Typing with Gaze: An Interaction Design Perspective

Björn Yttergren
Interactive Institute
c/o Umeå School of Architecture
90187, Umeå, Sweden
bjorn.yttergren@tii.se

Daniel Fallman
Interactive Institute
& Dept. of Informatics
Umeå University
90187, Umeå, Sweden
daniel.fallman@tii.se

CHI'12, May 5–10, 2012, Austin, Texas, USA.
Workshop on Designing and Evaluating Text Entry Methods

.

Abstract
In this position paper, we discuss some early in-house experiments in designing eye gaze input systems for text entry. We have focused primarily on improving the interface's feedback for dwell time and selection, using both sound and visual feedback. While carrying out these design experiments, we have become interested in the potential of factors such as Read Text Events (RTE). While these are generally seen as disruptive, we have instead chosen to regard them as a natural part of the process of typing with gaze and tried to use RTE as a design element rather than a problem. In this paper, we present some initial ideas for how to design eye gaze input around it rather than against it.

Author Keywords
Eye gaze, Eye tracking, Text entry, Interaction design

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Design, Human Factors

Introduction
From as early as the late 19th century, but certainly increasingly so during the last few decades, eye and

gaze tracking have been of interest to researchers from a variety of perspectives; from psychology and linguistics to market research and product design.

When it comes to the application of this technology in Human-Computer Interaction (HCI), we find that from the early 1980s onwards, substantial effort has been invested in ideating, prototyping, and testing various techniques of producing text by gaze [7].

A large number of novel designs have been suggested, but a typical use scenario is that a user gazes at individual letters on a virtual keyboard. A tracking system, either head mounted, built into the screen, or somehow attached in front of the user then tracks the user's gaze and transforms it to the corresponding screen coordinates. Typically, some kind of 'dwell time' mechanism is applied by the system to allow the user to select a particular letter by gazing at that letter for some amount of time. When the dwell time has passed, the letter is selected by the system and users can move on to gaze at the next letter in the word being inputted. The longer the dwell time, the less chance there is that the user makes a false selection. However, as dwell time increases, so does the theoretical speed of typing.

Our Story

Interactive Institute Umeå is an interaction design research studio focusing on design-oriented experimental interaction design [3], primarily for use in non-office type environments. In a recent project, we have designed and implemented an eye gaze and gesture-controlled prototype system for monitoring, controlling, and manipulating the status of industrial processes together with ABB Corporate Research. In carrying out this project, the need to find a simple and relatively effortless way of entering small amounts of text surfaced. When we started to try out our initial ideas and review the state of the art in this area of research, we soon found that most of the ideas came off the top of our heads on how to provide appropriate feedback (e.g. selection sounds, visual feedback) and speed up selection (e.g. adjustable dwell time, snap to target, alternative keyboard layouts) have already been implemented and tested to some degree.

When beginning to look at previous research in the area, we noted two things. First, the interest in producing text by eye gaze is typically justified from a relevance perspective by pointing at the assumed needs of "people with severe disabilities" [8, p. 15]. As practiced by the community, the lab test sessions and other kinds of assessments of typing speed, error rates, and so on seem however rarely to be carried out either with or for such users. The collected data is also very seldom analyzed with regard to specific qualities, needs, and desires of this user group. Rather, most research in this area seems to take on a quite traditional lab study perspective where 'subjects'—typically students—are treated as a crowd with certain rates of production, error, and satisfaction; not as individual users with distinct challenges and opportunities that can inspire future design. We found this tendency somewhat surprising given some recent trends in the field of HCI [3], but also in that in reviewing the state of the art in text input by gaze, Majaranta [7, p. 4] notes that even "small improvements in the interface design can lead to significantly improved user performance and satisfaction".

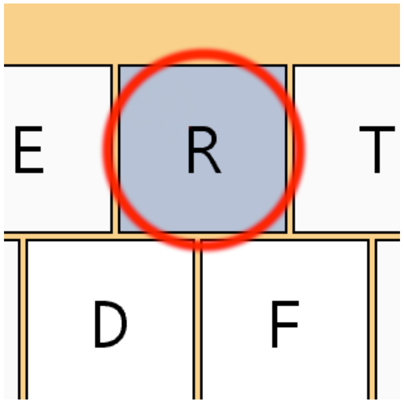

Figure 1.

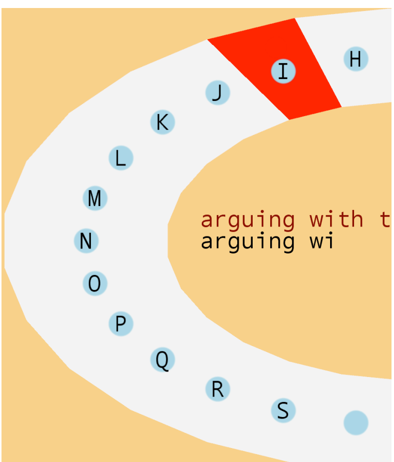

Figure 2.

Second, we were also a bit surprised that comparatively little effort seem to be invested into crafting the look and feel of the actual interfaces that are put in front of users. After all, users exposed to HCI prototypes are not testing a general idea of a keyboard layout, they are testing a particular, concrete instance of this idea (the prototype) that someone have designed and implemented in a particular way. Rather than radically changing the basic way in which text is inputted, maybe we can improve existing ideas by tweaking interface elements and feedback styles?

Our Experiments

We have been playing around with how already existing ideas in text input by eye gaze can be further enhanced. Specifically, we are interested in enhancing what is nowadays usually referred to as the 'user experience' of eye gaze text input, i.e. a combination of user satisfaction and engagement, which are different from, yet not necessarily dissonant with, more traditional usability-derived terms such as words-per-minute and error rates. A good user experience would also enhance the user acceptance, the latter of which was an important requirement from our client.

Different kinds of eye gaze-tracking systems have their individual pros and cons: some are very accurate; some allow their users to move their heads more than others do; some provide a very steady gaze point; some have longer lag times; etc. Thus, the user experience of an eye gaze-based text input system is never just a piece of software on a computer; it is the sum of a whole system where all parts need to work in harmony. In accordance with [7], we think that significant improvements in user experience can be made by relatively simple enhancements in interaction design. Below, we present some very early examples, which we have found to have promising qualities but which obviously are yet to be tested empirically:

First, we have carried out some basic experiments with various ways in which to improve and extend existing ways of giving feedback about the critical progression from *dwell time* to *selection* (Figure 1). Here, the user's eye gaze snaps to a particular letter (here the letter "R"), represented by a gray color. The dwell time, i.e. how far in time the user is from selecting R, is embodied in the red circle, which is animated and shrinking towards the R. When the circle 'touches' the R, it is selected. Feedback about that a selection has been made is given both by audio (a small click, cf. [7]) as well as visually (the button pulsates once in red). We are also experimenting with such things as adjustable dwell time based on user performance (dwell time is gradually reduced until problem occur, e.g. that the user uses the delete key) and letter occurrence in the alphabet (less common characters such as 'Q' have longer dwell times).

Second, we have also been interested in factors such as Read Text Events (RTE) and Re-Focus Events (RFE). As discussed by [7], Read Text Events (RTE) refers to the user looking away from the virtual keyboard to review the typed text. Re-Focus Events (RFE) is a measure of the average number of times a user's gaze leaves a key before it is selected. These aspects, which are typically thought of as distractions, seem to influence the user experience in a deeper way as they contribute to an overall character of the system and we are interested in how that character influences use. We believe these aspects are important to consider not least since text input by eye gaze can be very tiresome and straining to

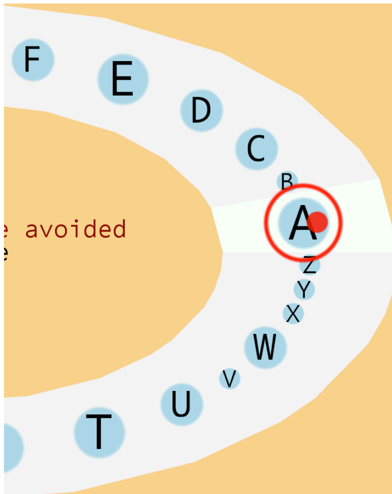

Figure 3.

the eye. If users experience a system where there is no obvious place to rest their eyes, they might be annoyed, nervous, and/or quickly become tired. This hardly improves performance, either. When it comes to RTE, according to [1], inexperienced users tend to review their texts more often than experienced users. While this on the one hand is sound and not surprising, one might question it from a methodological position: what were the users doing? Where they writing haiku poems or repeating to them meaningless phrase sets (such as those suggested by [6])? If we agree that different tasks might yield slightly different kinds of behavior, even an entirely different attitude to the whole system, is it then perhaps possible to design the system to learn and adapt to its user's behavior? (see [5]). Yet, the field retains a strong focus on WPM, which methodologically requires blindness towards what people actually do when they write and how different kinds of text input systems might best fit into those practices. In our view, disabled users may or may not have entirely different requirements and desires than lofty poets, which in turn may or may not be quite different from undergrads in computer science.

Along these lines, and as a result of our very early experiments in this area, we have become increasingly curious if RTE can be used more creatively for design. If the user's gaze is more on the written text than on the keyboard, maybe the keyboard could withdraw into the background, change form, disappear entirely, or otherwise realize it does not appear to be the center of attention right now? An example of such a keyboard that we are currently experimenting with is shown in Figure 2, which is a circular keyboard floating around the text being entered. The user's gaze then naturally passes over the text when moving in between different letters. Finally, Figure 3 depicts a slightly more advanced version of this keyboard where the gaze target sizes have been weighted (in three levels) according to the letters' occurrence in the English language [2]. The feedback mechanisms described above are also applied, and we are experimenting with various keyboard layouts (note that the 'reverse alphabetical' layout in Figure 2 and 3 is just an example).

References

[1] Bates, R. Have Patience with your eye mouse! Eye-gaze interaction with computers can work. *Proc. of CWUAAT'02*, (2002), 33-38.

[2] Beker, H. & Piper, F. *Cipher Systems: The Protection of Communications.* Wiley (1982).

[3] Fallman, D. Design-oriented Human-Computer Interaction, *Proc. of CHI 2003,* (2003), ACM Press.

[4] Fallman, D. (2011) The New Good: Exploring the Potential of Philosophy of Technology to Contribute to Human-Computer Interaction, *Proc. of CHI 2011,* (2011), ACM Press.

[5] Kristensson, P.O. Five challenges for intelligent text entry methods. *AI Magazine,* 30.4, (2009), 85-94.

[6] MacKenzie, I. S., & Soukoreff, R. W. Phrase Sets for Evaluating Text Entry Rate Development, *Proc. of CHI 2003,* (2003), ACM Press.

[7] Majaranta, P. *Text Entry by Eye Gaze.* Ph.D. Thesis, 11, University of Tampere (2009).

[8] Majaranta, P. & Räihä, K.-J. Twenty years of eye typing: Systems and design issues, *Proc. of ETRA'02,* (2002), ACM Press.