Voice Annotation of Visual Representations in Computer-Mediated Collaborative Learning

Christine Steeples
C SALT (Centre for Studies in Advanced Learning Technology)
Department of Educational Research
Lancaster University, Lancaster, UK
Email: C.Steeples@lancaster.ac.uk

Abstract

New computer-based communications technologies—such as electronic conferencing, electronic mail, communal hypertext, and communal hypermedia databases—make it possible for people to collaborate in their learning, even when separated from one another in space and time. The development of computer environments supporting collaborative learning in enriched ways provides the goal and context for the research outlined here. This chapter explores the potential value of voice annotations in computer-mediated collaborative learning. Some purposes for voice annotations are examined. Finally an outline of a small-scale study where voice annotations have been used to support collaborative writing processes is given.

1   Introduction

New computer-based communications technologies—such as electronic conferencing, electronic mail, communal hypertext, and communal hypermedia databases—make it possible for people to collaborate in their learning, even when separated from one another in space and time. The development of computer environments supporting collaborative learning in enriched ways provides the goal and context for the research outlined here.

This chapter begins by unpacking the notion of enriched computer-mediated collaborative learning. This is done to give the context of this research and to explain how a number of key threads have been drawn together in this work. Details of a recent small-scale study of computer-mediated collaborative writing are then given. This study has examined the potential value of voice annotations in collaborative writing processes.

2   Computer-Mediated Multimedia Communication Developments

There are a number of advantages to be derived and exploited through the use of asynchronous computer-mediated communications (CMC) in learning settings. Using CMC, participants are not required to be working simultaneously, so everyone has the freedom to work at times which suit them best, within agreed periods. This flexibility—breaking free from the constraints of fixed meetings—should not be underestimated. More subtly, CMC can support opportunities for broader discussion in terms of both available time and the range of participants able to join in. For the individual learner, reflection and self-pacing are thus supported.

Such communication is at present largely text-based (e.g., electronic mail and computer conferencing systems). Developments to enable richer communication through computer systems, allowing the integration of not only text, but also images/graphics and audio, have been piecemeal and fragmented thus far. Software designers are still seeking to establish coherent standards for multimedia development and delivery platforms. Concurrently as new systems with enhanced functionality emerge, new ways to use these systems, making new demands of them, begin to surface.

Multimedia-type communication facilities are now available within localized mail systems and word-processing software. Software developers have incorporated the facility to annotate a document using both audio and textual annotations. It is possible to integrate graphics into documents that can be attached to mail messages and sent over a local area network. The near future should see integration and enriched communication facilities operating over wider networks. The implication of these developments is that there will be a future where remote learners can collaborate on, and communicate with each other about, multimedia documents or courseware. In other words, a future for "learning together, apart" [1] that draws strength from the flexibility that computer-mediated communication offers us (e.g., in terms of timing and location) but that also takes advantage of the developments towards broadband communication.

Technological developments need to be realized in systems that support users (learners) in naturalistic ways, rather than circumscribing their practices. This research is therefore focused on understanding more clearly the potential value that multimedia communication might offer to remote learners and in particular focuses upon the role of audio information in flexible patterns of communication.

3 Why Collaboration?

Collaboration can be examined from two perspectives. Collaboration in learning tasks or activities and the skills acquired and developed in learning to collaborate.

Collaboration in learning activities can offer a number of benefits in respect of distributed learners. The salient features of collaborative activities advocated by a number of researchers [2, 3, 4] include that collaboration:

- encourages active and deep engagement in the learning process

- focuses upon the learner as an active participant in learning activity

- supports the sharing of individual experience and context

- requires articulation to make manifest an individual's understanding and expose deficits

- promotes the development of meta-cognitive skills of concept refinement and revision

- supports the human characteristic of working (and learning) with others

Learning to collaborate is part and parcel of the process of learning to work with others, demanding the integration of subject-specific skills with more generally applicable communication and co-ordination skills. These are skills required for effective co-operation and communication in work and increasingly in learning settings.

4 Enriched Communication Support—The Research Focus

This research is interested in the potential of audio in learning settings: to investigate the ways in which voice annotations can be used to enrich, but also to simplify communication between learners and learners (and between learners and tutors) over computer networks. It is also interested in whether the use of audio in computer-mediated settings offers or requires significantly different ways of working (and learning) from communication that is purely textual.

Let's say a little about voice annotations. Sound and vision are complementary forms of communication. One can explore how audio may be used to enrich or to ease the use of visual information in learning settings.[5] One potential to imagine for using audio in IT-based learning is to develop voice-annotated representations. Voice annotated representations may offer a means of allowing the user to focus his/her attention on the visual representation while listening to the commentary of another user thus giving the learner different perspective(s) on the information while retaining attention on the visual representation. Voice-annotations have the potential to convey information that is difficult or awkward to display in text or graphics. Audio may be used to provide information that is difficult to visualize. Using audio may also reduce visual clutter in the computer interface. It is also an alternative means to present information and one can investigate how and when audio might be appropriate as an alternative but complementary support to visual information, whether text or graphical representations [6].

The next section outlines how voice annotations might be used. This consideration of purposes has been a useful first step to realizing a set of investigations. It has enabled the formulation of a voice annotation matrix. The purposes are grouped around two dimensions:

- the use of voice annotations in collaborative production processes, and

- voice annotations used to enrich a shared resource.

5 Potential Applications of Voice Annotations

5.1 Production of a Document and Collaborative Writing

Much of the learning activity in Higher Education revolves around, or is made manifest in, written documents.[1] It is also generally agreed that writing can be an extremely difficult task requiring tremendous mental effort. One potentially valuable purpose for voice annotations could be in easing the cognitive load in short-term or working memory in the process of writing. Whether working alone or collaboratively, the writer could use the audio annotation facility to record his ideas while struggling to find the exact wording he wants to use in the text. Voice annotations could be used to rapidly capture a whole set of ideas. It could be a quick and easier way of making notes while working through conflicting ideas.

Collaborative writing tasks involve considerable social and intellectual complexity. Authors need to find mutually acceptable ways to communicate and share their ideas in working towards the common goal of a coherent text. They need to negotiate and assume roles and responsibilities to coordinate their activity. Collaborative writing processes have been studied extensively, particularly in their relationship to the technologies employed to support those processes [7, 8, 9].

---

[1]This section is only considering audio annotations. Some of the purposes or applications might make use of other (say textual) annotations. The interest here is what is different about using (and hearing) audio annotations rather than any other type of annotation. However there are some purposes envisaged for audio annotations where the audio might make a unique contribution that could not be replicated by (say) a textual annotation.

In deciding how to organize and coordinate a collaborative writing task three main patterns of practice have been identified:

1. Each person takes responsibility for a section and their efforts are then exchanged and blended into the final document.

2. Joint responsibility is maintained throughout—it is written entirely together.

3. One person assumes control, writing the first draft and coordinating the task—taking comments and revisions from the others in the group.

In any of these practices, voice annotations might be useful. They could be used to aid the creative process and to support the coordination of the activities. Voice annotations might provide a more personal way to communicate and receive complex, equivocal or controversial information as Kraut et al. [10] suggest. For example, it is argued that using voice not only allows the speaker to express his thoughts in a personal way but it does cause the speaker to consider the audience— the person(s) who will listen to their message. Conversely they argue that in writing there is a tendency to concern oneself with the content more than the audience, which could impact upon the process of writing with others.

Voice annotations could provide an easier and quicker way to express, explain, or justify a revision made. Easier in the sense that it is often easier to say rather than write, and often it can be quicker to speak than to type in text.

Voice annotations could be used to give contextual or background information. It is information to share with collaborators but not necessarily include in the actual text. An annotation allows the separation of the additional or contextual information from the actual document. By using a voice annotation the flow of reading through the text is not disturbed in the way that an inserted textual annotation might interrupt the flow of reading. Giving textual comments at the end of a document or section requires the reader to switch attention. A voice annotation could allow the reader to maintain visual attention to the actual text.

Writers have commented upon the need to maintain a global view of the text they are trying to achieve. Working with a word processor can make difficult that sense of the whole. A possible use for audio annotation could be to give high level or meta-comments that are independent of specific points in the actual text. These comments could be perhaps about the goals, purposes, and structuring of the document.

Uninterrupted flow in reading through a text can be important when creating a group-based, collaborative document. When the reader is not the author it can be useful to read through without breaks in the text, before attending to the comments of others. In a larger group where many might be adding in comments or suggested revisions, multiple textual annotations or insertions in the text would make difficult a coherent reading of the document. Voice annotations would be visually referenced to give the visual signal at the point of issue (or close to it) that there is an added comment. The reader can note this as he reads and can make his own choice about accessing the annotation.

Using voice annotations to support creative processes might lead us to assume that the annotation has a limited value, that once the document is completed the annotations are no longer needed. The development processes, including the annotations, could be made available to other learners, giving those new learners a rich learning resource. They can give cues to how the document evolved, how ideas were developed, and how conflicts were resolved.

This leads into a second dimension for voice annotations where the annotation could be considered as an enduring and enriching feature of the resource, rather than only of value in the process of creating that resource.

Alternatives/Contrast	to allow expression of alternative, possibly multiple viewpoints, or a contrasting viewpointe to th resource
More	to expand upon the resource—giving a wider context, or frame of reference to the resource
Emphasis/Redundancy	to emphasize agreement or to give a stronger message of the same information but presented via different communication channel
Questioning/Challenging	to pose a question or ask for an explanation relating to the resource
Localizing	to give a "localizing" interpretation to a resource, e.g., information that is relevant to a specific set of learners/course

Table 1: Potential purposes for voice annotations.

5.2 Using Voice Annotations with Shared Resources

This second section of potential applications of voice annotations considers their use as attachments to a resource that is shared between a group of people. Here is a set of purposes that voice annotations might perform when used as an extension to a computer-based resource. The list is not exhaustive and there may be considerable overlap between these. The term "resource" denotes a range of media types: not only written documents or text, but images (both still and moving), diagrams or tables or figures, screen shots and animation sequences. Of course this raises the need to consider *how*, and in the case of a moving image resource, *when*, is the user to be alerted to an annotation. It also begs the question how is an annotation attached to a resource particularly when the resource is a moving image, or an animation sequence?

Using audio as an extension to a resource offers the potential benefits of retaining attention to the actual resource. A voice annotation may support or enrich the visual information or it may help to convey what is difficult to visually explain. Using voice may allow a richer communication because of the subtleties of meaning that may be carried in tone and expression.

Table 1 categorizes a set of purposes for voice annotations to resources.

6 Three Investigations

The first of a set of investigations in the use of voice annotations in authentic collaborative learning tasks has recently been completed. The studies focus around 3 main themes:

- Voice annotations to support collaborative document production processes

- Voice annotations to aid learning from visual (e.g., diagrammatic/pictorial) information

  Leading to hypotheses for a third investigation, namely

- Voice annotations used to support collaborative multimedia courseware production

It is from these empirical studies that a basis to suggest appropriate applications of audio (voice annotations) to visual representations in learning environments will develop. This chapter outlines the first study of voice annotations to support collaborative document production processes.

6.1 Voice Annotations to Support Computer-Mediated Collaborative Writing Processes

The goal of this first study was to investigate voice annotations used in computer-mediated collaborative writing. It looked at the value of voice annotations in terms of process support, ease of communication and resolution of conflicts. The study gathered data on how individual users choose to use audio and textual annotations. The study compared audio or textual annotations in their ease of use, their speed of creation, their effect on reading through drafts with comments and/or revisions. It examined the effectiveness of audio in supporting and resolving equivocal communication. The study looked closely for examples of audio used to quickly capture ideas and ease cognitive load in working memory. It explored the users' experiences and the value they might suggest for annotations in collaborative learning processes.

6.1.1 Task

The task required groups of students to collaboratively author a review of a reading used in their course or module. Each group selected a reading from their course to use in the study. The review required a descriptive synopsis of the reading, a summary of the key issues and ideas raised, together with their own reflection upon those themes and ideas. The groups were asked to write their review as an aid to future students studying the same reading.

A deadline was set for completion of the review and each group was assigned nine hours over a three week period to use the computer. It was entirely up to each group how they used or divided out the computer time, but they had to use it on their own, and for all group communication.

6.1.2 Subjects

The study involved eight groups of three students who completed the single writing task together. Each group was made up students studying on the same course or module. Groups from several university departments were recruited.

The groups were randomly assigned to one of three conditions. This was a small sample so there could be no expectation of an even distribution of practices. Rather, the study intended to identify the practices of interest that emerge and the variables attached to those practices.

The three conditions applied were:

1. using textual annotations only for all communication/comments/ revisions—T

2. using textual and voice annotations with uses for the types of annotation given—T/A

3. using audio and textual annotations with free choice in use—T/A+

In condition (2) above (T/A), the groups were given suggestions of appropriate uses for the text and voice annotations, derived from the value suggested in the literature for spoken and written communication in collaborative writing processes. For example, it was suggested that text annotations are useful for simple and localized comments such as sentence structure or grammatical errors. It was suggested voice annotations would be useful for comments that may be complex and relate to global issues in the writing. It was suggested that voice annotations might help resolve conflicts or differences in opinion amongst the group because the spoken communication would allow the use of expressive language and tone that can be more difficult to convey in a textual annotation.
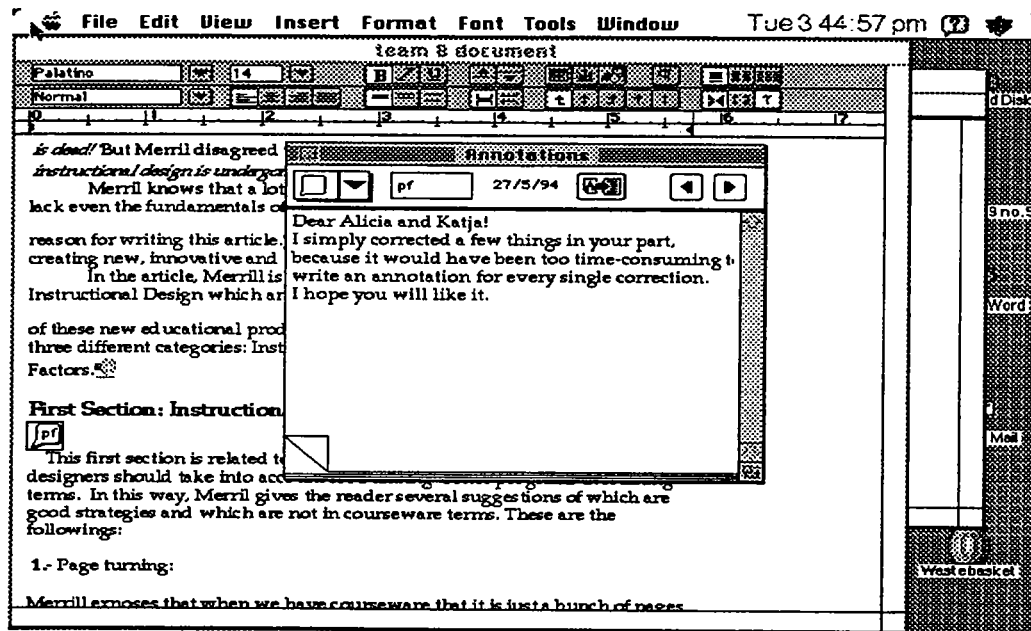
Figure 1: Text annotation.

6.1.3 Technical

The groups created their reviews using word-processing software on an Apple Macintosh computer. After the initial face-to-face meeting of the group, all communication was computer-mediated using Word 5.1 and Microsoft Mail. Word 5.1 provides the facilities to insert text and voice annotations. Microsoft Mail is integrated with Word so that documents may be attached and sent with mail messages directly from the word processor.

An introductory technical training session was given on using the computer and software. Training in how to use the text annotation facility was given to all participants. Training in using the microphone and recording voice annotations was given to all the students in conditions T/A and T/A+. Appropriate, simple guides to the software and recording equipment were provided to each student. Technical support was constantly available during the study.

Figure 1 illustrates the text annotation window as it appears on screen. The bubble icon in the document (with creator's initials in it), alerts the reader to the text annotation.

Figure 2 illustrates the control panel in Microsoft Word for listening to a voice annotation. The panel gives information about the creator, and date of creation of the voice annotation. It gives information about the size of the audio file and how many voice annotations are attached to this document.

6.1.4 Data

Each participant completed a questionnaire to gather data on their background (e.g., course details, age, computer experience and group skills). Each group then began the actual task by arranging a face-to-face meeting to organize the writing and to negotiate and coordinate their specific responsibilities. Each initial meeting was recorded on videotape.

In each computer session the subject was asked to "think aloud," that is, to talk through what they were doing and why, particularly why they chose to use an annotation. These sessions were videotaped. As the document was circulated among the members of a group, a copy (with its annotations) was automatically routed to a central store as part of the data capture. A copy of

team B document

Palatino
Normal

Voice Annotations

▶ Play ■ Stop ❚❚ Pause   Time: 0:00:11   Bytes: 84,992

Annotator: aa,26/5/94
Number: [   ] out of 10

Next
Previous
Find
Cancel

presented bef

4.- Attention F

This statement has the main idea of using devices for focusing the attention of the user. These devices could be timing, animation, flashing or inverse text, sound, colourful displays, etc... But all of them must be created under an instructional purpose, because we could have extremly colourful and flashing screens but without an instructional purpose and then all of the courseware

programmer's effort would be useless.

What Merril exposes is that programers should use all the attention focusing resources in an instructional way. They should think carefully what they are using and what for, instead of just creating fashionable and nice programs but, in instructional purposes, useless.

5.- Active Mental Processing:

In this point, Merrill is summarizing the idea of promoting mental processing and constructive feedback. He is really concerned about the idea of trying to create something else than just a book. In this way, the computer should be the perfect tool.

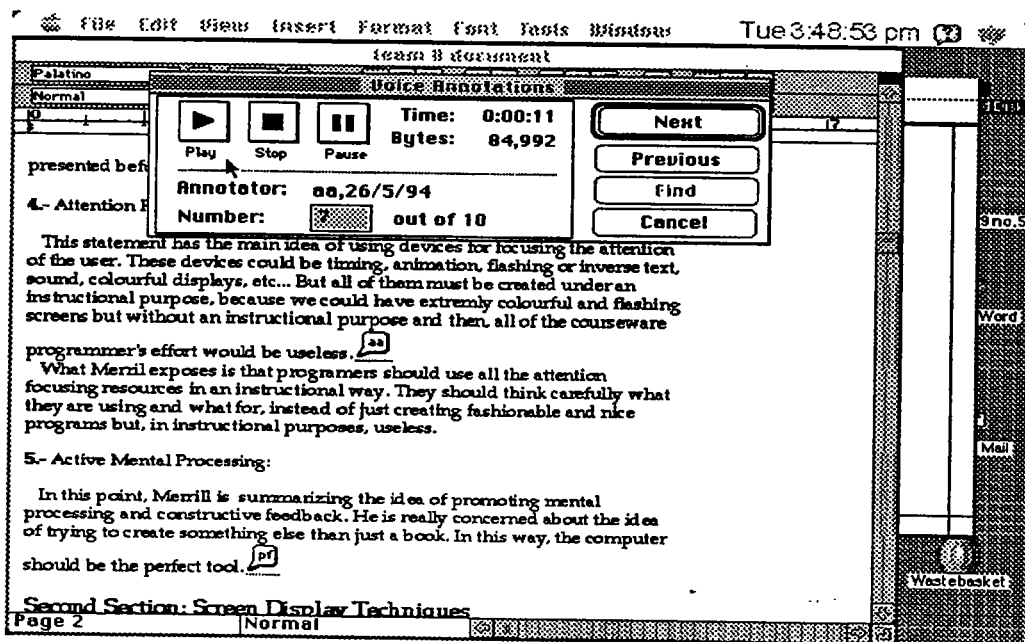Second Section: Screen Display Techniques

Page 2   Normal

Figure 2: Voice annotation control panel.

the completed review was also collected.

A post-study interview schedule was used to gather experiences and comments from the participants. It allowed the opportunity to clarify any issues raised by the questionnaires or immediately apparent from the videotaped working sessions. It also allowed an immediate follow up to any significant comments made.

6.1.5 Initial Findings

From the post-study interview it is already known that reaction to using voice annotations has been mixed. There was some enthusiastic use through novelty, as well as some dislike to recording their own voices from at least two of the subjects. But there were participants who discovered valid purpose for voice in collaborative writing processes. The comments they have made do support some of the suggestions that were put earlier in this chapter—that voice annotations can be useful for articulating complex ideas and that they can be useful to complement and ease understanding of visual information.

One subject said she had valued the voice annotations, *"to help her find the right words"* with the aid of her collaborators. Another subject had found the voice annotation invaluable to help him explain a visual procedure (creating a diagram in the document) to his colleagues. Equally, his colleagues could listen to his explanation while visually attending to the diagram.

In assessing the ease of use of the voice annotation facility it was found that there had been no significant problems in creating an annotation or being able to listen to an annotation. All voice annotations were used in this study to speak to others in the group, rather than used for one's own purposes specifically. The voice annotations were used to communicate and support working as a group, to help coordinate the task of collaborating in writing. They were most used at the stage of revising and developing first drafts.

It has been interesting to assess the listener's reactions to voice annotations. In this study it was found that all the annotations were understandable by the listener in terms of both clarity and content. Some of the listeners did, however, find they needed to re-listen to an annotation

and this was mainly attributed to the fact that in the document there was no visual clue to the annotation's creator or its content to remind them that it was an annotation they might have previously heard.

This suggests the need for voice annotations to be visually distinctive in their visual referent (i.e., their icon). Voice annotations could be visually differentiated by icon type to alert the reader to the content of the annotation. Alternatively, or additionally, an icon could indicate information about the creator of the annotation.

Retaining visual attention while listening to an audio track seemed to be one of the prime interests in developing voice annotations. The study examined whether listeners were in fact able to retain attention to the document while listening to a voice annotation. It was found that in this case it was not easy for listeners to attend to the document because the (visual) control panel for the voice annotation (see Figure 2 earlier) required the subject's visual attention. A voice annotation needs to be controllable when a listener plays it but it also needs to be transparently operated if our purpose is to assist the listener in maintaining visual attention.

7 Summary

This chapter has introduced this research about stored voice annotations used to support collaborative learning processes. The first study of voice annotations for computer-mediated writing processes has looked for examples of practice where collaborating learners have found value in working with enriched communication facilities. The analysis is still in progress but positive purposes for voice annotations in learning settings are beginning to emerge that will inform the development of this research.

Acknowledgments

References

[1] Kaye, A. R., ed. "Learning Together Apart." In *Collaborative Learning through Computer Conferencing: the Najaden papers.* Berlin: Springer-Verlag, 1992.

[2] Harasim, L. "Online Education: an Environment for Collaboration and Amplification." In *Online Education: Perspectives on a New Environment*, edited by Linda Harasim. New York: Praeger, 1990.

[3] Brown, A. L., and A. S. Palinscar "Guided Cooperative Learning and Individual Knowledge acquisition." In *Knowing, Learning and Instruction: Essays in Honour of Robert Glaser*, edited by L. B. Resnick. Hillsdale, NJ: Lawrence Erlbaum, 1989.

[4] Scardamalia, M., and C. Bereiter. "An Architecture for Collaborative Knowledge Building." In *Computer-Based Learning Environments and Problem Solving*, edited by E. de Corte, M. Linn, H. Mandl, and L. Verschaffel. Berlin: Springer-Verlag, 1992.

[5] Aarntzen, B. M. H. "Audio in Interactive Tutorial Courseware: Methodological and Design Guidelines." Doctoral dissertation, Enschede, The Netherlands, 1994.

[6] Kramer, G., ed. *Auditory Display: Sonification, Audification, and Auditory Interfaces*, 1–77. Santa Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading, MA: Addison-Wesley, 1994.

[7] Chalfonte, B., R. S. Fish, and R. E. Kraut. "Expressive Richness: A Comparison of Speech and Text as Media for Revision." *ACM/SIGCHI 91 Human Factors in Computer Systems* (1991): 21–26.

[8] Kellogg, R. T. "Computer Aids that Writers Need." *Behaviour Research Methods, Instruments & Computers* **17(2)** (1985): 253–258.

[9] Rimmershaw, R. "Collaborative Writing Practices and Writing Support Technologies." *Inst. Sci.* **21** (1992): 15–28.

[10] Kraut, R. E., J. Galegher, R. S. Fish, and B. Chalfonte. "Task Requirements and Media Choice in Collaborative Writing." *Human-Computer Interaction* **7** (1992): 375–407.