

# Auditory and Visual Feedback During Eye Typing

Päivi Majaranta, I. Scott MacKenzie, Anne Aula, and Kari-Jouko Räihä

Unit for Computer-Human Interaction (TAUCHI)  
 Department of Computer and Information Sciences  
 FIN-33014 University of Tampere, Finland  
 {curly, scott, aula, kjr}@cs.uta.fi

## ABSTRACT

We describe a study on how auditory and visual feedback affects eye typing. Results show that the feedback method influences both text entry speed and error rate. In addition, a proper feedback mode facilitates eye typing by reducing the user's need to switch her gaze between the on-screen keyboard and the typed text field.

## Keywords

Eye typing, text entry, feedback modalities, disabled users

## INTRODUCTION

For people with severe disabilities their eyes may be the only means for communication. Even though eye typing has been studied for many years, there is little research on design issues [2]. Our goal was to study how feedback could facilitate the tedious [1] eye typing task and make gaze-based computer-aided communication more practical for those who need it.

## Feedback Modes

During eye typing the user first focuses on the desired letter. To select the focused letter she continues to fixate on it thus using dwell time as an activation command. Feedback is given for focus and selection. The following four feedback modes were tested.

**Visual only.** In the Visual only mode, the key is highlighted on focus (the 2<sup>nd</sup> key on left in Figure 1) and its symbol shrinks as dwell time elapses. The shrinking draws the attention in, helping the user focus on the center of the key. On selection the letter turns red and the key goes down.

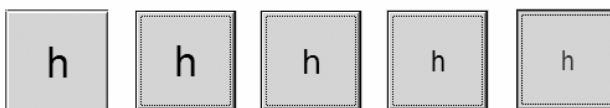


Figure 1. Animation for Visual only feedback mode

**Speech only.** The Speech only mode did not use visual feedback. The symbol on the key was spoken on selection.

**Click plus visual.** The Click plus visual mode uses two modalities; it has the same visual feedback seen in Figure 1

and, in addition, a short "click" is heard on selection.

**Speech plus visual.** The Speech plus visual mode again uses the same visual feedback plus the symbol on the key is spoken on selection.

The dwell time was the same for all modes: 400 ms before the start of the focus and 900 ms before selection. A summary of the modes follows.

Feedback mode	While focused	When selected
Visual only	shrinking letter	red letter, key down
Speech only	none	letter spoken
Click + visual	shrinking letter	red letter, key down, click
Speech + visual	shrinking letter	red letter, key down, letter spoken

## METHOD

Our study used 13 participants (5 females, 8 males, mean age 23 years). All were able-bodied with normal or corrected-to-normal vision. None had previous experience with eye tracking or eye typing but all were familiar with desktop computers.

The setup combined two computers and an eye tracking device (SensoMotoric Instruments iView X RED-III, 50 Hz sampling, 1-degree gaze position accuracy, see Figure 2).

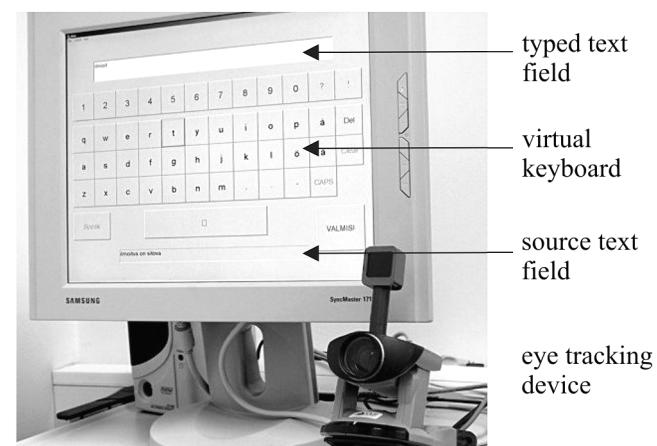


Figure 2. SMI eye tracker and experimental software

The software had an on-screen keyboard, a typed text field above, and a source text field below. The user first read the source text and then eye typed it letter by letter by gazing at a letter for the predefined dwell time. The typed text appeared in the upper field. The qwerty layout was chosen over alternatives based on pilot users' comments. For the

experiment, a special “ready” key was added. Activating this key cleared the typed text field and loaded a new source sentence.

We collected three types of data: fixation data, raw eye data, and event data logged by the experimental software. We also videotaped all trials.

The experiment was a  $4 \times 4$  repeated measures design with 4 feedback modes and 4 blocks of sentences. The order of administering the feedback modes was randomized across blocks and participants to minimize asymmetric learning effects. Each block involved the entry of the same five short phrases of text. The user was instructed to memorize the source sentence and then eye type it as fast and accurately as possible. There was a short pause after each block. Each participant came to the test four times. In the last visit we interviewed the user and gave a questionnaire. The total number of phrases was 1040 (13 participants  $\times$  4 feedback modes  $\times$  4 blocks  $\times$  5 sentences).

## RESULTS AND DISCUSSION

The grand mean for entry speed was 6.97 words per minute. This is typical for eye typing [1, 2], but it is still too low for fluent text entry. As evident in Figure 3, participants improved with practice, as a significant main effect for block was found ( $F_{3,36} = 10.92, p < .0001$ ).

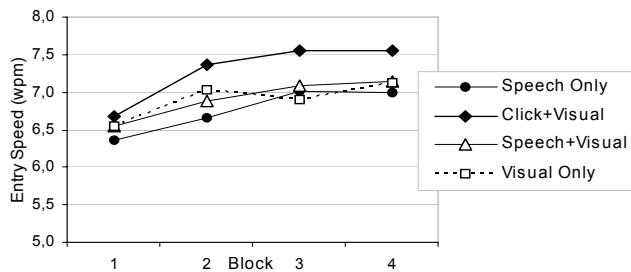


Figure 3. Entry speed (wpm) by feedback mode and block

The main effect for feedback mode was also significant ( $F_{3,36} = 8.77, p < .0005$ ). Overall, the combined use of Click plus visual feedback yielded the fastest entry rate, with participants achieving a fourth block mean of 7.55 wpm. The other fourth-block means were 7.14 wpm (Speech plus visual), 7.12 wpm (Visual only), and 7.00 wpm (Speech only). The dwell time was constant for all feedback modes. The entry rate will naturally speed up with a shorter dwell duration. This may be possible as a user develops proficiency with the apparatus.

Participants' accuracy also improved significantly with practice ( $F_{3,36} = .09, p = .005$ ), as seen in Figure 4. Character-level error rates were quite low overall with a grand mean of 0.54%. Participants proceeded quite cautiously to avoid a loss of calibration, which occurred occasionally and necessitated re-typing a phrase. There were significant main effects for feedback mode ( $F_{3,36} = 5.01, p = .005$ ). Eye typing with Speech only feedback was the most accurate technique throughout the experiment with error rates under 0.8% on all four blocks.

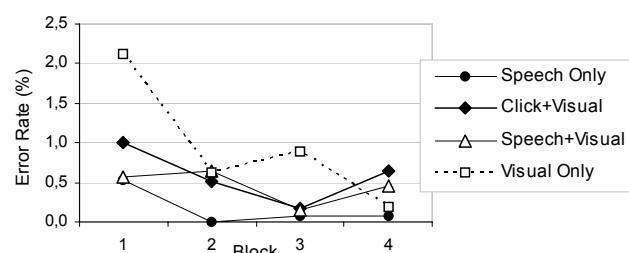


Figure 4. Error rate (%) by feedback mode and block

Our experimental software logged various events of interest. One such event was “read text”, referring to a participant switching their point of gaze to the typed text field to review the text typed so far. The values analyzed were the mean number of such events per phrase of text entered.

The overall mean was 1.63 read text events per phrase. By feedback mode, the means were 1.17 (Speech only), 1.28 (Click plus visual), 1.24 (Speed plus visual), and 2.77 (Visual only). The mean for Visual only feedback mode was significantly higher than for the other modes ( $F_{3,36} = 30.06, p < .0001$ ). The users' gaze behavior shows (Figure 5) that auditory feedback (click or spoken) significantly reduces the need to review and verify the typed text.

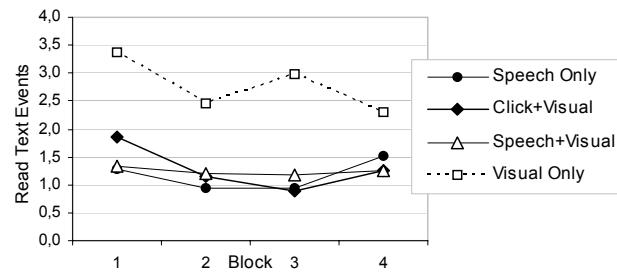


Figure 5. Read text events (mean per phrase) by feedback mode and block

## CONCLUSIONS AND FUTURE WORK

Our results show that the feedback mode affects typing speed, error rate, and the user's gaze behavior during eye typing. In particular, proper feedback may dramatically reduce the need to switch gaze between the soft keyboard and the typed text field thus reducing the entry time. The results also suggest that auditory feedback (click or spoken) is a more effective indication of selection than visual feedback alone.

The data analysis is ongoing; there are other interactions to analyze, such as gaze path, the types of errors, and the results of the questionnaire.

## REFERENCES

1. Frey, L.A, White, K.P. Jr., and Hutchinson, T.E. Eye-gaze word processing, *IEEE Transactions on Systems, Man, and Cybernetics* 20 (4), 1990, 944-950.
2. Majaranta, P., and Räihä, K.-J. Twenty years of eye typing: Systems and design issues, *Proceedings of ETRA '02*, New Orleans, LA, ACM Press, 2002, 15-22.