Gaze'N'Touch: Enhancing Text Selection on Mobile Devices Using Gaze

Radiah Rivu
Bundeswehr University Munich
Werner-Heisenberg-Weg 39,
81739 Munich, Germany
sheikh.rivu@unibw.de

Mariam Hassib
Bundeswehr University Munich
Werner-Heisenberg-Weg 39,
81739 Munich, Germany
mariam.hassib@gmail.com

Yasmeen Abdrabou
Bundeswehr University Munich
Werner-Heisenberg-Weg 39,
81739 Munich, Germany
yasmeen.essam@unibw.de

Florian Alt
Bundeswehr University Munich
Werner-Heisenberg-Weg 39,
81739 Munich, Germany
florian.alt@unibw.de

Ken Pfeuffer
Bundeswehr University Munich
Werner-Heisenberg-Weg 39,
81739 Munich, Germany
ken.pfeuffer@unibw.de

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
http://dx.doi.org/10.1145/3334480.3382802

Abstract
Text selection is a frequent task we do everyday to edit, modify or delete text. Selecting a word requires not only precision but also switching between selections and typing which influences both speed and error rates. We evaluate a novel concept, extending text editing with an additional modality, that is gaze. We present a user study (N=16) where we explore how the novel concepts, referred to as GazeButton, can improve text selection and compare it to touch-based selection. We also tested the effect of text size on the selection techniques by comparing two different text sizes. Results show that gaze-based selection was faster with larger text size, although not statistically significant. Qualitative feedback show a preference for gaze over touch, motivating a new direction of gaze usage in text editors.

Author Keywords
Gaze Selection, Text Editing, Interaction, Gaze and Touch

CCS Concepts
•Human-centered computing → HCI design and evaluation methods; Interaction techniques;

Introduction
Gaze as an input modality has a long history in HCI research [1, 5]. Recent advances in eye-tracking technology for low cost and small form factors may soon enable gaze

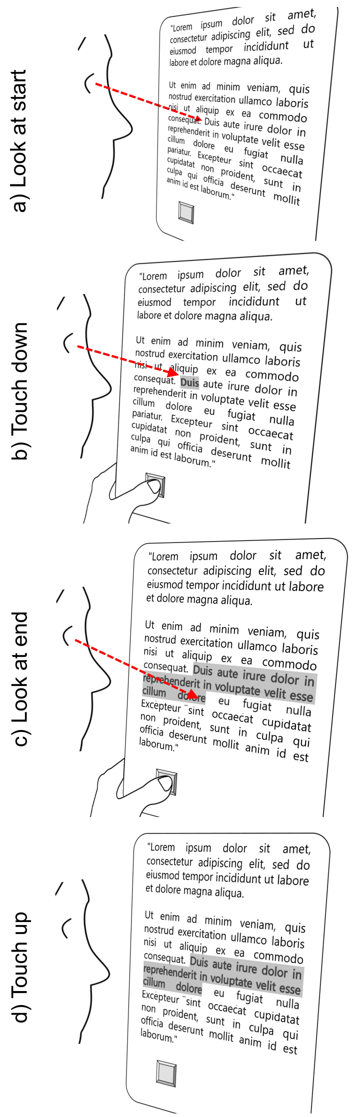

Figure 1: Illustration of the Gaze'N'Touch concept.

for mobile HCI [2]. This raises new research opportunities on how gaze can potentially advance touch interaction in mobile user interfaces. We investigate how gaze can support text selection on touch devices.

Text selection is a frequent mobile activity. When reading, we often mark a text for annotation, to look up words in the web, or to copy it to another application. Surprisingly, such a frequent and simple task is yet cumbersome to perform. This can be attributed to the fat finger problem [4]. The low resolution of fingers and the occlusion of content make it difficult to precisely select a text area. Current mobile interfaces use designated UI widgets, as anchors that users drag to the start and end position of the text. This however complicates the UI, and divides a unified task into multiple substeps. In most instances, the interaction includes at least three steps: 1) Hold your finger on text to activate the selection mode, 2) drag the 1st anchor to the start position, and 3) drag the 2nd anchor to the end position.

We investigate the *Gaze'N'Touch* interaction technique for text selection. The technique is complementary, as users can employ it in addition to the typical direct touch interactions in a simpler way. Figure 1 illustrates the interaction. The user looks at the text from the start to the end position of the text to be marked. To clarify what is the start and end, the user press a button at the start (b) and releases it when looking at the end (d).

Gaze and touch provides interaction benefits for touch surfaces [10, 18]. We believe it can be particularly useful for mobile text selection. At any time users can utilise this method to quickly select text by a glance. Gaze'N'Touch is similar to drag and drop of WIMP (Windows, Icons, Menus, Pointers) interfaces, where users can efficiently select text in a single mouse movement. Here users perform eye movement as they would naturally do for marking text.

We present a user study that compares Gaze'N'Touch to baseline touch for a text selection task. As eye-trackers are not accurate enough on mobile phones yet, we use a setup with a larger device (tablet) and easier task environment (increased text font). Mainly this compromise allows us to gain first insights into the user performance and experience of such an interface, but there are also other benefits. Such a setup furthers the understanding of gaze interaction on tablets, as well as provides insights into assistive use cases when the hands cannot be engaged as usual and for vision impaired users that use a larger font by default.

The study uses a counterbalanced design with two technique $\times$ two font size conditions with 16 users. Overall the study showed that the technique indeed has potential, as all users easily adopted using the technique. Results indicate that for larger text, users are indeed faster with Gaze'N'Touch (although not statistically significant). We also find 82% of users preferred Gaze'N'Touch selection over the baseline approach. These are promising findings and indicate how Gaze'N'Touch can improve mobile text selection.

Related Work

While touch based text selection provides many advantages, challenges such as the fat finger problem, one-handed reach, or occlusion through the use of direct finger input remain [7, 9, 16, 19, 21]. For example to allow one-handed use on the full screen with ease, indirect methods such as Bezel Swipe were introduced by Roth et al. [14]. The authors used the Bezel idea which is a swipe gesture made from the first line to the last line of a text area to be selected. However, the method would not allow the user to select a smaller text portion from a paragraph, and in addition it was not formally evaluated for user performance. Nonetheless, using alternative modalities

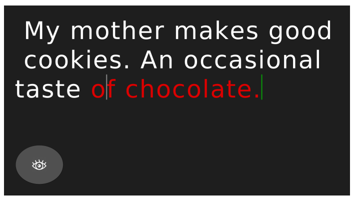

Figure 2: Example showing big text size with text to be selected in red color.

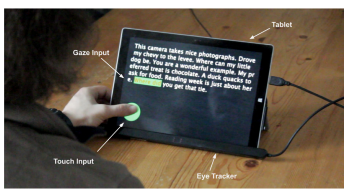

Figure 3: Experiment Setup.

rather than touch showed promising results. Pen and haptic feedback, grip detection, or pressure based touch can be used to account for out of reach targets [3, 6, 7, 12].

Gaze input has been extensively studied in HCI [5]. Gaze can be faster than manual input devices [15]. However, for a more intuitive interaction, it has been argued that gaze can be overloading the visual senses, therefore, the combination of gaze and manual input is key to gaze-enabled interfaces [22] and to avoid the Midas Touch problem [5]. Zhai et al. introduced MAGIC [22], a gaze-enhanced cursor pointing where the mouse cursor moves close to the gaze position to reduce major pointing overhead. Researchers also explored gaze with touch input, initially for interaction with remote [18] and transfer between devices [20].

Pfeuffer et al. propose gaze and touch interaction on the same interface [10], showing how gaze can provide new types of touch gestures avoiding touch occlusion and precision issues and enabling whole-display reachability. They also applied this concept to tablets [11]. They evaluated the method in a task for app selection on a homescreen interface. The results show that gaze and touch was slower, however most users preferred the method because of less physical effort. Similar work using gaze has been done by Sindhwani et al. [17]. Our work extends their work, focusing on the special but common use case of text selection. Rivu et al. [13] presents a novel concept called Gaze Button enabling users to perform selection using gaze. In this work, we evaluate this concept through a user study in a text editing application.

Evaluation

The goal of the evaluation is to gather insights into Gaze’N’Touch feasibility for which we compare it to its touch only variant.

Task

The task is to select a part of the text displayed on the screen. The target text is marked by a red font. When users select text during the task, its background is highlighted green. We created the selection paragraphs using 500 phrases from MacKenzie and Soukoreff with least possible redundancy [8]. To ensure approximately the same visual amount of text for all conditions, phrases are grouped and the text to be selected is counterbalanced. We combined two phrases with large text size conditions and eight for those with small text. The target text that has to be marked is randomly chosen. Only words have to be selected. The interface is shown in Figure 2.

Study Design

We use a within subject design with repeated measures to minimize learning effects. Touch-based selection for our study has been implemented to imitate the feature available in common devices. We therefore use a dwell time of 400 ms for anchors to appear, which can be dragged by the participant to change the selection area. For every repetition, the participant selects a coherent area of the displayed text that is selected randomly with red font colour as an indication. After each selection is marked correctly, the next repetition starts automatically. To investigate the influence of font size on both types of selection tasks, we used two sizes ("large text size" and "small text size"). Selection techniques and font sizes were the independent variables in our experiment. The study conditions are:

- Technique: Gaze’N’Touch, touch anchors
- Text size: small, large

Apparatus

The application software was implemented using Java with Processing on a Surface Pro 3 tablet (i3 core with 1,5 GHz and a 4 GB RAM). To record gaze data, a Tobii 4C eye

a) b) c) d)

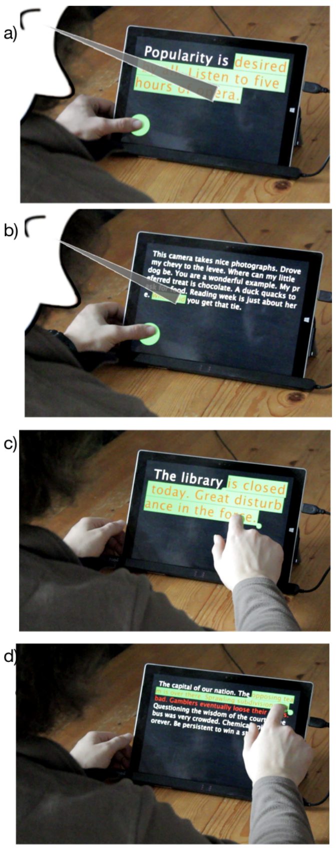

Figure 4: a) Gaze selection: Large text. b) Gaze selection: Small text. c) Touch selection: Large text. d) Touch selection: small text.

tracker (90 Hz) has been used, placed at the bottom of the tablet's screen. We use a moving average over the last four gaze samples, to ease gaze selections in addition to the smoothing that comes with the Tobii's software by default. Text size has been set as one-tenth of the screen size. The experimental setup is shown in Figure 3.

Participants
16 people (8 male) aged 17 to 52 years participated in the user study. 3 participants wore glasses and 3 participants used contact lenses and 14 participants were right-handed.13 participants stated to have experience with touch interaction(at least 4 on a scale of 1=no experience at all through 5=strong experience) and 13 participant stated to have no or very little experience with eye gaze interaction (2 or less on a scale of 1 through 5).

Study Procedure
Initially, each participant was explained the purpose of the study, then they completed the demographics questionnaire. After participants sat in front of the tablet in a relaxed position, the eye tracker was calibrated with respect to their eyes. After calibration, they first had 5 text selections with instructions to learn the task. Then, users conducted 50 text selections per condition with voluntary breaks after every 10 repetitions where participants were informed about their progress. At last, users filled out a questionnaire for evaluation of the tasks after the study has been completed.

An example picture for each task is shown in Figure 4. After the participant finished all four tasks they filled out a final questionnaire in which they ordered the four tasks by preference and provided justification for their decision.

Results

Task Completion Time
We conducted a repeated measures ANOVA to examine how font size and interaction technique influence the average time needed per repetition/text selection in the study. Figure 6 shows the task completion time results.

For interaction techniques there were no statistically significant differences at $F(1,13) = .645$ and $p = .436$. For font size significant differences have been found at $F(1,13) = 24.520$ and $p < .001$. As expected with larger font size participants were significantly faster. Between the effects of interaction technique and font size there was a statistically significant interaction at $F(1,13) = 10.068$ and $p = .007$. Because of this significant interaction pairwise post-hoc comparisons with Bonferroni corrections have been conducted and led to the following results:

1. In the selection tasks with small font size, participants were significantly faster with touch selection than with gaze selection ($p = .045$).
2. In the selection tasks with large font size, there was no significant difference between the interaction techniques ($p = .087$).
3. In the gaze selection tasks, participants were significantly faster with the large text font ($p = .001$).
4. In the touch selection tasks, participants were significantly faster with the small text font ($p = .038$), whereby this effect was less pronounced than in gaze selection tasks.

Questionnaires and Feedback
Participants filled out a performance questionnaire for each of the four tasks in which they rated learnability, ease, physical effort, eye tiredness, precision and speed they experienced (on a scale of 1= very low to 5=very high). The results are shown in Figure 5.

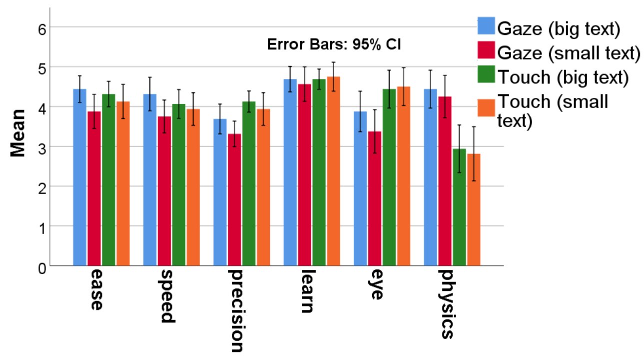

Figure 5: Likert scale results

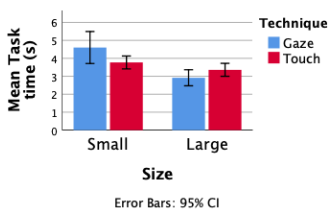

Figure 6: Mean task completion times

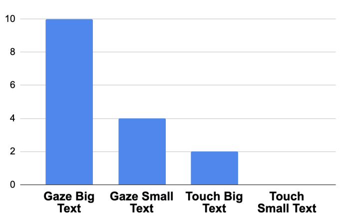

Figure 7: Frequency of most preferred conditions

Figure 7 shows the number of times every condition has been ranked between 1 (best) and 4 (worst) by participants. While the gaze selection task with large text is favoured over all other tasks by 10/16 participants, the gaze selection task with small text has been placed on every of the four ranks exactly the same amount of times. Furthermore, the gaze selection task with large text is the only one that has never been placed on rank 4 and the touch selection task with small text is the only one that has never been placed on rank 1. Overall, the ranking shows that participants prefer gaze selection with large font size.

A non-parametric Friedman test for the Likert questions showed statistical significance for precision $\chi^2(3) = 12.03$, $p = 0.007$ between Gaze Large text (Mean = 3.69 ; SD = 0.07), Gaze Small text (mean = 3.31, SD = 0.06), Touch Large text (mean = 4.12 , SD = 0.50) and Touch Small text (mean = 3.93 , SD = 0.77). In addition, it also showed statistical significant in the eye tiredness responses with $\chi^2(3) = 18.52$, $p < 0.005$ and physical effort $\chi^2(3) = 13.07$, $p < 0.005$. No other statistically significant differences were found for ease of use, speed and learnability.

Discussion

The user study has showed the potential to use gaze to make text selection easier, but yet with reservations about eye tracking accuracy. In the large text condition, gaze showed a 13% higher performance although not statistically significant. Users prefer gaze selection over touch selection with large text size. One likely possibility of this preference is due to the physical ease and comfort obtained using gaze selection with large text.

Nonetheless, in the small font condition, users were faster with the touch technique. This can be accounted to the prevalent gaze inaccuracy that we believe can be improved in future devices. Although most users preferred gaze, users reported the need to improve gaze estimation precision. Notably, participants had little training compared to touch input. Hence, gaze selection performance may improve in terms of speed once users are used to the technique.

The qualitative measures allow to better understand the reasons for user preference and performance. In particular for perceived effort, we find there is an interesting trade-off between manual and eye interaction effort. With gaze, eye tiredness was perceived as significantly higher, indicating that aiming with the eyes may have a toll. On the other hand, with touch the physical effort was reported as higher. Most users preferred gaze interaction, which indicates that physical effort may be more important to the user in the evaluated task.

Conclusion

Prior research attempted to extend conventional touch interactions with progressive gaze-based interactions. We analysed one such concept through a user study to discern acceptability and performance. To do so, we developed a text editing app and compared the new gaze-based text selection to the conventional touch selection technique, showing promising results. In future, we aim to evaluate how users perform under different task conditions during movement. Furthermore, we plan to explore how the new techniques might be adapted to smart phones where the screen is very small and text interaction is especially cumbersome. Our contribution is valuable for researchers exploring gaze-based interaction techniques on mobile devices.

Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant agreement no. 316457582 and 425869382. It was also supported by the Studienstiftung des deutschen Volkes ("German Academic Scholarship Foundation").

REFERENCES

[1] Richard A. Bolt. 1981. Gaze-Orchestrated Dynamic Windows. *SIGGRAPH Comput. Graph.* 15, 3 (Aug. 1981), 109–119. DOI: http://dx.doi.org/10.1145/965161.806796

[2] A. Bulling and H. Gellersen. 2010. Toward Mobile Eye-Based Human-Computer Interaction. *IEEE Pervasive Computing* 9, 4 (October 2010), 8–12. DOI: http://dx.doi.org/10.1109/MPRV.2010.86

[3] Christian Corsten, Marcel Lahaye, Jan Borchers, and Simon Voelker. 2019. ForceRay: Extending Thumb Reach via Force Input Stabilizes Device Grip for Mobile Touch Input. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 212, 12 pages. DOI: http://dx.doi.org/10.1145/3290605.3300442

[4] Christian Holz and Patrick Baudisch. 2010. The Generalized Perceived Input Point Model and How to Double Touch Accuracy by Extracting Fingerprints. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 581–590. DOI: http://dx.doi.org/10.1145/1753326.1753413

[5] Robert J. K. Jacob. 1990. What You Look at is What You Get: Eye Movement-based Interaction Techniques. In *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*. ACM, New York, USA, 11–18. DOI: http://dx.doi.org/10.1145/97243.97246

[6] Topi Kaaresoja, Lorna M Brown, and Jukka Linjama. 2006. Snap-Crackle-Pop: Tactile feedback for mobile touch screens. In *Proceedings of Eurohaptics*, Vol. 2006. Citeseer, 565–566.

[7] Markus Löchtefeld, Phillip Schardt, Antonio Krüger, and Sebastian Boring. 2015. Detecting Users Handedness for Ergonomic Adaptation of Mobile User Interfaces. In *MUM '15*. ACM, 245–249. DOI: http://dx.doi.org/10.1145/2836041.2836066

[8] I Scott MacKenzie and R William Soukoreff. 2003. Phrase sets for evaluating text entry techniques. In *CHI'03 extended abstracts on Human factors in computing systems*. ACM, 754–755.

[9] Dan Odell and Vasudha Chandrasekaran. 2012. Enabling comfortable thumb interaction in tablet computers: A windows 8 case study. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 56. SAGE Publications, 1907–1911.

[10] Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-touch: Combining Gaze

with Multi-touch for Interaction on the Same Surface. In *Proc. 27th Annual Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, USA, 509–518. DOI: http://dx.doi.org/10.1145/2642918.2647397

[11] Ken Pfeuffer and Hans Gellersen. 2016. Gaze and Touch Interaction on Tablets. In *Proc. 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, USA, 301–311. DOI: http://dx.doi.org/10.1145/2984511.2984514

[12] Ivan Poupyrev, Makoto Okabe, and Shigeaki Maruyama. 2004. Haptic Feedback for Pen Computing: Directions and Strategies. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems (CHI EA '04)*. ACM, New York, NY, USA, 1309–1312. DOI:http://dx.doi.org/10.1145/985921.986051

[13] Sheikh Rivu, Yasmeen Abdrabou, Thomas Mayer, Ken Pfeuffer, and Florian Alt. 2019. GazeButton: Enhancing Buttons with Eye Gaze Interactions. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications (ETRA '19)*. ACM, New York, NY, USA, Article 73, 7 pages. DOI: http://dx.doi.org/10.1145/3317956.3318154

[14] Volker Roth and Thea Turner. 2009. Bezel Swipe: Conflict-free Scrolling and Multiple Selection on Mobile Touch Screen Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 1523–1526. DOI: http://dx.doi.org/10.1145/1518701.1518933

[15] Linda E. Sibert and Robert J. K. Jacob. 2000. Evaluation of Eye Gaze Interaction. In *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*. ACM, New York, USA, 281–288. DOI: http://dx.doi.org/10.1145/332040.332445

[16] Katie A. Siek, Yvonne Rogers, and Kay H. Connelly. 2005. Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs. In *Proceedings of the 2005 IFIP TC13 International Conference on Human-Computer Interaction (INTERACT'05)*. Springer-Verlag, Berlin, Heidelberg, 267–280. DOI: http://dx.doi.org/10.1007/11555261_24

[17] Shyamli Sindhwani, Christof Lutteroth, and Gerald Weber. 2019. ReType: Quick Text Editing with Keyboard and Gaze. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI âĂŹ19)*. Association for Computing Machinery, New York, NY, USA, Article Paper 203, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300433

[18] Sophie Stellmach and Raimund Dachselt. 2012. Look & Touch: Gaze-supported Target Acquisition. In *CHI*. ACM, 2981–2990.

[19] Matthieu B Trudeau, Paul J Catalano, Devin L Jindrich, and Jack T Dennerlein. 2013. Tablet keyboard configuration affects performance, discomfort and task difficulty for thumb typing in a two-handed grip. *PloS one* 8, 6 (2013), e67525.

[20] Jayson Turner, Andreas Bulling, Jason Alexander, and Hans Gellersen. 2014. Cross-device Gaze-supported Point-to-point Content Transfer. In *ETRA*. ACM, 19–26.

[21] Daniel Vogel and Ravin Balakrishnan. 2010. Occlusion-aware Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 263–272. DOI: http://dx.doi.org/10.1145/1753326.1753365

[22] Shumin Zhai, Carlos Morimoto, and Steven Ihde. 1999. Manual and Gaze Input Cascaded (MAGIC) Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*. ACM, New York, NY, USA, 246–253. DOI: http://dx.doi.org/10.1145/302979.303053