Internet Multimedia Conferencing - Results from the ReLaTe Project

Jane Hughes
Department of Computer Science
University College London

Angela Sasse
Department of Computer Science
University College London

Abstract

ReLaTe (Remote Language Teaching over SuperJANET) is a real-time system for small group distance language learning. It uses ordinary computer workstations and runs over shared packet networks, such as the multicast backbone of the Internet. Incorporating real-time audio, video and shared workspace, the ReLaTe system has already proven to be an effective tool for small-group tutorials in foreign languages. Interaction through .the system allows simultaneous development of the four skills (reading, writing, speaking and listening) and instant error correction. The system has proven to be easy and fun to use, and its immersive environment looks to be highly conducive to effective learning. This paper reports and discusses the results of field trials and outlines further functionality, such as online dictionaries, grammar tables and personal notebooks and digital storage of lessons, which would enhance the pedagogical effectiveness of the system.

1. Introduction

ReLaTe (Remote Language Teaching over SuperJANET) started as a joint project between University College London and the University of Exeter in the UK. Its aim was to determine whether it feasible to teach languages using desktop multimedia conferencing over the Internet; further objectives were to assess the pedagogical implications and identify requirements for further development of hardware, software and network technology capable of supporting live language tutorials.

1.1 The educational context

Interest in and practice of open and distance learning has expanded in recent years. Motivating factors include financial pressures, the desire to improve student access, and acceptance that education and training are now a part of working life. At the same time, awareness of teaching quality issues and of accountability has increased in Higher Education [Fry 1995]. Whilst economic arguments may be persuasive, most educators feel that distance learning should offer students something comparable with the best experiences that can be found in traditional learning situations, not a poor substitute. Quality of learning was therefore a major concern in the ReLaTe project.

Most successful distance learning programmes include strategies for encouraging peer group contact and collaboration. Often this is computer-mediated. Electronic bulletin boards and electronic mail can provide a forum for discussion and give students access to help and advice. In some programmes, such as Open University courses in the UK, real face-to-face contact is incorporated through summer schools and the creation of local study groups. This suggests that such contact is perceived as beneficial to learning. It may also be worth mentioning that a strong positive correlation has been shown to exist between students' perception of the degree of their interactivity (with peers, tutor and study materials) and their satisfaction with a course of study [Fulford & Zhang, 1993]. Yet achieving such interactivity in a distance learning programme is far from straightforward and may well be costly.

1.2 The technological context

As mentioned above, computer-mediated communication is not new in education. Nor is the idea of using interactive multimedia, via CD ROM. The interactions referred to here are generally between student and course material. Where real-time interactions are needed between people at a distance, multimedia conferencing can be used. Multimedia conferencing, as provided by telephone companies, is often associated with specialised conference rooms. The service provided is a point-to-point link; this means the user pays for a pre-booked line between any two sites that have to be connected, the cost increasing with distance. If simultaneous, multiway communication is required, the number of lines needed increases faster than the number of sites, which is why such systems are often said to be expensive to scale. They also tend to be inflexible. The ability to book the connection does, however, provide certain quality assurances.

A more flexible and less expensive alternative emerged in the 1990s. Prior to this it was assumed that computer networks such as the Internet were not capable of transmitting real-time audio and video, since information carried by shared packet networks is subject to congestion and delay.

A new approach to routing in shared packet networks, known as multicast, has overturned this assumption. Real-time video and audio are now routinely sent over the internet via the multicast backbone, or Mbone [Macedonia & Brutzman, 1994]. Multicasting is a form of routing that scales up economically and efficiently, to offer multiway communication at reasonably low cost. It has been used successfully to transmit lectures to hundreds of people worldwide. Another advantage is that individuals can participate in a conference using a desktop computer, rather than having to go to a specialised facility. Whilst this may promise much for education there are still some unanswered questions about whether the quality achievable at present is adequate and about how best to implement a conferencing system to support teaching and learning. In order to address these questions, the ReLaTe project developed a working demonstrator of a multicast-based desktop conferencing system for distance language tuition and used this to provide remote small-group tutoring in a field trial with teachers and students.

2. The ReLaTe field trial

The aim of the field trial was to evaluate a prototype multimedia conferencing system and establish whether multicast conferencing can support remote language tutorials. The field trial took place from October to December 1995, with tutors and students from the UCL Language Centre and the Foreign Language Centre at Exeter University, over the 10 Mb SuperJANET SMDS service, the UK academic network. Four ten-week courses were offered, Advanced French, French for Business, Beginners' Portuguese and Reading Latin for Academic Purposes, in which students (at Exeter or UCL) were taught by tutors at remote locations (UCL or Exeter). Apart from Latin (one hour per week) all the courses were taught in once-weekly, two-hour sessions.

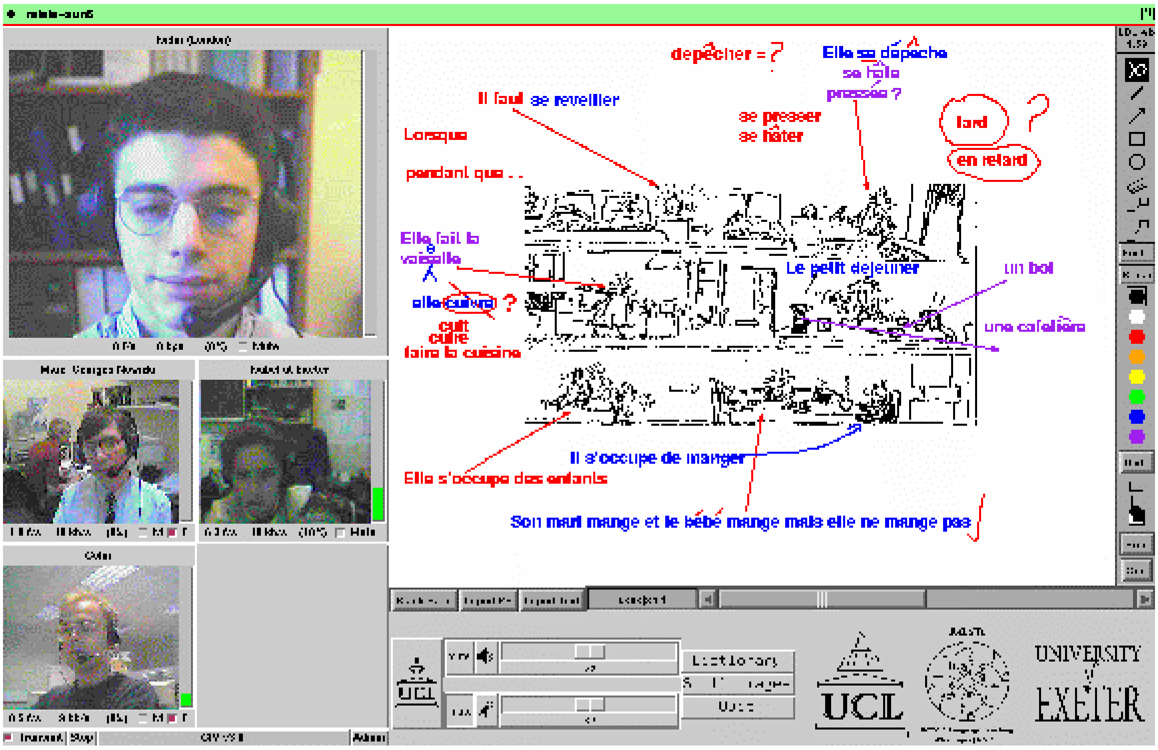

Fig 1 The ReLaTe User Interface

2.1 Students and tutors

Information about participants was obtained by means of questionnaires and interviews. There were two or three students in each tutorial group. They were predominantly Arts and Humanities students but only one was a language specialist. Tutors worked at the Language centres at Exeter and UCL where they also taught face-to-face classes. They ranged in age from mid twenties to mid sixties.

None of the participants had significant experience in computing, but most were familiar with word processing applications. The students had more experience of using a mouse and, generally, better keyboard skills than the tutors. None of the users was a touch typist, however, and one of the tutors had never used a mouse before.

2.2 Training and practice

Tutors and students were introduced to the system on an individual basis shortly before their first tutorial. There was no formal course of training. The amount of time devoted to this introduction varied according to the needs of individuals from fifteen minutes to about two hours. The tutorials themselves constituted the only practice for participants in the trials. The training was considered by all of them to have been adequate but results discussed in section three suggest that more training might have been beneficial.

2.3 The ReLaTe system

The ReLaTe system uses multicast conferencing as its underlying technology. Multicast tools are specific to a medium (audio, video, shared workspace). Inexperienced users have problems managing different tools in parallel - especially managing a multiplicity of windows. The system developed during the first year of ReLaTe [Buckett et al, 1995] featured an integrated user interface (Figure 1).

On the left hand side is the video area, where four images can be displayed with a choice of size. Individual participants can select the image to be displayed in the larger rectangle by clicking on the name bar above it. The shared workspace tool, wb, occupies most of the right hand side of the interface. This tool allows displayed text and graphics to be seen and annotated by all participants, and all may write or draw on it simultaneously. Below this are the audio controls which allow users to adjust incoming and outgoing volume.

During the tutorials tutor and students sat at individual UNIX workstations (in this case Silicon Graphics Indy machines). A small video camera was mounted on each workstation and all participants wore headsets.

2.4 Data collection

This was an exploratory study, a field trial with real users, not a controlled experiment. This influenced the data collection methods [Watson & Sasse (2), 1996]. With limited resources and a relatively small group of subjects, the main aim was to obtain qualitative data.

The principal method of evaluation was observation. Since the project was concerned with both system design/usability and pedagogy the tutorials were observed both by usability experts, sitting next to teachers and students, and by language teachers acting as expert assessors, who watched unobtrusively from a different workstation. Although resources did not permit a formal comparison with face-to-face classes, observers did watch tutors working in a classroom situation, in an attempt to increase their understanding of the teaching and learning tasks being undertaken.

Other data were obtained via questionnaires and from informal interviews with all participants during the trials. After the trials had ended, teachers and students were brought together for a group workshop in which many sessions were taped.

3. Results

After presenting the general conclusions, the contribution of the individual conferencing components (audio, video and shared workspace) is discussed, followed by other findings relevant to teaching and learning.

3.1 General conclusions

The answer to the main question posed by the project appears to be that small group language tuition is feasible using such a system. Considerable time was spent in the first year of the project drawing up a list with teachers of key teaching and learning activities that the system must support. They were identified as: conversation; grammar exercises; spelling; fill-in-gap exercises; reading; writing; vocabulary; pronunciation. Observations demonstrate that the system is able to support all these tasks and a number of others. Conversation, vocabulary and pronunciation tasks were undertaken in virtually every session, reading, writing and grammar in the majority. Spelling and gap-filling activities were less frequent. Observers also noted frequent use of translation, role-play, prepared talks, word games, drawing and labeling.

In addition to this, teachers and expert observers were keen to stress that the system encourages the integrated development of reading, writing, speaking and listening skills, something that is not always easily achieved in conventional teaching situations. It appeared also that users also found the system easy to learn and fun to use.

3.2 Audio

A language teaching application had been chosen precisely because it would make demands on the audio component. There were two main questions to be answered. Would the audio quality be good enough to support language learning? Would the design of the audio component allow the full range of activities to take place?

A number of factors affect audio quality, including the characteristics of the microphone and headsets used. With an application using the Internet, however, the main threat to quality is the network itself. Packet loss, which can occur when the network gets too busy, can degrade speech significantly. Whilst a small amount of loss is barely perceptible, with losses of about 10%, users notice small gaps or glitches in the audio. At higher loss rates (over 20%) gaps become more frequent and eventually, in very poor conditions, the speech can become almost unintelligible. The nature of the network means that conditions change constantly, which can lead to rapid and extreme fluctuations in speech quality.

To cope with this the development team added a packet loss repair mechanism to the audio tool, RAT [Hardman et al, 1996], which significantly improves perceived audio quality at low to medium rates of packet loss. Results of separate testing suggest that audio quality will be acceptable to most users even when the rate of loss reaches 30%.

Encouragingly, most students believed they spoke at least as much of the target language as they would have done in a traditionally taught course. Nevertheless, when asked to rate the system, both teachers and students cited audio quality as the feature most in need of improvement. As might have been predicted, it was usually network, rather than local problems which affected quality.

Instances where poor audio quality affected learning included the beginners' class (Portuguese) where the teacher wanted to practise the pronunciation of short sounds and a French class where students were unable to hear the difference between similar-sounding words, leading to confusion and delay in making a relatively straightforward point.

It might also be worth investigating the observation that the need to compensate for variable audio quality seemed to affect style of speech; teachers tended to speak slowly and clearly and one observer noted over-emphasis of words. This may have had an impact on the pace of lessons.

Participants were tolerant and persistent when packet loss occurred, especially once they came to realise such problems were likely to be transient. At the same time, this very unpredictability was a source of uncertainty and irritation. Not a single lesson was abandoned because of poor quality audio but this is an area where considerable work is still needed.

Apart from the effects of packet loss, users occasionally complained of "crackling" or "echo", which are more likely to be caused by faulty equipment, such as headsets. It is also evident that the choice of headset needs care; wearing a headset for up to two hours can cause discomfort and a short lead restricts movement.

Full duplex audio cannot operate successfully without headsets, which prevent the output from speakers leaking back into the microphone. A lightweight headset with minimum leakage and a good quality microphone would be popular with users. Alternatively, incorporating effective echo cancellation facilities into audio tools would remove the need for headsets. Although there were also a few problems over regulating the volume of different participants, users made relatively little use of the facility to adjust incoming sound levels and very rarely altered the volume of their own outgoing audio. Otherwise, there were few usability problems with this component.

3.2 Video

Key questions here were: how would tutors and students use the video channel? What would the video element contribute to the tutorials? It was also hoped to draw conclusions about whether the image sizes offered were appropriate for language learning.

Whilst television quality video runs at 25 frames per second, it is common practice in desktop videoconferencing to operate at 4 or 5 frames per second in order to minimise loading on the workstation and the network. In some of the ReLaTe sessions the frame rate was lower still. It is sometimes assumed that video at such low frame rates can make no significant contribution but it appears from the results collected that this is not the case.

Users felt that video was a valuable component of the system, and they made use of the images in many ways. As detailed in [Watson & Sasse, 1996], these included: to clarify meaning; to check reactions; to provide a means of common reference; to check whether anyone was speaking during an unusually long silence; to give psychological reassurance that the other participants were actually there.

It had been hoped that students would be able to use lip movement as a comprehension aid and an important part of the project's technical development work was implementing lip synchronisation. In multicast conferencing, audio and video are separate streams of data, routed separately through the network. There is therefore no guarantee that packets transmitted simultaneously will arrive simultaneously at their destination. By enabling the two streams to communicate with one another, the developers ensured that they could be synchronised when played out.

Unfortunately, the processing power of the workstations bought at the beginning of the project was not sufficient to support encoding and decoding of three to four simultaneous video streams at a frame rate which would support lip synchronisation (8-10 frames a second - [Kouvelas et al, 1996]). Hardware upgrades at the very end of the trial period did, however, allow synchronisation to be tested in the last session. Whilst this single session does not provide a base for conclusions, preliminary results are promising.

The lack of synchronisation may have meant users did not perceive an advantage in selecting the larger image for the current speaker. In any case, this facility was rarely used. It was therefore not possible to draw conclusions about the optimum size of video image for language learning.

It seemed clear that participants were aware that communication differed from that in a face-to-face situation. There was evidence that virtually all were conscious of the need to position themselves for the camera, though they rarely altered the position of the camera itself. There were still some problems, however. The most frequent involved moving partly out of shot. This was often related to space management; using a text book or making notes on separate paper meant leaning sideways and out of the camera's range. Teachers found that the computer desk provided allowed insufficient space for ease of work and recommended that book rests should be provided as part of the physical workspace. Another occasional problem concerned the use of gestures to illustrate a point - but in front of the screen rather than the camera and hence not seen by the students.

As with the audio component, the role of peripheral devices and the environment proved to be important. After a few sessions the lighting at one site was improved, and users responded positively to what was perceived as better quality video. They also reported that it was helpful to get an impression of the other participants' surroundings.

3.3 Shared workspace

Students in a face-to-face class rarely get the opportunity to use the whiteboard since it is usually in the teacher's domain. This may account for the excitement it generated and for the fact that both teachers and students rated it one of the best features of the system.

The tool was used in every session, in innovative ways and for varied tasks. These included matching and gap-filling exercises, verb conjugation, spelling, reading comprehension and writing exercises. Tutors routinely imported text files or postscript images scanned prior to the lesson, which gave students access to very varied learning materials. Amongst these were strip cartoons, advertisements, photographs, newspaper articles, literary extracts and grammar exercises. The fact that all were able to write simultaneously and see each other's contributions meant that students received immediate feedback on their work and that errors were corrected instantly. Students believed this "made grammar more fun".

Activities which required students to read extended prose extracts on the whiteboard were less successful. They said that they found this tiring, although good quality, appropriately sized print did help. Problems could also arise from different reading speeds, if one student wanted to change page before others had finished reading the current one.

Some effort appeared to be needed to adjust both from normal classroom practice and from familiar computer applications. There was a reluctance to select a new page (users preferred to delete text and reuse the same page, as one does with a traditional whiteboard). Users familiar with word processing applications experienced some difficulties since wb has no cursor, no automatic carriage return and one cannot select or edit single elements of a block of text.

Limitations in the tool itself also led to a few frustrations. There were some difficulties with page-changing and occasional confusion over the three possible ways to erase entries in wb. Some people found the mouse too crude as a drawing tool. It was also unfortunate for this application there was no facility to type accents on text. Participants often drew them in after typing the text or commented that an accent was needed. This could easily be solved by using target language keyboards, or by reprogramming keyboards to generate the desired characters' special key combinations.

3.4 Other findings related to teaching and learning

Tutors and students agreed that the ReLaTe system generated a very high degree of concentration. The screen is the focus of attention and the headphones cut out extraneous noise so that, as one student said, "You can't ignore the lesson". Not only is the system immersive but it appears to demand active participation. "You know you have to speak because of the equipment".

All agreed that these factors facilitated learning which one expert assessor described as "detailed and profound" All participants agreed that this high level of concentration, whilst it promoted deep learning, was also exhausting. The two-hour teaching sessions were found to be manageable only with a mid-session break.

Allied to this was a feeling amongst the teachers that being immersed in the system caused them to lose their sense of time and hence the ability to pace a lesson. They suggested that being unable to move about the room or hear external sounds led to a sense of isolation from the environment. For this reason there was a request to incorporate a clock or timer into the interface.

Expert assessors reported that "close and effective working relationships" between tutor and students were fostered by the interactive nature of the tasks and the element of play and strengthened by "effective and rapid teaching and learning" and the sense of "shared endeavour" generated by equal access to the whiteboard.

The group discussions revealed subjective perceptions about the relationships. Perhaps surprisingly for a remote learning situation, teachers spoke of the "intimacy" the system promoted and the "very personal" nature of their relationships with the students. Students considered the relationships more open and friendly than those they had experienced in conventional classrooms. Amongst the suggestions made to account for this were: that the system acted as a leveller, since teacher as well as student was learning to use it; that participants' faces were always visible; that first names appeared above the video images.

The observers recorded information about interactions between students. Interaction varied according to teaching style. In some classes the lesson was structured to include activities such as role play; in others interaction occurred casually. Student activity during breaks was also monitored. It was usual to hear them talking to one another over the system and they routinely drew or wrote together on the whiteboard. This was extended to providing Christmas decorations at the end of the course!

Group size probably contributed to the quality of relationships as did the enthusiasm and motivation of students and tutors. Expert assessors reported that their observations revealed relationships similar to those found in a traditional classroom, and that students seemed to have a sense of belonging to a group.

One concern was to determine whether a similar quantity of material could be covered teaching via the ReLaTe system as one would expect in a face-to-face course. Answers are not conclusive, largely because of the difficulty in finding appropriate face-to-face classes to act as controls. One point raised was that the need to speak more slowly and clearly and the slight end-to-end delays might slow the pace of a lesson, compared with that of a face-to-face class, so that slightly less material would be covered in a given time. Further studies would be needed to substantiate this and it might be counteracted by what the expert assessors described as "more effective learning"

4. Further development

ReLaTe2, a follow-up project funded by JISC started in September 1996. It aims to address a number of issues raised by the original project and, as with the first ReLaTe, enquiry into pedagogical issues will run alongside technical research and development.

A primary aim is to widen access to the technology by providing a similar system for PCs. Additional on-line resources will also be incorporated into the system. Another major area of research is audio. Apart from continuing work to improve quality, the intention is to enhance the reality of the sound and to improve lip synchronisation. Limitations in the shared workspace tool have begun to be addressed already. A shared text editor, NTE [Handley 1996] has been incorporated into the system. Users can select (by means of a button) the tool most appropriate to a given task. The graphical capabilities of the whiteboard are seen as essential for some tasks, however, and NTE is regarded as a useful addition, not a replacement tool. In the long run it is likely that a completely new shared workspace tool will be needed.

The implications of desktop conferencing for remote tutorials will be explored further. Although the original project demonstrated that effective language teaching can be supported by a system such as ReLaTe, it is clearly not safe to assume everything will be done in exactly the same way as with a face-to-face class. A new, longer trial has begun. Using this as a testbed, the intention is to develop teaching methods and materials and to investigate some of the changes to which tutors in particular may have to adjust if they are to work comfortably with such technology. Areas of particular interest in this project include communication strategies, lesson management, preparation and assessment and requirements for training and technical support.

The project team has always regarded the ReLaTe conferencing system as a useful tool in effective distance education, not the whole solution. There is therefore a strong interest in exploring the system's relationship with other kinds of support for teaching and learning, including asynchronous communication tools and self-study materials. The new ReLaTe project will also be exploring the use of record and replay facilities, both to enhance provision for students and as an aid to assessment.

ReLaTe2 will take place in a context of continuing technological change. Networks are being upgraded all the time and techniques for compressing data are improving. Work is also in progress on a system that would enable users who need consistent quality to reserve bandwidth. These developments should make it feasible for many distance learning programs to use the technology effectively. Like its predecessor, ReLaTe2 is working with language teachers and learners, but the findings discussed in this paper will be relevant to many other areas of education.

Acknowledgements

The authors would like to acknowledge the following contributors to the work reported in this paper: at UCL, Vicky Hardman, Isidor Kouvelas and Anna Watson. At Exeter, John Buckett, Mark Pack and Gary Stringer. Special thanks are due to teachers and students from the Foreign Language Centre at Exeter University (Director Elizabeth Matthews) and the UCL Language Centre (Director Dolores Ditner) and to the expert assessors, Marie-Danielle Chamary, Marc-Georges Nowicki and Dr. Carmo Ponte.

ReLaTe (Remote Language Teaching over SuperJANET) was a BT/JISC funded SuperJANET Applications project from September 1994 to December 1996. ReLaTe2 is funded by JISC for two years from September 1996.

References

[Buckett et al 1995] Buckett, J., Campbell, I., Watson, T., Sasse, M.A., Hardman, V., Watson, A. "ReLaTe - Remote Language Teaching over SuperJANET": Proceedings of NetWorkshop, 1995.

[Fry 1995] Fry, H. "Quality Issues in Higher Education": 'Viewpoint', University of London Centre for Higher Education Studies, June 1995

[Fulford and Zhang 1993] Fulford, C., and Zhang, S. "Predicting student satisfaction from perceptions of interaction in distance learning": 'Teleteaching' (A-29), Elsevier Science Publishers, 1993

[Handley 1996] Handley, M.J. "Network Text - A Scaleable Shared Text Editor for the Mbone": UCL Computer Science Department Research Note, RN/96/13.

[Hardman et al 1996] Hardman, V., Kouvelas, I., Sasse, M.A., Watson, A. "A Packet Loss Robust-Audio Tool for use over the Mbone": UCL Department of Computer Science Research Note, RN/96/8

[Kouvelas et al 1996] Kouvelas, I., Hardman, V., Watson, A. "Lip Synchronisation for use over the Internet: Analysis and Implementation": Proceedings of Globecom'96, London, UK, November 96.

[Macedonia and Brutzman 1994] Macedonia, M.R., Brutzman, D. P. "Mbone provides Audio and Video across the Internet.": IEEE COMPUTER, April 1994, pp 30 - 36.

[Watson and Sasse 1996] Watson, A, Sasse M. A. "Assessing the Usability and Effectiveness of a Remote Language Teaching System": Proceedings of ED-MEDIA'96, Boston, July 1996

[Watson and Sasse (2), 1996] Watson, A., and Sasse, M.A. "Evaluating audio and video quality in low-cost multimedia conferencing systems". Interacting with Computers, Vol 8, No. 3 (1996) pp255-275.