Fast Gaze Typing with an Adjustable Dwell Time

Päivi Majaranta, Ulla-Kaija Ahola, Oleg Špakov
Tampere Unit for Computer-Human Interaction (TAUCHI)
Department of Computer Sciences, FIN-33014 University of Tampere, Finland
{firstname.lastname@uta.fi}

ABSTRACT
Previous research shows that text entry by gaze using dwell time is slow, about 5-10 words per minute (wpm). These results are based on experiments with novices using a constant dwell time, typically between 450 and 1000 ms. We conducted a longitudinal study to find out how fast novices learn to type by gaze using an adjustable dwell time. Our results show that the text entry rate increased from 6.9 wpm in the first session to 19.9 wpm in the tenth session. Correspondingly, the dwell time decreased from an average of 876 ms to 282 ms, and the error rates decreased from 1.28% to .36%. The achieved typing speed of nearly 20 wpm is comparable with the result of 17.3 wpm achieved in an earlier, similar study with Dasher.

Author Keywords
Gaze typing, text entry, gaze input, longitudinal study.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces – Evaluation/methodology; Input devices and strategies

INTRODUCTION
*Gaze typing (*or *eye typing)* is a means of entering text by gaze. In the most typical setting, the user points at the letters on the screen by looking at them. A gaze tracking system tracks the user's gaze direction and transforms it to screen coordinates. Such gaze control is most needed by people with severe motor disabilities, who have no or little control of muscles. If the user is not able to press a manual switch or blink to select the focused letter, a command can be distinguished from casual viewing by *dwell time*, a gaze duration that is longer than the normal viewing time.

A long dwell time is good for preventing false selections but a long fixation on the same target can be tiring to the eyes. The dwell time also sets a limit for the maximum typing speed because the user has to wait for the dwell time to elapse before each selection. Majaranta and Räihä [4] provide a review of text entry by gaze. According to them, most gaze typing evaluations have been conducted with novices using a constant, fairly long dwell time (450-1000 ms). The reported typing speed has typically been fairly slow, from 5 to 10 words per minute (wpm). The text entry speed of real experts has not been measured for any of the gaze controlled systems [4].

In a more recent study, Wobbrock et al. [9] compared dwell time based gaze typing with gaze gestures in a longitudinal study (14 sessions with 8 trials in each). They used a short dwell time of 330 ms, nevertheless, their result of 7 wpm is in line with previous research. For experimental reasons, Wobbrock et al. restricted the size of the on-screen keyboard to match the fairly small window of the gaze gesture based system, which may explain the comparatively slow typing speed; small buttons are hard to hit by gaze.

Špakov and Miniotas [6] studied automatic adjustment of dwell time. Even though their results were encouraging there was some delay and involuntary variation in the automatic adjustment. The participants would have wanted to change the speed more quickly. Therefore, they suggest a trade-off on the extent of the automatic control to let the user decide when the duration is convenient.

We conducted a longitudinal study to find out how fast novices learn to gaze type when they are allowed to adjust the dwell time at will as they wish. The method and a summary of the results are reported below.

METHOD

Participants
Eleven able-bodied university students volunteered for the experiment (3 males, 8 females; from 18 to 30 years of age, with normal or corrected-to-normal vision). All were native speakers of Finnish and familiar with the QWERTY keyboard layout but novices in gaze typing. All participants were rewarded with four movie tickets. To motivate the participants in the ten-day experiment, we informed them after the first session that the participant who would learn to gaze type the best would receive an extra prize.

Apparatus
The Tobii 1750 gaze tracking device, integrated into a 17 inch TFT color monitor (with 1280 x 1024 pixels resolution), was used to track the participants' gaze.

Figure 1. Experimental keyboard. The closing circle on the letter 'i' indicates the elapsed dwell time.

The COGAIN ETU Driver, with a plug-in for Tobii, was used to implement the experimental keyboard and to save data. The stimulus phrase was shown on top of the experimental keyboard (illustrated in Figure 1). The transcribed text written by the participant appeared in the text input field below the stimulus. Letters were organized into a QWERTY-like layout, including keys for the most common punctuation. Space, Shift (for upper case letters) and Backspace were located below the letter keys. The last row included the keys for adjusting dwell time (in the middle) and a 'Ready' key (on the right).

We decided to use a speed meter as an indicator of the typing (or selection) speed instead of (a numeric) dwell time adjustment because it was considered more natural and easier to understand for the users. The gaze operated minus key decreased the speed by increasing the dwell time (max. 2000 ms) and the plus key increased the speed by decreasing the dwell time (min. 150 ms). When the speed indicator's hand was in the middle, the dwell time was 600 ms. Thus, the steps to adjust the speed became smaller as the indicator moved to the right. This enabled a rapid increase of speed with long dwell times and fine adjustment of speed with very short dwell times. The formula for the dwell time adjustment was based on pilot tests: DTadjusted = 300 * EXP(X/12) - 150, where X = {0,1,..,24}. X is the step controlled by the user using the minus and plus keys that change the X by 1. In the lower end, when the dwell time duration is long, the step is 160 ms and in the higher end (with very short dwell time) the step is only 25 ms.

An animated closing circle was shown on the key to indicate the progression of dwell time (see 'i' in Figure 1). The color of the animation was chosen so that it disturbed as little as possible but was still easy to see. When the dwell time run out (the circle closed), the key was visually pressed down and a 'click' sound was heard. The participants were told that they can ask the experimenter to remove the animated feedback if they find it disturbing.

The active selection area was bigger than the visible key in order to minimize potential problems caused by inaccuracy in calibration. Thus, the key was selected (and feedback shown on the desired key) even if the measured point of gaze was somewhat outside of the key.

Procedure and Design

Each participant was first briefed about gaze control and the motivation of the study. Before the actual test, the participants practiced gaze control briefly by playing three rounds of a simple board game (Tic Tac Toe) by gaze.

The participants were seated so that their eyes were approximately 50-60 cm from the monitor. They were instructed to sit fairly still but their movements were not restricted in any way. The gaze tracker was calibrated at the beginning of every session. Re-calibration was done if needed but we tried to do it between phrases. If the tracker had to be re-calibrated in the middle of a phrase, the phrase was ignored in the analysis.

The task was to type as many phrases as possible within the 15-minute time limit. The phrases were the

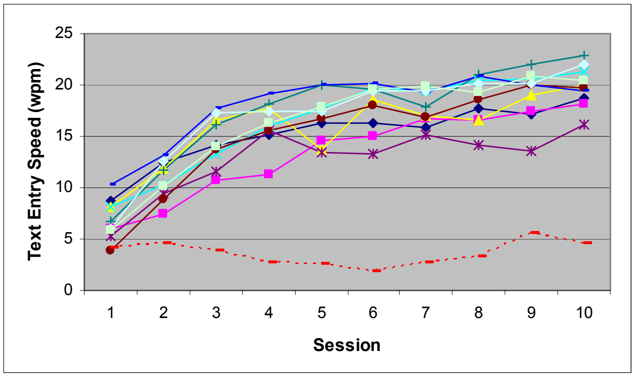

Figure 2. Text entry speed in words per minute. (The outlier is marked with a red dashed line.)

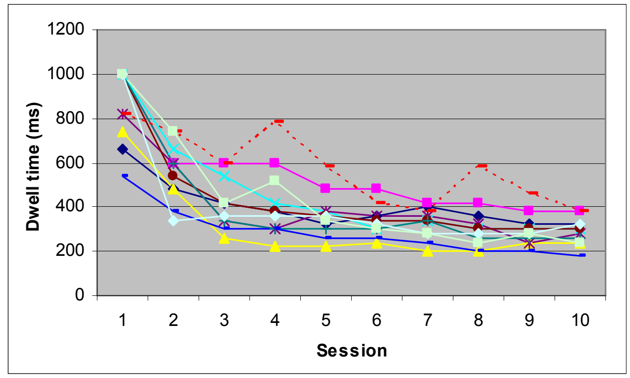

Figure 3. Dwell time duration in milliseconds.

Finnish translation [1] of the 500 phrase set originally published by MacKenzie and Soukoreff [3]. The phrases were easy to remember, neutral everyday sentences. Some of the phrases contained capital letters and punctuation; some only had lower case letters.

The phrases were shown one at a time. After finishing the phrase, the participant selected the 'Ready' key which loaded the next phrase. The software was set to stop after the 15-minute time limit had passed and the participant had finished typing the last sentence. The timer ran only during active typing, starting from the entering of the first letter and ending with the selection of the 'Ready' key.

Participants were instructed first to memorize the phrase and then to write it as quickly and accurately as possible. They were instructed to correct errors only if they detected them soon after the error occurred (within the last word).

The dwell time duration was initially set to 1000 ms for all participants. Participants were instructed to adjust the dwell time between sentences but they were able to adjust it anytime they wanted. The use of special keys and the rules of correcting mistakes were explained in every session.

Each participant visited the laboratory ten times. The sessions were organized so that there never was more than 2 days between the consecutive sessions, with a few occasional exceptions with 3 days between the sessions.

The first and the last session took about an hour with initial preparations, instructions and final interviews. Other sessions lasted about half an hour, including preparations and a short questionnaire before and after each test. In total, each participant gaze typed (10*15 min) two and half hours.

RESULTS

The results are based on 10 participants. One participant was a clear outlier that was excluded from the statistics but included in the figures (marked with a red dashed line). Due to technical problems, we lost the data for one session of one participant. The missing values were replaced with an average of the previous and the next session. Analysis of a phrase started from entering the first character and ended to the selection of the 'Ready' key.

Typing Speed and Dwell Time

The typing speed was measured in words per minute (wpm) where a word is defined as 5 characters, including space and punctuation [2]. The grand mean for the text entry speed was 6.90 wpm in the first session and 19.89 wpm in the last, tenth session. Thus, significant learning ($F_{9,81} = 93.60$, $p<.0001$) had happened (see Figure 2).

The dwell time duration was initially set to 1000 ms. The grand mean for the dwell time was 876 ms in the first session and 282 ms in the last session. The decrease in dwell time was especially rapid during the first three sessions (see Figure 3). Nobody used the minus key (to slow down typing by increasing the dwell time) in the first session but its use increased in later sessions when participants made minor adjustments both ways to find the highest manageable typing speed. In total, the plus key was selected 229 times and the minus key 96 times.

Accuracy

The error rate was measured by comparing the transcribed text (entered by the participant) with the presented text (stimulus) using the improved minimum string distance (MSD) error rate suggested by Soukoreff and MacKenzie [5]. The MSD error rate does not take into account corrected errors. The grand mean for the MSD error rate was 1.28% in the first session and .36% in the last session. Thus, the error rates decreased even though the typing speed increased. Overall, the error rate remained quite reasonable (below 5%) throughout the experiment.

Keystrokes per character (KSPC) [5] measures the average number of keystrokes used to enter each character of text. Ideally, each key press produces one character, meaning KSPC is 1. If the participant corrects mistakes during text entry, KSPC is greater than 1 (e.g., when entering `helx[backspace]lo`, the error rate is 0 but KSPC is 7 / 5 = 1.4). Thus, KSPC measures the overhead incurred in correcting mistakes. The grand mean for KSPC was 1.09 in the first session and 1.18 in the last session (excluding keystrokes for the Shift key). The increase in the KSPC suggests that people had to correct more errors when the typing speed increased, however, the increase was not statistically significant ($F_{9,81} = 1.13$, $p>.2$).

Subjective Impressions

We analyzed the subjective ratings with the nonparametric Wilcoxon Matched Pairs Signed Ranks Test. In order to measure eye fatigue, we asked the participants how tired their eyes were before each test, and again after the test using a scale from 1 to 7. The fatigue level was calculated by subtracting the first (before test) value from the latter value. There was no significant difference between the average level of the subtracted tiredness, which was .6 in the first session and .8 in the last session.

We also measured perceived speed, ease of use, and general fatigue after each session using a questionnaire with a scale from 1 to 7 (from 'very slow' to 'very fast' etc.). There was an increase in the perceived speed (from 4.2 to 5.5, $p<.005$), which is in line with the increase in the measured speed. The perceived ease of use (with average rating of 5.3) and general fatigue (aver. 3.5) remained approximately on the same level, showing no significant change over time.

Finally, we interviewed the participants after the last session. Participants felt that typing by gaze was fairly easy, easier than they had imagined, but clearly slower than using a conventional, hand operated keyboard. Participants appreciated the QWERTY layout because of its familiarity. All participants felt they had improved in gaze typing over the sessions, especially in the beginning.

All participants felt the typing speed adjustment was clear and easy to use. Participants felt they had enough feedback of the gaze controlled selection of a key. Auditory feedback was considered as either more important (by 6 participants) or equally important (3 participants) as the visual feedback. Participants also appreciated the animated feedback and wanted to keep the closing circle even with very short dwell times; five participants tried gaze typing without the circle but only two had it turned off in the end of the last session.

Half of the participants experienced problems in using the Shift key: with short dwell times, participants experienced a delay in screen refreshing when the lower case letters were changed to upper case versions. This caused disorientation and difficulty in selecting the next letter.

DISCUSSION AND CONCLUSION

All participants (including the outlier) adjusted their dwell time so that it was below 400 ms in the last session (min. 180 ms, max. 380 ms). The decrease was rapid during the first few sessions so that already in the fourth session, the average dwell time was down to 378 ms. Correspondingly, the average typing speed had increased from 6.9 wpm to 16.2 wpm in the fourth session, with a reasonable low error rate of .37%. Four 15-minute sessions equal to one hour of practice, after which the learning decelerated prominently.

Our study followed the method used by Tuisku et al. [7] in their Dasher study, with equal amount of practice and similar test procedures. Dasher [8] is considered as the world's fastest method to write by gaze [7]. The final typing speed of 19.9 wpm in our study is comparable with the results in the Dasher study with an average of 17.3 wpm in the tenth session. However, it should be noted that in the Dasher experiment, the speed curve was still growing rapidly after the ten sessions, suggesting a potentially significant increase in speed even after the two and half hours of practice. It is hard and probably unfair to compare these totally different text entry methods. However, our results do show that people can gaze type fairly fast and accurately using a simple, easy to learn on-screen keyboard, provided a fixed dwell time does not slow down the typing.

ACKNOWLEDGMENTS

We would like to thank all volunteers who participated in the experiment. We would also like to thank Kari-Jouko Räihä and Howell Istance for consultation and I. Scott MacKenzie for the use of his Java tools in analyzing the results. This work was supported by COGAIN, Network of Excellence on Communication by Gaze Interaction (IST-2003-511598, www.cogain.org) and partly supported by Academy of Finland (project 111658).

REFERENCES

1. Isokoski, P., and Linden, T. Effect of foreign language on text transcription performance: Finns writing English. *Proc. NordiCHI 2004*, ACM Press, 105-108.
2. MacKenzie, I. S. Motor behaviour models for human-computer interaction. In J. M. Carroll (Ed.), *Toward a Multidisciplinary Science of Human-Computer Interaction*. Morgan Kaufmann, 2003, 27-54.
3. MacKenzie, I. S. and Soukoreff, R. W. Phrase sets for evaluating text entry techniques. *Ext. Abstracts CHI 2003*, ACM Press, 754-755.
4. Majaranta, P. and Räihä, K.-J. Text Entry by Gaze: Utilizing Eye-Tracking. In I. S. MacKenzie & K. Tanaka-Ishii (Eds.) *Text entry systems: Mobility, accessibility, universality*. Morgan Kaufmann, 2007, 175-187.
5. Soukoreff, R. W. and MacKenzie, I. S. Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric. *Proc. CHI 2003*, ACM Press, 113-120.
6. Špakov, O. and Miniotas, D. On-line adjustment of dwell time for target selection by gaze. *Proc. NordiCHI 2004*, ACM Press, 203-206.
7. Tuisku, O., Majaranta, P., Isokoski, P., and Räihä, K-J. Now Dasher! Dash Away! Longitudinal study of fast text entry by eye gaze. *Proc. ETRA 2008*, ACM Press, 19-26.
8. Ward, D. J. and MacKay, D. J. C. Fast hands-free writing by gaze direction. *Nature 418(6900)* (2002), 838.
9. Wobbrock, J. O., Rubinstein, J., Sawyer, M. W., and Duchowski, A. T. Longitudinal evaluation of discrete consecutive gaze gestures for text entry. *Proc. ETRA 2008*, ACM Press, 11-18.