Päivi Majaranta and Kari-Jouko Räihä

Twenty Years of Eye Typing: Systems and Design Issues

Proceedings of Eye Tracking Research and Applications [ETRA2002], pages 15-22, New Orleans LA, ACM, 2002.

This copy is posted by permission of ACM and may not be redistributed. The definitive copy of the paper can be downloaded from the ACM's Digital Library: ETRA Proceedings.

ACM COPYRIGHT NOTICE. Copyright © 2002 by the Association for Computing Machinery, Inc. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Publications Dept, ACM Inc., fax +1 (212) 869-0481, or `permissions@acm.org`.

Twenty Years of Eye Typing: Systems and Design Issues

Päivi Majaranta and Kari-Jouko Räihä
Human-Computer Interaction Unit
Department of Computer and Information Sciences
FIN-33014 University of Tampere, Finland
{curly, kjr}@cs.uta.fi

ABSTRACT

Eye typing provides a means of communication for severely handicapped people, even those who are only capable of moving their eyes. This paper considers the features, functionality and methods used in the eye typing systems developed in the last twenty years. Primary concerned with text production, the paper also addresses other communication related issues, among them customization and voice output.

Keywords

Eye typing, eye tracking, alternative communication.

1. INTRODUCTION

Eye typing means producing text by using the focus of the gaze. There may be several reasons for using eye typing, but it is most needed by people with severe disabilities. Control of the eyes may well be the only option for the severely handicapped, thus their need for a communication system is acute.

Communication is not synonymous with typing: in this paper, we are mostly interested in eye controlled text entry systems. However, considerations of communication are strongly present, since most of the systems described here are designed for use as communication systems.

Because gaze has certain specific features, we will start by briefly discussing gaze as an input method. We will then continue by describing the eye typing procedure. Since eye typing systems are usually implemented using an on-screen keyboard, we also need to consider keyboard layout.

Furthermore, we will discuss some special issues arising when the eyes are the only means of communication and interaction with the outside world.

2. GAZE INPUT

Used as an input method, gaze has both advantages and disadvantages. People use gaze to obtain information about the surroundings. The direction of gaze shows the focus of our attention [Just and Carpenter 1976].

We look at things by fixating on them, *i.e.*, holding the gaze relatively still for a short while (typically 200-600 ms [Jacob 1995]). Between fixations our gaze jumps rapidly from one point to another. The jumps are called saccades. Saccades are ballistic movements lasting about 30-120 ms. Once a saccadic jump has been initiated it cannot be interrupted nor can its direction be changed. A gaze path consists of fixations and saccades. Information is only acquired during fixations.

For a description of other types of eye movements, like microsaccades and pursuit, see [Jacob 1995].

Eye movements are both conscious and unconscious. We do not usually think about eye movements, they are more or less automatic. On the other hand, we can, if necessary, control the gaze at will. Moreover, the gaze does not always tell where our attention is: we may stare at something but our attention is elsewhere. This creates problems when the same modality is used for both controlling and acquiring information. A problem called the Midas touch arises when the user fixates on some item solely to obtain information and the program interprets it as an input command. Things may then be selected against the user's inclination. [Jacob 1991]

Gaze can be considered as a natural mode of input. It is quite easy to focus on items only by looking at them. It is so intuitive a method that it requires very little or no training at all [Stampe and Reingold 1995]. However, this is not always the case with disabled people. Severely disabled from birth are not used to controlling anything physically. They first need to understand that they can control the cursor by moving their eyes. It may take from minutes to months to master a gaze-controlled system [Gips et al. 1996].

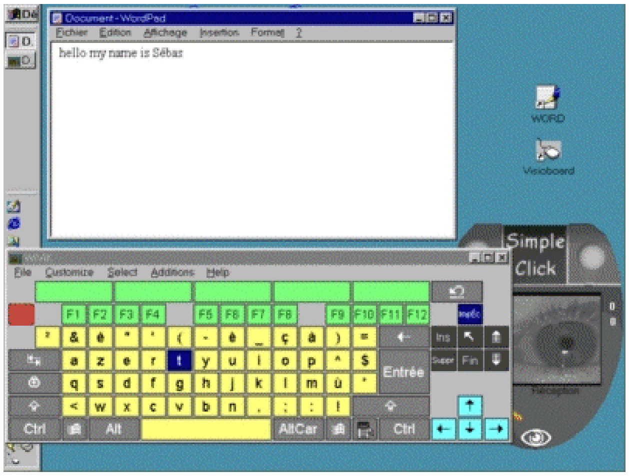
Gaze is also a very fast pointing device, faster than a mouse or certain other pointing devices, assuming the targets are large enough [Sibert and Jacob 2000; Ware and Mikaelian 1987]. Gaze is not as accurate as the mouse, because the size of the fovea restricts the accuracy of the measured point of gaze to 0.5 degrees. Another major problem is the accuracy of the eye tracking device. The newly calibrated device may be quite accurate at first but after a while the calibration starts to drift. The consequence of drifting is that the measured point of gaze is a few pixels off the actual point of gaze. Despite efforts to develop dynamic techniques to correct the drifting (*e.g.*, by Stampe and Reingold [1995]), it remains one of the biggest problems of gaze input, together with the Midas touch problem.

3. TYPING BY GAZE

Most of the eye typing systems are implemented by having a virtual keyboard on the screen. An eye tracking device tracks the user's point of gaze and a computer program analyzes the gaze behavior. Based on the analysis, the system decides which of the letters the user is focusing on, and whether the user wants to type it. There are many ways to type by gaze, to be described later in this paper, but first we present a simplified example of how eye typing typically proceeds.

First, the user decides which letter s/he wants to type and *focuses* on it by looking at one of the virtual keys on the on-screen keyboard. The system gives *feedback* by highlighting the focused letter, for example, by showing a cursor on the virtual key. If the user wants to *select* the focused letter, s/he continues to fixate on it thus using time as an activation command. Alternatively, s/he may also trigger a switch to select the focused item. The system may give auditory feedback indicating that a virtual key press was successful. Finally, the typed letter appears in the text field located above or below the virtual keyboard.

For example, the VISIOBOARD [2001] system illustrated in Figure 1 uses the WiViK© on-screen keyboard (http://www.prentrom.com/access/wivik.html). The text appears in a Microsoft© WordPad window. The cursor position indicates the point of gaze. Selections are made either by using time as an activation command, or by using an eye blink or some manual device (*e.g.,* a button) as a switch.



hello my name is Sébas

Figure 1. The VISIOBOARD [2001] communication system.

We will describe the eye typing process (focusing, getting feedback and selecting) in more detail in the following subsections.

3.1 Focus

Before the user can select anything on the screen, the item to be selected must be focused on. As with the conventional mouse, the user first needs to move the cursor over the relevant item, before s/he can make the selection.

Many of the eye tracking devices can actually be used to emulate a conventional mouse; with the eyes replacing the mouse. For example, the EagleEyes [Gips et al. 1993] and EyeWriter [Lileg et al. 1999] systems can be used to move a cursor around the screen. Both EagleEyes (Figure 2) and EyeWriter are based on electro-oculographic potential (EOG) indicating the position of the eye relative to the head. The EOG changes when the user moves either the eyes or the head, thus enabling the user to control the cursor position by moving the eyes, or head, or both. In video-based systems, moving the eyes alone moves the cursor.

Figure 2. Eye painting with EagleEyes, electrodes measure the EOG [Gips and Olivieri 1993].

The focus can be shifted from item to item by a brief fixation. The duration of continuous fixations on an item is called *dwell time*. Experienced users may use a shorter dwell time than inexperienced users. The maximum dwell time is usually about one second. If dwelling is used for focusing, the system usually provides the user with an indication that the predefined dwell time has elapsed. The feedback methods are reviewed later in this paper. Mouse emulation and dwell time are the most used methods for focusing on an item, using the point of gaze to define the location of focus.

Some people may have difficulties in fixating because of their state of health. They cannot sustain their gaze still for the duration needed to focus. The user may also be able to move his or her eyes in one direction only (*e.g.*, in locked-in syndrome [Chapman 1991]). In such cases, other methods for selecting an item are needed. The eyes can be used as simple one- or two-way switches and the focus can be shifted from one item to another by using a method called scanning [Frietman 1984; Kahn et al. 1999; Kate et al. 1979; Shein 1997].

The scanning method is widely used among the disabled in various computer-aided communication systems. Scanning techniques allow the user, for example, to use one switch to change the focus from one item to another, and another switch to select the item in focus. This is called step scanning.

Automatic scanning is used if the user has only one switch. The focus automatically shifts from item to item after a definable time and all the user has to do is to activate the switch when the desired item (*e.g.*, a letter) is in focus. There are also a number of more advanced scanning methods [Nisbet and Poon 1998; Shein 1997] outside the scope of this paper. Most of them are variants of the basic methods described here.

Scanning can proceed from item to item, focusing on one single item at a time. Scanning is quite a slow procedure, and therefore it is advantageous to scan larger groups of items before focusing on the single items. A common approach is to scan rows first and then proceed to the individual items in the focused row (see Figure 3).

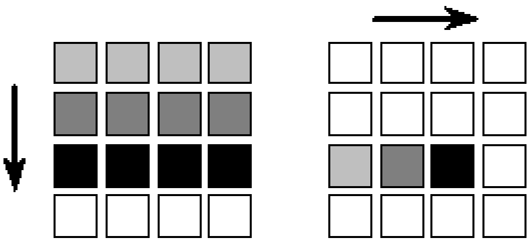


Figure 3. Sequential row-column scanning [Shein 1997].

Scanning has been successfully combined with an eye tracking system, for example, in the VisionKey system [Kahn et al. 1999] and the Eye-Switch Controlled Communication Aids [Kate et al. 1979], requiring only coarse vertical or horizontal eye movements.

In the Eye-Switch Controlled Communication Aids the letters and function items are arranged into a matrix. The user can initiate automatic scanning by glancing 15° to the left. The system automatically scans the columns in order, until the user stops the scanning by glancing to the right. The focused column is highlighted. Next the system starts scanning the rows. When the focused and highlighted row intersects with the column with the desired letter, the user can stop the scanning by glancing to the right.

3.2 Feedback

When a traditional keyboard is used, the user can feel the keys and hear the clicks while typing. S/he can also observe the changes on the screen as the letters appear. Before the user of an eye typing system can select the focused item, the system should give clear feedback. Istance *et al.* [1996] observed that the users sometimes misunderstood where the focus was and operated on the wrong items. Lack of feedback also caused the same key to be inadvertently selected twice.

Adequate feedback is especially important for a disabled person who possibly has never had any control over anything. For example, the developers of the LC Eyegaze System [Chapman 1991] observed that initially most users experienced difficulties in sensing where their gaze was fixed on the screen. A small red cursor was added to show the location of the gaze point. They report that most of the experienced LC Eyegaze users simply "dragged" the cursor around the screen, not consciously fixating (with dwell time) on each key [Cleveland 1994]. Most of the gaze-controlled systems have an option to show an explicit cursor on screen. In Figure 2, the user is painting with her eyes. In this case, the feedback is quite clear as the eye cursor leaves a bold line that follows the user's gaze path.

Moving the cursor as the point of gaze moves may cause problems [Jacob 1995]. A moving cursor may distract the user's attention. If the calibration is not accurate, the cursor will not be exactly where the user is looking. If the user starts to chase the cursor, s/he will never get to accomplish the task. If the cursor is visible, the system should recheck the calibration often enough – possibly even at every successful selection – to dynamically correct the drift.

A better way to show the focus is to highlight the item, for example, by changing the background color (as in Figure 1) or by drawing a thick colored border around the selected item (as in Figure 8). There may be a short delay in the highlighting so that not every item is highlighted as the user's gaze is moved around the screen. In many cases, the slight delay indicates that the predefined dwell time has elapsed and the item is now focused.

The feedback should clearly reflect the selection process: which item is focused, when, and when the item can be selected. The cursor (its color, border or shape) or the highlighting should change as an indicator of the stage of the selection process. The system may warn the user that the dwell time is about to end and the item will soon be selected. A shrinking rectangle is used in the Eye-gaze Response Interface Computer Aid system (ERICA) [Hutchinson et al. 1989] to indicate how the selection is proceeding. First a rectangle marks the key the user is fixating upon. If the user maintains the fixation, the rectangle starts to shrink. When the rectangle ceases to shrink, the key is selected and the action associated with the key is performed [Lankford 2000].

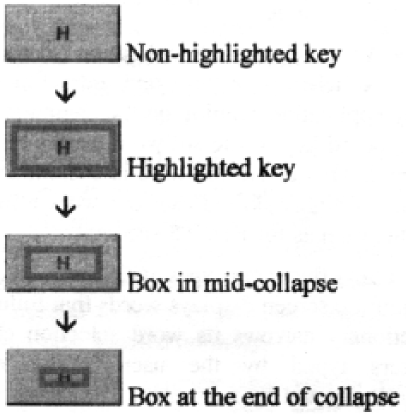


Figure 4. The shrinking box in ERICA [Lankford 2000].

Some systems also provide auditory feedback. For example, the system can produce a simple sound (like a "beep" or "click") when an item is focused and/or selected. Alternatively, the clicking sounds of a standard keyboard can be emulated. The sound effects are often used together with the visual feedback [Hutchinson et al. 1989].

3.3 Selection

When the desired item is focused (and highlighted) the item still has to be selected. In many cases, the focusing, feedback and the actual selection are linked in such a way that the user may not perceive them as separate operations.

Dwell time, for example, can be used for both focusing and selecting: first the item is highlighted and subsequently selected. Dwell time is the most common method both for focusing and for selecting an item. A typical dwell time is between 600 and 1000 milliseconds. The selection process (dwelling) can be interrupted by simply looking away from the focused item before time runs out.

An alternative is to use *blinks* or *winks* [Rasmusson et al. 1999; VISIOBOARD 2001] for selection. If a blink is used as a switch to select the focused item, it must be separated from the natural blinks occurring unconsciously [Jacob 1991]. *Wrinkles* or some other facial muscle activity could also be used to select a focused item, as suggested by Partala *et al.* [2001], but this can be tiring or even impossible for some users.

Additional switches could, of course, be used [Rasmusson et al. 1999], but only if the user has some control over his or her muscles. In that case, the user might not need an expensive eye tracker at all. There are numerous inexpensive and easy-to-use communication devices for users who can operate switches by rough limb or head movements [Beukelman and Mirenda 1992]. However, if the body movements require a lot of effort, the eye movements would probably be less tiring.

The eyes themselves can also be used as switches; a horizontal or vertical glance can trigger a command. Glancing left, right, up or down can be interpreted as a command to select or deselect an item. The advantage of this method is that it can also be used by people with impaired eye control.

Off-screen targets could also be used for triggering commands as suggested by Isokoski [Isokoski 2000]. In his system gaze target points are placed around the screen. The user gazes at them in a certain order to produce a sequence of eye glances (*e.g.*, north-south-west) to compose a command.

The VisionKey [Kahn et al. 1999] system applies a somewhat different and interesting selection method. It uses fairly coarse eye movements to shift the selection around the screen. Letters are displayed in a grid. One cell contains four characters. Normally, the screen is in a "rest and plan" state. The user can look around the screen without selecting anything, avoiding the Midas touch problem. To start a command (see Figure 5) the user has to glance at a corner of the key chart and then (1) glance back at the center. Then the user (2) looks at the top right corner, indicating that the desired letter (in this case 'G') is located in the top right corner of a cell. The top right cell will be highlighted. Then the user (3) glances back at the center, and from there (4) the gaze moves to the target cell, in which 'G' is located. The 'G' cell will be highlighted until a 'beep' is heard to indicate a successful selection. [VisionKey 1995]

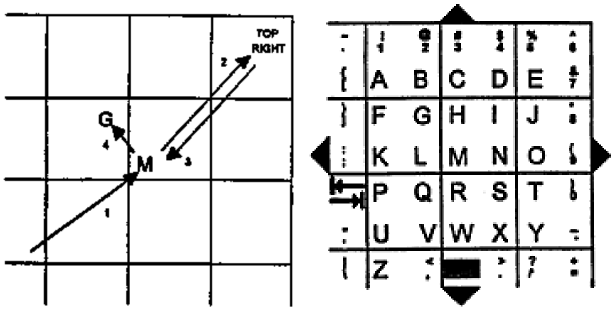


Figure 5. The VisionKey selection method [VisionKey 1995].

If a wrong key is highlighted, the user can drag the highlight (the cursor) to another cell, which means there is no need for exact accuracy.

As VisionKey demonstrates, the standard QWERTY key layout is not the only option. Various alternatives are described next.

4. KEYBOARD LAYOUT

For disabled users the standard QWERTY key layout is not always best. The user may not be familiar with it and it takes time to learn where all the keys are. Alphabetic order may be faster to learn. The keys could also be organized so that the most common letters are grouped together. For systems based on scanning, the most frequently used letters are often placed close to the item from which the scanning starts [Nisbet and Poon 1998]. The order of letters depends on the language. For instance, in the eye-switch communication system developed by Kate *et al.* [1979] the letters were arranged according to their frequency in the Dutch language.

Because of the inaccuracy of the gaze, the selectable items on the screen must be large enough. Generally, as the size of the items increases, so does the selection accuracy. However, after a certain point (4 degrees) there is no longer any increase in accuracy [Stampe and Reingold 1995]. The convenient size of the target items depends on the eye tracking device, screen resolution and the distance between the user and the screen.

The demand for fairly large items on the screen leads to lack of screen space. If all the needed keys of the virtual keyboard are to be simultaneously displayed on screen, little space will be left for other elements (see Figures 1 and 8).

Luckily, the screen space is no longer the problem it was some years ago, when the monitors and the resolution of the screen were smaller. For example, in the early stages of the ERICA system [Hutchinson et al. 1989], only six selectable items were available on screen at a time. Character entry was accomplished using a tree-structured menu hierarchy. It took from two to four menu selections to select a single letter. First, the user selected a group of letters, then either another group of letters or the single target letter. The letters were arranged to minimize (on the average) the number of steps needed to write a word [Frey et al. 1990].

A similar hierarchical method is also used by the EagleEyes system [Gips and Olivieri 1996]. It has only two levels (see Figure 6) and few special keys. The upper row consists of groups of letters. The letters of the selected group appear in the boxes below the text field. The bottom row of boxes includes a space key, a key to speak the written text and a delete key, or a return key, depending on the state of the program.

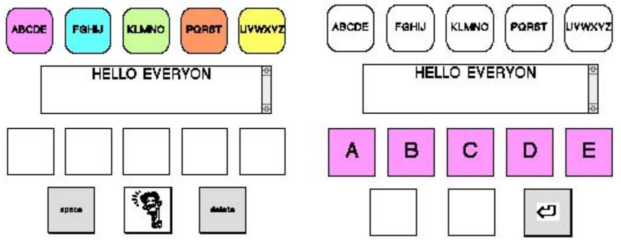


Figure 6. Two screens from EagleEyes' two-level speller [Gips and Olivieri 1996].

There are also other reasons to use only a few keys on the keyboard layout. Hansen *et al.* [2001] have been developing a system that tracks eye movements with a standard web camera. They emphasize the need for an affordable eye communication system, because the current eye trackers are quite expensive. The accuracy of a web camera is only about 4 degrees. That is why there cannot be many selectable items on the screen at a time. The type-to-talk system developed by Hansen *et al.* has 12 large on-screen keys. The system utilizes text entry methods developed for mobile phones; it is a hybrid of several different methods including the so-called "T9 input method" and a probabilistic character layout strategy. Every time a character has been entered, the 6 most likely characters to follow the current one and the 6 most likely complete words are shown. Initial tests have shown that users can produce up to 10 words per minute once they get used to the probabilistic character layout.

5. CUSTOMIZATION AND CONTEXT OF USE

We have seen that there are many alternative ways to select items and to arrange the layout of the keyboard. The best method depends on the skills, abilities and past experience of the user. Therefore it is highly desirable that the system allows the user to change the various properties.

If possible, the size, color, location and the contents, as well as the functioning of the on screen buttons should be customizable. The user should be able to define what the buttons look like and what they do.

Another target for customization is the cursor. Whether to show it or not, in what size, and whether it is always visible or only after a period of fixating, should be customizable attributes.

Experienced users may also want to adjust the dwelling time, both for focusing on an item and for selecting. Also, the feedback attributes should be adjustable; the user should be able, for example, to control the volume or turn the sound off.

Using an eye typing or communication system may be fatiguing, even though some users have reported using an eye communication system for hours. An able-bodied user of a conventional mouse can always let go of the mouse, but the user of an eye cursor cannot, except by shutting his/her eyes. The possibility to disengage eye-control is important, because it gives the user a chance to rest and to look around without issuing commands [Istance et al. 1996; Jacob 1995].

In addition to communication, the user can use an eye mouse to control other applications. The level of integration with the operating system (like Windows or Macintosh) determines whether the system can be used to control just a few predefined applications or any standard application. The latter usually means that the eye control system must be able to emulate the conventional mouse or keyboard or both.

There are usually many small elements included in the standard graphical environment. Gaze alone is not accurate enough for operating such small items as menu items. If we want the eye typing system to have access to all the functionality of any standard word processing program, the system should provide a method for selecting the tiny objects on the screen. Zooming, screen magnifiers and fish-eye views have been used in Quick Glance [Rasmusson et al. 1999], ERICA [Lankford 2000], and Viserg Eye Mouse ([Istance et al. 1996], see also http://www-rcf.usc.edu/~wdutton/comm533/EYEM-WU.htm).

6. TYPING VS. COMMUNICATION

How fast is eye typing? An eye typist can type about one character in a second [Gips and Olivieri 1993; Kahn et al. 1999; Kate et al. 1979]. Even though the user can theoretically eye type 60 characters per minute, in practice the typing rate is often slower [Frey et al. 1990], possibly only one [Istance et al. 1996] or a few words in a minute. This is very slow compared to the normal typing rate of an able-bodied typist. The slow typing speed is an even greater problem in everyday face-to-face communication. It is hard to communicate if one can only produce a few words per minute while the normally speaking interlocutor can easily speak over 100 words a minute.

Some efforts have been made to speed up eye-based communication. Salvucci [1999] developed advanced methods where the user can look around the virtual keyboard without the dwell time delay. The system analyses the gaze path and tries to map the fixations on letters. It uses a dictionary and predefined grammar of how the letters follow each other when deciding whether the fixations belong to the word or not. In the latter case the user is probably just gazing around, searching for the correct letter. The problem with this method is that it takes time to deduce if the fixations belong to a word or not. Moreover, the accuracy of the system decreases when the number of words in the dictionary increases.

A more common approach is to set eye typing in the more general context of communication. The goal of the user is not just to type letters but to produce phrases and sentences. Writing a message letter by letter is slow. Phrases for every day usage could and should be included in the program. The system should support editable phrases, because the needs of the disabled users vary a lot. Since not all the phrases can be visible at a time, they can be arranged into a tree structure, as in the phrase selection menu of LC Eyegaze [Chapman 1991] (see Figure 7).

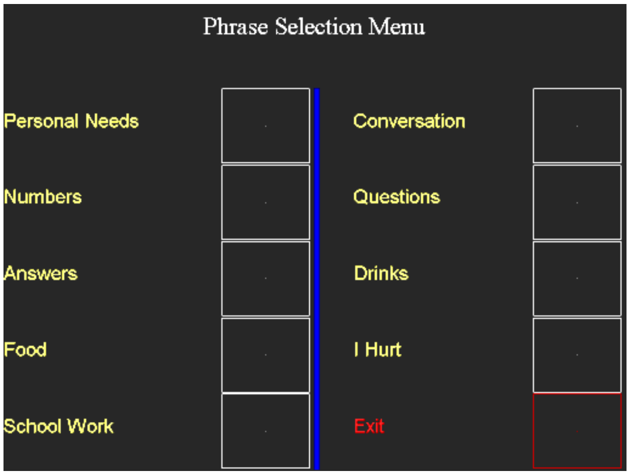


Figure 7. The phrase selection menu of LC Eyegaze [Chapman 1991].

The gaze communication system may also have a sentence buffer for predefined strings that can be joined together. The Eyetracker communication system [Friedman et al. 1982] provides the option to select first a standard phrase (*e.g.*, "please give me") and then complete it with another ("a drink of water").

Not all disabled people are able to read or write, but instead use pictures and icons for communication. Therefore, an eye gaze based communication system should support pictures in addition to text. There are several kinds of communicative pictures in use, including PCS (Picture Communication Symbols), Rebus, Makaton, Minspeak, Picsyms and Bliss [MacDonald 1998]. To meet the various needs of disabled users, the choice of symbol set should also be customizable.

In addition to flexible forms of input, computer-based communication systems typically also make use of several forms of output. In particular, they are able to speak the typed message aloud. Voice output is important for many reasons. The user needs a means to attract attention in order to communicate. Besides, the interlocutor may not see the computer display.

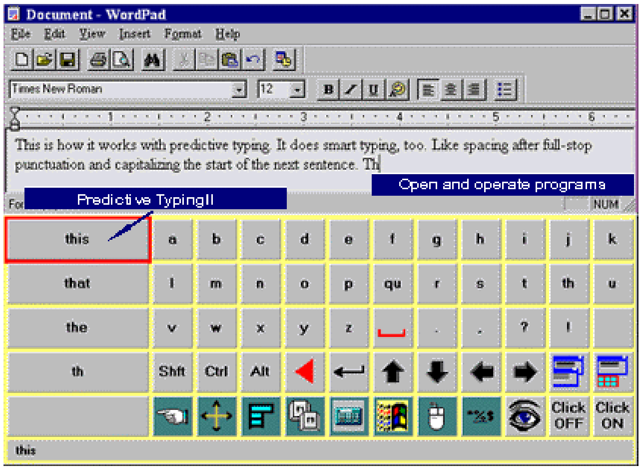
The quality of voice output affects the intelligibility of the message, but there are other things to be taken into account, too. Friedman *et al.* [Friedman et al. 1982] found out that even though the able-bodied children liked synthetic speech, the speech impaired did not. They preferred to use a speech system that was sex and age-specific, because they wanted their communicator to sound the way they should sound.

7. SYSTEM LEVEL ISSUES

So far we have concentrated on the design of interaction and communication. We will now turn to implementation issues. What kind of software components can be used, so that each new eye typing system does not need to be implemented from scratch? What are the hardware requirements set by the operation environment?

7.1 Virtual Keyboards and Communication Aids

There are a number of virtual keyboards aimed at users unable to use a standard keyboard or mouse. The on-screen keyboards can typically be operated by a conventional mouse, or by an alternative input device that can trigger mouse events. Therefore, any eye tracking system that can emulate the mouse can probably be used to control such keyboards. There may be advantages in using existing on-screen keyboards avoiding the need to redo well-done work like implementing a word prediction system, icons that activate various window handling commands, and so on. Many of these have already been tested with real disabled users and have advanced setup features for customizing the keyboard to meet the user's individual needs.



This is how it works with predictive typing. It does smart typing, too. Like spacing after full-stop punctuation and capitalizing the start of the next sentence. Th

Figure 8. HandsOFF! supports predictive typing.

Microsoft Windows, for instance, contains a simple character chart that can be operated by mouse or by some pointing device through a touch screen, but it does not support alternative input devices very well. Nor is it customizable; the user cannot resize it or change the order of the keys. In contrast, several virtual keyboards have been implemented especially for disabled users. Examples include WiViK$^{©}$2 (http://www.prentrom.com/access/wivik.html), SoftType (http://www.orin.com/) and HandsOFF! (http://www.zygo-usa.com/handsoff.html, see Figure 8). They emulate the standard keyboard and can be used to control several (if not all) standard applications.

Many of the on-screen keyboards include a set of on-screen keys to emulate mouse control, too. For communication purposes, many of them include support for synthesized speech output. Most of the on-screen keyboard programs can display graphics, pictures and symbols for those who have difficulties with text [Nisbet and Poon 1998].

The on-screen keyboards that are specially made for disabled users also include many special features like "sticky keys" for selecting key combinations such as Ctrl+C, support for scanning methods, word predicting, predefined phrases, and support for environment control. Many of them are fully integrated into the operating environment, meaning that they can be used to control most of the standard applications like WordPad in Microsoft Windows (see, for example, Figure 8 of the HandsOFF! on-screen keyboard).

7.2 Portability

Even though the user cannot move independently, it is advantageous if the system is portable enough to be easily moved from place to place. Some systems can be attached to a wheelchair so that the user can control the wheelchair by the eyes (*e.g.*, LC Eyegaze [Chapman 1991], VisionKey [Kahn et al. 1999], and EagleEyes [Gips et al. 1993]). For preliminary results of using EagleEyes to control a wheelchair see http://www.cs.bc.edu/~eagleeye/papers/paper3/paper3.html.

Some issues affect the portability of the system. First, the size matters. It is not easy to move a heavy system around, not to mention mounting a heavy system on a wheelchair. The system's sensitivity to body movements also merits consideration; the system may not be able to cope with compulsive movements. The lighting conditions may change radically. Many video based systems, cannot function properly in bright daylight.

The VisionKey system, for example, is easy to transport: it is attached to eyeglasses (see Figure 9), and it has a portable control unit with an LCD display for the text output. Thus, no computer is required and the system can be used on its own.

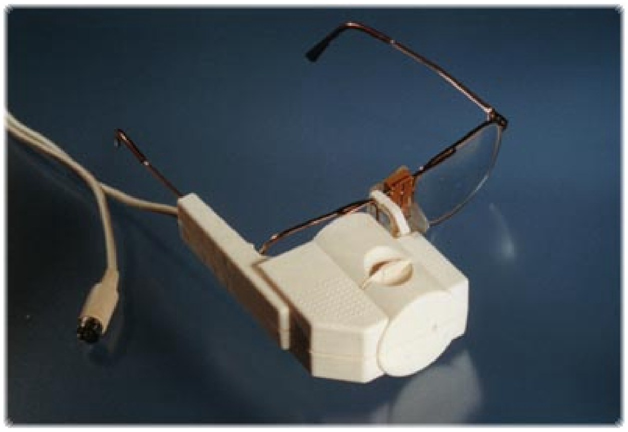


Figure 9. VisionKey mounted on eyeglasses.
(http://www.eyecan.ca/pages/faq.htm)

Finally, it is important to remember to consult the potential users in the early stages of development. Even if the able-bodied test subjects give positive feedback on an eye communication system, one really cannot tell how usable it will be until it is tested with the intended end-users [Hutchinson 1989].

8. DISCUSSION

Even though there are a number of systems that support gaze based typing, further development is still needed. The eye tracking devices and the algorithms that track the gaze still need to be improved – many of the most critical problems arise from the fact that the system does not retain its accuracy for a long period. If the user can only control his or her eyes, how is s/he supposed to be able to start up the calibration process, if the system no longer reacts to his or her gaze? Even though many systems dynamically correct the calibration drift, only few have a procedure for re-calibration that can be initiated by the user.

From the research perspective, even though there are many eye-based communication systems actually in use by disabled people, the typing issues have not been studied in much detail. There are many interesting issues arising from the typing task that have largely been overlooked. For instance, methods for editing text, selecting a chunk of text (word, line, or paragraph), scrolling the text, and undo methods need to be improved. Also, the typing rate is still quite slow – too slow for efficient communication.

The interaction between the virtual keyboard and the text field should also be studied more carefully. How does the gaze move between the virtual keys and the field? Also, the feedback may not give much information about what is happening, compared to an ordinary keyboard that gives both tactile and auditory feedback, while at the same time the user can observe the changes on the screen.

Eye typing provides a rich set of issues for study, both from the practical point of view in order to develop more usable systems, and from the basic research point of view to understand better the properties of gaze in communication tasks. Much of the previous research is scattered in sources that are not commonly known or readily available, thus surveys like this, and the preceding work by Istance *et al.* [1996], are needed. The goal of this paper is to stimulate research in this field and to provide a jump-start for researchers wishing to concentrate on the various design problems.

9. ACKNOWLEDGMENTS

We would like to thank Dr. Edward E. E. Frietman and Dr. Jacques Charlier for providing material for this survey. Päivi would also like to thank Antti Aaltonen for constant encouragement.

10. REFERENCES

BEUKELMAN, D.R. AND MIRENDA, P. 1992. Augmentative and Alternative Communication: Management of Severe Communication Disorders in Children and Adults. Paul H. Brookes Publishing Co, Baltimore.

CHAPMAN, J.E. 1991. Use of an eye-operated computer system in locked-in syndrome, in *Proceedings of the Sixth Annual International Conference on Technology and Persons with Disabilities* (*CSUN'91*), Los Angeles, CA. Available at http://www.eyegaze.com/doc/csun91.htm (28.9.2001).

CLEVELAND, N.R. 1994. Eyegaze Human-Computer Interface for People with Disabilities, in *Proceedings of 1st Automation Technology and Human Performance Conference*, Washington, DC. Available at http://www.lctinc.com/doc/cathuniv.htm (28.9.2001).

FREY, L.A, WHITE, K.P. JR., AND HUTCHINSON, T.E. 1990. Eye-Gaze Word Processing, *IEEE Transactions on Systems, Man, and Cybernetics 20* (*4*), 944-950.

FRIEDMAN, M.B., KILIANY, G., DZMURA, M., AND ANDERSON, D. 1982. The Eyetracker Communication System, *Johns Hopkins APL Technical Digest 3* (*3*), 250-252.

FRIETMAN, E.E.E., BRUGGEMAN, F., AND VAN SPRONSEN, C.J. 1984. Eyeball-Position-controlled Communication System EPCOS, *Microprocessing and Microprogramming 14* (*3*), 173-175.

GIPS, J., OLIVIERI, C.P., AND TECCE, J.J. 1993. Direct control of the computer through electrodes placed around the eyes, in M.J. Smith and G. Salvendy (eds.), *Human-Computer Interaction: Applications and Case Studies*, Elsevier, Amsterdam, 630-635.

GIPS, J. AND OLIVIERI, P. 1996. EagleEyes: An Eye Control System for Persons with Disabilities, in *Proceedings of the Eleventh International Conference on Technology and Persons with Disabilities* (*CSUN'96*), Los Angeles, CA. Available at http://www.cs.bc.edu/~eagleeye/papers/paper1/paper1.html (28.9.2001).

GIPS, J., DIMATTIA, P., CURRAN, F.X., AND OLIVIERI, P. 1996. Using EagleEyes – an Electrodes Based Device for Controlling the Computer with Your Eyes – to Help People with Special Needs, in J. Klaus, E. Auff, W. Kremser, and W. Zagler (eds.), *Interdisciplinary Aspects on Computers Helping People with Special Needs,* Oldenburg, Vienna, 77-84. Available at http://www.cs.bc.edu/~eagleeye/papers/paper2/paper2.html (28.9.2001).

HANSEN, P.J., HANSEN, D.W., AND JOHANSEN A.S. 2001. Bringing Gaze-based Interaction Back to Basics, in *Proceedings of Universal Access in Human-Computer Interaction (UAHCI 2001)*, New Orleans, Louisiana.

HUTCHINSON, T.E., WHITE, K.P., MARTIN, W.N., REICHERT, K.C., AND FREY, L.A. 1989. Human-Computer Interaction Using Eye-Gaze Input, *IEEE Transactions on Systems, Man, and Cybernetics 19* (*6*), 1527-1534.

ISOKOSKI, P. 2000. Text Input Methods for Eye Trackers Using Off-Screen targets, in *Proceedings of the Eye Tracking Research & Applications Symposium 2000* (*ETRA'00*)*,* Palm Beach Gardens, FL, ACM Press, 15-21.

ISTANCE, H.O., SPINNER, C., AND HOWARTH, P.A. 1996. Providing motor impaired users with access to standard Graphical User Interface (GUI) software via eye-based interaction, in *Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies* (*ECDVRAT '96*)*,* Maidenhead, UK. Available at http://www.cyber.reading.ac.uk/P.Sharkey/WWW/icdvrat/WWW96/Papers/istance.htm (28.9.2001).

JACOB, R.J.K. 1991. The use of eye movements in human-computer interaction techniques: what you look at is what you get, *ACM Transactions on Information Systems 9* (*3*), 152-169.

JACOB, R.J.K. 1995. Eye Tracking in Advanced Interface Design, in W. Barfield and T.A. Furness (eds.), *Virtual Environments* *and Advanced Interface Design*, Oxford University Press, New York, 258-288. Available at http://www.cs.tufts.edu/~jacob/papers/barfield.html (28.9.2001).

JUST, M. A., AND CARPENTER, P. A. 1976. Eye fixations and cognitive processes. *Cognitive Psychology 8*, 441-480.

KAHN, D.A., HEYNEN, J., AND SNUGGS, G.L. 1999. Eye-Controlled Computing: The VisionKey Experience, in *Proceedings of the Fourteenth International Conference on Technology and Persons with Disabilities* (*CSUN'99*), Los Angeles, CA. Available at http://www.dinf.org/csun_99/session0263.html (28.9.2001).

KATE, J.H. TEN, FRIETMAN, E.E.E., WILLEMS, W., TER HAAR ROMENY, B.M., AND TENKINK, E. 1979. Eye-Switch Controlled Communication Aids, in *Proceedings of the 12th International Conference on Medical & Biological Engineering*, Jerusalem, Israel.

LANKFORD, C. 2000. Effective Eye-Gaze Input Into Windows, in *Proceedings of the Eye Tracking Research & Applications Symposium 2000* (*ETRA'00*)*,* Palm Beach Gardens, FL, ACM Press, 23-27.

LILEG, E., WIESSPEINER, G., AND HUTTEN, H. 1999. Evaluation of the EOG for communication through eye movements, in *Proceedings of 10th European Conference on Eye Movements* (*ECEM 10*), Utrecht University, the Netherlands. Available at http://www.phys.uu.nl/~ecem10/Abstracts/Lileg_E.html (28.9.2001).

MACDONALD, A. 1998. Symbol Systems, in Augmentative Communication in Practice: An Introduction, (2nd ed.), 19-26. Available at http://callcentre.education.ed.ac.uk/acpsbook/introbook.pdf (28.9.2001).

NISBET, P. AND POON, P. 1998. *Special Access Technology*, University of Edinburgh. Chapters of the book are available at: http://callcentre.education.ed.ac.uk/satbook/satbook101-124.pdf (Ch. 7: "On-screen Keyboards"), and http://callcentre.education.ed.ac.uk/satbook/satbook125-150.pdf (Ch. 8: "Switch Access using Scanning and Encoding") (28.9.2001).

PARTALA, T., AULA, A., AND SURAKKA, V. 2001. Combined Voluntary Gaze Direction and Facial Muscle Activity as a New Pointing Technique, in M. Hirose (ed.), *Proceedings of the Eighth IFIP TC13 Conference on Human-Computer Interaction* (INTERACT2001), IOS Press, 100-107.

RASMUSSON, D., CHAPPELL, R., AND TREGO, M. 1999. Quick Glance: Eye-tracking Access to the Windows95 Operating Environment, in *Proceedings of the Fourteenth International Conference on Technology and Persons with Disabilities* (*CSUN'99*), Los Angeles, CA Available at http://www.dinf.org/csun_99/session0153.html (28.9.2001).

SALVUCCI, D.D. 1999. Inferring intent in eye-movement interfaces: Tracing user actions with process models, in *Human Factors in Computing Systems: CHI 99 Conference Proceedings*, Pittsburgh, PA, ACM Press, 254-261. Available at http://www.mcs.drexel.edu/~salvucci/CHI99/ (28.9.2001).

SHEIN, G.F. 1997. Towards Task Transparency in Alternative Computer Access: Selection of Text Through Switch-Based Scanning. Ph.D. Thesis, Dept. of Industrial Engineering, University of Toronto. Available at http://www.wivik.com/documents/wiv97d1.pdf (28.9.2001).

SIBERT, L.E. AND JACOB, R.J.K. 2000. Evaluation of Eye Gaze Interaction, in *Human Factors in Computing Systems: CHI 2000 Conference Proceedings*, Den Haag, the Netherlands, ACM Press, 281-288. Available at http://www.eecs.tufts.edu/~jacob/papers/chi00.sibert.pdf (28.9.2001).

STAMPE, D.M. AND REINGOLD, E.M. 1995. Selection by looking: A novel computer interface and its application to psychological research, in J.M. Findlay, R. Walker, and R.W. Kentridge (eds.), *Eye movement research: Mechanisms, processes and applications*, Elsevier, Amsterdam, 467-478.

WARE, C. AND MIKAELIAN, H.T. 1987. An Evaluation of an Eye Tracker as a Device for Computer Input, in *Human Factors in Computing Systems: CHI+GI'87 Conference Proceedings*, Toronto, Canada, 183-188.

VISIOBOARD 2001. A demonstration CDROM by Metrovision. More information is available at http://www.metrovision.fr/mv-vi-notice-us.htm (28.9.2001).

VISIONKEY 1995. VisionKey Communication Aid, in *Closing the Gap,* June/July, 1995.