Access to the Text Component of Multimedia Conversation Services for Non-speaking People with Severe Physical Disabilities

Hine N.A, Beattie W., McKinlay A., Arnott J.L.

MicroCentre, University of Dundee, Dundee, UK.

e-mail: nhine@mic.dundee.ac.uk

Abstract. This paper discusses an investigation of text communication as an element of interaction mediated by a broadband telecommunication system. This study is part of a wider investigation being undertaken by the EEC RACE IPSNI II project. The focus of the study was to determine if the typing rate achieved by non-speaking people with additional severe motor impairments could be improved by refining a scanning array. Whilst some improvement was achieved, the rate was still far lower than that of an able bodied typist. Other factors that affect the text production rate are discussed. The results suggests that a scanning text selection method could cause excessive expense for users seeking to undertake real-time text-based communication over broadband telecommunications networks, and alternative approaches are presented.

1. Introduction

The RACE IPSNI II project is a Pan-European collaborative research project investigating the issues governing accessibility of future broadband multimedia services by people with special needs. The goal of the project is to demonstrate that it is technically feasible to enable people with a wide variety of impairments and disabilities to use future multimedia telecommunications services. As such, the project is working with experimental services and terminals in order to ensure that the underlying infrastructure that is currently being developed throughout Europe will be able to meet the needs of people with disabilities.

The project partners identified a set of users that would be considered. These were people with blindness, deafness, motor impairments and those who were non-speaking. For users with these impairments, disability issues were identified, through a series of experiments, when using one or more of the following services: Video Telephony, Multimedia Database Access and Collaborative Working.

The target user groups considered by the team at the University of Dundee were non-speaking people and the service chosen was videotelephony. Experience in a series of preliminary studies confirmed that for non-speaking users the videotelephone needed to be adapted by including a text telephone service component. Various strategies were explored to enable these users to exchange text. An important consideration here was to investigate well established techniques for using computers, to see if they can be utilised in a situation where information is to be exchanged via computer-based systems, between remote users in a real time "conversation". These techniques included alternative keyboards, eye gaze tracking, on-screen keyboards and on-screen scanning arrays [1]. This paper will describe the study of the on-screen scanning arrays to illustrate the issues that are raised when these services available to these users.

2. Study

2.1 Goal

A number of different techniques have been tried and tested to enable people with motor impairments or people who are non-speaking to have conversations using a videophone adapted by the addition of a text telephone. The focus of this study is on those non-speaking people that have severe motor impairments with additional uncontrolled tremors or head movements. These people have a very limited set of possibilities for data input into the text telephone. One of the few techniques of data input available to them is to employ a switch driven scanning keyboard. An initial study at the beginning of the IPSNI II project found, however, that scanning is a very slow method of using a text telephone, where information is exchanged in real time. The results from that study are summarised in Table 1 below. [2]

	WPM
Able Bodied Typing	25.46
Foot Typing	20.59
Scanning Keyboard	0.58
Eye Gaze Typing	6.41

Table 1: The word per minute typing rate achieved by subjects

Because there are few alternatives available to severely disabled, non-speaking people, however, the technique was revisited and a revised scanning array was developed in an attempt to improve the data production rate.

A simple method that could yield useful results involved optimising the location of the letters so that the most frequently used letters were closest to the scan starting point. Alternatively, a method of reducing the effort involved in producing text that has been employed successfully with membrane keyboards to include bi-gram and tri-gram letter combinations [3]. A scanning array was constructed that incorporated bi-grams and tri-grams into the array.

2.2 Method

For the purpose of this study, three scanning array layouts were constructed and tested. The first had the letters of the alphabet arranged in alphabetic order. The second had the letters arranged with those most frequently used in the English language [4]] closest to the point where the scanning started. The third had an alphabetic arrangement of letters, with the 25 most frequently occurring bi-grams and tri-grams occurring in the English language [5] added to the array. The bi-grams and tri-grams were placed in a position in the array at a location close to their first letter. This array is shown in figure 1 below.

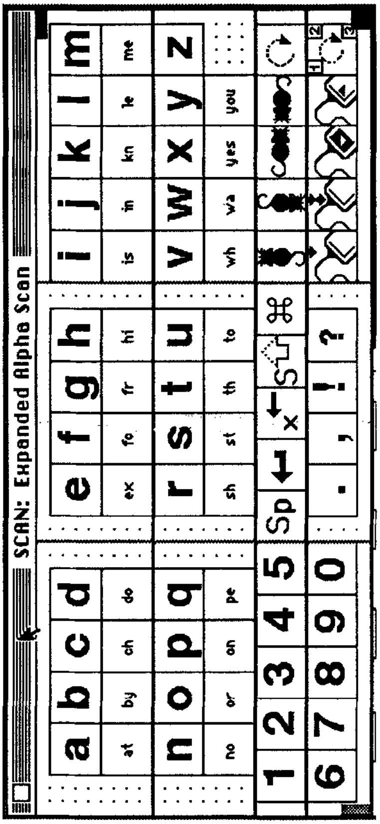

Fig. 1. Scanning Array enhanced with bi-grams and tri-grams

The underlying hypothesis was that by optimising the placement of letters relative to the point where the scans starts, the overall text production rate should improve. In addition, by providing bi-grams and tri-grams, scanning time would be further reduced since fewer characters would need to be scanned and selected.

The software used to construct the scanning arrays was the KE:NX software which allows the scanning to be broken up into blocks. First the rows are scanned, then, when a row is selected, the blocks on the row are scanned, and then, when a block is selected, the elements of the block are scanned, firstly across the top row of the block and then the next row of the block. Each array had a common set of numeric, punctuation and mouse control elements in addition to the letters.

The arrays were tested over a six week period at a cerebral palsy residential and day care centre, using the following procedure. The first array was installed on the host computers, and the subjects had two weeks in which to practice using them, the practice sessions being administered by the training staff at the centre. All of the subjects had previous experience of using KE:NX scanning arrays. The amount of time spent on practise was logged. At the end of two weeks, subjects held a scripted "conversation" via a text telephone, using the scanning array. The script represented an imaginary conversation between a patient and a doctor's receptionist. The time of each keystroke event was recorded by the computer, and the entire conversation was logged on each of the two machines being used for the conversations. Then the next array was installed and the procedure repeated. This was followed by the third array. In all cases the arrays were installed and examined in the order shown in the figures above. Although this did not allow order effects to be eliminated, it was considered that to try to have different subjects practising and being tested on different arrays at any one time would have added too much extra work to already busy centre staff.

2.3 Subjects

Four subjects took part in the study. Each subject was familiar with operation of the Ke:nx software. In addition, each subject had difficulty speaking and needed to enter text with a scanning array operated with a switch. Subject 1 pressed a switch fixed to the back of his wheelchair by moving his head, Subject 2 operated a foot switch, subject 3 held a switch box in one hand and pressed the switch plate with the other and subject 4 pressed a switch that was fixed to a desk. All subjects were adults.

2.4 Apparatus

The system used in the study consisted of two Macintosh LC II computers connected via an Appletalk/Localtalk link. The text telephone was developed using the SuperTalk project development environment. The scanning arrays were referred to as "conventional" where the elements were in alphabetical order, "optimised" where they were in most frequently used order and "enhanced" when bi-grams and tri-grams were added.

2.5 Measurements

The system recorded the time for each keystroke event, and recorded the complete conversation. From this, time between keystrokes and values for typing speed, expressed as words per minute (wpm) could be calculated. Because the scripts that were used were the same as those followed in the earlier study [1,2], the wpm figure was calculated taking the average number of characters-per-word figure from that study. This figure was 5.3 characters per word for these conversations. By using the same scripts for all three conditions, the three arrays could be directly compared. In addition, the performance could be compared with the earlier study.

In addition, the number of errors that were corrected by the subjects using the delete key and typing a preferred character were counted and the percentage of errors per total number of keystrokes were calculated.

3. Results

The average time taken per character selected and the average word per minute (wpm) typing rate is shown in table 2 below, along with the percentage of errors that occurred in each session. The change in typing rate is shown in figure 2 below.

	wpm	errors (%)
Subject 1		
Conventional	0.83	11.89
Optimised	0.86	12.41
Expanded	0.91	6.99
Subject 2		
Conventional	0.39	10.20
Optimised	0.47	23.44
Expanded	0.54	20.18
Subject 3		
Conventional	0.52	16.67
Optimised	0.58	7.69
Expanded	0.70	23.26
Subject 4		
Conventional	0.52	2.60
Optimised	0.55	2.70
Expanded	0.57	5.97

Table 2: The word per minute typing rate and percentage of errors achieved by subjects

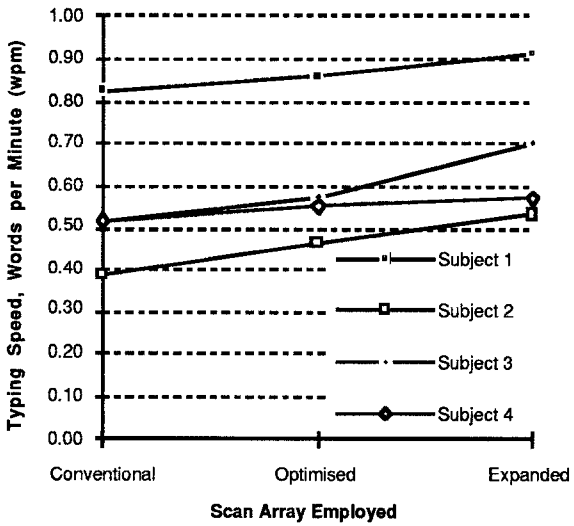

Fig. 2. Change in typing speed per subject when using the three different scanning arrays.

Because of the small sample size and its hetrogenous nature, it was not considered appropriate to analyse the results by means of statistical tests.

4. Discussion

The empirical results show that there was a measurable improvement in the text input as the subjects moved from the conventional to the optimised and then to the enhanced scanning array. Whilst this is true, the improvements were small compared with the vast difference between the performance of an able bodied typist, or even a person typing with eye gaze.

Two trends in the proportion of selections that had to be corrected were observed. The first was that change from a conventional alphabetic layout introduced more errors. This was particularly problematic if the wrong bi-gram or tri-gram was selected as this introduced two or three characters that had to be removed. The second was that the addition of bi-grams and tri-grams reduced the number of errors because the second or third letter was supplied automatically and could not be miss-selected.

Three of factors were observed to be affecting the performance of the subjects which may have a confounding influence on the measured results.

The first factor was that the subjects had very poor ability to handle text, although all had been selected for the study on the strength of their apparent reasonable reading abilities. All subjects used a communication board or book. One pointed to the board with his foot, two used their hands to point to a board that was held for them and the last made noises to confirm or reject selections made from a book on his behalf by someone wishing to communicate. All subjects had exposure to Ke:nx scanning arrays and knew how to operate them. Because of this, they were considered to be suitable candidates for the study, and likely to benefit from it. The study exposed, however, that their use of a communication board was based more on a knowledge of a small set of words and phrases and the patterns necessary to generate those words and phrases rather than on a genuine ability to read or generate text. This poor text handling ability was most evident when the subjects were using the enhanced scanning array. On a number of occasions, the subject did not use a bi-gram or a tri-gram when it was appropriate to do so, particularly if the letters were not at the beginning of the word. It seems that the subjects were not sufficiently fluent in their ability to handle text to recognise and match letter pairs and triplets from the words that they were seeking to type.

It is possible that this problem was compounded by the fact that although the 25 bi-grams and tri-grams on the array were the most commonly occurring ones in the English language generally, they did not occur very frequently in the text used by the subjects. It may be that a larger set would have encouraged the subjects to look more carefully for suitable pairs and triplets to use. This solution is rather problematic, however, in the context of the experiment. The scanning array was only one of a number of function represented on the screen of the computer terminal. Space needed to be left for the text telephone, and in a more complete experimental set-up of the project, for a videophone window and a text prediction window. This limits the space available for a scanning array, which in turn limits the number of elements that can be represented in that array without the label on each scan cell becoming too small to be readable. An obvious solution would have been to use a larger screen, but, for cost reasons, none of the screens in the day centre are larger than 14". This is typical of the type of practical constraint that is encountered when investigating technological solutions: the eventual customer will often not have the resources to be able to afford an ideal solutions and is therefore forced to compromise.

When these observations were discussed with the care and training staff, they offered a further explanation for this poor text handling skill. Technology to assist in the learning of written language skills was not available when the subjects were young and so they had not received systematic training in this skill. Another explanation was that, when they were younger, these people were not considered ever to be likely to be able to read a newspaper and so training in this skill was not considered to have been of primary importance for them. The recommendation from the care staff was that extra attention is paid to the written skills training of children with disabilities so that they will not be handicapped to the same extent.

The second factor was the suitability of the switches being used by the various subjects. The switches that were operated by hand or by pressing with the foot were far from ideal, in that frequently errors were made and cell selections missed because the subject did not press the switch properly. The position of the person relative to the switch was critical. As the subject shifted position slightly their use of the switch could improve or deteriorate significantly.

The third factor that seemed to have a significant influence on the performance of the subject was their motivation and practical situation on a day to day basis. All of the subjects professed to have been interested in the experiments and enjoyed being involved. Over the duration of the study, however, the subjects had to continue to deal with the consequences of having severe disabilities. This meant that their level of concentration and application varied depending on other events that they were dealing with at the time.

5. Conclusions

This study exposes the problems faced by non-speaking people who have an additional severe motor impairment as they seek to take part in real time conversations. Despite improvements in text production rates by all subjects, the rates achieved were very much slower than that of an average able bodied typist. In the context of telecommunications, this would mean that a conversation would take longer and consequently cost more. For this reason, three alternative approaches should be tried depending on the task to be accomplished.

1 Where conversation is possible by selecting from a board or a book, experiments should be conducted to see if it is possible to conduct a conversation with a videophone, perhaps with two cameras. One camera could show the board or book and the other show a view of the person.

2 Where conversation does not need to be in real time, electronic mail should be tested to allow conversation blocks to be created off-line and transmitted when complete.

3. Where a conversation would involve the exchange of a high degree of factual data, a form of computer based interview could be implemented, where the factual data is gathered off line, transmitted and then discussed. in this way, conversation duration could be reduced.

6. References

1. IPSNI II, “Access solutions for people with special needs to telecommunication”. Deliverable to the CEC from RACE Project R2009 IPSNI II, Workpackage (WP 1.4), CEC Deliverable Number R2009/IRV/WP1/DS/P/008/b1, 1992.

2 N. A. Hine, W. Beattie, A. McKinlay and J. L. Arnott, (1994) “Consideration of Scanning Keyboard and Text Prediction in the context of access to Telecommunications Services”, ISAAC: 6th Biennial Conference of the International Society for Augmentative and Alternative Communication, Maastricht, The Netherlands, 9-13 October 1994.

3 Mathy-Laikko, P., West, C. and Jones, R., (1993) “Development and Assessment of a Rate Acceleration Keyboard for Direct-Selection Augmentative and Alternative Communication Users”, Technology and Disability, 2(3), pp. 57

4 Ferguson, T.J. and Rabinowitz, J.H., (1984) "Self-Synchronizing Huffman Codes", IEEE Transactions on Information Theory, Vol. IT-30, pp. 687-693

5 Rubin, D.C. (1978) Word-initial and Word-final ngram Frequencies”, Journal of Reading Behaviour, X, 171-183