Study Of Picture Annotation As A Means Of Assisting Non-Speaking People To Use Telecommunications Services.

Mr. Nicolas Hine, Dr. William Beattie & Dr. John Arnott,
MicroCentre, Applied Computer Studies Division, University of Dundee, Scotland.
Phone. +44 1382 344711, E–mail: nhine@mic.dundee.ac.uk

Abstract

People with speaking difficulties have been at a social and practical disadvantage for many years because of the difficulties that they have in using a telephone. The increasing role that telecommunications will have in our lives in the future makes the search for solutions for these users all the more urgent. The fact that multimedia communication will be possible, however, holds the key for many people, as alternative methods for expressing and communicating ideas will be possible. This paper highlights the limitations of text as a means of communication for users with speaking difficulties, and describes work to investigate the value of pictures and graphics as an alternative.

Key words: Disabled non-speaking users, Communication devices, Distributed services

1. Introduction

Telecommunications is becoming increasingly important in all aspects of life today. In business, in addition to conventional telephony, many people regularly use fax services. Studio and desktop based videophone services are now being used as an alternative to face-to-face meetings. At home, the conventional telephone is being used to provide a wide range of services such as 24 hour banking and 24 hour access to insurance. More powerful multimedia services will become more prevalent as the number of cable TV installations increases. Other trends include a greater penetration of the Internet, increased use of personal computers for accessing information and for remote communications, and increased bandwidth available in telecommunications networks. All these trends point to a world where the use of telecommunications services will become essential in many aspects of daily life.

With this in mind, it is obvious that a user with a communication impairment will be at a severe disadvantage. They will encounter difficulties in participating in activities that most members of the population will take for granted and depend on.

Many interpersonal interactions depend on speech. As a substitute, text is often used in low bandwidth situations (for example text telephone facilities on Internet videophones). A logical step would be to propose text as an alternative media for people with speaking difficulties. This is not a realistic solution in many cases as a

high percentage of non-speaking people have an additional motor impairment that makes typing difficult, or they have reduced language abilities. In Augmentative and Alternative Communication (AAC) devices symbol systems are often provided as an alternative to text. However, many non-speaking people, particularly adults, do not use these as they are not understood by many able bodied people. The more powerful symbol systems are effectively a language, and require the same degree of tuition and practice as traditional orthography.

In conversation/communication, pictures and graphics can be used to convey complex ideas, to highlight specific facts, or even to set a general context. If a communication device allowed communication parties to annotate pictures, attention can be drawn to detail within the picture. This would allow quite specific information to be exchanged without needing to employ specific textual explanation or other language like constructs.

This paper will report on an experiment to explore this approach as an alternative to text. The results will be explored in the light of previous studies [Hine et al (1995a & 1995b)] and the implications for users in terms of call cost will be highlighted

2. Hypothesis & Goal

Some users can exchange information at a significantly faster rate by displaying and annotating pictures than they can by typing the same information.
This will be tested by comparing the time taken to exchange information in text form and in picture form.

3. Subjects

8 user with speaking difficulties. A number of these have additional motor impairments affecting their ability to manipulate a keyboard and/or mouse.
All subjects were familiar with the computer systems being used, but not with the software. Training was provided as described in section 4 below.

The subjects had varying degrees of reading and language abilities. In general they were not proficient in the use of maps. For this reason, the maps were marked clearly with the locations and routes, and the text for the text exchange sessions was copied from sheets written in a large font.

Table 1: Subjects, Impairments and Adaptations

Subject	Impairment	Adaptations Required
1	Impaired speech and reduced manual dexterity	Mouse and Keyboard
2	Impaired speech, poor dexterity and uncontrolled head movement	Rollerball and on-screen keyboard
3	Impaired speech, no manual dexterity	Chin manipulated rollerball and on-screen keyboard

4	No speech, reduced manual dexterity	Concept keyboard with keyguard
5	No speech, reduced manual dexterity	Mouse & Keyboard
6	No speech, no manual dexterity	head switch driving a scanning keyboard and mouse
7	mildly impaired speech and manual dexterity	Mouse and Keyboard
8	mildly impaired speech, reduced manual dexterity	Rollerball and on-screen keyboard

4. Method

The scenario used involved an experimental subject seeking to explain to a remote user (a "travel Agent") that they want to visit a number of locations on a holiday. In some cases the information will be discussed using a text phone, and in other cases the information will be exchanged displaying a map and annotating the map. Each conversation consisted of a preamble where the travel agent asked which state was to be visited. The reply was either a typed reply or the display of the state map on each terminal. This was followed by a question asking the for first town to be visited. The answer was either typed or a marker selected and a box placed on the map to outline the town. Confirmation of the town name was requested by the travel agent by repeating, in the text window the name of the town. The subject confirmed, either by typing yes or placing an arrow within or close to the box on the map. If the travel agent did not give the correct town name, the subject either re-typed it, or placed a new marker box on the picture.

This preamble was followed by five episodes where the travel agent asked for the next town to be visited, and the road to be used to get there. Each time, the subject gave a response, and confirmation took place as described above. The time taken for each episode was measured from the time that the question about the town was received by the subject until the time that the subject sent a confirmation that the road number identified by the travel agent was correct.

In each experiment session, the state, towns and roads were unknown by the "Travel Agent". No subject visited the same state more than once, although different subjects visited the same state. Every route of towns and roads was unique, even in the case where the state had been visited by a different subject. In this way, the travel agent was not able to predict a route or the towns that would be visited.

The first phase of the experiment was a adaptation phase. Each subject was already familiar and experienced in the use of the computer. The adaptation that they normally used was installed in the test system and optimised to their individual abilities. This phase was also used as an informal training phase.

This phase was followed by a training phase for each where they formally performed 1 text based conversations and 1 picture based conversations.

This was followed by the experimental phase consisting in a set of experiments involving 2 text based conversations and 2 picture based conversations. The order of text and picture sessions was different between subjects in order to eliminate order effects.

4.1 Training:

1 picture and 1 text conversations per subject - all subjects use the same 8 example scenarios

4.2 Experiment:

8 subjects - 4 experimental sessions per subject - 2 picture and 2 text conversations per session

5. Apparatus

Two Macintosh computers were connected with data transfers via a dedicated section of Ethernet network. Video and audio, used for setting up the experiments, were connected directly (see Figure 1).

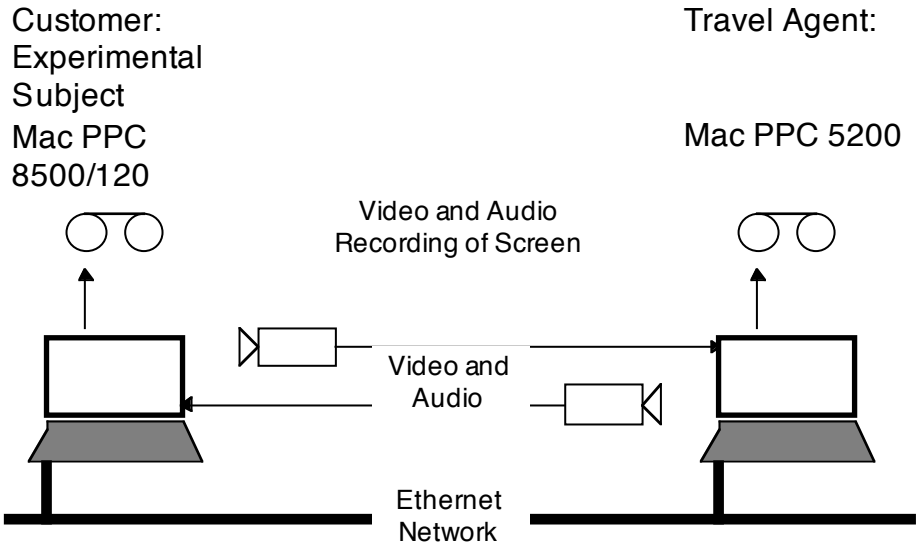


Figure 1 Experimental Setup

The following materials were prepared for the experiments

- Handout explaining the experiment
- 24 text conversation scripts (8 for training, 16 for experiments)
- 24 picture conversation scripts (8 for training, 16 for experiments)
- Timetable of training and experiment sessions
- Experimental Scripts

A typical layout showing all the components of the communication interface and an on-screen keyboard are shown in Figure 2 below.

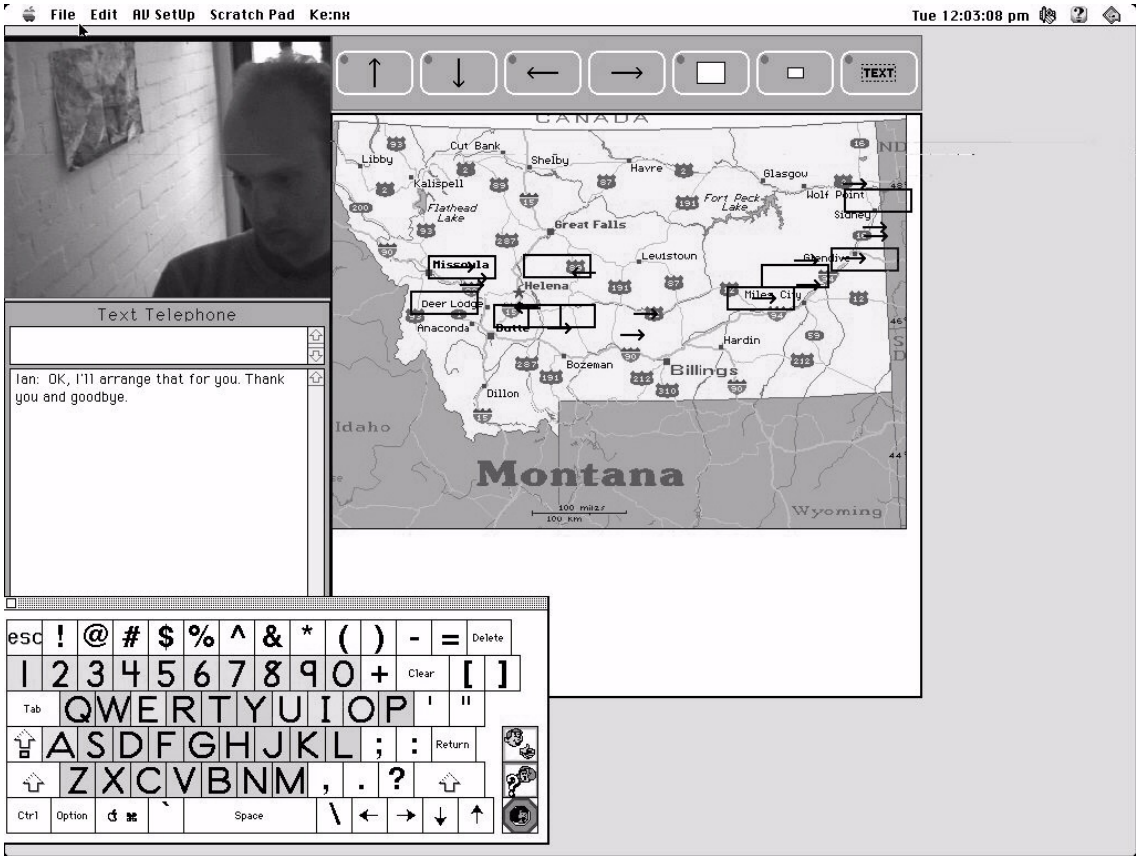


Figure 2: Screen layout of the experimental system, showing communication interface and on-screen keyboard

6. Results

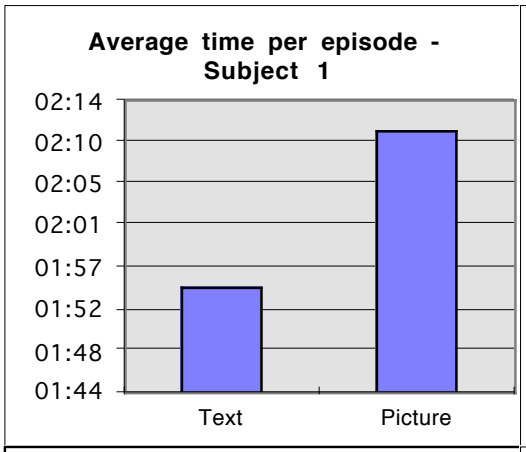
The following records were taken:

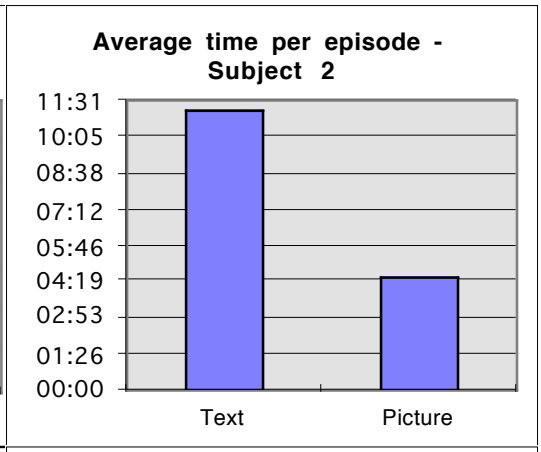
Time coded video recordings of the monitors of the subjects and the remote "travel agent"

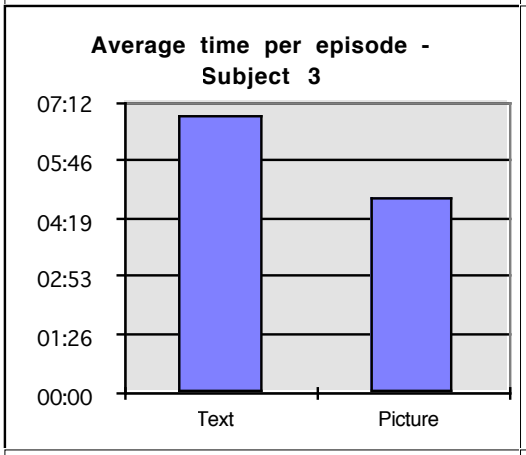
Time coded records of all data generated in each terminal

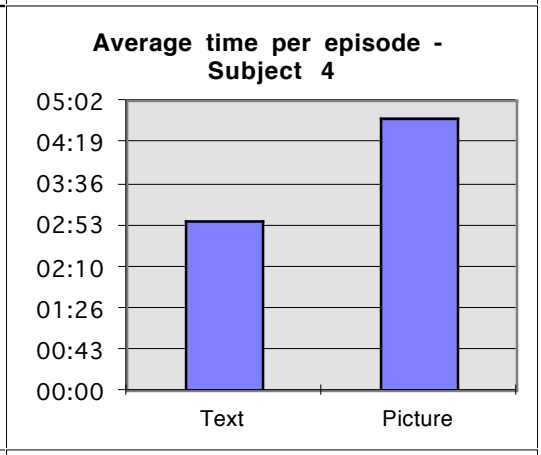
A description of the adaptations made to the terminal for each subject, and the screen layout of the components of the videophone interface.

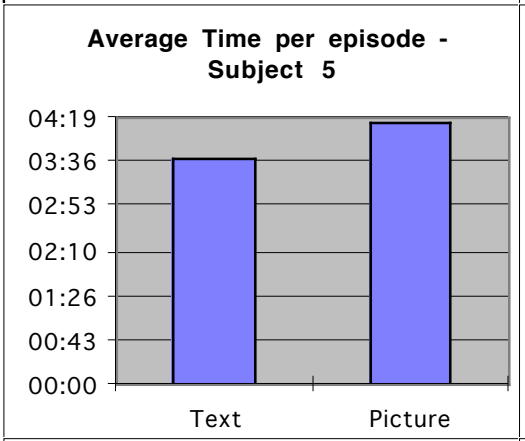
From these records, the average time taken to complete a town/road episode were calculated for both the text and picture annotation methods. These are plotted as a graph and are shown in subject order below.

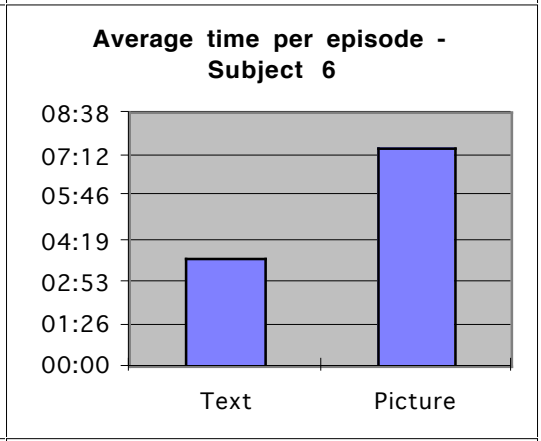

Average time per episode - Subject 1
02:14
02:10
02:05
02:01
01:57
01:52
01:48
01:44
Text
Picture

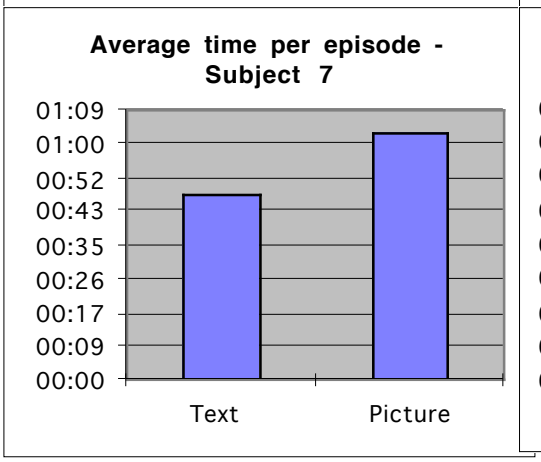

Average time per episode - Subject 2
11:31
10:05
08:38
07:12
05:46
04:19
02:53
01:26
00:00
Text
Picture

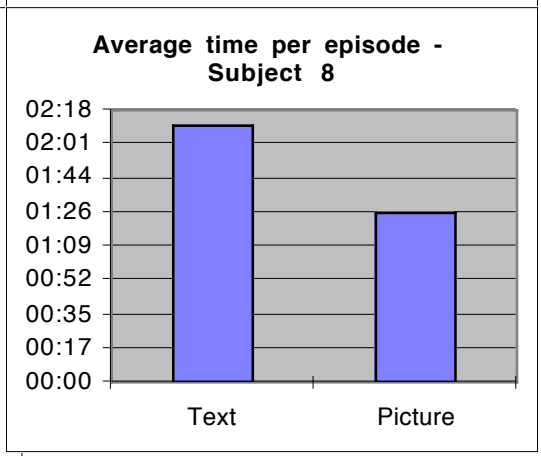

Average time per episode - Subject 3
07:12
05:46
04:19
02:53
01:26
00:00
Text
Picture


Average time per episode - Subject 4
05:02
04:19
03:36
02:53
02:10
01:26
00:43
00:00
Text
Picture


Average Time per episode - Subject 5
04:19
03:36
02:53
02:10
01:26
00:43
00:00
Text
Picture


Average time per episode - Subject 6
08:38
07:12
05:46
04:19
02:53
01:26
00:00
Text
Picture


Average time per episode - Subject 7
01:09
01:00
00:52
00:43
00:35
00:26
00:17
00:09
00:00
Text
Picture


Average time per episode - Subject 8
02:18
02:01
01:44
01:26
01:09
00:52
00:35
00:17
00:00
Text
Picture

Figures 3: Graphs comparing average typing and annotation times for each subject

In general, 5 out of the 8 subjects were slower t marking the map that in typing the equivalent information. 3 out of the 8 subjects were faster at annotating the map than they were at typing the equivalent information.

7. Discussion

The initial hypothesis stated that “Some users can exchange information at a significantly faster rate by displaying and annotating pictures than they can by typing the same information”. This was not the case in 5 of the 8 subjects, but was clearly the case of three subjects. The reason for this result can be found in the nature of the task and the type of adaptation used by the subjects.

The task to be performed had two conditions: a text based information exchange, and an information exchange based on picture annotation. The picture annotation condition required the user to select a tool from a toolbar, move to the location within the picture and place the annotation object. This had to be repeated twice for each episode.

In the case of text, once the text typing cursor had been placed in the text input window, the only movement required is to select the text character to be typed. Movement takes time, so the more movement involved in each episode increase the time taken to complete the episode. This is clearly shown in the case of subject 7 who has only slightly reduced manual dexterity.

When an adaptation introduces further delay in moving the cursor across the screen, as in the case of subject 4 and 6, the annotation technique becomes increasingly proportionately slower than text production. Subject 1 and 5 who have reduced manual dexterity but still use mouse and keyboard illustrate the same phenomenon.

Subjects 2, 3 and 8 all used rollerballs and on-screen keyboards. In this case the work involved in selecting annotation tools and placing them was less demanding than that involved in moving around an on-screen keyboard. Because of this, the time taken to complete the picture episodes was reduced in comparison with the time taken to type the equivalent information.

8. Conclusions

These results provide an insight into the work involved in conveying information to another person at the end of a communication link. Improvements in information communication rates can be achieved by some people when they annotate pictures rather than type the information. For the technique to be useful for other users with severe manual dexterity impairments, the procedure for selecting annotation tools and placing them at the appropriate location on the picture will need to be refined to avoid the time delay associated with the computer scanning the cursor across the screen.

9. Acknowledgements

This work is undertaken within the ACTS UMPTIDUMPTI Project AC027 with financial support of the Commision of the European Communities.

10. References

Brodin, J., Magnusson M. (1993). Using New Technology in Communications for People with Speech and Language Disabilities in Sweden. in Technology and Disability, 1993, 2(2) pp 63-70

Brodin, J., (1995) Still Picture Telephony, VideoTelephony and Fax for persons with Mental Retardation, in A Bagdonas (ed.) Learning Difficulties and Disabilities, Vlinius, Lithuania, Vlinius University Press, 1995, pp 20-29

Hine, N.A., Wilkinson, D, Gordon, I.A.S & Arnott, J.L. (1995a). Picture Annotation As A Communication Method For Non-Speaking People. In Proc. ECART 3, 1995, pp 146-149.

Hine, N.A., Wilkinson, D, Gordon, I.A.S & Arnott, J.L. (1995b). An Adaptable User Interface To A Multimedia Telecommunications Conversation Service For People With Disabilities. In Proc. Interact '95, 1995,Chapman & Hall, pp394-397.

IPSNI ii (1995). CD-ROM Integrated Telecommunications For People With Special Needs.