/10
D2000	12 (75%)	4 (25%)	16/18
A1997	7 (64%)	4 (36%)	11/15
A1998	9 (90%)	1 (10%)	10/13
A1999	3 (75%)	1 (25%)	4/4
A2000	8 (53%)	7 (47%)	15/16
Total	67 (71%)	31 (29%)	95/118

Generally, automatic authentication requires that users remember a login and a password; the intention was to free users from remembering them. Users did not report that his/her name had been used by another person. The list of the participants' names was available to us so it was possible to determine if the comment was indeed sent by one member of the group or if it was sent by an outsider.

It was interesting to know how the users identified themselves. Did they use their own name or a nickname? Or both? How many used always the same name? Table 2¹ shows how many users adopted only one name (or nickname) or at least two different names to identify themselves in the domains, for example, José Rivas or J. Rivas. To build up this table, we took into account in every case study only those users who sent two or more comments. The comparison of the names was not case sensitive, e.g. Martin - martin, were considered as the same entry. Additionally, obvious misspelling errors in their names; Helgsreand - Helgstrand, Mikael- Miakel, for example, also were counted as the same entry. At least 29% (31) of all the users who sent two or more comments adopted at least two names. Although the number of users who falls in this category might not be alarming, the same would not be true in cases where the exactness of the names is of vital importance; for example if an automatic filtering feature would be implemented. This situation would be more remarkable if we consider misspelling or case sensitive letters. This suggests that it is important to provide users with an automatic authentication system while they communicate in a text-based asynchronous mode.

An automatic authentication feature would provide:

- The basis for filtering functions
- Save time; user will need to write their name once per session and not every time a comment is sent.
- A security policy could be implemented
- Private comments might be supported by indicating to whom the comment is sent

In the version used by the case studies R-group, D1998, and A1997, if the user did not write any name at all, intentionally or not, and sent a comment, it was still accepted by the system. There were several comments (9) with anonymous sender in all domains, e.g. "anonym", "ghfg", or no identification at all but in most of the cases these anonymous comments were classified as not relevant (see section 3.2). To let users send comments anonymously could lead to some problems but we did not experience any embarrassing or serious situations.

However, not having an automatic authentication feature in our system has some advantages. For example, two or more users may sign and/or write together one comment. Note that in example 1 Maria and Juan are the authors of the comment and that they used several times *we* in the body of their comment. This might increase the sense of collaboration. Moreover, requesting to sign a comment every time might increase the trust that nobody is tracking them when sending anonymous comments. The examples are

1. The case study D2001 was not taken into account to build up this table

presented in the same format in which they were shown in the DHS except for the italic font used in these examples. The first line labels a comment with a number, the date, and the time. The second line identifies the author's name. The other lines contain the comment itself. The name of the users has been changed.

1) ***Comment # 44 97/05/13 14:37***

Maria & Juan:

We think that it would be nice We agree with Pedro when it comes to.... Moreover, we think...

One possible solution would be to have a modifiable authentication feature. That is, users would automatically get their names in the Add comment window, but be able to edit it and if needed, change it.

3.2 What is considered a relevant comment

In all the domains we found 29 comments that we considered as not relevant comments: e.g. "54t6yrethr", or, "this is a test". These comments were not taken into account in the analysis. We, thus, consider "relevant comments" those that made sense in their context and expressed an opinion or gave new information: a question, a suggestion, or a reply to a previous comment.

3.3 Restrictions to enter into the domains

To enter a domain, users need to have the domain's URL (Web address). This URL, in most of the cases, was sent via email to the group and in few cases it was given during the class. In the R-group and discussion context (the D-samples case studies), no security system was provided to ensure that those who visited the domain, or sent a comment were members of the group working with the domain. In the annotation context (the A-samples case studies), members asked explicitly for a security protocol so we provided a password policy for this domain. This was because students were told that they could submit homework related to their current research work. They considered that this information might be published in the future.

3.4 Members of the groups and participants in the domains

Members of the group in these case studies were defined as follows: a) in the R-group consisted of all those who were part of the research department; b) in the D-samples members were the students who were taking the course and the teacher; c) in the A-samples members were the students who were taking the course, the teacher and external reviewers.

There were some members of the groups that entered no comments. For the analysis of the communication in the domains, we consider as participants only those who entered at least one relevant comment.

4 Case Studies

This section presents the three studies in which the DHS was used. The pilot study was made in a research department. After the pilot study, we decided that the DHS could work well for distributed, asynchronous discussions around a Web document. In the educational field, students work often around a document. For example they have to read and comment on other students' writing. Therefore, we decided to use the DHS in a traditional classroom context. Thus, the second study was made to investigate how the DHS supports discussion about the content of a document (about one page long). The third study was made in a course where students were supposed to learn to write in academic style. In this case, the aim was to investigate how the DHS supports annotations to a document.

4.1 Pilot study: the R-group

This was the first study in which the DHS was used. This study had an explorative nature and aimed at disclosing the possible uses of the DHS.

4.1.1 Method

Participants

Members of an interdisciplinary research group (R-group) participated in the study. There were about 25 members at R-group. All members of this domain had a Human Computer Interaction (HCI) background and were computer literate. No reward was given for participation.

Apparatus

The version of DHS used had the following characteristics: The Add comment window presented two input fields (for the user identification and for writing the comment) see figure 2. Participants had different Web browsers (Netscape, IE), versions (Netscape 3.1, Netscape Gold, etc.) and platforms (Mac, PC, Sun).

Procedure

R-group members were asked to make a draft project description of their research activity. These reports in a finished form were supposed to be published on R-group's Web site later, and each member would get to know about other members' projects. The R-group's members were able to make comments on each other's project descriptions during the drafting stage.

A domain for R-group's topics was created. The index frame presented a list of the research topics going on at the R-group. R-group's members could read about other colleagues' work and make comments on this description. This domain was on line for 97

days. This period is calculated taking into account the date of the first comment and the very last comment registered in the system.

We had 15-20 minutes long interviews with the users about the DHS's interface and about the different uses that could be given to this tool

4.1.2 Results

A total of 15 project descriptions were presented in this domain. Despite the request made by the head of the group members to use the system, very few contributions in the form of comments were made in the domain related to the topic that were submitted. Instead, participants readily made comments on the interface of the prototype. No active discussions about the topics were registered except for the one in the ATP document. The ATP document got the major amount of comments, 48 of 77. Using the unit words-in-comment, we found that almost 80% of the discussion made on this domain was related to the ATP document, in other words, to the design or the use of the DHS.

Twenty-nine comments were found in the other documents that formed the R-group domain. These comments were related to the content and style or layout of the document but they were rather brief; 24 of them contained less than 40 words. Example 2 is a comment that we found in the R-group domain.

2) **Comment #1 97/03/18 21:18**

Soledad:

This text could be improved. There is too much about the author (this fits better on a home page), too much background and too little about the results of the project.

The interviews showed that all the users welcome the idea to get comments on their work from their colleagues. However, some of them said that this situation might create conflicts among them. Users also reported that it was hard to find the URL, which was sent via email, for the domain. What kind of comments the users were supposed to make was also an issue that was not clear according to the interviews.

Authors in most of the cases did not reply to the comments made to their project descriptions. When replies were made it usually took more than 8 days to respond. Example 3 shows such a situation.

3) **Comment # 1 97/03/18 22:18**

Diana:

I think it is best to remove all the "forthcoming" references. It is much nicer to include them when they get published!

next comment on the same document was made: 26 days 11h 28 min. later

Comment # 2 97/04/14 09:46

Oliver:

I agree with Diana. It seems to me as if the text is an old version. Where should I put the corrected version?

The main concern for the participants was that they could not make a change by themselves after reading a comment that suggested that change in the project description. The system did not support users to revise the documents but they could ask the editor to

upload a new version of their document. However, it was in only two cases that users requested the editor to do so.

4.1.3 Discussion and conclusion

In this case study, we found that the users had some problems to accomplish the task. It was not easy to find the entry to the system and also it was not clear what was expected from them. Users with a HCI background were, of course, willing to make comments on the interface of the system and most of the discussion was centred on this issue. Additionally, members of the group knew that we were developing a new system and it was common that they wanted to comment on this issue.

The name of the system “Domain Help System” was also creating confusion among the users. Most of the users asked during the interviews “but why do you call this a help system?” The name of the system was interfering with the accomplishment of the task because users were trying to compare the DHS what for them was a help system. One person said: “I do not even see the question mark button”, probably referring to other help systems known by the user. For administrative reasons, we decided to keep the name.

Before starting this case study, we did not perform any social study to determine whether the users were really willing to get comments on their work, especially in public. We were introducing a new tool. Did the users need this tool? Why is a new channel of information needed? Which are the new benefits, if any? The fact is that getting comments, specially if they are good and fruitful, is something most would wish. On the other hand, to which extent participants in this study were able to make a comment on others’ work was difficult to determine. Users claimed very enthusiastically that they thought it was a great idea to get comments from their colleagues on the ongoing work. However, the result of the interviews shows that commenting on others’ work could be very sensitive. “This might create a lot of problems” one of the users said. The reason given by this user was that it could publicly confront colleagues at work. Another user reported that it would be acceptable to get comments from those colleagues that knew about the subject area of their work but not from those who were ignorant in the subject. The main issue here is the public characteristic that a domain has in DHS- it is accessible to anybody, any time, anywhere who has access to the URL.

Another issue that arose was the lacking possibility for the author of the document to decide when and from whom the comments should be made. Yet, another problem the user confronted was very basic: where can I find the system? We sent the URL via email, so that the users could enter the system, but users reported that it was hard to find the URL in their email-boxes later on. Generally, I observed that communication via email might not be so effective for instruction. Users were somehow ignoring the requests the head of the group sent by email three times. Besides, members of the R-group were co-located and could meet daily. Probably, they did not find it natural to discuss the project description through the system when they had the chance to do it face-to-face.

One finding was that users were making comments on the interface instead of doing the task they were supposed to do. Almost 80% of the words that were registered in this

domain were found under the ATP document. This might suggest that users found it difficult to make a comment about a document (e.g. its content or formatting) using electronic format and the Web. Probably, some people thought that the task was to make comments on the system. However, since the very beginning of the domain creation there were some comments that could be used as model by them to make others comments. The problem of awareness here is also evident. To know whether a document had got a comment, users had to click on its link, so that its comments appeared in the commenting space. The user had to click every document to explore the whole domain. This exploration might be unfruitful.

Furthermore, people who use a computer to write are used to their word processor. Asking them to move an activity as personal as writing from their word processor to the DHS is doubtful of success as long as the DHS does not contain most of the functionalities of their word processor. The DHS offers a very simple editing interface in which only plain text can be entered and it is only given for making comments. The editor of the domain received mainly documents in a ready-to-upload stage. Otherwise the editor produced the HTML version using other means as the “Save as HTML” feature of some word processors.

The discussion that evolved under the ATP document demonstrated that the DHS was appropriate for supporting asynchronous discussions.

From this pilot study we could draw the following conclusions:

- At first, R-group’s members welcomed the idea of using DHS to receive feedback from other colleagues about their work. However, during the interviews many expressed a negative attitude mainly because of the public character of the system and the risk to get comments from people who were not qualified to do so. Some indicated that they wanted to have control on when and by whom their work would be read.
- It might be that asking via email for something is easier to forget, or to avoid, or the commitment is not as strong as in face-to-face communication.
- The DHS worked well for the meta-discussion about the prototype itself. This suggests that this kind of tool might be used for asynchronous discussions within a group.
- An awareness system should be available for asynchronous discussions. Its lack might decrease users’ participation in the discussion and users could respond inopportunistically, that is, they could reply when the information is not needed anymore or too late. This could unfavorably affect the collaboration.

4.2 Study one: the DHS as a discussion tool

The DHS was used during four periods (in 1998, 1999, 2000, and 2001) as a Web-based discussion tool in an educational setting, for a course in computer supported collaborative work. The experiences from this use will be presented below.

4.2.1 Method

Apparatus

Each instance of this case study used a different version of the DHS (see table 1). Instead of a list of topics in the index-frame as we had in the R-group domain, for this case we used the students' names and under each name the title of his/her contribution. The contribution's title was really a hypertext link such that when activated, users could see its content as well as its comments.

Participants

The participants in this group were registered in the course CSCW given to graduate and undergraduate students. The use of the DHS was mandatory. There were 68 participants (including the teacher and the designer of the DHS) in these case studies distributed as follows: 26 participants in D1998, 11 in D1999, 18 in D2000, and 13 in D2001. Most of them had a technical background and most of them had completed at least one course in HCI. All of them had at least two years of experience using the Web and were familiar with other Web-based discussion tools (e.g. USENET news, on-line editions of newspaper, text-chat, etc.) However, very few reported to have had experience with Web-based tools to discuss the content of a document on-line.

Procedure

One of the course assignments included using the DHS, which was introduced as a prototype designed to support discussion on Web documents. Students were first asked to write a text about one page long. They were supposed to tell their experiences while working in a team, focusing on issue of collaboration and technology. The D1998 class could submit their homework assignments in groups of two, but for the other case studies, the task was individual. They submitted their contributions to the domain's editor who included them in the domain. It was required for partial fulfilment of this course that everyone must read all the documents in the domain and enter at least two constructive comments on others' work. The students were supposed to discuss the content of the documents in the domain and not the form of the documents.

The lab was supposed to be done over a period of 15 days. However, the DHS was on-line for a longer period and some students continued to use it. The URL of the site was given in the homepage of the course and sent via email to the students. No password protocol to enter the domain was given. Participants could communicate with the domain's editor via email, phone, or face-to-face if desired.

Students were told that they could make any comment on the DHS system under the ATP document. For every course the ATP was updated, indicating the new features that were included in the new version and all the previous comments attached to it were put aside. The ATP document indicated how to overcome some of the problems that were reported in a previous use of the DHS.

Participants of the D2000 (18) and D2001 (13) had to fill out a Web-based survey at the end of the lab. This survey aimed to evaluate the functionality and the interface of the DHS system. The survey was basically the same for both groups but the latest one included questions about the new features that had been added to the version used by this group. All the students (21) for these courses filled out the survey.

4.2.2 Results from the interaction

Quantitative results

The 68 participants submitted 63 documents and produced 327 comments in these four case studies. All the documents in these domains received at least one comment. A total of 284 comments were made on these documents, averaging 4.5 comments per document (s.d.=1.86). These comments hold 26,796 words. The 43 remaining comments, including 4,224 words, were made on the ATP document. Most of the comments made on the ATP document were an error or bug report on the system, a design suggestion, a new idea by the designers of the system, or a reply from the designer. One point worth noting here is that 58% (25) of the comments made to the ATP were made during the D1998 course. The reason for this, very likely, was that this was the first version used in these studies and it was normal to find bugs, errors, or better solutions for the interface.

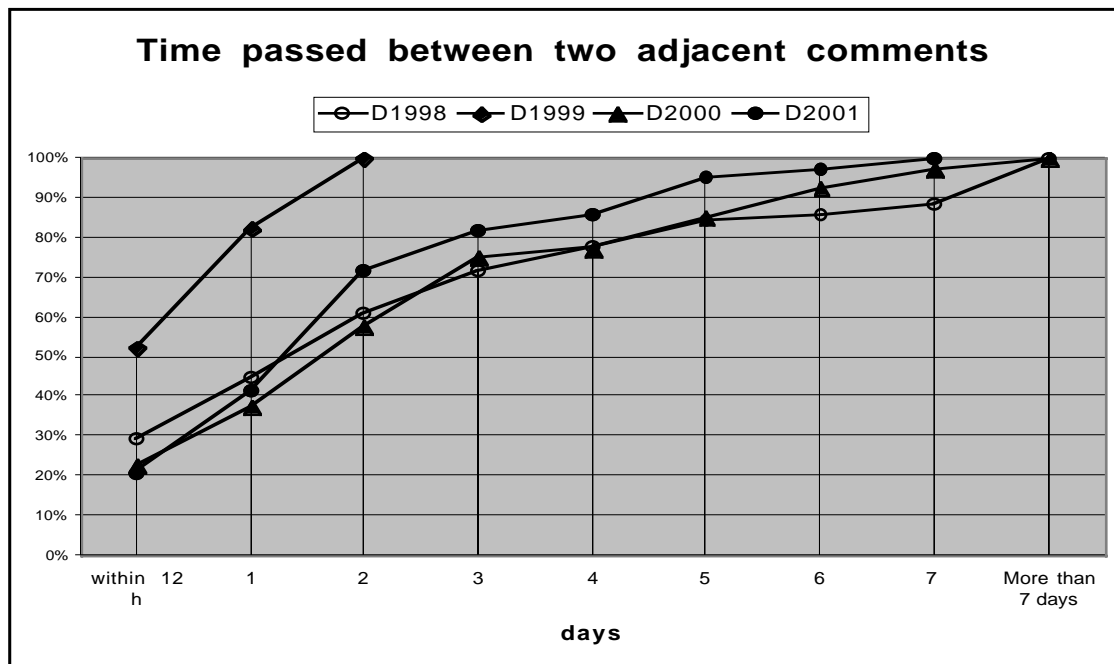


FIGURE 4. Times passed between two adjacent comments for the use of DHS in the discussion context.

I computed the time that passed between two chronologically adjacent comments in the same document ¹. Observe that the next comment is not necessarily a reply to the previous

1. I excluded the ATP from this analysis because I was taking part in that discussion and it was part of my job to monitor it.

one. Notwithstanding, from this data we can appreciate the interactivity of the participants in this system. We cannot draw any conclusions, however, about how the email awareness feature influenced this result because the comment is sent only to the author of the document in question and not to the whole group. The only group that did not have access to the email awareness feature was the D1998. Figure 4 shows that in all these case studies, at least 20% of the comments submitted were made within 12 hours after the previous comment and at least 50% within two days. The graphic shows also that the results for the D1998, D2001, and D2001 are similar in relation to the time passed between two adjacent comments. However, in the sample D1998, 14% of the comments (9) were submitted at least one week after the previous comment. Comments sent after one week are practically absent in the others samples that had the email awareness feature (except for one comment in D2000 that falls in the category more than seven days). Yet, in the D1999 course all the comments were sent within two days time. There are many factors that could explain these differences, for example, the interest that the students had in the course might differ for every sample. I observed that all the comments sent in the D1999 course were submitted four days before the deadline of the lab, so it was normal that all the comments had been sent within two days. Also, the size of the group can be a factor to consider. In D1998, the number of participants were almost twice of that in the other samples and the interactivity is very likely higher than for smaller groups.

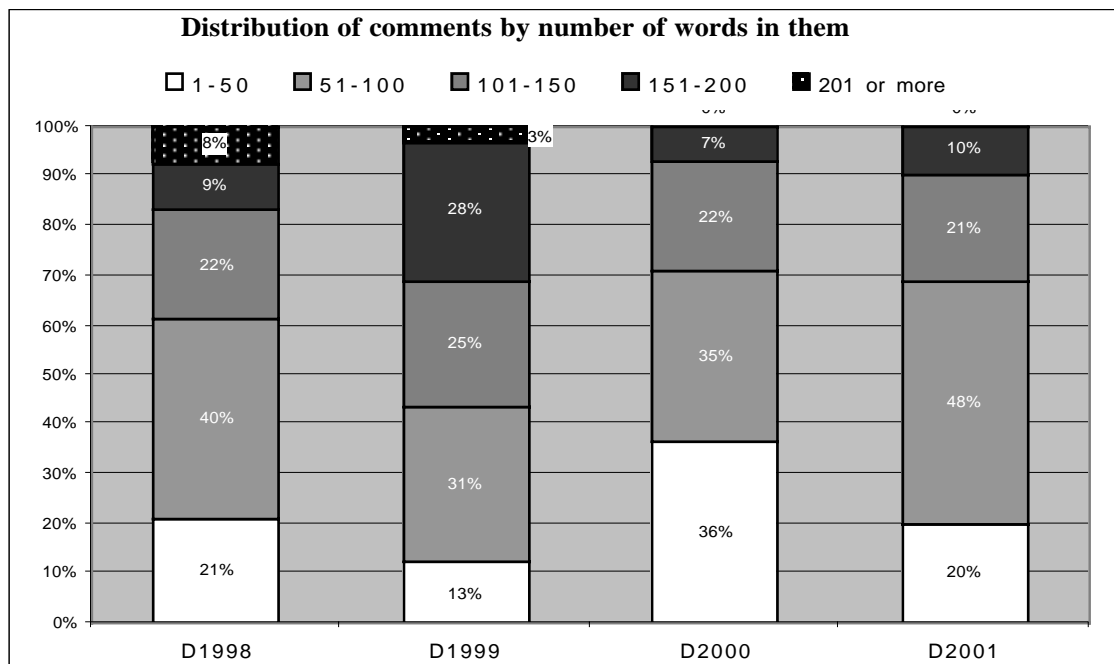


FIGURE 5. Distribution of the comments in relation to the number of words they content

Figure 5 shows the distribution of the comments in relation to the number of words the comment holds. I decided to group them in intervals of 50 words because in the pilot study most of the comments were not longer than 40 words. About 60% of the comments were not longer than 100 words in all of the samples but D1999 that has 44%. We can also see that comments in the interval 51-100 words were the most common ones.

Results from the survey

Twenty-nine students from the D2000 and D2001 filled out a Web-based survey. The questions concerned background information in relation to the use of the Web, issues about sharing their contributions and its advantages and disadvantages, security issues, problems in the system, quality of the awareness given by the DHS, the use of the DHS in the educational field, the use of some specific features of the system, and their attitude toward the DHS.

In relation to having a password or not to enter the system a majority of 55% of the respondents think that it is important to have a password control. The main reasons given for having a password protocol to enter the system were: 1) to prevent unwanted visitors to enter the system, 2) to prevent names usurpation, 3) not to disclose confidential information. One of the respondents said “...*I don't want the whole world to read my draft*”. This was also a result in our pilot study. Many people are very sensitive to the idea of anybody reading their draft. Those who did not welcome the idea to have a password to enter the system said that they did not want to have one more password to remember. They also considered the information they presented in this task not confidential, so no security was needed. However, many of them suggested that it depends mainly on the degree of confidentiality of information that is included in the system. For those who were not in favour of having a password the motivation was mainly because they did not want remember “one more” login and password or the information they provided was not confidential.

As for sharing their documents through the system, the most common advantage reported was the easy accessibility to the documents and its comments. Respondents valued the fact that all the information was in the same “place” and that they could read and comment whenever they got the time to do it. Several respondents compared DHS with email and pointed out that one advantage was that they did not have to look for the documents in their “inbox”, as one respondent said. In relation to email, they also said that it was very worthy to have an overview of the document and its comments at the same time without jumping from one window to another, something they cannot do with email. A few disadvantages were also mentioned. Internet connection requirement, sharing publicly the contribution and its comments without control of the document owner, were mentioned by the respondents as disadvantages.

Many of the respondents indicated that the main problem with the system was related to the structure in which the comments were presented. Some of them, 29%, found it difficult to associate a comment with the part of the document it referred to and 41% of them found it difficult to relate the comment to a previous comment if that was the case. Additionally, they reported that it was difficult to follow the conversation, especially when a reply to a previous comment was not adjacent to it. All those who pointed out this problem proposed threading as a potential solution (see example 4 in the next section).

Most described the experience of getting a comment on their documents as “fun”. Several also expressed that getting a comment made them feel that their contribution was interesting. One respondent wrote: “*At first, when not receiving any comments, it felt a*

little as a failure. As if the contribution was of no interest to anyone.” All but two thought that sending the comments via email to the document owner was a good reminder. As one respondent wrote “*That was good, because otherwise maybe I would have missed it*”. They also said that in this way they could avoid to run the system just for reading a comment because they had got it in the email. Several expressed the need to alert not only the document owner but also those who previously made a comment on the document in question. However, they said that it could be annoying to receive a lots of emails if the interactivity was high. A few of the respondents also reported that getting the comment via email sometimes did not make any sense (for example, when a comment was a reply to a previous comment that was not made by the document owner). In fact, the document owner is then getting a reply to something he/she has not said.

As for awareness, one question was to find out how much users knew about the changes that the content of the system had gone through since the last time they used the system, both for the system content as a whole and for those contributions in which they were particularly interested. More than 60% thought that the general awareness about what has happened to the whole content of the system was poor. Only one respondent thought it was good and the rest regarded it as enough. On the other hand, the awareness about the contributions in which they were particularly interested was rated as poor for only 38% of the respondents. Only three respondents thought it was good and the rest regarded it as enough. In other words, the system provides a better support for awareness of those contributions in which users are particularly interested. The reason for this might be that users, very likely, are interested in their own contributions and every time a comment is made on it they receive an email message.

At the end of every comment the system provides a navigation bar “First-Previous-Next-Latest”. Only one respondent reported to use it often, 15% seldom, and more than 80% never made use of this navigation bar. Instead they just scrolled the comment frame. They were also asked if they felt the need to print a contribution or its comments and if they had printed it. Only 24% indicated that they had felt the need to print a contribution and about 7% had felt the need to print a comment. None of them printed the comments and only two respondents said they have (seldom) printed a contribution.

Most of the participants reported to have good but not excellent knowledge of HTML. Also, all of them reported to know that their word processor could automatically convert text into HTML format, though seven of them had never used this feature. However, only six participants used HTML tags in their comments and only 28 of the 327 comments included HTML tags. Furthermore, these HTML tags were used for simple formatting purposes as presenting one word in italics or bold font. If users wanted to include an HTML tag in their comments, they had to do it in plain text. Another possibility to do this is to write their comment in a word processor, convert this formatted text into HTML format using the automatic converting function it provides and copy-paste its source into the DHS. Clearly, it is too troublesome to perform.

In D2001, the paste feature was provided in the Add comment window. None of the respondents in this group made use of this feature.

4.2.3 Discussion and conclusion

A few discussions, but not very interactive ones, turned out. Some participants indicated this situation; one respondent said “...*But it never became any real discussions, just comments.*” A discussion is an opportunity for everyone to defend, explain, and modify their viewpoints. It should be a way to negotiate positions. A comment could hold just a position that concedes defeat, probably the user referred to this kind of comment. We also have to consider the nature of the task. The students submitted a story telling an experience they had while working in a team focusing on collaboration. They were just indicating the points in which the collaboration went wrong or right. This situation is not very suitable to generate a discussion as no negotiation is needed. Most of the comments were asking for more details of the story or were just telling another story similar to the one they commented. We have learned from this that we should expose students to controversial issues that ensue in a real discussion, for example, a topic in which they are forced to show their personal opinion. The most active discussion was around the ATP document in the case study D1998. One reason was because students were giving their opinion about the system and I was trying to justify or motivate the system features. That is we were negotiating our positions on the system.

In the commenting dialogue of DHS, comments are chronologically ordered, and no subject line is used when creating a new comment. Users have indicated several times that threading would be necessary in the comment space of the DHS. Users manifested this in the survey and also using the commenting function of the system on the ATP document. Example 4 is a comment made on the ATP document in the D1998 sample.

4) **Comment #23 98/02/20 14:21**

Enrique:

I think threading of comments will be important as their number [of comments] grows.

Although this suggestion recurred several times, we decided that threading was not suitable for our goal with the system. Threading refers to arranging the sequence of contributions to a computer-mediated discussion according to their subject, by using the “reply-to” relationship as ordering principle. In newsgroup readers, for example, users can select to read a particular thread, and the threading of discussions is emphasised by graphical means such as indentation. Threading requires an initiating post with a subject or topic that users consider a representative one. In our case, the subject of the discussion was given by the Web-document. By creating another subject the discussion deviates from the original purpose. Another reason was that we were not expecting a large number of comments for each Web-document.

One possible solution would be to keep the linear structure but with the possibility to place a new comment adjacent to the comment it replies to. This, however, implies other problems, for example: when a new comment refers to two previous comments, should it be placed adjacent to the first comment it refers to or adjacent to the second?

Whittaker, Terveen, Hill & Cherny (1998) found in a quantitative study on Usenet newsgroups discussions, that messages on average referred to two previous messages, which indicates a lower amount of interactivity than might be expected (p. 262). This could mean that for a commenting space with a moderate number of comments, say ten,

threading would not be necessary. In the DHS case studies, users were often using the identification number to make reference to a previous comment. We also observed that if a response to a comment was adjacent to it, the author, usually did not bother to make a reference to it, which increased the sense of a written dialogue.

An important aspect was that during the discussion, most of the time participants made an explicit and direct reference to the content of the document in discussion. For example, “what you wrote”, “the situation you describe”, “the people in the examples”, “I found your text to be interesting”. That means that the conversation was around the document presented in the content-frame, namely, the document was the centre of the ongoing discussion. In other Web-based forum tools, like newsgroups for example, the original text that started the discussion is not easily reached by users nor on focus and the discussion often gets far away from the original topic suggested in the first text. The tendency to lose the initial focus gets more pronounced when time passes. In DHS, we found that the central topic of the discussion, in this case, the content of the document, was always recalled in the comments. Even if the new comment was made a long time after the very first comment; the new comment was, somehow, touching the document that originated the discussion. This is, most likely, related to the fact that users can always view and read the original text while using DHS, and the discussions are not threaded. Also in our system, the document generally has a more important role than the start message of a newsgroup discussion.

Finally, in DHS users may refer in the same comment to another comment and to the Web-document, which means that this is a dual context for the commenting dialogue. The reference to previous comment was usually made by its identification number. Example 5 shows such a situation. Paola makes reference to comment #1 and gives her opinion. Immediately after this, she mentions “the first example” but this time she refers to the Web-document around which the discussion is held. This would not be so easy to understand in a threading environment, which rather emphasizes the coherence relationships of the dialogue of messages. Moreover, in a threading discussion, participants cannot see at the same time the conversation as a whole and the message that originated the discussion. The lack of an overview of the replies and the original message might cause the next contributor to concentrate more on the last read reply than in the original message. As a result, the focus of the discussion may be lost.

5) **Comment # 3 98/02/19 12-42**

Paola:

Javier,

comment #1

You wrote about what you can "gain" from collaborating. You didn't mention how much does it "cost" to you!

In the first example, what kind of big effort could it be to say to the one who sits next to you "feet up, the cleaner is coming"?

4.3 Case study three: the DHS as an annotation system (A-samples)

The DHS was used as a Web-based document annotation tool in an educational setting. This domain was created for a course that aims to teach academic writing style in English to graduate students (non-native English speakers). The DHS has been used four times in

this course since 1997, each corresponding to a case study here. It is important to point out that in these cases the students were supposed to discuss the form of presentation of the documents in the domain and not the content of the documents. There were thus four case studies: A1997, A1998, A1999, and A2000. Each took place in a different year. The use of the DHS was not mandatory for these courses. However, all the students agreed to use the system.

These studies were done because we were interested to see how users made annotations to web-documents using an editing interface that offered only plain text. Also, we were interested in designing a tool that could support some of the activities in a writing course as sharing material, communication among members, indicating improvement of the text, and sharing examples. In this writing course it was very likely that users were going to make annotations to suggest a change in the text; e.g. wording, or indicate an error, e.g. misspelling. Generally, to indicate these changes, it is very common to use proofreader's marks or at least writing comments in the margin. They are simply symbols adopted for expediency, to save writing out instructions that take time and space. Proofreader's marks vary somewhat, but all sets of marks have the same general intent. These marks were conceived bearing in mind that they would be handmade, on paper, and by ink (or carbon). HTML documents pose several problems to use proofreader's marks. The first problem is that these documents lack a margin. The second problem is that proofreader's marks are too complex to be represented in digital format. Moreover, they are not available on the Web.

We are fully aware that the DHS provides a simple interface for making annotations which very likely will not entirely satisfy users' needs. The data collected, annotations made by third party reviewers in digital plain text, could be considered as "raw data". This is because the system does not offer any mechanism to incorporate topographic elements so users have to create their own resources to create a mutual understanding between the author and the reviewer. This raw data can be used to identify the needs reviewers have to communicate with the authors while making their annotations.

The evolution of the tool has been to some extent guided by participants. On the other hand, the teacher's influence was minimised. The teacher did not give any guidance on how to write a comment using this tool. The importance of this simple interface is that it did not offer any sophisticated option for helping the students to make their comments. HTML tags could be used, though very few students reported having good knowledge of it. Those students made use of it, but only sparingly, probably because including HTML in plain text was tedious and time consuming. Most of the participants mentioned in the survey that they knew about the automatic HTML conversion feature in their word processors. Using this feature they could have written their comments using topographic elements but they had to send the source of the HTML file and it was probably too much work or they simply did not know how to do it. The only alternative available to them was to write in plain text.

This section was written in collaboration with Sandra Brunsberg, who was the teacher of the course.

4.3.1 Method

The students used the DHS to share their homework and to make annotations on other students' homework. In total, 287 annotations made by 42 Ph.D. students and two teachers were recorded by the DHS. The intention of this study was to observe the strategies and patterns that the reviewers used to communicate suggested changes to the author, and to develop the DHS taking into account the writer/reviewer needs. As in the comments in the D-sample, the annotations could also include HTML tags. No support to edit the HTML tags was given, so if the user wanted to include them in their annotation, he/she had to introduce them in plain text. After each case study, using the users' responses and the observation, a new version of the DHS was released. In table 3, we show the features of the DHS for every case study. The table provides information about the number of comments and words produced in every case study by students and the teachers.

Table 3. The case studies in which the DHS was used as an annotation tool and their characteristics

Case Study	Participants	Comments made		Words in comment		Relevant Features of the system				
		Teachers (%)	Students (%)	Teachers (%)	Students (%)	Email awareness	Log file	Aligned layout	Comment counter	Paste-text function
A1997	15	19 (26%)	55 (74%)	5094 (30%)	11625 (70%)	No	No	No	No	No
A1998	13	16 (20%)	63 (80%)	2377 (26%)	6850 (74%)	Yes	No	No	No	No
A1999	4	10 (91%)	3 (9%)	1913 (84%)	351 (16%)	Yes	Yes	Yes	No	No
A2000	16	29 (24%)	94 (76%)	5765 (32%)	12208 (68%)	Yes	Yes	Yes	Yes	Yes
Total	48	74 (26%)	213 (74%)	15149 (33%)	31034 (77%)					

When the DHS did not have the “paste” feature, the users had to copy-paste the text of interest using the mouse if they wanted to quote from the original text.

Eight members of the A1997 course were interviewed (15-20 min.). The interviews were done in their work place from where they use to run the system. During the interviews, I asked them to run the system, so that we could discuss about the interface “in live”. Additionally, I was part of that course and when we met in the class, I used to ask them if they had any problem with the system. For the A1998 and A2000 a web-based survey was carried out with two and eleven responses respectively. The teacher of the course was interviewed after each case study. Both the survey and the interviews aimed to assess participant's knowledge about a few aspects of the Web, their views on the DHS system's features and the impact it had on the course. In 1999 the DHS was introduced two weeks after the start of the course and it was hardly used at all. Therefore no further studies on this group were made.

Participants

The students that participated, 14-16 for each group, were non-native English speakers. They were Ph.D. students that had different engineering and scientific backgrounds, e.g. chemistry, metallurgy, and physics. The teacher of the course also took part in the study and wrote several annotations. An outside teacher took part as a reviewer on a voluntary basis in A1999 and A2000. She was an English teacher as well and was located in Boston, USA during the course. Both teachers were native English speakers. All the participants had access to the Internet and students reported that it was the first time they had taken a writing course. Participants reported that they had little experience in using the Web as a tool for supporting course activities though all of them had used the Web for at least two years. They reported that the Web had so far mainly been a means to retrieve information, e.g. the homepage of the course. All participants had used email to communicate with the teacher and, very infrequently, with classmates in other courses.

Apparatus

The DHS was the Web-based tool we used to collect the annotations made by reviewers. In our case studies, the index-frame displays the student's name and his/ her assignments. Each assignment is a hypertext link. The DHS has evolved through the case studies that we report here. See table 3 (above) for details.

Procedure

Each course had a duration of 10 weeks and the class met once a week for three academic hours. During each class, a new topic was discussed, and homework was assigned to the students. The homework normally involved writing a short text, about one page long, using the techniques and writing styles that were discussed during the class. The class was divided into groups of three students, who reviewed each other's texts, so that each student's work was read by two others. Each student sent their homework to me via email, I made an HTML version of it, and included it in the DHS system. Once the homework was in the DHS, the students had to read and make annotations on other students' homework using the DHS. This was done during the week before the next class was given, so that the authors could read the annotations made on their text. In the next class they were supposed to discuss the annotations face-to-face with the students who played the role of reviewers.

We included one document in the DHS describing some conventions developed informally by me to provide a model that could help the students make corrective annotations using digital plain text. A corrective annotation refers to those annotations that intend to represent a change on the text suggested by the reviewer to the author. These conventions were limited by the impossibility to use traditional proofreader's marks that are based in graphic symbols. It should, nevertheless, be borne in mind that the conventions that were given to the students might not be the most fitting ones. However, the teacher of the course and I considered important to give student a model so that they could start to make annotations. The conventions I suggested (see table 4 for one example) were used by most of the participants but several of them developed their own method for making annotations

or they adapted some of the conventions to their personal preference. For example, some students used only the parenthesis without the asterisk. This document also included an explanation of how they could copy-paste the original text into the text input area (the input field in HTML used to enter the annotation in DHS), and eliminate unnecessary line breaks that the cut-paste originated, see section “Total copy of the original text”.

Table 4. One example of the marking suggestions given to participants in the DHS

Case and syntax	Original	Your correction	Revised
To report a misspelling, grammar error or opinion use (* <i>text</i> *)	it has help me	a) it has help (*helped*) me b) it has help(*ed*) me	it has helped me
	Two of them (pi 0; pi 5,3) were separated	Two of them (pi 0; pi 5,3) were separated (*Why do you use were in this case *)	
	in Latin America Spanish is spoken. Despite the powerful gift that history has given	in Latin America Spanish is spoken. Despite the powerful gift (*does this refer to the language. if so, how is relate to the rest of the sentence*) that history has given	

I have analysed all the annotations that were entered by the students in these four years to find a common pattern. Additionally, I developed a simple parsing program that counts words or phrases in the text of the annotations recorded in DHS. Before counting the phrases or words, we corrected some obvious misspellings that were entered in the annotations. For example, sentance, sentece, sentencs were changed to sentence; firts, firsat to first; etc. This was done because we wanted to know how many times a specific phrase was used, for example “This sentence”. In the following section we will present and discuss the findings of this analysis.

4.3.2 Results and Discussion

The annotations entered by the students in the DHS are the basis of this study. We will present portions of some selected annotations to highlight some observations. The text that is written in italics is taken literally from the reviewers' annotations in the DHS.

Strategies used by reviewers to write their annotations

In general there were three ways in which the reviewers made their annotations: total copy of the original text, partial copy of the original text, and straight annotations. The first two ways in one way or another use quoting to preserve the context. Severinson-Eklundh (1998) has approached quotation as a “pragmatic and interactive device frequently used both in email and newsgroup corpus.” In this context of annotations in the case studies, quoting was used mainly as a structuring mechanism.

Total copy of the original text

One strategy was to copy the entire original text and work on it. This was observed in all the case studies, even those in which the “paste” function of the system was not available.

Copying text from a Web browser window usually has the inconvenience of adding an extra line break at the end of each line relative to the browser window's width. The DHS conserves the line breaks that the user might enter. So the extra line breaks added by the reviewer when he or she performed a copy-paste from the browser window were also conserved. When these extra line breaks were not deleted the DHS could show the annotation in a strange layout. An example is shown in figure 6.

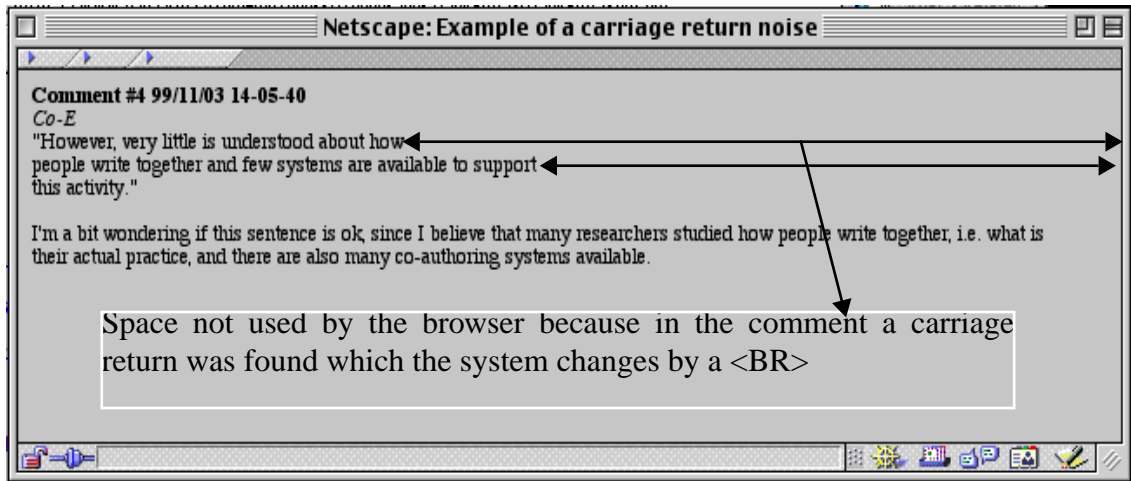


FIGURE 6. Example of how the system could present a copy-pasted text from the original braking the flow and bad using the space.

Despite the browser window being wide enough to show a sentence as long as the width of the browser, the DHS breaks the line where a line break (added by the copy-paste action) is found and does not use the whole space of the browser window. To solve this problem, instructions on how to eliminate extra line breaks were included in the DHS in the same documents in which the conventions about how to write digital annotations were given. Eliminating the extra line breaks required a great deal of effort on the part of the reviewers. When they used this strategy they made interlinear annotations. This topic will be discussed in chapter IV, section 5.2.1.

Partial copy of the original text

Reviewers did not always copy-paste the whole original text. Sometimes they copied only the portion they wanted to annotate, for example, a sentence or part of it. We do not know exactly how this was done. They could have copy-pasted the interesting portion or they could have typed it. In these cases the correction context was given by the copied portion and the reviewer did not mention to which part of the text his/her revision was related to. The annotations were put in the next line after the text (example 6), aligned with the text (example 7), or were interlinear (example 8). When this strategy was used, we found no cases in which the annotation was made before the copied text. Below are three examples of partial copy.

For the readers, it is also difficult to distinguish between the original content and the comment. Different conventions used by different reviewers make this even more difficult.

Some support is given, however, as the system always shows the original text with no changes in one window (in the content-frame) and in the other the comments (in the comment-frame). For the authors of the commented text, however, distinguishing between the original content and the comment was easier.

- 6) *whether genetically modified (GM) food is hazardous to human health or not.
/you could leave out “or not”, gives a better flow I think but it's not necessary/*
- 7) *The process are the repeated >> The process is then repeated*
- 8) *These (*+factors*) include(*+:*)flow velocity of mobile phase,...*

Straight annotations

Reviewers who made no copy at all of the original text clearly felt the need to clarify the correction context. As can be seen in example 9, the reviewer used the ordinal position of the paragraph and the sentence to recreate the correction context.

- 9) *Paragraph 2:
You give some “solutions” here. Could you make sentence 2 clearer by starting with something like “One way is by using...”*

Defining the change representation

In many cases, reviewers explained the notation of their annotations and their meaning as we can see in examples 10 and 11:

- 10) *Mathematical signs were used:
+ means adding a word and
- means take a way a word*
- 11) *Color was used:
What is in RED means to delete
What is in BLUE means to add what is in brackets [] or my annotation
What is in Violet means that I am not sure if that should be used or could be replaced by something else.*

The words RED, BLUE, and Violet were coloured with the colour they described. The reviewer in this case used HTML tags to achieve the colour effect.

As we can see here, reviewers used signs or colours to represent the actions add and delete which are the most common ones used for corrective annotations. In example 11, the reviewer decided to assign a colour for those cases in which he or she was not sure whether his/her annotation was valid/right. The most important point here seems to be that the reviewer is trying to reach a common understanding with the author. The reviewer tries to guide the author on how to read his/her annotations. This explanation was found at the very beginning of the corrective annotation sent by the reviewer.

Some writing programs support the communication about the changes of the document among collaborators. In a recent study on on-line reviewing with change representation tools, Kim (2000) found that reviewers felt that a commenting function should be used to complement the limitations of suggestions of suggestions by the change representation alone. That is, it might not be enough just to suggest a change but to motivate it.

This situation was found: one of the student, Carlos, acting as a reviewer made his first corrective annotation to Juan. Carlos used his own change representation (e.g. colour or signs meaning), and explained to Juan how to interpret the annotations. For the next homework it is Carlos who has to revise for the very first time Pedro's work. Carlos assumed that Pedro had read the explanation of his own change representation (made to Juan), therefore did not bother to explain the change representation to Pedro. This is very likely because of the use of the shared space that the DHS offers.

Need to represent corrective annotation and to copy original text

For corrective annotations it is very important that the reviewer has access to the original text and is also able to copy and paste it. This will save time typing (or copying-and-pasting) the text that will be annotated. In the next example, the reviewer quotes the sentence he/she wants to refer to and then makes the annotations:

- 12) *“But the most important thing is to change the mentality”. I feel this sentence is in the wrong position as the word “but” puts it in the opposite of the previous sentence.*

It might be important to support the common actions for corrective annotations, namely: add, delete, change, and change the place of a word to avoid extra effort by the reviewer (see example 13).

- 13) *in a sandwich structure (I would suggest you add a comma here, and possibly also “however;”) the faces take the place*

In this case, the reviewer had to type 65 characters to indicate the changes. Authors and reviewers should agree on the change representation in the text or the system should support it. One way would be to include in the system information about these annotation conventions as we did in these cases studies using the DHS. Using, for example, the annotation conventions that we have talked about before in the Procedure part of the section 4.3.1, the same reviewer's annotation would look like:

- 14) *in a sandwich structure (~however;) the faces take the place*

that is, 53 fewer characters. Unfortunately, reviewers and authors have to learn the conventions, an effort which would be justified for long revisions. However, this solution could have a secondary effect, especially on the author. It does not encourage a dialogue between author and the reviewer because it is impersonal. For example, example 14 does not involve any subject. Observe that in example 13 the pronoun *I* is used. Many readers would agree that example 14 feels more remote and cold than example 13. On the other hand, a corrective annotation in which the reviewer addresses the author as is done in example 13, in a more personal and friendly manner, might result in better communication between the reviewer and the co-author, but this topic is outside the scope of this thesis.

We observed that some reviewers abided by the convention that we suggested and produced annotations that were hard to read. and to produce as well. Example 15 is the whole annotation that one of the students submitted to the system.

Example 15 also shows the amount of effort the students were willing to put into writing the annotations in the DHS. Counting the characters used to follow the conventions we

suggested for making annotations: (,), *, =, ~, :, +, -, we found that the reviewer made 182 keystrokes just to type these marking annotations. If we separate the annotations that were written by the reviewer from the original text we find that the minimum number of keystrokes necessary to make the annotations is 604. In other words, little bit more than 30% of the keystrokes were made just to produce the characters that the convention suggested.

- 15) *Figure 9 shows the totals of Canadian and Japanese students in U.S. colleges and universities (*universities*) from (*from*) 1955 to 1990. As can be seen, both(*-both*) the number of (*+ Japanese and Canadian*) students (*-from Japan and the number from Canada*) seem to have an upward trend although the =Canadian student number= (*:=number of Canadian students*) had a steep (*+of*) fall in 1975. =The student came from Japanese was about 1,000 in 1955=. (*:= The number of Japanese students was about 1.000 in 1955=*). This number =remained a low steady increasing rate= (*:=steady increased*) until 1970, when (*~and as result of this*), there was (*~were*) about 3,800 student(*s*), (*~;* after (*- after*) then it rose at an accelerating speed. By (*~In*) 1990, the number of which (*What do you mean by-"the number of which", Japanese or Canadian students? It's difficult to understand.*) had reached 28,000. In contrast, the growth of Canadian student (*s*) was much moderate. Despite the fact that 7,500 student (*s*) already (*-already*) studied in U.S. in 1965, =it only grewed to about= (*:=the number of students has reached*) 9,000 in 1965. There was (*~were*) a obvious upsurge in 1970, (~following*) a (*-a*) steep (*+of*)fall in 1975, and then a sharp rise which expanded the number (*+of students*) to 18,000 in 1980. After this dramatic (*speling*) change, the growth rate was similar with that of between 1955 and 1965. In fact, there were about 22,000 canadian (*Canadian*) students in 1990, which (*~and this number*) was on 6,000 less than (*+ number of*) Japanese student (*s*). It might be possible that the number of Japanese students keeps increasing at the average growth speed (*-growth / speed*) from 1955 to 1990, which is about 770 increment every year. In the meanwhile, the number of the (*-the*) canadian (*Canadian*) students might remain (*+at*) a very (*-a very*) slow growth, most probably, (*-most probably, *) with a increment of 260 every year.*

In the following example, the reviewer writes the original text first and his or her corrective annotation in the following line. This makes it simpler for the authors to compare what they have written with what the reviewer suggests.

- 16) *in most bigger towns in sweden we have something called "tidnings insamling".
in most big towns in Sweden we have a system for newspaper collection called "tidningsinsamling".*

Use of parenthesis to highlight annotations

Parenthesis was the most commonly used sign reviewers utilised to make a distinction between their annotations and the original text. However, the use of parenthesis is not a good procedure because it could be used in the original text as well and might be confusing. The following example is a case in point:

- 17) *(But) Unlike to (omit) the concentrated narrow Web in an I-beam*

In this case, the reviewer suggests some changes but it is not clear which ones. The sentence could be understood in different ways. Below, words that are struck-through

mean “to be deleted” e.g. ~~word~~, and underlined words mean that they should be added to the text e.g. word.

Table 5. Four possible interpretations of example 17

	Potential interpretation	It would read
a	But Unlike to omit the concentrated narrow Web in an I-beam	But to omit the concentrated narrow Web in an I-beam
b	But Unlike to the concentrated narrow Web in an I-beam	But the concentrated narrow Web in an I-beam
c	But Unlike to the concentrated narrow Web in an I-beam	But to concentrated narrow Web in an I-beam
d	But unlike to omit the concentrated narrow Web in an I-beam	But unlike to omit the concentrated narrow Web in an I-beam

The reader might interpret the word “omit” as a new word to include and not as an action, as is shown in case a) of table 4. On the other hand, if the reader interprets the word “omit” as an action, it is not clear on which word the action should be taken: it might be on the word “to” as shown in case b), or it might be the word “the”, a common error for speakers of languages that lack the article, as shown in case c).

Reviewers need to express uncertainty

The question mark (?) was the second most commonly used sign by reviewers in their annotations. The question mark was used in the following format: **word?**, see example 18. In this way the reviewer showed uncertainty about the understanding of the context and whether the writer should accept the remark as valid; that is, the reviewer is inviting the writer to reflect on the context based on the word marked with the question mark. Here are some examples:

- 18) *During (in?) the sixties Sweden was in*
 19) *between the word's appearance and its meaning, which is the (a?) typical characteristics (remove) of the Chinese language*

Sometimes, however, reviewers explicitly mentioned some of the changes with words like “perhaps”, “maybe”, “I would rather”. In this case the reviewer suggested that the author, in a very subtle way, accept his/her annotation.

- 20) *Actually (*maybe you don't need to use it*), virtual crashes are*
 21) *I don't know if it's right or wrong, but I think it sounds better to say “suitable for mapping, monitoring...”*

In some cases, the question mark was repeated more than once in a row (e.g. ???), which suggests that probably the reviewers need to express the extent of their doubts.

The primary reason for this could be that all students, in our case studies, were non-native speakers of English. Second, as Ph.D. students are highly specialised i.e. it is hard to write authoritatively about other's work if not in a closely related area (see example 21). A third interpretation could be that questions are regarded as more like suggestions for change, rather than instructions (see example 18). They are also typical features of interpersonal

communication rather than transactional and could serve as a means of establishing a dialogue between writer and reviewer. Furthermore, they signal respect for the writer and acknowledge that there might be other possible interpretations of the text. Keh, cited by Smith (1997), found that her students considered the question mark a very helpful form for commenting because of its interactive nature.

Global annotations appended

Not surprisingly, general annotations were mainly given at the end of the text that the students had reviewed. The natural sequence in which this was done (first read and then comment) is probably the main reason for this behaviour. An interesting observation is that the reviewers clearly labelled these annotations with such words as “General annotations”. When the reviewer did not label them, blank lines (two-three lines) were left between the original text and the annotation so that readers could identify the annotation as special. Example number 22 was appended after the original text.

- 22) (**General: Clear and easy to understand. Would it be of interest to annotation the shape of the decline in spontaneous polarization with increasing temperature**)

General corrective annotations related to a specific paragraph were also written immediately after it and were delimited by parenthesis.

Justifying the corrective annotation

In many cases, the reviewers mentioned not only the error, but also a solution or the rule to be applied.

- 23) *Sides come in contact with English in several ways, (I would use colon here instead. You are introducing a series) through movies, TV (here must be a comma,) and recently also over the internet.*

In the next example, the reviewer indicates the action and then supports the suggestion made given the rule that concerns to the case. In this example, the action to be taken is represented and explicitly indicates the rule that supports the correction:

- 24) *Some examples of such words used in Swedish are “site”, “mail” (+,) and “freestyle”. Rule: Use a comma before and, or, nor in a series.*

Helping the author to find the correction context

Reviewers use invisible location marks to help the writer enter the context. They used the spatial context to make reference to their annotations. It was very common to find annotations in which the spatial context was important in understanding the annotation. Some of the situations found are given in table 6. From this table, we can also see that in 12 of the different fragments found the reviewer uses the word “sentence” while only five times the word “paragraph.”

Table 6. Fragments used to relate the annotation to the original text

Fragment used	Found (times)
This sentence	33
The last sentence	29
The first sentence	19
First paragraph	14
Next sentence	11
Last paragraph	9
The second sentence	6
The final sentence	4
Opening sentence	4
Previous sentence	4
Next paragraph	4
Following sentence	2
Opening paragraph	2
This paragraph	2
The same sentence before and after the table	1
Before the sentence	1
A few sentences later	1

Relative Position

One common strategy was to use ordinal position to locate annotations. Expressions such as “the last”, “the beginning”, and “the opening” were adjectives commonly used by reviewers.

- 25) *The last paragraph needs rewriting*
- 26) *The first sentence in the last paragraph I would write*
- 27) *in the middle of the paragraph you start three sentences with “The”*

Specific position

Being unable to point to the text or to link an annotation to a specific part of the text during the revision, reviewers help the reader to locate the annotation-related text by enumerating the sentences or paragraph.

- 28) *Your sentences are too long. Paragraph 2 and 3 are one sentence each, containing 3 resp. 5 lines!!!*
- 29) *You introduce sentence two by “in other words”*
- 30) *in sentence 2, your verb describes you could probably combine it with sentence 3*
- 31) *the first 6 sentences in paragraph 3 with references to*

Topic related location

Another way to locate the sentence was to mention what it was about: its content or theme. Reviewers usually wrote the words “the sentence about”, but in some cases they just use the structure sentence-topic (see example 34). For example:

- 32) *in the sentence about diffusion I have doubts*
- 33) *The sentence about that method could be left out*
- 34) *I would like to change the sentence about how enthalpy*
- 35) *The only sentence I don't like is the one with the explanation to why?*
- 36) *Then you continue with a sentence about knowledge*

Quoting the beginning of the sentence

We observed that reviewers identified a whole sentence just by quoting its beginning. In none of the cases did we find that the number of words copied by the reviewer was less than five.

- 37) *the sentence starting with The increasing sensitivity of NMR.*
- 38) *Might I suggest a minor change to the sentence starting /Due to the fact that. / Can you reduce this*
- 39) *Finally, the sentence:
> The mechanical pulps are of course bleached in order to
doesn't sound very formal.*
- 40) *Just ignore the sentence “The scientific community is clearly”*

Positive comments and their position

A teacher can make “comments to motivate, explain, or chastise her students” (Smith, 1997). In our study of the comments made by the teacher or the outsider reviewer 84%, 86% and 100% made a positive evaluation (in 1997, 1998, 2000 respectively). Similar results were presented by Smith (1997) who made studies on end comments written by teachers on papers produced by students in a rhetoric course. In her study more than four out of five teacher evaluations of the entire paper are positive. She wrote (p. 253) that teachers may be reluctant to write a negative global evaluation because they might consider it more important to pinpoint failings which can be corrected or to protect the self-confidence of the students.

In our case studies, we found that students, though not so often as the teacher, included positive sentences in their comments. In total 51%, 34%, 70% (in 1997, 1998, 2000 respectively) of the students' comments encourage the author's work. This behaviour might also have been influenced by the public nature of the DHS.

We also observed the position in which these positive sentences were found in the comment. Here, we find a significant change over the years. In the case study in 1997, 93% (15 of 16) of the teacher's comments that had a positive sentence were placed at the end of the comment. Only 7% were at the beginning. In contrast, the same teacher changed this pattern in the case study in 2000, in which 92% (27 of 29) were placed at the

start of the comment (8% and the end). Students behaved in a similar way. In 1997 these positive comments were placed 50% at the start and 50% at the end. On the other hand, in the case study in 2000, students placed 92% (53 of 57) of these comments at the start. This contrast might be explained by the nature of the system. In 1997 the DHS did not have the feature “paste” which was present only in the case study 2000. It is very likely that users, who had already the intention to make an annotation, the first action he/she made was to use the “paste” button and *then* write his/her comment.

Few comments, three of 287, presented a positive sentence both at the beginning, and at the end of the comment. In one case one of the students sent a comment just to say “*I forgot to mention that I found the text well structured and easy to understand*”

On the other hand, we observed that the only person who sent greetings (e.g. hi, hello) was the outside teacher who did so in every comment sent. This, of course, is related to the personal manner of every individual. However, one reason could be that this was the only person who never met the students face to face.

4.3.3 Student perceptions of the DHS

All the participants reported that it was the first time they had taken a course in writing and that they had used the Web for at least two years though their knowledge about HTML was never more than basic. All respondents but one had not used a Web-based tool in a course. The use of the Internet was limited to searching for information and to email. None of them thought it was difficult to submit a comment using the DHS interface.

The common space that DHS supported for sharing the homework was extensively used by the students. Most reported that they used to read others' texts. This information is supported by the log file which shows that they were reading not only those contributions they were supposed to comment on but also others. The main reason, reported by most of the respondents, was that they wanted to see how the others wrote and made comments and in this way compare or have a pattern for doing the homework. Another reason was that they wanted to know whether they had “*understood the assignment correctly*”.

Once a comment was submitted, it was sent via email to the author of the homework (this feature was, however, not available in the 1997 version). Respondents appreciated this feature, saying that it allowed them to save time as they did not have to visit the system to check if a new comment had arrived on their homework. They also reported that having the homework on line saved time because they did not have to “*run after*” other students (as one student put it) and everything was in the same place.

As for getting comments, respondents liked when this happened; several of them described it as “*fun*” to receive comments on their writing. One student even wrote about this question “*I was glad*”. In contrast, they reported that it was a waste of time writing a homework assignment and getting no comment on it. Students from the 1998 and 2000 case studies reported that the comments they got helped them in their academic writing. However, in the interviews made with the group of 1997, students felt that the tool had not helped them to “*learn more*” about academic writing itself but it had helped them take a

more active part in the course. Sharing the homework assignments had been easy and fast, and they had had access to others' text, that is, to many examples. Going through a great number of examples is very important in any learning process.

On the other hand, students could observe the comments made by the teacher or outside teacher made and to whom. One student complained openly indicating that comments from experienced people were missing and that so far neither the teacher nor the outside reviewer had commented on his/her text. However, others mentioned that the lack of comments could be a question of time; as one student stated "When I don't receive any comment it must be due to lack of time."

All of the respondents thought the DHS was a good tool and all of them encouraged the work and the use of such a tool in an educational environment. As one respondent expressed it, "*I felt more motivated to study*". However, the response was not wholly positive. Some experienced a conflict; they felt that the texts were being discussed twice. Others seemed to take the discussion a step further in the classroom. It is important to point out that in this case study of the DHS very few comments show evidence of a dialogue. Only five times did students react to a suggestion by trying to explain themselves or by posting a new version. This was probably because that was not part of the homework and because they met on a regular basis in the classroom. Also if the content of the annotation was not such that a new version was required, users might not see the need to send a new version, for example if the annotation referred only to misspelling.

Example 41 shows that students became used to working with their peers, which may have improved team work productivity.

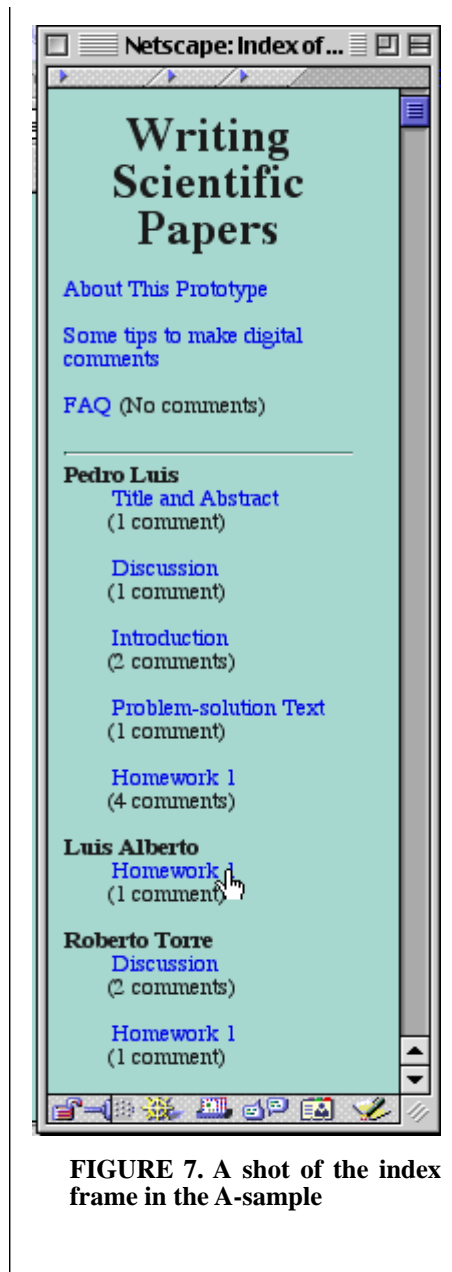
41) *Excellent writing as usual.*

Students could browse others' homework. In our survey we found that all of them except one regularly checked the system. The main reason reported was that they wanted to compare what they had done with the others and they were also interested in reading the comments received by others. One respondent said "*Interesting to see how others have written and how they make comments*". Another said "*I think we learn a lot looking at what other people do*". This was also a form of collaborative learning,

Teacher perceptions of the DHS

The teacher of the course was interviewed after every case study focusing mainly on usability issues in the system. We also discussed the impact of including the DHS in a traditional classroom from the teacher's perspective. We would like to point out that the same teacher ran two traditional courses in parallel with each DHS course.

The interview was itself used to involve users in the design of the Web tool. The teacher described situations she had experienced and in many cases proposed features that were subsequently incorporated into the system.



The teacher felt that the layout of the system afforded a general overview of how the course was evolving on two levels. One was a more general level, in which she could see how the course as a whole was developing. The fact that all the homework assignments were available in the DHS meant that she could monitor how well the students had understood the assignments; access to the student-reviewer comments allowed her to observe the way students approached this task and to offer guidance as required. The ability to move so quickly and easily from one student's assignment to another was appreciated; it enabled her to study a student's homework assignments cumulatively, or to compare a particular assignment by a student with the same one written by others. Figure 7 shows a shot of the index frame in one of the courses. The teacher could, for example, see the assignment "Homework 1" for all the students one by one consecutively.

The other, more specific, level is given by the index frame itself. The teacher could monitor at a glance each student's progress in the course in terms of the homework assignments submitted. Looking at the index-frame, the teacher could for example see that Luis Alberto is behind in relation to the other students. This information is useful as an measure of a student's commitment to the course, or it could be an early indication of a variety of other problems. In either case, it gives the teacher an opportunity to act promptly. Also, the last version in particular showed clearly which students had received comments and how many. This allowed her to give priority to those who had received fewer or no comments.

Another area that came into focus in our discussions was that of the teacher's role. Being able to read all of the comments made on a student's work gave the teacher valuable insight into those aspects of writing that students considered to be important. This became a particularly useful learning experience for the teacher as a "non-expert" in the highly specialised fields of Ph.D. students. Moreover, since all of the comments were posted chronologically, the teacher's comments were less obvious: she felt more as a participant than a teacher in these circumstances.

It was convenient to be able to work on the texts from home and outside normal working hours. However, occasionally the work load felt excessive: while some students sent in their contributions immediately after a class, many of them waited until the day before the

following class. This meant there was extra pressure to read the work of as many students as possible before meeting them, with the risk that the comments would not be as useful as they might be. With traditional classroom teaching, the teacher is perhaps more in control of correction time, having an interim period between meetings. Thus, using a Web-based tools not only places extra demands on teachers to respond quickly to assignments, it can also affect the scheduling of correction time. Teachers need to factor in these inherent features of the medium when planning their courses.

The DHS was used as a complement to a traditional writing course, and we were looking for indications that working with texts on the Web had somehow changed the character of the original course. However, the teacher noticed no clear difference in behaviour during classroom discussions. The students still discussed their texts at length in groups, which suggests that they had not exhausted all of their ideas in the Web comments. Nevertheless, in the light of comments from students who considered the classroom sessions to be superfluous, she felt that the two discussion spaces could be better utilised and integrated.

4.3.4 Discussion and conclusions

The main objective of this study was to observe how reviewers devise means to communicate their suggestions to the writer in electronic format using the Web. Another aim was to evaluate and develop the DHS system so that it would supplement and support traditional classroom activities in an academic writing course.

We found that few change representation artefacts or symbols could be satisfactorily expressed using plain text. Furthermore, these representations might be weak and misleading. Thus, a common system or standard for change representation needs to be created among reviewers and writers if a simple interface like DHS should be used.

Reviewers clearly needed to have access to the original text on the screen, both when commenting and when making corrective annotations. For example, they frequently quoted the beginning of the sentence referred to; wrote interlinear comments, even when the paste function was not available; invested great effort into cut-pasting using their mouse; and even cut-pasting by dragging. The last-mentioned might be a problem because unless the reviewer removes the carriage return it involves more work.

In order to locate relevant sentences, reviewers often referred to their position in the paragraph (First, second, third, etc.) Particularly the adjectives “first” and “last” were extensively used, “last” more than “first”. It is not clear whether this was because these sentences more often contain key information in a paragraph, or because reviewers found it difficult to allocate the correct number to the intervening sentences.

The public nature of the DHS provides several advantages over the traditional classroom environment. The shared space gives students immediate access to many more examples of both writing and comments than they would have in a classroom. Although allocated to groups for reviewing purposes, the students frequently ventured outside these groups to view the work of others. This allowed them to compare the nature and quality of their own work with a broader selection of peers. It also provided rich examples of reviewing

comments by teacher and students alike, which might contribute to the transfer of tacit knowledge within the reviewing process.

The openness of the system means that it is immediately apparent which students are active and which are not. The overview it provides the teacher is valuable for monitoring at a glance both the development of the course and the progress of the individual participant. It might also serve as an extra incentive to students to participate since their absence from the Web page is quickly noticed by all. At the same time, membership of a smaller group within the class, clearly displayed by the grouping in the index frame, may give students a sense of security and foster a team spirit which could improve their productivity.

On the other hand, knowing that their comments are visible to all might inhibit both students and teachers. They may be too polite and not be willing to point out shortcomings or errors. This might reduce the credibility of the reviewer.

Similarly, students who have received no comments on their work are likely to feel more keenly that they have been neglected. They might think that people do not make comments because they do not find their work interesting. Teachers need to be aware of this and allocate time to ensure all students are given equal attention. They can also prompt students to review each other.

As a Web-tool in the traditional classroom, the DHS constitutes a valuable supplement to the usual channels of communication. It is accessible from any computer with an Internet connection, enabling students and teachers to work flexibly both in terms of time and location. Furthermore, it gives students a quick response to their work, avoiding the time lapses of scheduled lessons. The use of a simple Web-based tool that has taken shape on the basis of users' needs in traditional classrooms might be the key for the acceptance of Web-based tools in education.

IV COL•LABORACIÓ: A WEB-BASED COLLABORATIVE WRITING TOOL - SUPPORTING DIALOGUE

1 Introduction

From the experiences of the DHS prototype, many ideas for more refined collaborative tools were brought out. We decided to go on using the document as the main entity, while the goal was now specifically to support collaborative writing. Col•laboració is a collaborative writing tool that uses the WWW as its infrastructure. It supports document sharing and takes into account the need for dialogue which co-authors might have during the writing process. In this system, the communication among co-authors is text-based. Col•laboració inherits several features from the DHS, among the most relevant ones we find making annotations to Web documents, sending email notification to co-authors for relevant events, and the screen layout of the system.

It is important to bear in mind that in the system Col•laboració the term document is treated in a particular way. A document is defined in this study as follows:

A document is a set of HTML files related to each other. Each HTML file represents a distinct self-contained portion of the document called section. A section, being a separate file, can be added, changed, or deleted independently of the other sections. Merging the sections produces the document as a whole.

Col•laboració presents the document to co-authors displaying the titles of its sections which are hypertext links that show the section's content. It also allows co-authors to work on different sections simultaneously. It should be stressed that the system is not designed for the transcription of the document to be produced. Rather, it supports sharing of the document and the discussion among co-authors that the production of the document might demand.

The structure of a document could be complex. Representing the structure of a document as a graph, we can assume that every section of a document is a node in the graph. A document's hierarchical tree structure could be deep and broad. However, there are certain kinds of documents that have a relatively simple hierarchical tree structure, that is, shallow and narrow. Research papers belong to this type of document. They usually have a fixed content format i.e. introduction, methods, results, and discussion. I have restricted myself to studying collaborative writing tasks of small groups using the WWW to produce documents that have a simple hierarchical tree structure. This study addresses (a) the communication needs co-authors might have during the writing process, (b) how they made off-side annotations in digital format using a split-screen interface, (c) the characteristics of a text-based communication in a collaborative writing task. The intention is to gain new insights into the design of the development of Web-based collaborative writing tools.

2 The Author-Document-Author Workspace.

My approach to building a collaborative writing tool on the WWW needs a shared workspace for co-authors. A graphical representation of the shared workspace consists of the illustration of nodes of two types: 1) document nodes, and 2) author nodes (see figure 1). These nodes might be related to each other but two nodes of the same kind cannot be related directly. In other words, a document node can be related to an author node but never directly to another document node. Document nodes must have at least one relation, namely, related to at least one author. Author nodes can be directly related to document nodes or stand alone. A stand alone author node corresponds to an author who has not (yet) written a document. This would be the case when a new author node is added to the workspace. The documents are therefore the elements shared between co-authors.

Col•laboració uses the concept of the tree, a well-known data structure representation, to personify different views of the shared workspace. The type of node at the root determines the process that Col•laboració uses to build up the tree. Consequently, there are two processes.

In the first, a given author node is the root of the tree. All the documents connected to this node are presented in the second level. Finally, all the authors' nodes connected to these documents form the bottom level of the tree. That is, we have a three-level tree counting the root as level one. We named this the author-document-authors tree, see figure 1. The system uses this when a co-author enters the system.

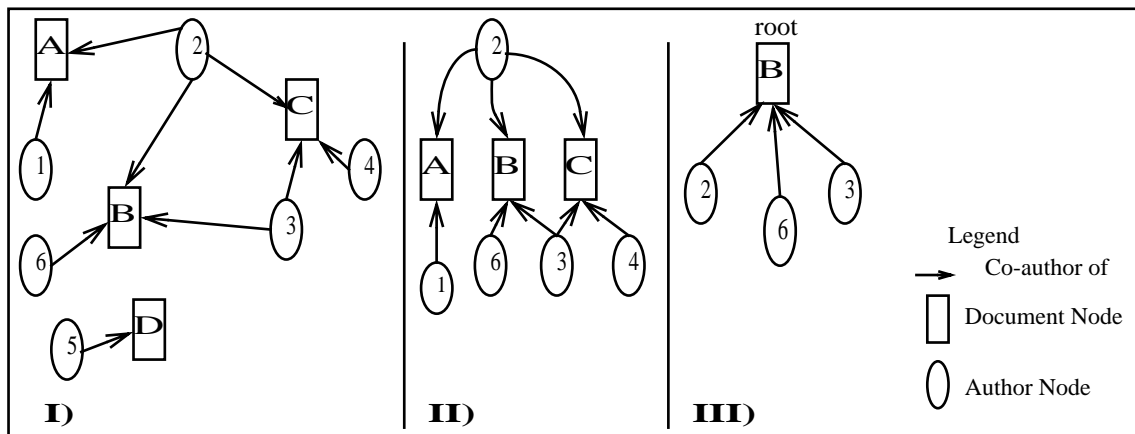


FIGURE 1. I) The graph representation of the shared space. Documents and authors are represented as nodes. II) A three-level tree built up from the shared space having an author node (2) as the root. III) A two-level tree built up from the shared space having a document node (B) as root.

The second process assumes a given document node as the root of the tree. In this case only a two-level tree is built up as author nodes become the second generation. We named this the document-authors tree. This is used by the system to determine who are the document's co-authors, see figure 1.

3 Description of the system Col•laboració

In the next few sections I will describe the system Col•laboració with two modules:

1. **The document development module**, which is directly related to the development of one document (adding a section to a document, editing the document, and adding a comment to a section, etc.). This module has been implemented, tested, and evaluated.
2. **The author management module**, which handles the authors' objects in the shared workspace (i.e. creating a new document, joining as co-author to a document, and setting preferences). This module does not handle changes in the document itself. This module has not been fully implemented yet nor evaluated.

3.1 Description of the document-development module

Col•laboració borrows the layout from the system DHS. There are two important differences here (see figure 2). Generally speaking, the index frame contains links to the sections of the document. In other words, the index-frame reflects the structure of the document. The other change is in the command-frame that has several commands in addition to the "Add comment" command it..

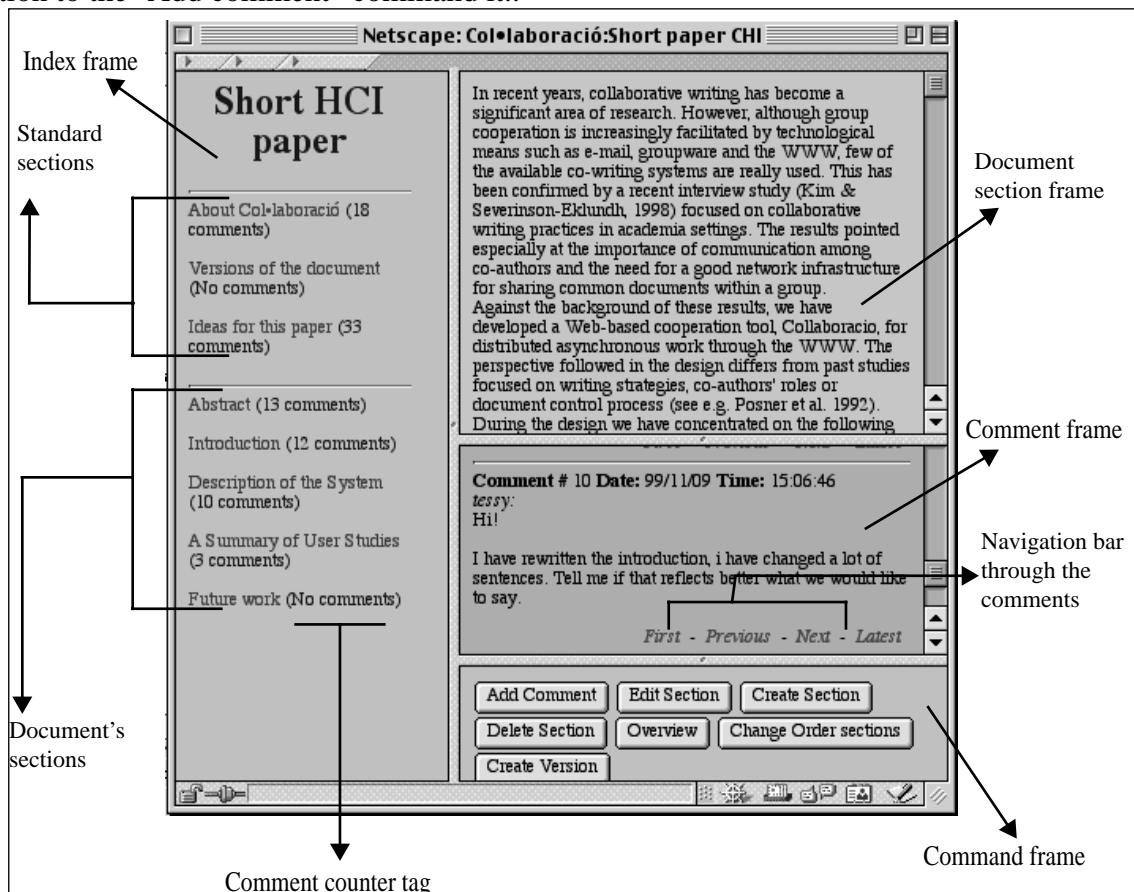


FIGURE 2. The screen layout of the document-development module

3.1.1 The index frame in Col•laboració

In the system Col•laboració, the index frame is divided into three parts: (1) the title of the document; the only item of the list that is not a hypertext link, (2) the standard areas part, and (3) the document's sections part, made up from the different sections' titles. The links in the standard areas part are added automatically by the system when a document is created and the sections in the document's link are added by the authors as the document is developed. Also, every hypertext link shows the number of comments a section has received so far. Next, I will describe the three elements that constitute the standard area part and discuss the reasons why we decided to include these elements here.

The standard area part

From our initial experience using a draft of the system, it was clear that there were some common areas that should be present for every document that was going to be produced using this system. I have called this the standard area part and it is made up of the "About this prototype", the "Version of this document" and the "Ideas for this paper" link. In this part we have three items of the hypertext link list.

The "About this prototype" link

The "About this prototype" link shows a document that briefly describes the features of the system and how to use it. It is a default section that is loaded when the user enters the system. It also provides a list of known problems related to the system and how to overcome them in order to support the users. Because of our experience in the case studies for the system DHS, we know that many users feel the need to make comments, not on the document content, but on the system they are using, in this case Col•laboració. A similar finding was made by Mhashi et al. (1992) in a study about the use of a discussion system as an annotation system. A great amount (75%) of the comments they analysed were of an administrative matter (see Ideas for this paper below) and system-specific type. Thus, the "About this prototype" section is intended to receive co-authors' comments on the system. In this way we could keep the comments relating to the system and to the contents of the document in separate files.

The "Version of this document" link

Using Col•laboració co-authors may work sequentially or concurrently on the document and all of the users may modify the document. Thus, there must be some form of version control. The system provides a facility to create a version of the document or any of its sections. We decided to separate the version management section from the rest of the document. The second link in the standard section part is the "Versions of this document". The activation of this link displays a document that has a list of the versions made by the co-authors.

The “Ideas for this paper” link

One of the results of the first study in which Col•laboració was used is consistent with the statement “communication about plans is needed to ensure that co-authors have shared goals, and that it (the communication) would have notably increased the likelihood of their producing compatible products” (Neuwirth et al. 1990). The third link in the standard sections part is the “Ideas for this paper” in which co-authors may have a meta discussion about the writing task. A meta-discussion of the document as a whole is often needed. In this section co-authors could discuss co-ordination matters, administrative aspects, as well as planning. The particular objective with this section is to improve performance by offering a share space single-minded for coordination in distributed decision making. Co-ordination is informally defined as the activity of independent agents making harmonious, nonconflicting decisions (Molene and Crowston, 1990). Interdependence means that, in cooperative settings, activities flow from one individual to another, in our case from one co-auhtor to another.

3.1.2 Commands of the system

Create a section. When this option is selected a separate window is opened in which the user can enter (1) the name of the section, (2) the email addresses of the co-authors that are responsible for it, and (3) the section content. The system will update the index frame, appending the new section to the document section list. The email address is used to broadcast when a comment/change has been provided to a particular section. If no email address is entered, the email notification will be sent to all co-authors.

Add a comment. Comments can be attached to every section of the document. When this option is selected a separate window is opened in which the user can enter (1) his/her name or nickname, (2) his/her comment. These comments can include HTML tags. When a comment is sent, its content is also sent via email to the co-authors.

Edit a section. Co-authors can make changes to the content of the document. This option will show in the document section frame the HTML source of the section. Any changes made to a particular section will be available to the rest of co-authors immediately. This function is more suitable for minor changes in the section.

Delete a section. This option updates the index frame and erases the link that was specified from the active document section. Co-authors cannot undo this action, only the editor of the domain can do it.

Change position of the sections. If co-authors need to change the sequential order of the sections they can do so by using this option. For example, if the “Abstract” section is created when the introduction section already exists, it can be moved to the first position of the document section.

Make an overview of the document. Co-authors can select different sections of any particular document to have a personal view of the document and their comments. The order of the sections will be kept in the same sequence they have in the index frame. The

“Ideas for this paper” section can be included here. Also, this overview presents the information in a suitable format for printing, e.g. it provides a margin in the hard copy.

Make a version. Co-authors can select different sections of the document and add their comments to create a personalized version of any active or specific document as per subject topic and content that might interest a particular group or co-author. A version can be also created for the purpose of saving a back-up copy of it.

3.1.3 Logging function

The system has a log file feature that registers every change that is made on the browser of the user using the system. Once a command is executed in the system (see previous section) it is registered in the log files.

4 The author management module (AMM)

In this section a complement module to the system Col•laboració is presented. Though it is not fully implemented, it might be interesting to supplement the work that has been done so far. This section assumes that the AMM would be integrated with the previously presented module, the document development module.

Once an author enters the system Col•laboració (login-password based) it will present a view of a shared workspace. This view is automatically generated using the author-document-authors tree that the system builds using the author node as the root of the tree. In other words, the first level of the tree is the author who entered the system (see figure 1). This can be compared to a folder in which each author has in their possession all the documents to which they are a co-author. Figure 12 shows a picture of an author’s management module.

4.1 Information given in the AMM

The following information is presented in the form of a table for each document to which the author is a co-author:

1. The title of the documents he/she is producing in collaboration with others. This text also works as a hypertext link that activates the document development module for that document.
2. When the document was created, the last time it was modified and the last time the author activated the document development module for this document.
3. The co-authors’ name list for that document. This list names the co-authors and activates a script that allows the author to select in a flexible way to whom an email will be sent as an awareness mechanism (this will be explained later).

Additionally, the commands (see next section) that the user can perform from the AMM are also available.

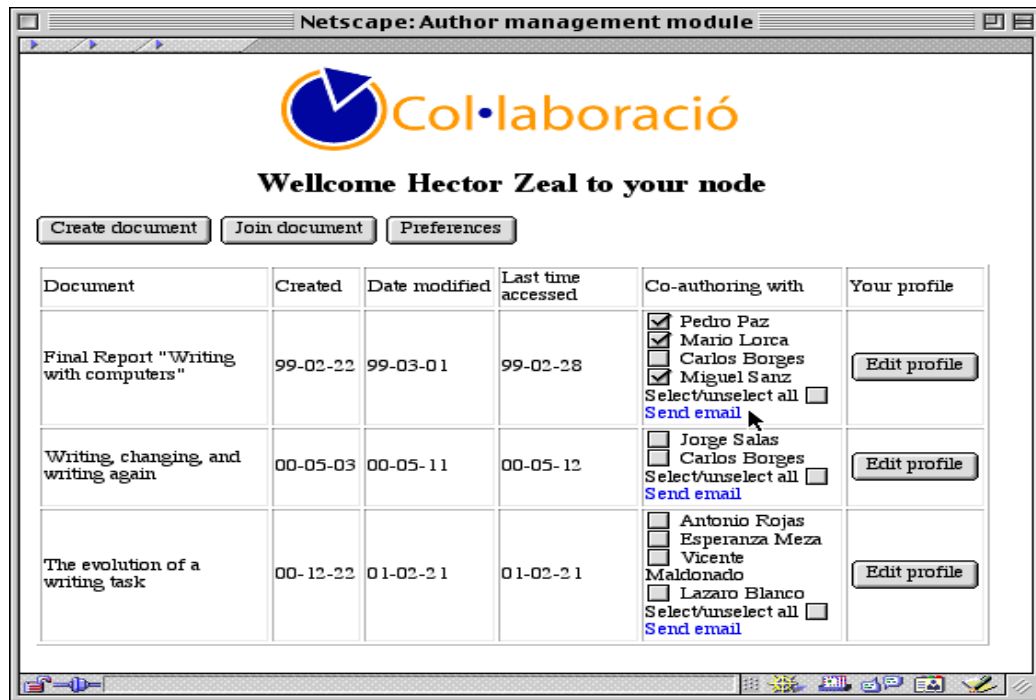


FIGURE 3. An example of a co-author's management module

4.1.1 Commands that can be made from the AMM

To create a new document

Graphically, creating a document in the system Col•laboració means adding a new row to the author management node table (see figure 12). An author creates a new document and 1) inputs the document's title and 2) selects the co-authors from the list of all authors registered in the system.

Once this is done, all selected co-authors will be sent an email inviting them to participate. Upon acceptance, the AMM will be updated for the rest of co-authors indicating his/her acceptance. When this confirmation arrives, the name of the co-authors will be labelled properly to indicate the waiting status.

Join a document

Authors can join active documents at any moment, that is, become co-author for that document. To take this step, one of the previous co-authors has to invite the new one.

Edit profile

Authors can edit their profiles and set up their preferences for each document. This includes to which email address they want to receive their awareness messages and which label the email address should have (see below), how often they want to get notification via email, which font/colour they would prefer to use for their comments, and the

preferred size of the pop-up windows. The system Col•laboració offers authors the possibility of putting an author-defined label on their email address..

The profile can be set up in three levels. The first is the general one. These values will be set up as the default for the author's preferences. The second level is set up for a document as a whole. Finally, the third level can set up preferences for a particular section of the document.

4.1.2 The use of email in the AMM

Supporting communication between distributed participants is important. Email is well known because it is “the most widely-used computer application today for person-to-person communication” (Bälter, 1998). Furthermore, using the WWW as infrastructure, as Col•laboració does, allows the members of the groups to use email as a channel of communication. In the next section, I will present two modifications to the use of email that I believe will improve communication, and thus the productivity of the distributed groups.

Multiple labels on the email address

Generally, the format of the email address could be structured as: label <email_address>. The email_address field is given in the following way user@domain, where the fields user and domain are mandatory. In most cases the email address label will be the name of the user (at least in institutional settings). However, the label might be any text. The email programs known to me do not support users to have at the same time different labels under the same email_address.

The idea presented here is to use the edit profile function that the AMM provides to supports co-authors organizing the email notifications that the system would send. Most email programs contain filtering functions that help users structure their email messages. Filters use the information contained in the message itself, either in the header of the email (e.g. to: from:). Filters aims at lowering the user effort in the identification of messages for later processing. Filtering can be triggered automatically for incoming messages with the help of the email program's filtering function. These filters can, for example, group the messages by automatically moving an incoming email to a specific folder.

The non sequential patterns of related messages (Palme, Karlgren and Pargman, 1995) is one of the problems that the use of filters faces. As we said before, Col•laboració sends automatically email notifications to co-authors. As this is done automatically, it can be assured that a sequential pattern of related messages will be satisfied. On the other hand, email notifications sent by Col•laboració may contribute to an increased stream of messages for the co-authors. According to Palme et al., applying a filter to a stream of messages reduces information overload. To illustrate the idea, consider this scenario:

An author in the shared workspace might be the co-author of several documents. Suppose Author A, is a co-author for document D1 and D2, and labelled his or her email address as “D1 <authorA@domain>”, and “D2 <authorA@domain>” respectively. Using the

filtering function, the incoming email messages can be placed automatically in different folders. At least the label D1 or D2 will notify to the recipient that the email is related to the co-authoring work that he/she has taken part in and further, which document the message refers to.

A flexible mailing list

The co-authors' list for each document in the AMM is a hypertext link list which activates the email application from the author management view (activating the mailing program connected with your Web browser). I have designed a flexible mailing list. The term "mailing list" derives from the conventional feature that most email programs have, in other words, sending an email to many under a single email address. It is flexible because, by selecting a checkbox, users can easily set up a subset of members from a pre-set participants list. It is worth noting that all members of the list can be selected with one click. Although selecting a large subgroup from the list could be cumbersome, it is important to bear in mind that this script is designed for small groups (2-9 participants). In figure 12 we find a co-authors group formed by Pedro, Mario, Carlos, Miguel and Hector Zeal. Using a flexible mailing list, Hector can send an email to a sub-set of the co-authors (e.g. all but Carlos - say the others are preparing a surprise party for Carlos, see figure 12). Using a mailing list it would not be possible, because the email would be sent to all the members of the group, not allowing the sender to select a subset of members. To do that, users would have to create a new mailing list, or search in their address book for the email addresses of the co-author and add them to the email. Note that this subset might be relevant for just one message and the effort that has to be put in doing that, using a mailing list or the address book, might discourage users from sending the email. In the example I gave, preparing a surprise party for Carlos, this could affect the social relationship of the group. Yet group members can use other channels, e.g. the telephone. It would become easier to build up a mailing list as they can be constructed "at a click".

5 Case Studies of Col•laboració

The Col•laboració system has been used to support collaborative writing in eight different tasks in which 2-9 co-authors took part. Table 1 present the case studies in which Col•laboració has been used so far.

The case studies are presented in chronological order. The duration of the case studies is counted in relation to the date of the first and last registered comments that were made to any of the sections apart from the ATP section. Another unit of measurement in this study is the number of words that the submitted comments contain. I decided that the number of comments is not sufficiently a representative unit because the length in words of the comment found in the system varies from two to 800. That is why I decided to complement the number of comments unit with the number of word in the comment. Also, the number of words closely represents how much was "said" in a text-based communication. In the result that I present, all the counting refers to the number of comments and number of words in the comments.

It is also important to indicate that in all the case studies the developer-programmer was following the discussion that was supported by the system and giving technical support to the co-authors or was one of the co-authors. Also, problems reported by the co-authors were tackled immediately after they were sent in the form of comments in the ATP document. Co-authors gave suggestions too, some of them implemented after the task, so the system was iteratively improved after each case study.

Table 1. Case studies in which Col•laboració has been used

Case Study	Co-authors	Duration (days)	Co-located or distributed	The task
Poster A	3	23	co-located	To design and write the text for a workshop poster.
Report Alpha	2	24	distributed	The author had to write a technical report. The author and the reviewer were located in different countries.
Report Beta	9	42	co-located	To write a 10-section technical report. Co-authors were responsible in different sections in sub-groups of 2-3.
Poster B	3	6	co-located	To design and write the text for a conference poster.
Masters' thesis proposal	2	22	distributed	The author had to write a Masters' thesis proposal. One reviewer. Both co-located in different places in the city.
Flyer	4	11	co-located	To write a flyer for a presentation in a conference.
Workshop	4	15	distributed	Co-authors had to write the final report for an international conference they had participated in.
Short paper	4	24	co-located	To write a short paper (two pages) for a conference.

In none of the cases, the final formatting of the document was done within Col•laboració. By mutual agreement one of the co-authors took the current version from the system and finished the document using a word processor. Email awareness for each case study was set up in different ways. Generally, all co-authors were coupled to all sections of the document. Thus, they received an email notification when a comment was added to a section.

In total there have been 14 participants who have used the system in real cases as co-authors of documents. All of them are experienced computer users and are familiar with the WWW and Web-browsers. They had experience in writing and most of them had previously published a report or a paper in collaboration with others in an academic environment. Two of them were undergraduate students and one was from the industry. All the others (11) had a Ph.D. or were graduate students. Three of the participants were part of the project that supports the development of Col•laboració. Two of them participated in six of the eight case studies and the third in seven.

Col•laboració uses JavaScript, a client-side scripting language. Using JavaScript the system can validate some entries and to determine the status of the document to be produced. We have restricted the use of this system by asking users to use Netscape because many problems were presented if the system was run on other browsers.

It has not been the purpose of this study to analyse the content of the document produced or its evolution during the writing task. Rather, I wanted to study the communication and the characteristics of the comments that the co-authors produced during these studies. Overall, we have divided the comments that were found in each case study into three categories. The first is related to the system itself. Most of them are found in the ATP section. However there were very few cases in which the co-authors placed a system-related comment in another section different to the ATP. I have included those comments as part of the ATP and the number of words for each section was counted after the “transposition” of the comment. The second category is the one related to the “Ideas for this paper”. Comments that were found in this section are related to the document as a whole. No misplaced comments for this category were found in those case studies (7 of 8) in which this section was included. The third category are those comments that were made under the context of the document’s sections.

The case studies will be presented initially in a brief way presenting the method and results of the studies. Later, in section 5, I will discuss the results and the communication process of all the case studies.

5.1 Case study 1: Poster A

5.1.1 Method

This was the first study in which we used Col•laboració. It was done in June 1998 and lasted 23 days. Three co-authors took part in this task, all based in the same research department. The task was to design and write the text for a poster about one of the projects in which they were involved. Additionally, the co-authors were interested in the design of collaborative writing tools and had experience with such tools. When this case study was performed, Col•laboració included only one of the standard sections: the ATP section. That is, the “Ideas for this paper” and “Versions of this document” were not part of the sections of the document in the system. Co-authors decided to communicate as little as possible using other channels, i.e. face-to-face, when they were going to discuss the task. At this time the system did not keep a log file of the actions co-authors performed in the system.

5.1.2 Results

The co-authors created seven sections and sent 38 comments that contained almost 3,000 words. One of them created a version of one of the sections as a section itself in the document. This was the only one that did not get any comment. The other six sections received at least one comment. However, only 18% (524) of the words were comments on the document's sections. Most of the conversations co-authors held using the system related to the system itself and 82% of the words they produced during this task were related to the ATP document.

As can be see from figure 3, the document's sections were almost un-attended by co-authors. Instead they were discussing the system itself. This is probably a consequence of

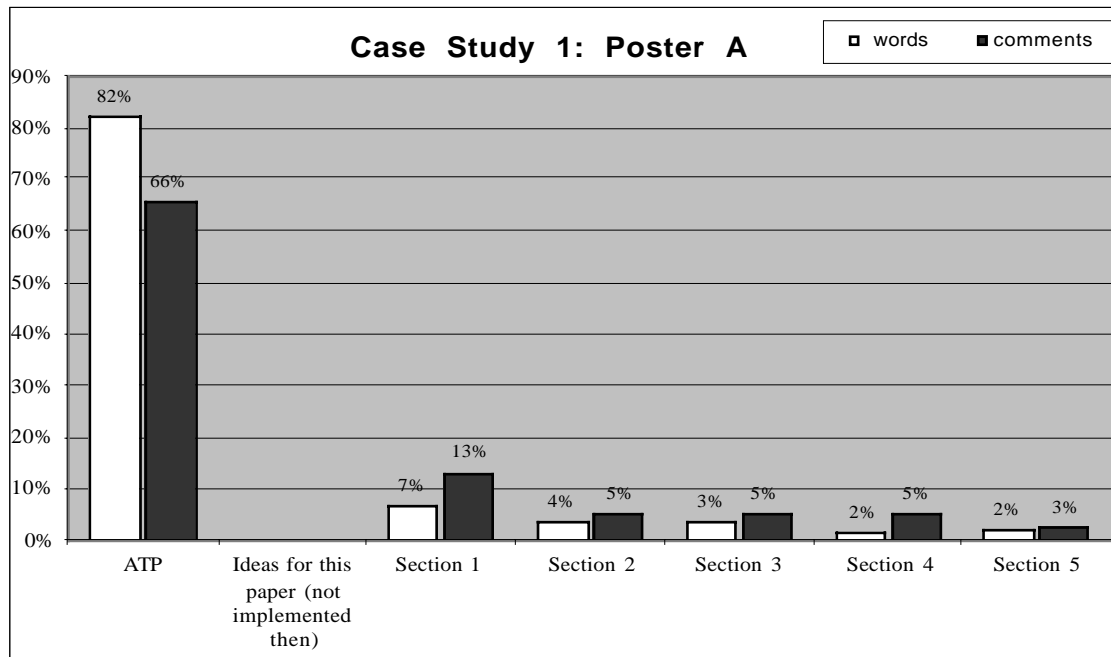


FIGURE 4. Distribution of the number of words and comments in the case study 1 their unfamiliarity with the system and certainly of the interface-related problems the system presented at that time.

5.2 Case Study 2: Report Alpha

5.2.1 Method

The task in this case study was to write the final report of a project that had run for two years. Co-authors (2) were located in two different countries in the same time-zone. Both co-authors had experience of the system. One of the co-authors was acting only as a reviewer of the report. All the task-related communication was done through the system. It lasted 24 days and in this case study we had all the standard sections mentioned in section 3.1.2. At this time the system did not keep a log file of the actions co-authors performed in the system.

5.2.2 Results

Co-authors produced about 5,700 words in the 70 comments that were sent in. The report consisted of 12 sections. Five of the sections did not receive any comments but this might be because the information they contained was not worth discussing at that time during the writing task. The title page, appendix, and conclusion were some of these non-commented sections. Seven of the sections were commented upon but this discussion occupied only 23% of the words (1,300). The ATP document was the most discussed topic with 43%. The “Ideas for this paper” section got 34% of the discussion. Two of the comments embedded pictures and some of the comments used HTML tags.

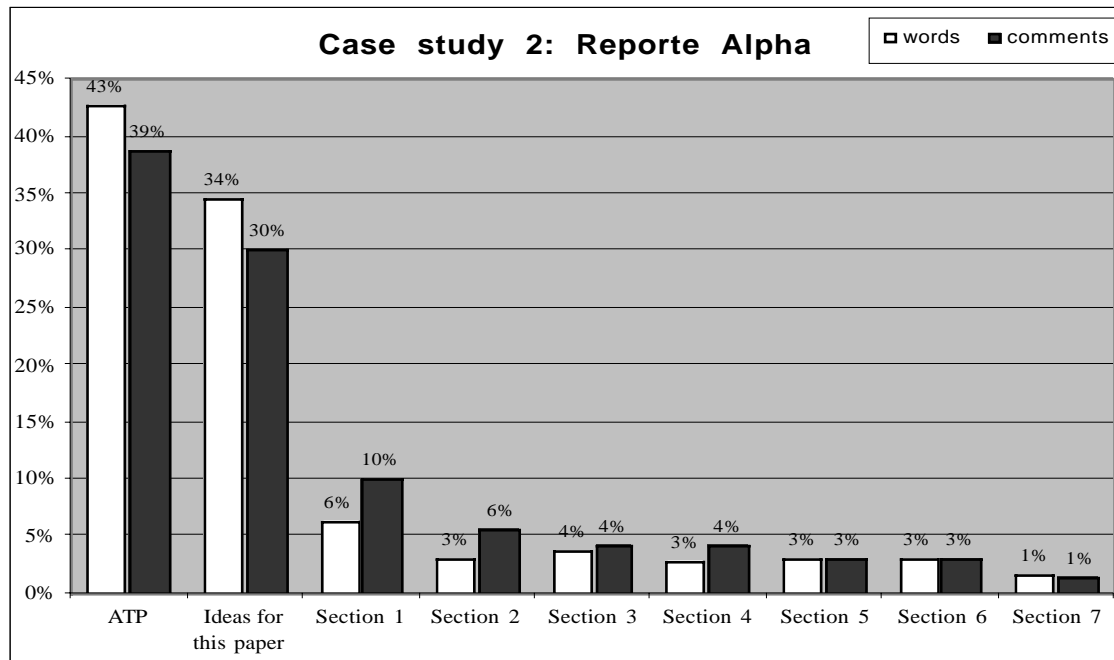


FIGURE 5. Distribution of the number of words and comments in the case study 2

5.3 Case Study 3: Report Beta

5.3.1 Method

Nine co-authors took part in the production of the annual report of a project they had been working on. This project focused on four different areas, each constituting a separate sub-project. Eight of the co-authors were located in the same department and only three of them had previous experience with Col•laboració. These three co-authors were involved in the same sub-project. The use of the tool was decided on a meeting that the group had, when they were also informed how to use the system. The report was divided into 10 sections. Four of the sections were the description of the sub-projects and the other six were more global, for example, the introduction, references, connection between the sub-projects, etc. Each section was supposed to have been written by two or three co-authors, according to the project that the section described. The awareness notification sent by email was set up in this way: those who were responsible for one section (or project description) were to receive an email when a comment was added or a change was made to the section for which they were responsible, whereas for the sections of the document for which they were not responsible they did not receive any email awareness. For the section “Ideas for this paper” the system sent an email notification to all co-authors when a comment was added to it

After the task co-authors were asked to answer a Web-based survey to evaluate the system, to determine which other channels they used to communicate, whether they had read others’ sections, and which tools they used in combination with the system to perform the task

5.3.2 Results

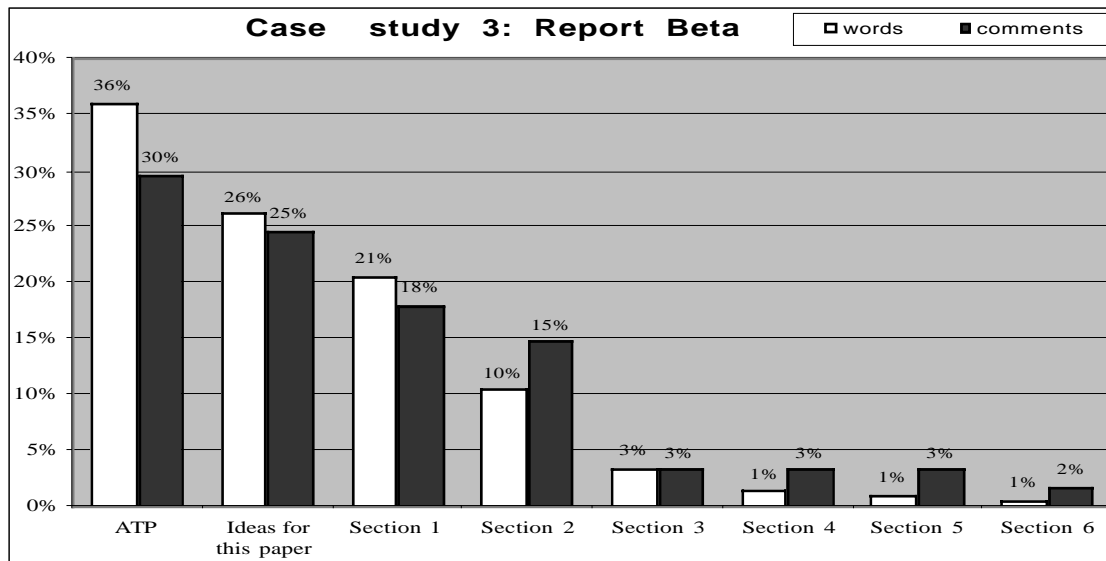


FIGURE 6. Distribution of the number of words and comments in the case study 3

Throughout 42 days this task was mediated by Col•laboració. Co-authors produced 63 comments in total with 3,300 words. Though the percentage of words relating to the ATP document was more than twice that of the first case studies, the sections were still the most commented on (in number of words and comments). Again, the “Ideas for this paper” was the second most commented section and included 26% of the discussion. The next most discussed section was that of the three co-authors who had previous experience. This section included 21% of the discussion. The “Ideas for this paper” section was now used as a co-ordination point. The leader of the project required actions, sending comments to this section. The project leader submitted nine of the 15 comments that were related to this section.

5.4 Case Study 4: Poster B

5.4.1 Method

Three co-authors that were familiar with Col•laboració used it to produce a poster for a conference. They came from the same department. One point worth noting here is that the deadline for the poster submission was close. Co-authors worked for six days on the poster text using this system.

5.4.2 Results

The ATP section included only 12% of the discussion, possibly because the co-authors were familiar with the tool. The comments on this section were mainly related to enhancements suggested by the co-authors. Only one error was reported which turned out to be a Web-browser problem. Co-authors produced 3,017 words in a period of six days. The “idea for this paper” section included 45% of the discussion while the rest of the

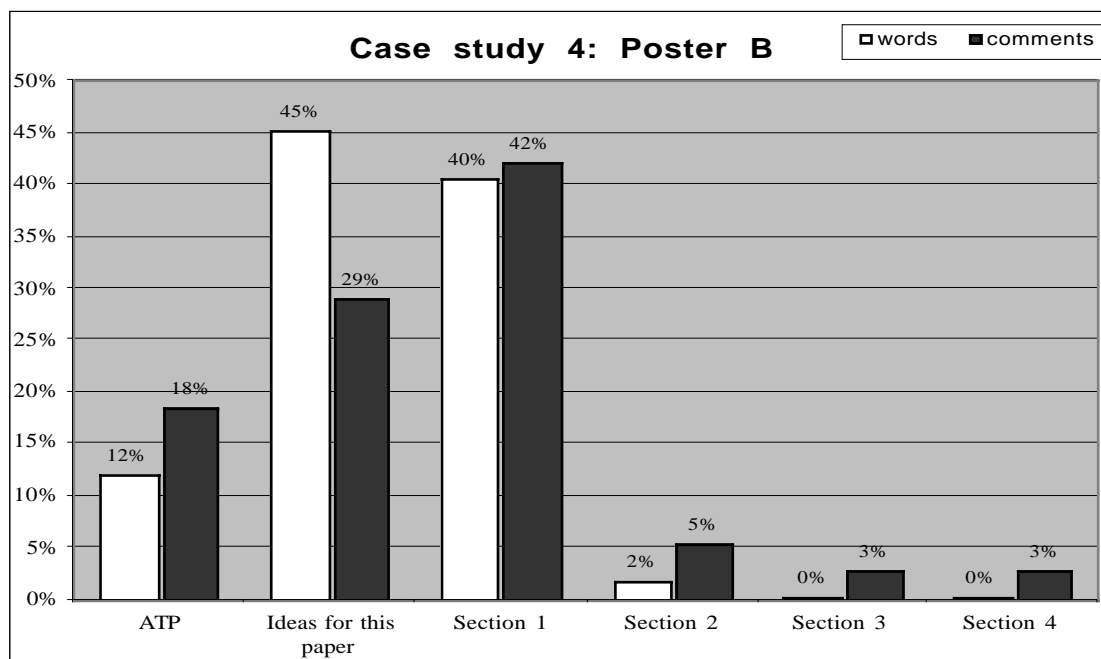


FIGURE 7. Distribution of the number of words and comments in the case study 4.

sections contained 43%. Five sections were created and the co-authors commented on four of them. Two of the sections were commented on very briefly, just to indicate that they were fine. Figure 6 reveals that co-authors concentrated more on the task than the system itself in this case study. A summary of one Web site was included as part of one of the comments (it was a guide about how to prepare a poster). Additionally, a hypertext link was also part of that comment. One point to make is that using Col•laboració, co-authors could easily include external information that might be helpful for the task's accomplishment and that a hypertext link might be seen as a “window” to the world of information that the WWW offers today.

5.5 Case Study 5: Masters' thesis proposal

5.5.1 Method

The task in this case study was to write a Masters' thesis proposal (about four pages long) at a technical university. The student had a technical background and was familiar with Web-based applications. The student was trained in the use of the tool for two hours, after which he claimed to understand how to use it. The supervisor had also used it several times before. The URL of the system was sent by email to the co-authors. The student had no reward for participating in this study and initially produced one draft which the supervisor revised. Later the student updated the section according to comments made by the supervisor, revised it himself and communicated his comments through the system. After the task the student was interviewed in order to evaluate the tool

5.5.2 Results.

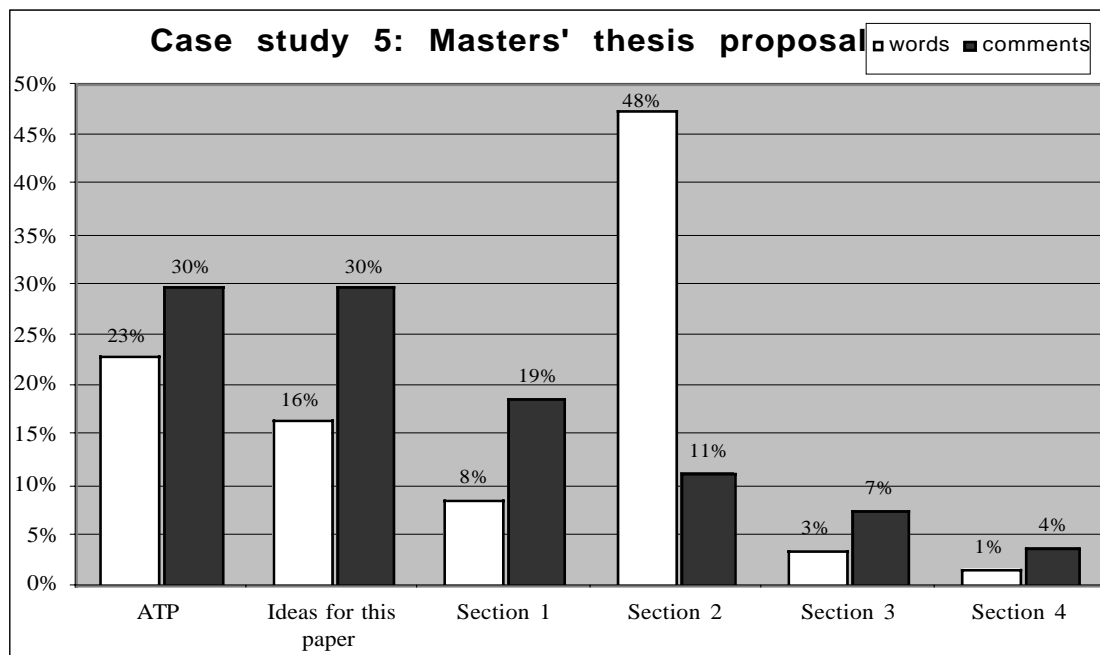


FIGURE 8. Distribution of the number of words and comments in the case study 5.

As shown in figure 7, the discussion around the “ideas of this paper” fell to 16%, the ATP section got 23% and the rest of the sections included 61%. The duration of this case study was 22 days. Co-authors produced about 2,700 words in 27 comments. Despite “Section 2” having only three comments, it was the section that included most of the discussion, with 48%. This supports my decision to use the number of words as the unit of data analysis for these quantitative results. In this case study we also found a hypertext link that pointed to a Web site that was related to a writing guide.

5.6 Case Study 6: Flyer

5.6.1 Method

Four co-authors together wrote a flyer for a poster that was going to be presented in a conference. Three of them were familiar with the tool and all of them were located in the same department. Co-auteurs had access to the document and comments they had prouced using the system. The new user was trained personally to use the system until a global understanding of the tool was reached. Co-authors started to rework the text of the poster that was produced in case study 4, that is, Poster B. In other words, co-authors were working on a text that was already written. Co-authors had to produce a one-page document that was going to be handled during the poster presentation in the conference. This flyer could include more detailed information than was possible to include in the poster.

5.6.2 Results

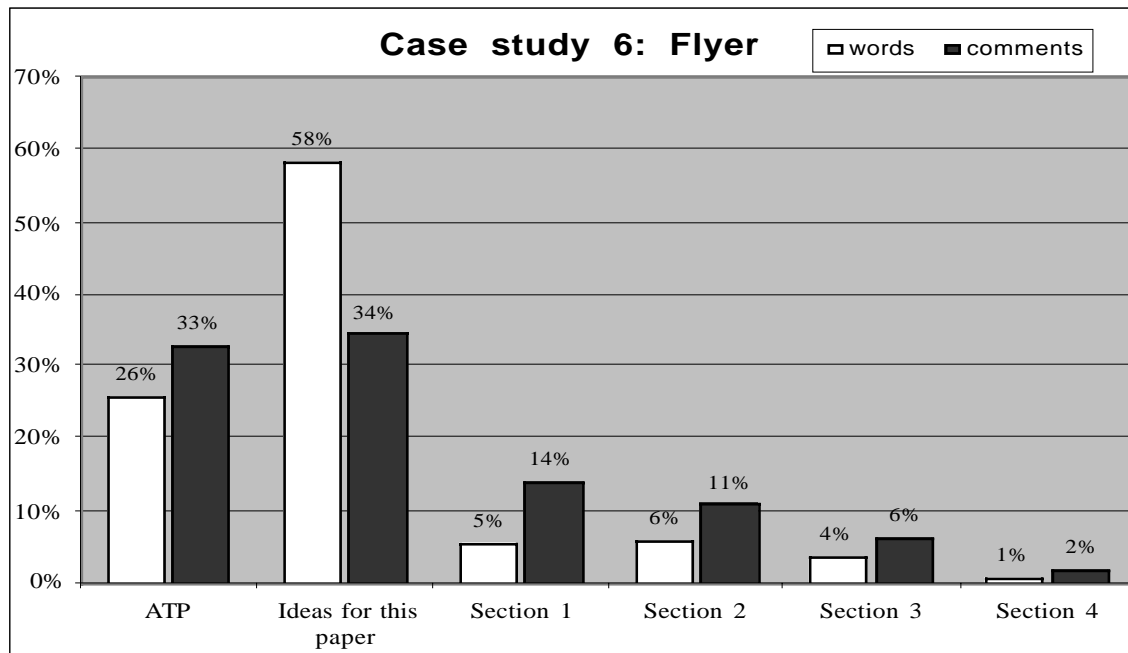


FIGURE 9. Distribution of the number of words and comments in the case study 6.

During eleven days the co-authors produced almost 4,000 words included in 64 comments. The most commented section was the “Ideas for this paper” with 58% of the discussion and only 16% of the discussion was dedicated to the rest of the sections. Closer examination of the comments found in the “Ideas for this paper” section reveals that a lot of effort was put into defining what information was going to be presented, and in which order. I found that an old comment, made during the previous task of the Poster B, was used in this task as well. One of the co-authors copy-pasted it. Hypertext links were used here, four in total. I also found two pictures embedded in the comments.

5.7 Case Study 7: Workshop

5.7.1 Method

Co-authors in this case study took part in an international workshop. Four people formed the group. Two of them were located in Europe and two in the USA. One of the co-authors had experience with Col•laboració and all of them had good knowledge of Web-navigation. They communicated via email. The leader of the team had been chosen during a face-to-face meeting while the users were in the workshop. Pre-questionnaires were done to determine co-authors' computer literacy, writing experience, and collaborative writing experience. Post-questionnaires were also used to understand what happened during the task, and to evaluate the tool. At first, the co-authors used the system for one week, to learn how it worked. After this period they started to write the report. Each team member was responsible for the writing of one major section of the paper. Participants were asked to use Netscape to run the system.

5.7.2 Results

This was the only study with “outside” participants, that is not from our department. This case study was the one that presented the least number of comments and number of words. The 21 comments included 1,838 words. Seven sections were created by the co-authors. The “Ideas for this paper” was the most commented section. The leader of the group was the only one who submitted comments in this section. Only 11% of the discussion was related to the document content. This might be because of the nature of the task: each team member was responsible for one part of the document, so they did not have to discuss the content of the sections. They were mainly exchanging information and co-ordinating their efforts. The number of emails sent by the group was 67. The leader of the group sent almost half of them (34).

In the email text there is no evidence that the document content was discussed. One of the co-authors reported that “we did not succeed in discussing the text, which is a sign of failure”. The ATP section included 15% of the discussion and two errors relating to the tool were reported, which were solved soon after being reported. Co-authors mentioned that face-to-face meetings are needed before the outset of a collaborative writing task, and that the motivation for performing the task was too low. The use of email was also a problem when exchanging files. They reported not being able to open some of the files sent by another member's team related to this task. An important error happened in this task: for some unknown technical reason Col•laboració sent the same comment 50 times via email to the same user, who found this annoying. The situation has not been reported again. One of the co-authors reported that Col•laboració was “dragging out the process a lot longer than needed” and another reported that using the system made the task even more difficult to accomplish because they had deal with a new system and eventually forced to use another technology.

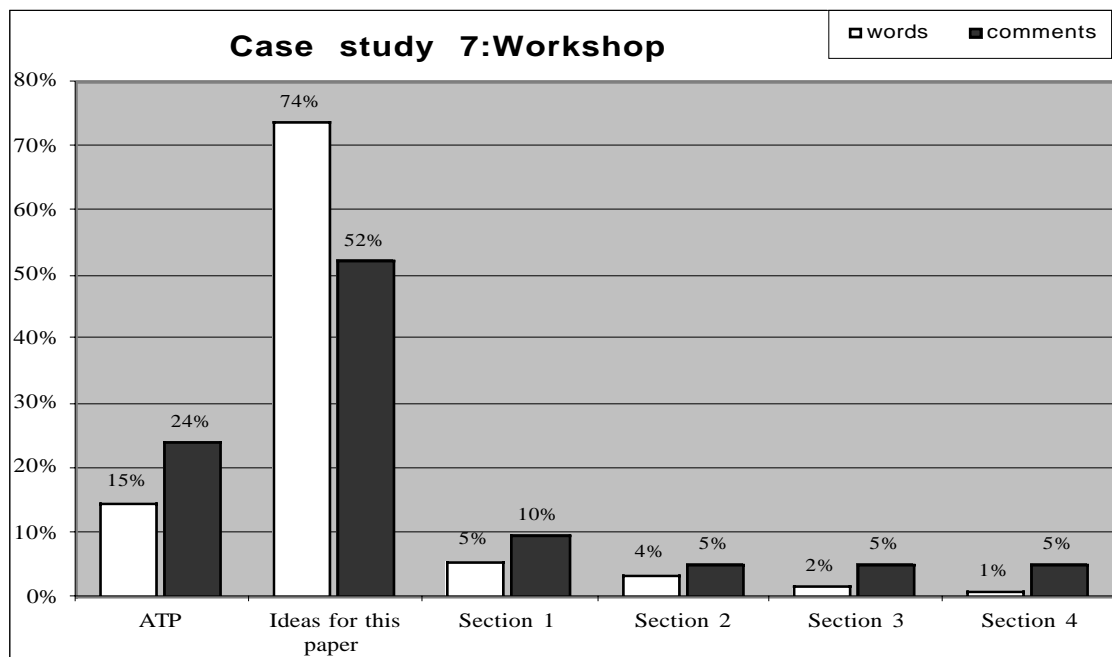


FIGURE 10. Distribution of the number of words and comments in the case study 7.

5.8 Case Study 8: Short paper

5.8.1 Method

A short paper (two pages) to be submitted to a conference was the task that four co-authors had to accomplish using the system. All of them had had previous experience and were located in the same department. One of the co-authors worked on the task for a week in another country in the same time zone. The comments made by co-authors were parsed to pick out quotations from the text of the short paper. The quotation text was identified as such if the text included a quotation mark (“;”) as it is shown in the case 1 (see below) or when the co-authors in the comment makes it clear that the discussion is about the content of the section. In case 2 the co-author refers to the use of the word *GroupWare* found in the text of the short paper. These are two fragments taken from the case studies.

1) *“When a comment is added to the system, it is sent also via e-mail”. This sentence seems interrupted.*

2) *I would not call GroupWare (Groupware like software, hardware, the capital W is not needed) a technology.*

5.8.2 Results

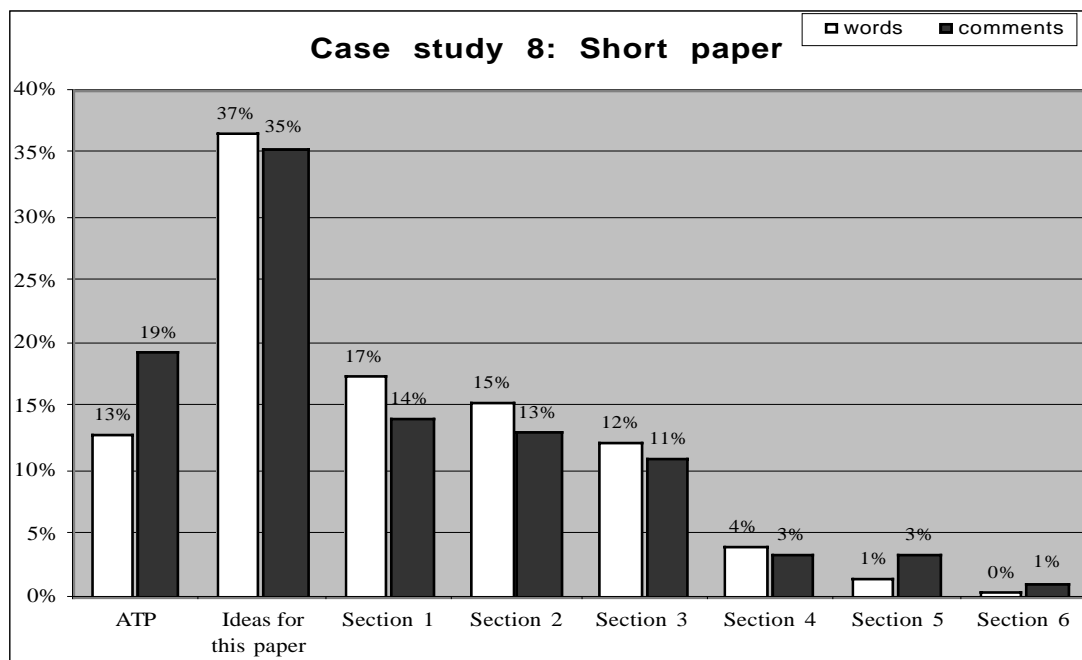


FIGURE 11. Distribution of the number of words and comments in the case study 8.

This was the most discussed document produced by Col•laboració in terms of words. Through the system co-authors exchanged 93 comments that included 6,808 words in 24 days. Five sections were created and four of them were commented on. In the comments were found images and formatted text with HTML tags, used to highlight suggested changes by co-authors when playing the role of reviewers. In the discussion of the

sections' content of the document, co-authors made 35 quotations of text. One of the comments contained a link to the conference Web site where detailed information about the requirements for a short paper was given. The number of words (for a section or paragraph) was an important factor that the co-authors used several times in their discussion. This was probably because of task-specific space limitations. I found also that co-authors were co-ordinating face-to-face meetings through the system.

6 Discussion

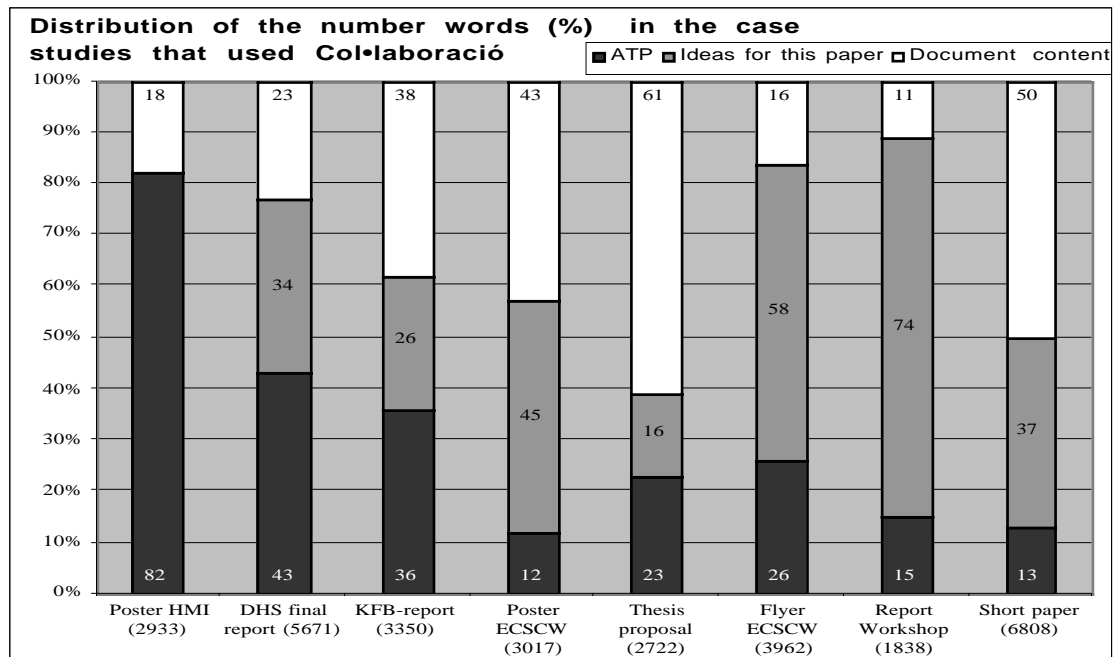


FIGURE 12. Distribution of the words in the case studies that used Col•laboració

Col•laboració, as it has been said, is a Web-based tool that aims to support communication among co-authors. From the case studies there are three important observations. The first one is that the amount of words co-authors dedicated to discuss the system decreases drastically if we compare the first case study (Poster A) with the last case study (Short paper), from 82% to 13%. This might be due to various reasons: The most obvious reason might be the fact that those who used the system knew that we were developing a Web-based tool and it is logical that participants give some comments to improve the interface. Additionally, the very first versions were very simple and not thoroughly tested, therefore it was easy to find bugs and errors. Improvements to the system occurred after each case study. Most of the errors that were reported by co-authors were solved and many of the enhancements were implemented as well. Also, the participants gained more experience with the system.

It is worthwhile to highlight that in the case studies “Masters’ thesis proposal” and “Report Workshop” new users took part in the writing task. However, the percent of the number of words produced on the ATP document did not exceed 23%. We also have to bear in mind that many of the comments related to the system were made to suggest an improvement of the system interface.

The second observation to be noted is related to the “Ideas for this paper” section. This section appears to be really needed in a collaborative writing context. In six out of the seven cases in which this section was included it contained more than 25% of the words produced during the task. In the “Masters’ thesis proposal case” it only reached 16%. This result would seem to stem from the role co-authors were playing. One was the reviewer and the other was the writer of the proposal.

In the case study “Flyer” we found that 58% of the words were dedicated to this section alone. This can be attributed to the nature of the task. Co-authors were working on a short text that had already been written and they were supposed to expand it. So they were more concerned about how to structure the text, discussing what to extend and how to coordinate efforts on the text as a whole than on the content of the document. Another relevant point is that in the “Report Workshop” where co-authors were distributed in different countries and time zones this section got 74% of the words produced during this specific task. One interpretation of this would be that the need to coordinate a collaborative writing task in such a condition requires more effort from participants. In contrast to this observation we have the case study “Report Alpha”. In the former study, co-authors (2) were also in different countries but in the same time zone. In this case the co-authors’ role were defined as reviewer and author. As in the case “Masters’ thesis proposal” the need for coordination seems not to be so demanding as in those tasks that where co-authors had to interchange ideas and all were playing the role of authors.

The third observation is related to the comments co-authors made on the document’s sections. From figure 11 we can observe in the first five case studies an increase in the amount of words included in the document’s sections, 18%, 23%, 38%, 43%, and 61% respectively. This seems to indicate that the system indeed was getting more stable and that co-authors could concentrate on the task and not on the system design. However, it is important to distinguish carefully between the tasks that co-authors were accomplishing. The tasks where the role of author and reviewer were clearly defined, that is the reviewer was acting just during the revision process (case studies “Report Alpha” and “Masters’ thesis proposal”) seem to be more suitable for discussion on the content of document. In contrast, in tasks where co-authors had to coordinate their efforts (case study “Report Workshop”) or discuss and negotiate their ideas (case study “Flyer”) the discussion of the content of the document’s sections is blurred by the meta-discussion of the document and task as a whole. In our system that means that the discussion is moved over to the “Ideas for this paper” section.

6.1 The comments in the case studies

The last case study, “Short paper”, is the most interesting study. First, at this time many improvements to the system had been made and most of the errors reported were eliminated. Those that were not solved were known to the co-authors (4) and they knew how to avoid them. Also, during the first week, one of the co-authors worked in distributed mode in another country in the same time zone. That is, for a period of time this case study the team was working in distributed mode. However, the group activity was not interrupted. Co-authors were supposed to produce the text from scratch. Each co-author was responsible for one or more sections in the short paper. However, all sections were

open to other co-authors and they were given suggestions and even changing others' sections. In this case study, a sort of equilibrium was reached. The documents' sections got 50% of the words and the "Ideas for this paper" received 37%. It lasted 24 days and was the case study that produced most words in comparison with the others (6808) and lasted 24 days. I will take this case study as an example to make an analysis of the comments exchanged by co-authors. The principles discussed here cannot, however, be generalized because of the diversity of the collaboration contexts studied.

6.1.1 Discussion about the text

Many of the comments were oriented to discussion of the document content. One interesting strategy that was used by co-authors was to present a text that possibly could be part of the document content embedded in a comment. In several cases a text was sent as a potential alternative. This text embedded in the comment was discussed while it was in the comment-frame and when a mutual agreement was reached this text was included as part of the section. This suits nicely with the hierarchical structure layout that is contained in the system Col•laboració

I remind the reader that the content-frame is above the comment-frame. So, the potential alternative text is discussed firstly in the comment frame and when accepted it is "elevated" to the document section frame to become part of it. This situation was found 14 times.

To illustrate what I am saying in this section I will quote some of the comments that were sent by the co-authors. In the example number 1, the co-author Luis announced first that a new version of the abstract is ready and embeds it as part of the comment. Luis is not asking others to read because he expects them to read when the comment comes to them in the form of an email. On the other hand, Luis decides to wait for others' comments or for the version to be discussed before the abstract section of the document is updated. This might be related to the problematic interface that the system provided to change a section (see in section 3.1.2 for details) and that the text in discussion was relatively short (one paragraph).

1) **Comment # 6 99/11/18 12:22:54**

Luis:

This is a new version for the abstract.

Col•laboració is a Web-based collaborative writing tool. It focuses on communication, awareness, and sharing information. Surprisingly, many collaborative writing tools are focused on writing strategies, document control protocols, or co-authors' role during the writing task.

They overlook the relationship between communication and writing despite its importance.

Experiences from user studies have shown that the support for a dialogue among co-authors gives users a sense of the current state of the document and its transformation during the task.

6.1.2 Trigger of actions

Another characteristic presented in this case study was that the opinion of the co-authors was requested by another co-author. There were 11 cases when the repetitive question "what do you think" was sent as a comment. An indicator that could illustrate that co-

authors were interchanging ideas is the number of times that the sentence “I think” was found: total 49 times. This indicates that a negotiation of their ideas was going on.

An important event in a collaborative writing task is when a co-author feels this his/her part is ready to be read by the others. Then, an invitation to read his/her contribution is often offered. The comments are sent via email, they were also used to trigger the act of reading/commenting from the other co-authors. Twelve comments that explicitly alerted that a co-author’s contribution was ready to be reviewed by others were found in this case study.

2) **Comment #6 99/11/09 11:48:43**

Luis:

The new version of this section is on line, please read it and let me know what do you think..

3) **Comments #10 99/11/09 15:06:46**

Maria:

Hi! I have rewritten the introduction, i have changed a lot of sentences. Tell me if that reflects better what we would like to say.

6.1.3 The dialogue among co-authors

It is very important to support a dialogue among co-authors. The interface that Col•laboració provides indeed gives the sensation that co-authors are “talking” to one another. To support this assertion I will present some of the comments that were found in this case study.

4) **Comment # 10 99/11/18 16:36:58**

Isabel

...

I think the following parts are important, but they can be ordered in different ways:

1. Collaboracio is a Web-based tool for cooperation during the writing process

....

5...

What have we forgotten in the abstract? [my font]

6 ...

...

I think we have to find a way of including the most important of those elements. All of them don't fit in an abstract!

In example 4 Lorena is making a list of points that are considered important. After a while, five points have been described. It is very likely that Lorena stops to reflect and think “What have we forgotten in the abstract.”

5) **Comment # 7 99/11/18 14:57:15**

Maria:

Luis,

this is a negotiation, right?

Well, to tell you the truth when I read...

In example 5, Luis is named by Maria as while talking and also to index the comment. Observe that the style in example 4 and 5 is colloquial. That is, co-authors are engaged in an asynchronous dialogue.

6) **Comment #15 99/11/08 17:22:38**

Isabel

Here are two possible titles:

1. "Collaboracio: a Web-based tool for cooperation during the writing process"
2. "Collaboracio: a tool for collaborative writing on the Web"

Comment #16 99/11/08 17:43:33

Maria:

Title number one identifies quite well the purpose of Collaboracio, I think....

Comment #17 99/11/09 11:41:30

Luis:

...I aslo think that the title number one is better.

Comment #18 99/11/09 12:54:16

Frank:

I have no objection on both titles. I think it is good whatever we take.

In example 6, all the co-authors are involved in the discussion to decide the title of the paper after one of them suggested two alternative possibilities. It is very important to notice that between comment #15 and comment # 18 more than 18 hours have passed. However, none of the co-authors makes references to the comments that initiated the debate in the first place. Furthermore, they use the identification Lorena gave to the possible titles (1 or 2) to mention their preferred choice. This shows that if the topic to be discussed is presented in the shared space, all co-authors may share and make use of that knowledge also. Co-authors can make reference to what others have said using the same code (for example the identification number Lorena used) that others have used before. Also, the time that has passed do not affect the "team" memory. In example 6 that shows comments #15 to #18 no recapitulation is found as the context itself is presented.

One way to reference a previous comment is accomplished by referring to the comment's number, see example 7. Co-authors used the comment number when they were making reference to a comment that was not contiguous. In example 7 comment #10 is shown which make reference to comment #7. Comment #8 and comment #9 are in between them so co-authors solve this problem by explicitly naming to which comment he/she makes reference.

7) Comment # 10 99/11/18 16:36:58

Isabel

I agree with Maria's comments in #7.

This example indicates that the chronological order in which the comments are posted might create difficulties to follow the dialogue among co-authors as it might graphically be interrupted by other incoming comments, comment #8 and #9 in this case. Another situation is shown in example 6 where comment #15 to #18 nicely form a continuous flow. This, certainly, is the reason why co-authors did not use any device to make reference to a previous comment. Note that even though comment #18 has 2 different comments in between, Frank does not make explicit reference to comment #15. The contiguousness of the comments gives the flow to the dialogue and frees the co-author of using a device so his/her comments are placed in the right context by the others.

6.1.4 Co-authors' activities in the dialogue.

One interesting topic in computer supported collaborative writing is the sense of awareness co-authors have of other members' activities. I found that co-authors in eight cases during this case study explicitly let other members know about their activities. Four such cases are shown in example 8.

8) **Comment #31 99/11/22 12:45:53**

Luis:

...I will write about CGI, JavaScript and WebDav on thursday and the same day I will send the lines about it ...

Comment # 28 99/11/17 15:34:40

Maria:

...I will check the short paper again, print it in its official format and give you a hard copy to read it before I send it ...

Comment #14 99/11/06 18:32:24

Luis:

I will make a new version of the section Description of the system according to Isabel's recommendations. I will not be in the office next week but you can reach me via email.

Comment #10 99/11/03 09:48:33

Luis:

I moved the section "Description of the system" just after the Introduction, I think that is the right place.

Note, that in the last case of example 8 (comment #10) Luis describes the transformation that the document has gone through so that others were aware of the change. This is very important for team working.

Using email it takes usually 6-48 hours from a statement to a response. As a result, processes that need many interactions might need a lot of time using email (Palme, 1999). For example, to schedule face-to-face meetings might be time-consuming using only email. The fact that messages in the Col•laboració are available immediately to all co-authors and that they are in the same space presented in the form of a dialogue might reduce the interaction time in this mode of CMC. By clicking to a link in the system, co-authors do not have to wait for the email to arrive to check if a new message has been made. There can be situations in which co-authors might be aware that in a certain moment other co-authors are connected to the system . In that situation the system might reduce the gap of the interaction time and work in a semi-synchronous way.

Face-to-face meetings were in fact scheduled using the system and its comment function. Example 9 presents one sequence of comment in which co-authors were arranging a meeting.

9) **Comment #3 99/11/01 19:40:11**

Maria:

...Luis has a meeting at 16:00 with Isabel, What do you think if we meet between 17.00 and 17.30 ? how does it sound to you ?

Comment #4 99/11/01 19:40:33

Frank:

That's fine for me.

Comment #5 99/11/01 19:42:30

Luis:

Maria,

Looks like my meeting with Isabel will take a long time, we should change the meeting time.

Comment #6 99/11/01 19:42:45

Maria:

no problem, we meet on monday, Frank came and said that it is better to meet on Monday instead

This sequence took less than three minutes.

6.1.5 A window to the outside world

As the system is Web-based all the resources that are in the WWW can be part of the writing environment. Furthermore, co-authors can include these sources in two different ways. One method would be to insert an object inside the environment. An object can be anything Web-based, for example a picture, text, video clip, sound track, etc. The other alternative method would be to attach an object in the form of a hypertext link. The current version of the system supports this practice in a cumbersome way and it also requires HTML knowledge from the co-authors. In example 10 Luis has included a link to a website outside the system.

10) *Comment #1 99/11/01 19:15:50*

Luis:

...Click on the link "Short paper for XXX-00" [XXX-00 inserted by me]

As it was said before, the task for this case study was to write a short paper for a conference which had a Web site. Some administrative aspects related to the paper's submission to the conference were given in this site. Luis sent a comment with a hypertext link that opened a new browser window with the conference site. Doing so, it is possible to have a window to the world given by the WWW infrastructure. Co-authors can take from the outside world and put objects inside their "world" or attach them to it.

6.2 Problems with text-based communication within the system Col•laboració

In this section I will focus on some problems with the text-based communication I found in the comments sent by the co-authors.

6.2.1 Carriage return "noise"

In section 5.1.1 it was mentioned that co-authors very often made reference to the original text. In doing so, it is logical that instead of typing they just use the copy facility provided by the computer system. Very often this is done by a drag of the mouse on the text. Drag-copy from a Web-browser has the particularity that a carriage return sign is appended at the end of every line of the browser window. On the other hand, when a comment is written by a co-author, he/she could just press the return key when he/she thinks it is necessary. This action also produces a carriage return that breaks the text line. Line breaks

independent of the browser's windows size are produced by a special HTML tag,
. So, the system changes the carriage return it finds in the comment text into this HTML tag in order to preserve the line breaks that the co-author commanded. While making a comment, co-authors can also paste text that they have drag-copied into the input text area that could also contain some "added" carriage returns. There is no way for the system to identify whether the carriage return was produced by the co-author's pressing the return key or was appended during the copy-paste. For the sake of preserving the co-author's line breaks, the system transforms all carriage returns into
 tags. This produces a sort of "noise" in the reading activity. Moreover, there might be empty spaces that are not used to present the information, as shown in figure 6 in chapter III, section 4.3.1, page 59.

6.2.2 Expired Validity of Comments

Once a comment is added to a section it remains in the comment window even after its validity expires. For example, one comment suggests changing the word "arcane" to "unusual" and one of the co-authors decides to do the change. If any other co-author did not get to see the text while the word "arcane" was used, then the comment will be meaningless and it might create confusion. Example 11 shows such a situation in a real case.

11) Comment # 2 99/11/03 13:02:41

Isabel:

I find it difficult to understand what you have done so far in the Introduction, because the comments seem to refer to some other text than the one displayed.

Comment #3 99/11/03 13:21:53

Maria:

that's right Isabel, the comments correspond to an old version of this section. After a meeting I decided to rewrite the introduction and the version you read is the last one. At the moment, we don't have any comments on this last version of the section.

We have discussed several times what we should do with those comments whose validity has expired, but the situation is complex. The first thing that comes to mind is to delete them but that would produce other problems. For example, deleting a comment implies re-numbering all the previous comments. This will probably invalidate those comments in which co-authors used the comment's number to refer to. Also, one comment might address several issues and not all of them might fall into an expired validity situation. Finally, deleting a comment might break the flow of dialogue that the comments form. For the sake of avoiding these problems we have decided to keep those comments whose validity has expired and in this allowing a "design rationale". We hope also that this conflict will be resolved quickly and easily. In example 11 we can see that only 19 minutes after Lorena expressed confusion a response was given in which the situation was clarified.

Cadiz, Gupta and Grudin (2000) in their studies pointed out that a key complaint was the orphaning of annotations. That is, when the document got changed, the annotations lost the link to it. This was according to their interviews the main reason that people stopped using the annotation system. They point out that it might be frustrating for annotators to comment on a document, only to see the comments become meaningless through orphaning. In our system, though this situation can be found, co-authors can see their expired comments more from a historical perspective.

6.2.3 Probably too much “Talking”

In this case study we found a lot of communication between co-authors. The finished document that was submitted to the conference contained 1290 words. On the other hand, the text-based communication for this case study contained 6808 words. This is almost four times that of the finished document. Furthermore, there were some face-to-face meetings. Generally, this problem is common in CMC communication. The main reasons are that there is usually no time limitations as in face-to-face and that there is no a mechanism to stop the discussion (Palme, 1999). One question that arises is: does this tool encourage too much discussion among co-authors making the collaborative writing process more complex? Much work remains to be done to find an answer to this question.

V DESIGNING FROM INSIDE: THE COMMENT SPACE OF THE ATP DOCUMENT AS A DESIGN TOOL

For every case study of the DHS or Col•laboració the ATP document was part of the domain. Users and designers used this shared space for making comments about the system. These dialogues evolving around the ATP document had a crucial role in the development of the DHS and Col•laboració. This space for comments was used in very different ways, reflecting the users' varying background and interest in system design.

The experience that I had during these years in which I have developed the systems DHS and Col•laboració has made me reflect on the design of collaborative tools. It is clear that users participation in the design of information systems is very important. There are several ways to involve users in the design process. Participatory design (PD) and user-centred design (UCD) for example, foresee the importance of involving users as early as possible in the design process. The core of PD is to get active participation from users. "Something more than being [the user] used as mere data sources" (Muller, Hallewell, and Dayton 1997). One of the claims PD makes is that the designer should have the knowledge of the end user's culture and needs. To know this, designers have to communicate with the users e.g. engage in a dialogue. However, getting users involved in the design is not an easy task. Often, users are not willing to co-operate for diverse reasons e.g. giving away their time or inconvenient time for asking. On the other hand, designers might find it difficult to communicate with the users for reasons e.g. of time schedule. All too often, designers use professional communication forms that users might, very likely, find difficult to relate to. Therefore, too much translation might be needed when designers and users meet.

During our pilot study for the DHS, users got involved in the design of this system using the commenting feature of the system. We call this method "designing from inside". This method has been used in all of the case studies described so far in this thesis. The basis for designing from inside is to include a shared space "inside" the tool being developed as part of the system. Users and designers can add comments to this shared space and hopefully these comments would result in a text-based dialogue among them. The comments are sent to the design group via email. Furthermore, adding a comment should be easy and quick.

Designing from inside is a method that supports a) unobtrusive users' participation, b) communication and discussion among designers and users, c) problem identification on different platforms, d) repository of data, e) visual prototyping.

The examples that I will present here are comments taken from the ATP document of different case studies. These comments were made by users or designers. I will identify comments made by the designers. The names of the users and designers have been changed.

1 Unobtrusive user participation

Probably, the best moment to report about a problem or a need is just after it is experienced. At this moment, users might be willing to report it. Example 1 shows a situation in which the user express his opinion towards the system after he experienced a failure.

1) **Comment #17 98/02/17 17:21**

Benito:

I don't like when the comment window erase all text I have written when I resize the window. I would also like to have a comment window that is a little bit larger than it is today.

Comment # 23 99/05/18 15:25

Josefina:

I don't like the behaviour of the frames and their relative size....

If the users have at hand the way to express or report a problem, they could do so. Users might feel that it is important to show that he or she has “discovered” a bug or a problem. Users can also experience needs that the system is not able to fulfil. The lack of this feature might ensue a feeling of frustration. Reporting this need might be a way to respond to this frustration (see example 2). Users can unobtrusively, opportunely, and spontaneously get in touch with the designers.

2) **Comment #2 00/02/20 11:48:15**

Maritza:

I really miss the possibility to respond to a comment. ... Some kind of hierarchy in the dialogue structure would be great!

2 Supporting communication

2.1 Communication among users

Another point is that users can get in touch with each other and share their opinion about the system. They can just as well evaluate a suggestion made by another user or by the design group. Example 3 shows a situation in which one user suggests a change in the system and other users evaluate this suggestion (in the version used, the button to add a comment had the label “send”). After this dialogue we took the decision to label that button with “Add comment” as it is today in the current version of Col•laboració and DHS.

3) **Comment #7 98/02/16 10:07**

Armando:

...About “send” I suggest “add comment” or “add”. and I'd move clear far away from the ok button, so it makes a difference...

Comment #18 98/02/18 15:25

Susana:

...I agree that the Send-button is inappropriate, Add would be better....

Comment #22 98/02/19 12-50

Lucia:

Add or Send ? I would say Add, so that you recall the previously chosen “Add comment”....

2.2 Communication among designers and users

Users need to communicate with the designer to express their opinion. Example 4 shows how the designer approves the user's suggestion and tries to justify the current interface. It is important to show interest in what users say. Explaining to the users the reason why the system works in the way it does is also important because the users are informed and they can be more benevolent when a problem arises.

4) *Comment #6 97/04/03 15:16*

Mario:
Wouldn't "next" be more suitable than "more"...

Comment #9 97/04/03 16:58

Pedro:[designer]
Mario, (about Ref #6)
Yes, you are right.
Next is more suitable for this prototype. Let me tell you why the present prototype uses 'more' instead of 'next'...

2.3 Communication among designers

It is important that designers communicate with each other as in any other work that is done in collaboration. Face-to-face meetings in a working environment might be difficult to set up. Decisions should be discussed within the group. Example 5 shows the very first discussion from where the method "designing from inside" arose. Observe that the comments in examples 4 and 5 were made in the pilot study and also that they are related. See the number of the comments to follow the sequence.

5) *Comment #10 97/04/04 16:47*

Carlos: [designer]
Probably "more"~is~ a good name when there are more levels of information. Right now, "more" gets you to the next comment. We probably have to foresee the need for both...

Comment #11 97/04/04 16:53

Magnolia: [designer]
The iceberg model applies to the domain information, but not to the comments, as I understand it. I think "next" is a more suitable term for this reason...

Comment #12 97/04/04 16:54

Magnolia: [designer]
Right now, this is developing into an interactive discussion. I feel like we need other tools for this, since e.g. there is no way to see that there is a new comment. But on the other hand, we are not developing a conference system. What do the rest of you think?

Comment #13 97/04/04 18:27

Pedro: [designer]
I think that 'next' is more adequate for this prototype ... I will change it.

It is important to notice that the discussion in example 5 was carried out in a distributed-asynchronous mode. Note that the trigger of the discussion presented in example 5 was initiated by comment #6 in example 4. Each of the designers sent their opinion in relation to that comment. Three hours later the design group had taken a decision without a face-to-face meeting. Most importantly, they were taking into account the user's participation.

This dialogue was also available for the users. Designers took the decision to perform a change in the interface following the user's suggestion.

2.4 Problem identification and error repair by mutual participation

The lack of compatibility between different platforms, browser versions, and user set-up preferences of their computers make the WWW sometimes an unstable user environment. To simulate all the possible environments in which users might work would take a lot of effort and time. Users, who have their own, probably unique, environment, might experience some errors that no one else experiences. Therefore, problems or bugs can come exclusively to a particular user. This situation is almost impossible to detect unless the user reports it.

6) *Comment # 27 97/04/17 09:51*

Leonardo:

My browser is set to check one per session, it still doesn't give me anything else to see but "About this prototype"...

Comment # 37 97/05/05 23:41

Carlos:[designer]

Using a simple 640x480 pixel PC at home, I find that the comment window has resize handles, but still cannot be resized. (It can, however, be maximised...) ...

Example 6 shows two technical problems that the programmer did not experience in his own environment. It would have been very difficult to detect these errors if they had not been reported. In many cases these users were running the system in a different environment than the one used by the programmer.

Using the dialogue space, it is also very likely that errors which are not seen by the designer might be found and communicated by another person. Example 7 illustrates this situation. Mark discovered an error that is reported via the DHS. About an hour later the error was repaired.

7) *Comment # 7 97/04/03 15:35*

Carlos:[designer]

I just noticed that the mail links to Magnolia and yours truly are missing: "The requested URL /~Pedro/mail was not found on this server."

I would also like to be able to resize the comment window, but perhaps there is a reason for it to be fixed that I have forgotten? (Hopefully this, rather than the window itself, can be fixed?;)

Comment #8 97/04/03 16:47

Pedro:[designer]

Carlos, Your Ref # 7 was a mistake of mine in the HTML tag. Now it works. Thanks.

About resize the comment window there is no way as

2.5 A designer's logbook

Members of the design group, because of the ubiquitous accessibility of the WWW among other things, used this space as a repository of ideas or logbook. During the design process, ideas for the interface could come at any time, especially during the interaction with the system to be developed. Designers could annotate their new ideas, describe an error, or a reminder in this shared space. The nature of the written representation of the

idea and its public nature, somehow force designers to be more elaborate when presenting it. They know that other designers will visit this shared space. We observed that when an individual idea was reported in the system by one of the designers, it was, generally, communicated in a clear and explicit manner. It seems that this was done because the designer wanted the rest of the group to understand the idea. As usual, if the ideas are only jotted down, it is very easy to forget what it was about, even for the person who jotted it down. A more elaborate and careful recording of the ideas could ensure accurate recall. If the idea was not understood by others, they asked questions until the issue was clarified.

Text-based communication is usually more elaborate than spoken. However, the language used in computer-mediated communication (CMC), that is the one we use in designing from inside, is more informal and speech-like than in other text-based communication e.g. memos, newsletters. Furthermore, it is well known that we read faster than we talk and in case we do not understand what we read we can always re-read (Chesebro & Bonsall as cited by Fåhraeus 2000). The nature of text-based asynchronous communication also requires one comment to be sent at a time. This reduces the case of extroverted personalities dominating the “talk”.

However, I do not wish to imply by this that the design process is improved just because it is public and text-based. Designers, for example, might find it inconvenient not to be able to jot down an idea or to present a draft to their colleagues. One of the results of my investigation is that people might be very sensitive to present in public draft version of their work in public. A number of aspects of the problem requires further investigation.

In example 8, one of the designers made an annotation. Pedro labelled the comment with “Problem to solve”. It is followed by a description of the problem. No automatic solutions were given in this comment. In any case, the problem was made public, users and designers knew about it. This resulted in two things. Firstly, it alerted designers to the problem. As no solution was given any of them could come up with one. On the other hand, if a solution was presented, they could evaluate it. Secondly, users reading the ATP document were aware of the problem and they could avoid it or the error would not take them by surprise. Note that the designer might give a temporary solution to this problem, as Pedro does in the last sentence of his comment. This temporary solution is also learnt by the users.

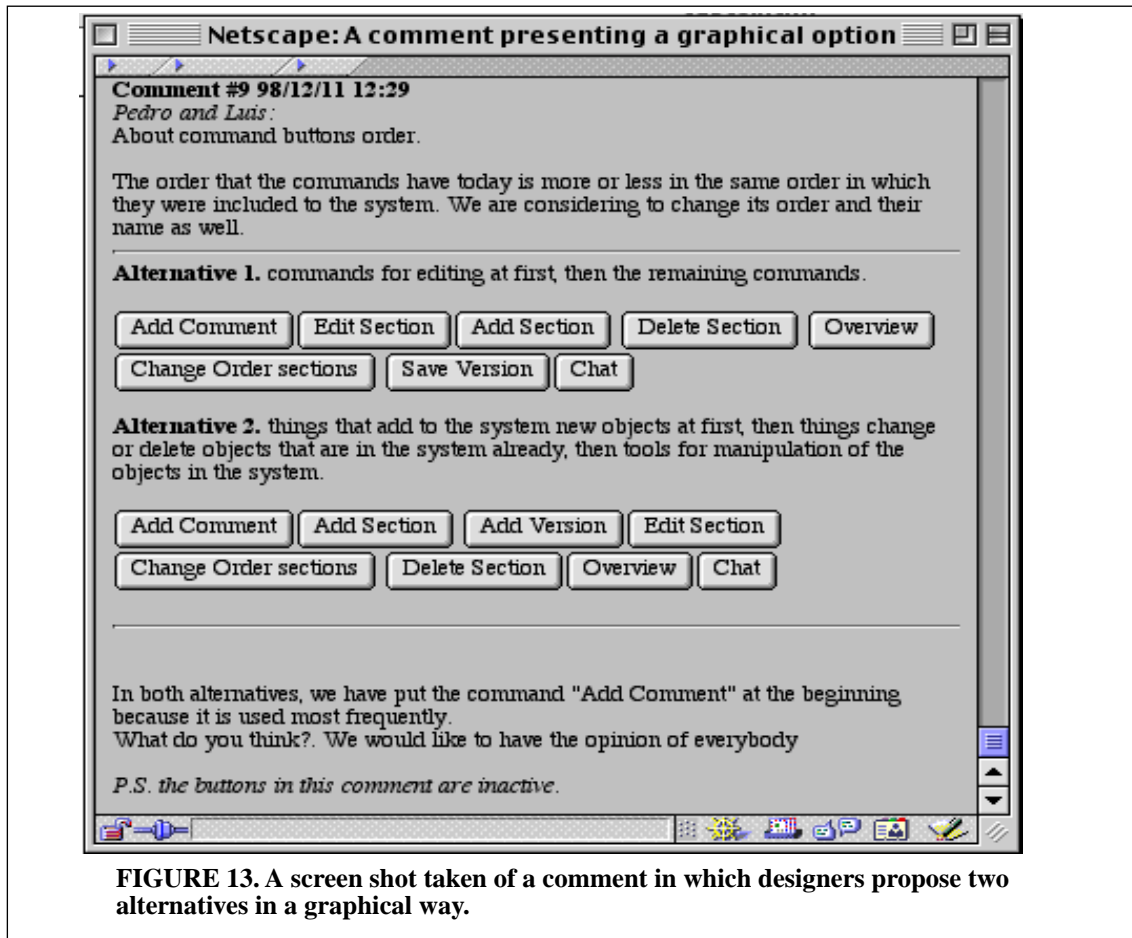
8) *Comment #14 98/12/15 12:59*

Pedro:[designer]

Problem to solve.

When a change is done to the left frame (add, delete, or order of section) and two or more users are working at the same time, changes are only reflected on the screen of the user who has made the change. For the other users this information will not be available until the left frame is reloaded again by them.

2.6 Visual prototyping



It is important to bear in mind that the comments could hold HTML tags, thus pictures could also be included in the comments. Also, as it was said before, this seems to work fine for Web-based tools. To prototype an interface using HTML might be easy and, most importantly, it would be very close to what the end-system would look like. Making a mock-up, on paper for example, would not be the same because designers and users have to “translate” this representation. As users and designers are running the system in what, very likely, would be their workstation, the principle WYSIWYG (what-you-see-is-what-you-get) is valid. Figure 1 shows a picture of one comment in which the designers, here Pedro and Luis, presented two alternatives and the rest of the design group were supposed to decide which alternative was the best. Observe that the comment presents the two alternatives in a clear way. The design group will see these alternatives in their own environment under the principle WYSIWYG. Therefore, the prototype is very close to what they will finally get, whatever the decision may be.

2.7 Social issues

There are other advantages that are parallel to the above-mentioned, but in this case more oriented to social issues.

Comments made on the system are public. A comment from a user might get an answer from the design group. This reply is presented in the shared space. A reply to one of the users, because of its public nature, is not only the answer to his/her comment but also an acknowledgement that the design group is “listening” to them. Yet, those users that observe the interaction will understand that it makes sense to comment on the system because comments from other users have been noted. This might encourage their participation in the design process. Moreover, a public question or suggestion expects a reply in a social context. As the comment is sent via email, a timely reply is likely to happen. A quick response, even if it gives no solution, could also show the users that the design group cares about them. This might create the feeling that there is a person “inside” the system who is taking care of the users and their problems. A feeling of trust from the users toward the designer might be developed, which might decrease the social gap between users and designers. This is very important to facilitate further collection of data from the users

2.8 A brief quantitative analysis of the users’ participation

We have counted and categorised the comments made in the ATP document in the pilot study. Additionally we have counted the number of words for the comment. The results were distributed in two groups in relation to who made the comment: a user or someone from the design group. In this way, we try to show the participation of the users in the design process of the DHS.

Table 1. Project member versus users’ participation

	Design group ₀		Users		Total
	Number	%	Number	%	
Comments	33	69	15	31	48
Words produced	1923	63	1106	37	3029
Error reports	6	43	8	57	14
Suggestions for the design	8	44	10	56	18
Statements/ opinions	9	64	5	36	14
Encourage the work	4	67	2	33	6
Question design related	4	80	1	20	5
Solutions/answers	10	83	2	17	12

The comments made by the users contained error reports, requirements, and also their impressions of the system (see table 1). Note that in the categories “Error report” and “Suggestion for the design” the participation of the users exceeds the one made by the

design group members. This, may be related to the face-to-face meetings that the design group had. However, it is important to emphasize that users suggested 10 different ideas for the design of the DHS and reported 8 different errors.

3 Conclusions

These experiences suggest that the quality of user-designer communication is an important aspect of successful design. User-participation in the design process of the DHS was in fact very active. This might be because they sensed that their comments in the shared space were going to be read, sooner or later. On the other hand, the public response from the designer also might have contributed to increased users' trust in the design group. Including a common dialogue space inside the system that is being developed is a simple way to increase users' participation in the design process. Clearly, collecting other data from users during the use of a system might be demanding or difficult for reasons of integrity. Giving the possibility to users of reporting their impressions of the system in an unobtrusive, voluntary way, and ensuring that what is said has been taken in account, might increase users' participation in the design, as well as promoting the ongoing discussion among designers.

VI CONCLUSIONS

In this thesis, I have presented two Web-based tools for collaboration: the Domain Help System (DHS) and Col•laboració. The main research issue has been how small or middle size groups work collaboratively to discuss, annotate, or produce a document on the Web. Also, as the DHS and Col•laboració have been developed from scratch, we have gained experience in issues related to the design and development of Web-based collaborative systems.

1 About the methods used

A common factor in all the case studies presented in this thesis is that the Web-based tools that users were utilizing were developed iteratively. That is, after each case study the tool was changed, trying to solve problems in the interface or to include new features. Changing the web-based tool for every case study, of course, makes it more difficult to generalize from the results; however, these changes have been motivated in relation to the users' interaction with the system. We have been exploring how people interact and communicate in order to accomplish a collaborative writing task using the Web as infrastructure. We have gathered information that helped shed light on a number of issues:

1. How people discuss the content of a document. This was done using DHS in the discussion context.
2. How people make digital annotations on an on-line document using an interface that allows only plain text input. This was done using the DHS in the annotation context.
3. How people engage in the production and collaborative reviewing of a short document. This was done using Col•laboració to accomplish different writing tasks.

The results and the conclusions of this thesis are based on subjective data; interviews and surveys, and objective data; the users comments made using the commenting function of the system and the log files. Although a great amount of data has been accumulated in the log files, they have only been used marginally. Participants in the annotation experience claimed that they were reading others' work and using the mark-up conventions that we gave. I browsed the data in the log file to confirm that claim.

I have to admit that it has been a problem to report on a project in which the tool we used has been changed in every case study. The main problem has been how to report the almost four years of the development history. It is also important to indicate that when we started the DHS project in 1996, the WWW was a relatively unexplored environment. At that time, few Web-based applications had been developed and programming on the Web was not an activity that I mastered. To this, we must add that the Web browsers in which we were expecting our prototypes to run were also changing. Web-browsers presented new features, therefore, new possibilities were available. A particular problem was that in many aspects the different Web browsers we used were responding in different ways, often divergently, to the same conditions, which made the programming a "headache". Furthermore, the different platforms in which users were running the system and their individual settings and preferences made programming on the Web very difficult.

Notwithstanding, many of these problems were solved and the current version of DHS and Col•laboració are more stable than the previous ones. I can state that because the use of the current versions have been characterized by:

1. A very small number of error reports
2. Users showed more interest on the task than on the tools
3. Results from the survey show that users have a positive attitude toward the systems.

The subjective data was collected just after the end of the tasks except for one survey made to one of the groups that used Col•laboració, which was made almost two weeks after the end of the task. Participants that supplied subjective data did that in only one way, i.e., they were either interviewed or fulfilled a survey. The only respondent that was interviewed several times was the teacher of the writing course in which the DHS was used as an annotation tool.

For the case studies in which Col•laboració was used, we decided not to have interviews. Only in two of the case studies a survey was used. The reason why we were reluctant to gather more subjective data for these case studies was simple. Most of the participants in the case studies were involved in the development of Col•laboració and this could create biased answers. One reason why we took part in many of the case studies in which Col•laboració was used was because experiencing the tool in a real task allowed us to understand its impact on collaborative writing processes and to identify the system's shortcomings and benefits. In one of the cases in which a survey was collected, we excluded those who were related with the project. In the other case, only one of the four participants for that writing task was involved in this project and completed a survey. Most of the data collected for the study of Col•laboració was therefore objective data.

One point worth noting is that all the case studies presented in this thesis were field studies. Grudin (1988) observes that "it is difficult or impossible to create a group in the lab that will reflect the social, motivational, economic, and political factors that are central to group performance" (p. 87). In the D-sample (the CSCW course) students had to discuss the content of a document, in the A-sample (the writing course) students had to make annotation to others' text. Both samples used the DHS. In the case studies in which Col•laboració was used participants engaged in a real writing task. This gives validity to the objective data collected in these cases.

2 The discussion experience

Web-based tools that support discussion are very common. Frequently, those discussions tools are text-based, topic-oriented, and embedded in virtual communities. Many of these forums are oriented for mass communication and support several topics at the same time, e.g. Usenet newsgroups. Others are oriented to support defined groups (as students from the same class) discussions, e.g. WebCT, FirstClass, and TopClass. The DHS can also be classified as a discussion forum but with the particularity that it is oriented to discuss the content of a document. The interface of the DHS has intentionally been kept simple. The main reason for keeping it simple was that we wanted users to concentrate as much as

possible on the task: to discuss the content of the document. Nevertheless, we understand the importance to support administration features for moderators, rating and voting, group calendars, etc. From the experience when the DHS was used as a discussion tool, we can draw several conclusions:

- The discussions were focusing on the content of the document. Users very often pointed out the content of the document. This might be because the system supports an overview of the document and its comments together. The document and the comments are presented in independent scrollable frames that are one under the other but the relation between the document content and its comments is always maintained. Having the document content visible and easy to access while making a comment on it, might keep the focus of the discussion.
- Though users could have an overview of the comments as a whole and navigate easily through them, they found it difficult to follow the discussion when a comment was a reply to a previous comment that was not adjacent to it. Threading requires commonly a subject line of the message and the relation “reply-to” this subject line. The origin of the problem in DHS is the chronological order in which the comments are presented. One solution might be to let users re-locate the comments using the number id of the comments. Users employed extensively this number when they replied to a comment that was not the very last one in order to help readers to find the context. Giving the possibility for users to re-locate the comments adjacent to its context, that is, adjacent to the comment it intends to respond, might help them to follow the discussion as a dialogue. However, the possibility to re-locate comments in the comment space merits further investigation
- The dual context for the comments is supported by the layout of the system. A comment may refer to another comment and to the Web-document at the same time.
- Email notification about comments on user’s document supports awareness. The email itself reminds users about the task.

3 The annotations experience

The DHS was used as an annotation tool in four successive academic writing courses. Annotations can be created to fulfil different functions (e.g. planning, revising) and directed to particular audiences (e.g. the next reader, a co-author, oneself) and have distinct uses (e.g. to serve as a reminder, to request a change). For example, during the planning stage, the writer will probably make a note of some books, articles, or situations he or she would like to refer to in the text. During the reviewing process, the reviewer role can be played by the author or by a third party. Clearly, an annotation by a third-party reviewer demands more elaboration than if the author-reviewer is one and the same person. Annotations made by third-party reviewers can also have distinct purposes. They could be used to advise on reading material, or schedule a meeting with the author. Some annotations can be classified as corrective. A corrective annotation aims to inform or persuade authors to perform a correction (change) in the text in question. It could be as simple as marking a misspelling or suggesting a language style change; making corrective annotations mainly implies suggestions to add, delete, move, or change the text.

The limitations of the annotation experiences become evident if one considers that in these studies only corrective annotations were encouraged. However, this does not mean that the study was not valuable. From the DHS experience used as a corrective-type annotation tool, we can draw a several conclusions:

- Reviewers need access to the original text on the screen, both when commenting and when making corrective annotations. This is in line with Neuwirth's (1990) recommendations.
- Private posting of corrective annotations should be available. The public nature in which comments are presented in the DHS might inhibit reviewers to point out shortcomings or errors in the document.
- When using plain text the parenthesis seems to be the most used symbol mark for representing recommended changes. It could be a good idea to avoid the use of parenthesis for making a corrective annotation since they also might occur in the text and might be confusing. If parenthesis is used, the suggestion has to be explicit, the action has to be verbalized, for example:
(replace Unlike with But) Unlike to (omit the word to) the concentrated narrow Web in an I-beam
- A common markup system or change representation needs to be created among reviewers and writers before they engage in the collaborative writing task. These conventions should be not only defined but also available during the whole revision process so that reviewers and writers could easily retrieve them when needed. In our system, the conventions that I defined, though far from optimal, were always available for the students through the system. Most of the students used these conventions and according to the log file they often retrieved them during the whole course. Not surprisingly, some students developed their own conventions for annotation marks but most of them were based on the ones we gave them or were similar.
- The ordinal position of the paragraph and sentence in the text was often used by reviewers to help writers to find the correction context. This implies that there should be a mechanism that automatically indicates the ordinal position of paragraph and sentences for reviewers and writers.

4 The collaborative document production experience

Col•laboració has been used to perform different collaborative writing tasks in which mainly small groups had to produce a document. Communication plays a crucial role in any collaborative writing task. Participants need to communicate and coordinate their actions and for the sake of general understanding he/she has to motivate them for the rest of the participants. Yet, they could ask for clarification of other participants' actions. In any writing process, very likely, a document goes through many changes. It is important for co-authors to be aware of these changes and, furthermore, to understand them.

Col•laboració is a prototype for collaborative writing and not a finished application ready for the market. There are points that need to be expanded, improved and examined more closely. However, Col•laboració is a working collaborative writing tool despite all its

shortcomings. The most important part missing is that related to the translation phase from ideas to text. In other words, editing the document text is clumsy and laborious. Version management is also very limited in the system. So far, the system allows a user to make a version of the document but automatic recovery is not supported. In fact, rather than a collaborative writing system, it is a document sharing system for collaborative reviewing, discussion and annotation.

Common ground refers to the mutual knowledge, beliefs, and assumptions of the participants in a conversation (McCarthy, Miles and Monk, 1991). Clark and Shaefer (1989) refer to the process of reaching a common ground as grounding. They present a descriptive model, in terms of presentation and acceptance phases that allow them to track the augmentation of common ground as the conversation proceeds. Also, they predict that common ground is more difficult to establish in communication based only in text. However, MacCarthy et al. argues that providing a common report space, significantly reduce the problem of achieving a common ground. All the links in the index-frame of system Col•laboració are a “common report space”.

Col•laboració focuses explicitly on the need for communication that co-authors have during the collaborative writing process. In other words, the need to have a dialogue is supported. Also, being text-based, the dialogue it supports is permanent. It can be retrieved, reused, and reflected on over and over again by any of the co-authors independently of time and place.

This dialogue could be related to both the document content and annotations made to the document. The dialogue could be associated with a meta discussion about the writing task. Certain facets of planning the writing task might be included here. For instance: the length of the document, global ideas of the document, coordination, etc. It is remarkable that co-authors extensively used the space “Ideas for this paper”, the space designated for a meta-discussion. Not surprisingly, in those case studies in which interactivity among co-authors was expected to be low, this space was not used very much. For example, in the “Masters’ thesis proposal” in which one of the co-authors played only the role of reviewer, this section received merely 16% of the words of comments.

Co-authors might need to “talk” about the technology they are using to accomplish the writing task. It is important that co-authors can discuss about 1) the document content, 2) the system and the technology they use, and 3) the co-ordination of they effort. These three aspects should be discussed independently of each other. If the collaborative system supports these aspects keeping them interrelated but separately, co-authors may easily not only decide where to place their comments but also speed-up a search of one specific topic. For example, suppose that all co-authors have been “talking” about how to overcome a shortcoming of the system in the space that has been designated for this. If one of the co-authors wants to recall that conversation later, it is very likely that he/she first might explore the space that has been designated to communicate about the system. As a result, the search is speeded up. This is supported in all topic-related discussion forums but not related to documents. Moreover, it allows co-authors to navigate easily and interchangeable between these topics.

Co-authors made use of the commenting space:

- To make others aware of their current activities. Co-authors indicated which task they were engaged in during the writing task.
- To ask co-authors to read and comment on their text.
- To share URL:s to other Web sites.
- For decision making in relation to the task and co-authors activities. For example who does what and when, coordination of face-to-face meetings.

Co-authors might feel the need to discuss the text before they make it part of a section. They sometimes submitted the potential text as part of their comment, that is, this potential text was presented first in the comment-frame and discussed. If approved, it was made part of the relevant section. One reason might be that the translation phase of the writing is poorly supported by the system. Users might want to minimize the amount of editing, avoiding insertion of a text that probably will change. It might be better to discuss it first in the comment space and later make it part of the section.

Col•laboració keeps the current version of the document available for all the co-authors. Co-authors can browse the whole document if desired, even those parts that they are not responsible for. This increases the sense of how much is done of the whole task and awareness of others' activities.

Sending email has been shown to be effective as a reminder of the task. However, users indicated that in case a lot of comments are made, then receiving every comment via email might be problematic and undesired.

5 Final words

The WWW offers an infrastructure that embraces many activities central to collaborative writing. It can be used as a shared space in which co-authors exchange documents, communicate, and co-ordinate their efforts, but also be aware of the changes that have taken place on the documents to be produced. The studies made here have shown that such communication, coordination, and awareness can be successfully maintained through text-based dialogue among co-authors, and that this dialogue is also a valuable instrument to develop collaborative systems "from inside".

REFERENCES

- Appelt, W., Busbach, U. (1996). The BSCW System: A WWW-based Application to Support Cooperation of Distributed Groups. In: IEEE (ed.): Proc. of the Fifth Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, June 19-21, 1996, Stanford, CA. Los Alamitos, CA: IEEE Computer Society Press, pp. 304-309.
- Baecker R., Glass G., Mitchell A., and Posner I. (1994). SASSE: The Collaborative Editor. Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems, Volume 2, pp. 459-460.
- Baecker, R., Dimitrios, N., Posner, I., and Mawby K. (1993). The User-Centred Iterative Design of Collaborative Writing Software Meetings and Collaborative Writing, Proceedings of ACM INTERCHI'93 Conference on Human Factors in Computing Systems, pp 399-405.
- Beck, E. and Bellotti, V. (1993). "Informed Opportunism as Strategy: Supporting Coordination in Distributed Collaborative Writing", in Proc. Third European Conference on Computer-Supported Cooperative Work ECSCW'93 (Milano, Italy), pp. 233--248.
- Bälter, O. (1998), Electronic Mail in a Working Context. Doctoral thesis, TRITA-NA-9820, Royal Institute of Technology of Sweden.
- Cadiz, J., Gupta A., Grudin J., (2000), Using web annotations for asynchronous collaboration around documents, Conference proceedings on Computer supported collaborative work (CSCW'00), pp 309-318
- Carroll J. M., Kellogg W. & Rosson M. (1991), The task-artifact cycle. In Designing Interaction: Psychology at the Human-Computer Interface Eds. Carroll J. M. Cambridge Series on Human-Computer Interaction pp74-102, Cambridge, UK Cambridge University Press
- Caudery, T. (1995). What the "process approach" means to practising teachers of second language writing skills. TESL-EJ, 1(4), A-3, URL: <http://www-writing.berkeley.edu/TESL-EJ/ej04/a3.html>
- Churchill, E., Trevor, J., Bly, S., Nelson, L., and Cubranic, D. (2000). Anchored Conversations: Chatting in the Context of a Document. Proceedings of the 2000 ACM Conference on Human Factors in Computing Systems (CHI 2000) Page 454-461.
- Clark, H. & Shaefer, E., Contributing to discourse, Cognitive Science, Vol. 13 (1989), pp. 259-294.

- Ellis, L. and Gibbs, S. J. and Rein, G. L., "Groupware: Some Issues and Experiences", *Communications of the ACM*, Vol. 34, Iss. 1, pages 38-58, January 1991
- Fish, R., Kraut R. and Leland, M. (1988), Quilt: a collaborative tool for cooperative writing; Conference Sponsored by ACM SIGOIS and IEEECS TC-OA on Office information systems, pp. 30-37.
- Fåhraeus, E. (2000), Growing Knowledge: How to support collaborative learning e-discussions in forum systems, Licentiate thesis, Department of Computer and Systems Sciences, Stockholm University, Stockholm, No 00-005, ISSN 1101-8526.
- Galegher, J., Kraut, R. (1992), Computer-mediated communication and collaborative writing: media influence and adaptation to communication constraints; Conference proceedings on Computer-supported cooperative work, 1992, Pages 155 - 162
- Grudin, J., 1988. Why CSCW applications fail: Problems in the design and evaluation of organizational interfaces. In: Proceedings of the Conference on Computer-Supported 28 Cooperative Work (CSCW88), ACM Press, New York: 8593. pp 85-93.
- Haake, J., and Wilson, B. (1992). Supporting Collaborative Writing of Hyperdocuments in SEPIA. In Proc. of the 1992 CSCW Conference, pp 138-46.
- Haake, J.M. & Wilson, B. (1992). Supporting collaborative writing of hyperdocuments in SEPIA. Proceedings of the Conference on Computer-Supported Cooperative Work. 31 Oct.- 4 Nov. 1992. (pp. 138-146). New York: Association for Computing Machinery.
- Haas, C. (1989). Does the medium make a difference? Two studies of writing with pen and paper and with computers. *Human-Computer Interaction*, Vol. X, 149-169.
- Haas, C. (1996). *Writing Technology: Studies on the Materiality of Literacy*, Lawrence Erlbaum Associates, ISBN 0-8058-1306-3
- Halasz, F., Moran, T. and Trigg, R., "Notecards in a Nutshell", Proceedings of the ACM CHI+GI Conference 1987, Toronto, Ontario, April 1987, pp. 45-52.
- Hayes J. R. and Flower L. S. (1980). Identifying the Organization of Writing Processes, Lawrence Erlbaum Associates, Hillsdale, NJ, pp 3-30
- Horstmann, T., Bentley, R. (1997). Distributed authoring on the Web with the BSCW shared workspace system; *StandardView* 5, 1 (Mar. 1997), pp 9-16.
- Hughes, R., Jake S., Okelberry, C. (1998). Ceilidh: collaborative writing on the Web; Proceedings of the 1998 ACM symposium on Applied Computing, Pages 732 - 736

- J. McCarthy, V. Miles, A. Monk, An Experimental Study of Common Ground in Text-based Communication. Proceedings of the ACM Conference on Human Factors in Computing Systems, 1991, 209-215.
- Kellogg, R.T. (1989). Idea processors: Computer aids for planning and composing text. In B.K. Britton and S.M. Glynn (Eds.), *Computer writing environments: Theory, research, and design*. Hillsdale, NJ: Erlbaum.
- Kellogg, R.T. (1994). *The psychology of writing*. New York: Oxford University Press.
- Kim, C-H, (2000). On-Line Reviewing with Change Representation Tools. In Proceedings of NordiCHI'2000, STIMDI, October 2000.
- Kim, H-C. and Severinson-Eklundh, K. (1998). How academics Co-ordinate their Documentation Work and Communicate about Reviewing in Collaborative Writing. Report TRITA-NA-P9815, IPLab-151, Royal Institute of Technology, Stockholm.
- Knister, M., and Prakash, A. (1990). DistEdit: A Distributed Toolkit for Supporting Multiple Group Editors. Proceedings of the conference on Computer-supported cooperative work, pp 343-355.
- Kraut, R. E., Galegher, J., Fish, R. S., and Chalfonte, B. L. (1992). Task requirements and media choice in collaborative writing. *Human-Computer Interaction*, 7, pp 375-407.
- Malone, T. and Crowston, K.: 1990, What is Coordination Theory and how can it help design cooperative work systems?, Proc. of the Conf. on Computer Supported Cooperative Work CSCW'90, ACM Press, Los Angeles, California.
- Marshall, C. (1997). Annotation: from paper books to the digital library; Proceedings of the 2nd ACM international conference on Digital libraries, pp 131-140.
- Mhashi, M., Rada R., Beck E., Michalidis A., Zeb A. (1992) Computer-supported discussion and annotation, *Information processing & management*, vol 28, N 5, pp 589-607
- Muller, M., Hasllewell, J. and Dayton T. (1997), Participatory parctices in the software lifecycle, In *handbook of human-computer interaction*, Helander, M., Landauer, P., Prabhu, P. (eds.), Elsevier Science B.V., 2nd edition, pp 255-269.
- Neuwirth C., Chinch R., Carney D., Wojahn P. and Kim L. (1994) Distributed collaborative writing: a comparison of spoken and written modalities for reviewing and revising documents. Proceedings of the CHI'94 conference companion on Human factors in computing systems, Page 20
- Neuwirth, C .M., Chandok, R., Kaufer, D.S., Erion, P., Morris, J., & Miller, D. (1992). Flexible diff-ing in a collaborative writing system. Proceedings of the Conference on

- Computer- Supported Cooperative Work. 31 Oct.-4 Nov, New York: Association for Computing Machinery, pp. 147-154.
- Neuwirth, C. M. Chandh, R., Charney, D., Wojahn P., and Kim, L. (1994). Distributed collaborative writing: a comparison of spoken and written modalities for reviewing and revising documents. Proceedings of the CHI '94 conference companion on Human factors in computing systems, 1994, p 202.
- Neuwirth, C., Kaufer, D., Chandhok, R. and Morris, J. (1990). Issues in the design of computer support for co-authoring and commenting, Proceedings of the conference on Computer-supported cooperative work, pp 183-195.
- Neuwirth, C., Kaufer, D., Chandhok, R. and Morris, J. (1994). Computer support for distributed collaborative writing: defining parameters of interaction, Proceedings of the conference on Computer supported cooperative work, 145 -152.
- Neuwirth, C., Kaufer, D., Chimera, R. and Gillespie, T. (1987). The Notes program: a hypertext application for writing from source texts, Proceeding of the ACM conference on Hypertext, pp 121-141.
- Nystrand, M. (1989). A social-interactive model of writing. *Written communication*, 6, 65-85.
- Ovsiannikov Iliia A., Arbib M., McNeill T, (1999) Annotation Technology international Journal of Human-Computer Studies 1999 v.50 n.4 p.329-36
- Palme, J. (1999), Support for decisions by E-mail, URL: <http://dsv.su.se/jpalme/ietf/JPMADS.html>
- Palme, J., Karlgren, J., Pargman, D. (1995) Issues when designing filters in messaging systems, Technical report No 95-008 at Department of Computer and Systems Sciences, Stockholm University, Stockholm
- Rimmershaw, R. (1992). Collaborative writing practices and writing support technologies. *Instructional Science*, 21(1-3), pp 15-28.
- Rodríguez, H. (1999). The Domain Help System. Technical report TRITA-NA-P9912, CID-56, NADA, The Royal Institute of Technology of Sweden. URL: http://cid.nada.kth.se/pdf/cid_56.pdf
- Rohman, G. (1965). Pre-writing: The stage of discovery in the writing process. *College Composition and Communication*, 16, 106-112.
- Romero, M. and Decouchant, D. (1997). Structured cooperative authoring on the world wide web, *CSCW: the journal of collaborative computing*, 6(2-3), pp 157-174.

- Roundy, N. and Mair, D. (1983), The composing process of technical writers: a Preliminary study. JAC 3, URL: <http://jac.gsu.edu/jac/3/Articles/10.htm>
- Severinson-Eklundh, K. (1990). The use of idea processors for studying structural aspects of text production. In Nordtext Symposium Text structuring - reception and production strategies, Hanasaari, Helsinki, .
- Severinson-Eklundh, K. (1992). Problems in achieving a global perspective of the text in computer-based writing. Interaction and presentation laboratory IPLab, Royal institute of technology, 1992, pp 73-84.
- Severinson-Eklundh, K. (1998). To quote or not to quote: Setting the context for computer-mediated dialogues, Interaction and Presentation Laboratory IPLab, Royal institute of technology, 1998, pp 1-31.
- Sharples, M. (Ed.). (1993). Computer supported collaborative writing. London: Springer-Verlag.
- Shneiderman, B. (1998). Designing the User Interface: Strategies for Effective Human-Computer Interaction: 3rd ed. Addison Wesley, 1992, ISBN: 0-201-69497-2.
- Smith, S. (1997), The genre of the end comment: conventions in teacher responses to student writing, College Composition and Communication, 48(2), pp 249-268.
- Whitehead, E. J. Jr. & Goland, Y. Y. (1999). WebDAV: A network protocol for remote collaborative authoring on the Web. In Proceedings of the Sixth European Conference of Computer Supported Cooperative Work, Copenhagen, Denmark, pp. 291-310.
- Whittaker, S., Terveen, L., Hill, W., & Cherny, L., (1998). The dynamics of mass interaction, Proceedings of ACM. CSCW'98 Conference on Computer-Supported Cooperative Work, pp 257-264.
- Wojahn, P. G., Neuwirth, C. M., and Bullock, B. (1998) Effects of interfaces for annotation on communication in a collaborative task, Conference proceedings on Human factors in computing systems, pp 456-463.