

A Comparative Study on Single-handed Keyboards on Large-screen Mobile Devices

Kunpeng Zhang

Department of Computer Science
University of Houston
Houston, Texas, USA
kzhang21@uh.edu

Zhigang Deng

Department of Computer Science
University of Houston
Houston, Texas, USA
zdeng4@central.uh.edu

ABSTRACT

Many questions regarding single-hand text entry on modern smartphones (in particular, large-screen smartphones) remain under-explored, such as, (i) will the existing prevailing single-handed keyboards fit for large-screen smartphone users? and (ii) will individual-customization improve single-handed keyboard performance? In this paper we study single-handed typing behaviors on several representative keyboards on large-screen mobile devices. We found that, (i) the user-adaptable-shape curved keyboard performs best among all the studied keyboards; (ii) users' familiarity with the Qwerty layout plays a significant role at the beginning, but after several sessions of training, the user-adaptable curved keyboard can have the best learning curve and performs best; (iii) generally the statistical decoding algorithms via spatial and language models can well handle the input noise from single-handed typing.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**; **Empirical studies in HCI**.

KEYWORDS

Mobile computing, keyboard interfaces, single-handed typing, comparative study, individual customization

ACM Reference Format:

Kunpeng Zhang and Zhigang Deng. 2022. A Comparative Study on Single-handed Keyboards on Large-screen Mobile Devices. In *AVI '22: Proceeding of International Conference on Advanced Visual Interfaces 2022, June 6-10, 2022, Frascati, Rome, Italy*. ACM, New York, NY, USA, 9 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 INTRODUCTION

Text entry is one of the core activities users often perform on mobile devices. It is one of the primary communication methods and the cornerstone for many other activities such as to reply texts, compose emails, and search for information. One-handed text entry is inevitable in some cases, such as, i) for disabilities of the hand,

Permission to make digital or hard copies of all or part 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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

AVI'22, June 06–10, 2022, Frascati, Rome, Italy

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00

<https://doi.org/XXXXXXX.XXXXXXX>

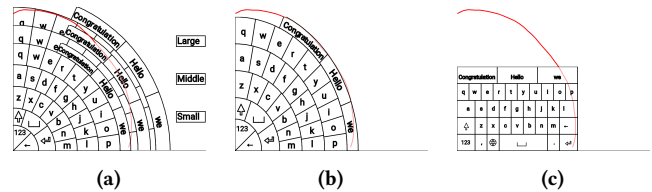


Figure 1: The prevailing single handed keyboard (a) compared to two user-adaptable single-handed keyboards (b)(c), the black lines in the keyboards' bottom are the screen width of the smartphone. The red curves are the touchable region drawn by the left hand. (a) the quarter-Circle Keyboard with three sizes (CK). (b) the user-Adaptable-shaped Curved keyboard (AC). (c) the user-Adaptable-sized Qwerty keyboard (AQ).

it's a must; and ii) single-handed text entry has distinct advantages to free the other hand for other urgent or parallel tasks.

Despite its importance, current smartphones, in particular, those with screen size bigger than 6.0", are innately unfriendly for single-handed text entry. The screens could be too big to be fully touchable for many users. Without considering the human hand structure, other well-known quarter-circle keyboards [24, 28], which specialized in single-handed use, also have some out-of-reach regions. Therefore, many questions regarding single-hand text entry on modern smartphones (in particular, large-screen smartphones) remain under-explored, such as, (i) will the existing prevailing single-handed keyboards fit for large-screen smartphone users? and (ii) will individual-customization improve single-handed keyboards performance?

Inspired by the above questions, according to the Fitts' movement theory[13], we hypothesize the user-Adaptable-shape Curved keyboard could achieve a good performance among various single-handed keyboards on large-screen smartphones. As such, in this work, we implemented and compared three single-handed keyboards for large-screen cellphones (refer to Figure 1): i) a user-Adaptable-shape Curved keyboard (AC), ii) a quarter-Circle Keyboard (CK), and iii) a user-Adaptable-size Qwerty keyboard (AQ) (The width and height ratio of the AQ is the same as the Google Gboard). Based on the above keyboards, we systematically studied and compared user behaviour of single-handed text entry through three inter-related user studies: (i) The first study is to acquire spatial models for different keyboard layouts; it is a preparation for implementing hint functions; (ii) the second study is to compare the performances and usability of the three single-handed keyboards;

and (iii) the third longitudinal study is to evaluate the comparative performance of different single-handed keyboards in three days.

We obtain the following main findings from our studies. First, we found that the AC performs best among all the studied single-handed keyboards (AQ, AC, and CK). At the beginning, users' familiarity of the Qwerty layout plays a significant role, thus AQ outperformed others. However, after several sessions of training, the AC achieved the best learning curve and performance among the three. Second, generally, statistical decoding algorithms via spatial and language models can robustly handle the input noise from single-handed typing, despite the typing noise from single-handed typing is clearly larger than that from two-hands typing. Third, user-adaptivity (i.e., individual customization) is a crucial feature for single-handed keyboards on mobile devices, in particular, large-screen smartphones. Our user studies also reveal that user-adaptable single-handed keyboards are significantly better received than others in terms of both objective measures and subjective user experience.

2 RELATED WORK

In this section we review previous related research efforts on single-handed text entry, and spatial and language models for statistical keyboard decoding, which inspires this research.

2.1 Fitts' Movement Theory

The well-known Fitts' movement theory [13] has been widely used in human-computer interface design. Its core idea is to design elements in an interface as big as possible, and to design elements as close to each other as possible. Many studies proved the theory's validity. Also, researchers have shown it can fit for the design of smartphones' keyboards [4].

2.2 Single-handed Text Entry

A substantial amount of previous research efforts have been done on single-handed text entry in both academia and industry [18, 19]. For example, Microsoft designed a windows phone app to support both relocating keys and modifying the shape to form a fan-shaped keyboard [22]. Similar keyboard layouts for other operating systems have also been developed [28]. Meanwhile, modifying the shape of the standard Qwerty keyboard was also proposed, including relocating all the keys for single-hand use [8]. Kwon et al. [15] pointed out that it is too difficult to use a shrunken Qwerty keyboard when shrunken to a certain degree because of the occlusion of the keys during typing. But they also mentioned that the occlusion problem could be partially remedied with the aid of regional error correction. We follow this theory and perform the study #1 in this work.

Other studies attempted to make full use of various existing sensors on smartphones [9, 10, 25, 31]. For example, during typing, the smartphone screen is tilted to shift the position of the keyboard such that all the keys could be within the reach of the thumb. But the tilting method needs certain training and coordination; otherwise, the typing speed could be slower than the conventional sliding method [6].

With the increase of the screen size, using smartphones by a single hand becomes increasingly difficult. Many previous works

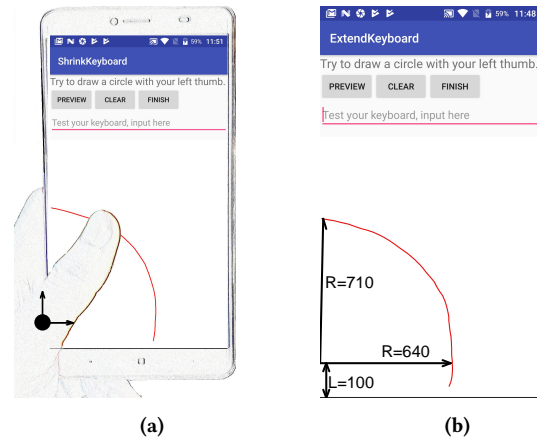


Figure 2: Calculated parameters of the quarter ellipse which is drawn by the left hand. From left to right: (a) the pivot of the thumb. (b) Measures of the variable radii of the thumb.

have been done to find a way to touch the out-of-reach region. When it refers to the touchscreen, some methods provide users an imaginary zoom in or zoom out glass, the users can touch the desired positions within the glass [17], which is more applicable from tiny to large screens, since when the users shrink a large screen to a touchable size, the keys could be too small to touch accurately. For other components of smartphones, such as Back-of-Device (BoD) touch panels or other sensors, their reachability and usability have been studied previously [16, 23]. In addition, Rateau et al. [26] proposed an innovative single-handed menu selection method called "EdgeMark Menu", by exploiting the curved display edges of smartphones.

2.3 Spatial and Language Models

One of the well-known keyboard decoding models is the spatial model with visible conditions [5, 34]. Specifically, its decoding is based on two models: a spatial model (SM) that relates a touch point to keys, and a language model (LM) that gives the prior probability distributions of words given a typing history. The touch points for tapping a key follow a bivariate Gaussian distribution. Therefore, the spatial model is parameterized with the mean and the deviation of each key related to touch points [2]. Since previous related research on single-handed typing is not available, we can only use their methodology as the reference. Other works attempted different ways to improve the keyboard efficiency through algorithm refinement [24] or the adjustment of key locations [11], typically for dual hand use.

Along with the spatial model, the language model is also essential to a lexicon containing most frequently-used words for input correction. Usually, if a lexicon contains more words, it would lead to a higher hit rate when searching for a word. But this relation is not linear. To balance the search time and the hit rate, many researchers used a lexicon with about 70K words [5, 34].

3 DESIGN AND IMPLEMENTATION OF SINGLE-HANDED KEYBOARDS

The prevailing single-handed keyboards include curved keyboards [18, 19, 22] and the qwerty keyboard. To do comparative studies, we modified both the curved keyboard and qwerty keyboard to their user-adaptable versions. The user needs to draw a region when holding the phone. Our designs analyze the features of the thumb's touchable region and then model it using discrete parameters. We further use the parameters to draw keyboards that can fit all keys into the user-touchable region.

Although previous studies predicted a quadratic formula for the thumb's touchable region [3], the smartphones used in their studies had regular-size screens [7, 32, 33] (e.g., 5.5" maximum), we found the grasp method of regular-size screen smartphones is different from that of large-screen ones. Since the verification of others' trajectories is not the focus of this work, we propose a feasible trajectory which also can fit all user-touchable regions. The model employed in our studies contains one-fourth ellipse and the elevation distance when the user's thumb joint is away from the left-bottom corner of the screen. Figure 2 illustrates the parameters of the used model.

Discussion: *Why do we choose left-handed typing in our experiments?* The text entry on smartphones generally involves both hands nowadays; both hands are equally important. According to the existing research [27], the text entry performance of the left-handed only is similar to that of the right-handed only. Therefore, in principle our user-adaptable design of the keyboards fit for both left-handed and right-handed users equally or at least without noticeable differences. In addition, in this work we choose the experiments of our keyboards with the consideration of the following use case: Assuming the user is a right-hand user and thus the left hand is mostly a spare hand when he/she is performing a sophisticated task. Also, assume his/her right hand is busy with something (e.g., write something, or do something while walking), then the user may need to use his/her spare left hand on the smartphone to perform some simple tasks, such as quickly reply SMS or search for certain information via keywords on the Internet. Therefore, without loss of generality, the single-handed keyboards used in our experiments are left-handed.

3.1 Keyboards Design

The **CK keyboard** is a prevailing keyboard with three different sizes for selection [22, 28]. Figure 1(a) shows the CK keyboards with different sizes, overlapped with each other. The user can choose the most comfortable one to use. From this figure, we can see that the largest quarter circle and the middle one often cannot be fully touched by the user, while the smallest one wastes a portion of the top region.

The **AC keyboard** treats the top radius plus the elevation from the thumb joint to the screen bottom as the ellipse's vertical radius, and assumes its horizontal radius is the same as its right radius. All the keys are located within the quarter ellipse of the keyboard, by deforming the keys to multiple concentric sectors. The sizes of the keys from a to z are nearly the same, while other control keys including space, enter, etc., are larger than alphabet keys (refer to Figure 1(b)). The main reason for the above design is that the

Qwerty keyboard uses the same mechanism, and thus we keep it in a similar way. Although the keys in the bottom-left corner are enlarged, they are still relatively small to comfortably touch. The values of the horizontal radius and the vertical radius in the AC keyboard may be different. Therefore, its touchable region may not be a perfectly quarter circle. Each key (except the bottom-left two) in the AC keyboard is enclosed by two elliptic sections and two straight lines. In each row, all the keys in elliptic sections have equal angles. But due to the difference between the horizontal radius and the vertical radius, the keys in different positions can have different sizes. As a result, the touchable region is fully utilized.

The **AQ keyboard** measures the quarter ellipse, which is drawn by the user's thumb, retrieves the right radius of the circle as the right limit (refer to Figure 2(b)). It shrinks the Qwerty layout with an appropriate ratio so that its rightmost keys are just within the reach of the thumb. All the labels on this keyboard are shrunken with the same ratio, so the AQ keyboard with different sizes still have similar visual appearances (refer to Figure 1(c)). The Google Gboard supports the single-handed typing mode. The width/height ratio of the AQ keyboard is the same of that of the Gboard.

Discussion: *Is it necessary to study smartphones with screen size 6.5" or larger?* One recent trend of smartphone screen size is increasingly bigger. Existing studies on single-handed typing were mainly based on smartphones with the screen size ranging from 4" to 6" [6, 11, 28, 34], arguably it is because large-screen smartphones did not exist or were not popular when these studies were conducted. But since 2018, all major smartphone vendors have launched their flagship products with screen size around 6.5" or larger, for example, the screen size of Apple iPhone XS MAX is 6.5", and the screen size of Samsung S10 Plus is 6.4". Due to the significant difference of screen sizes, single-handed typing methods and their expectations would be substantially different accordingly. There is a clear yet unresolved need to quantify and understand the single-handed text entry on large-screen smartphones.

4 STUDIES

We conducted a total of three inter-connected studies in this work. The first study is to acquire the spatial models of the single-handed keyboards, which is essential for the implementation of hint functions. The second study is the comparative study of the above mentioned keyboards. The third study is a three-days longitudinal study.

Apparatus. The experimental smartphone is an iRULU GeoKing 3 Max (3GB RAM, 32GB ROM, and Octa-Core CPU and 6.5" touchscreen), running Android 7.0 system. The resolution of its screen is 1920 by 1080; its performance is sufficient to fulfil this experiment's need.

Design. All the studies use a phrase transcription task, with a within-subject design. The independent variable is different keyboards. Before the experiment, participants are required to do a practice session to get familiar with the keyboards. Specifically, in the practice session, for each keyboard, each participant needs to input 15 practice phrases. The practice phrases are the same for all the participants.

The actual user study followed the practice session; each participant needs to input a certain amount of test phrases. The test

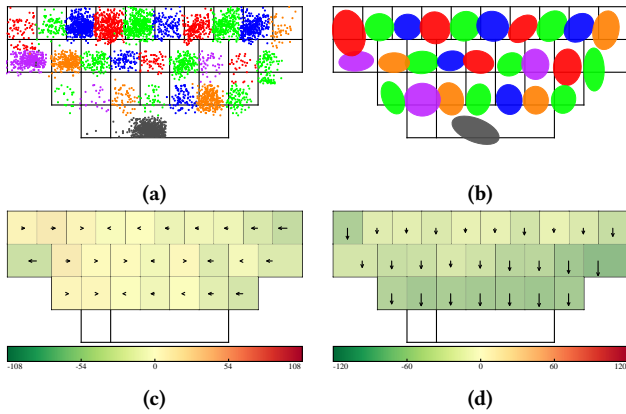


Figure 3: Data extracted from the AQ keyboard. (a) Touch points. (b) 95% confidence ellipses. (c) Horizontal offsets. (d) Vertical offsets.

phrases are in groups of 5. The keyboard is changed when a group of test phrases was finished. To prevent the participants from memorizing the test phrases, each keyboard takes the turn to be the first keyboard for a group. Both the practice and test phrases are randomly selected from the MacKenzie’s test set [20]; we also ensure the test phrases cover all the alphabets.

4.1 Study #1: Acquiring Spatial Models for Single-handed Keyboards

The purpose of this experiment is to learn spatial models for the AC and AQ keyboards, while ignoring the performance statistics of the two keyboards, including the error rate and the input rate.

Participants. 10 participants did this experiment (4 females, 6 males; ages from 25 to 50). All the participants have at least 5 years experience of typing on the touchscreen keyboard on mobile devices. Normally, all of them use both hands as the input method. One of the participants is left-handed. A total of 40 different phrases are used.

Discussion: *Why the first study is only about AC and AQ, not including CK?* Since the CK is a special case of AC keyboard where both top radius and right radius are same, the elevation is 0, referring to Figure 2(b). AC is chosen to represent the curved keyboard in this user study.

Experiment Setup. The participants are instructed to type on the experimental mobile device using the left hand at their normal speeds, given some reference phrases. In this process, assuming the keyboards would automatically correct the typos they could make, they do not need to go back to fix typos if any. Whenever the number of the inputted characters reaches the length of the reference phrase, the participants can press the Enter key to continue the typing of the next phrase. At the beginning of the experiment, all the participants need to first draw a quarter circle as shown in Figure 2(a), and then two user-adaptable keyboards (i.e., the AC and AQ keyboards) will be created for them to try and confirm. They can redo the keyboard customization process if they want.

Results. 122,356 valid touch points were collected in the first study. Figure 3(a) and Figure 4(a) show the touch points of the AQ and AC keyboards, respectively.

Variances. Figure 3(b) and Figure 4(b) show the 95% confidence ellipses of the bivariate Gaussian distributions of the touch points for the two keyboards. The bigger ellipse the bigger the variance. From the figures we can see the sizes of the ellipses from the AQ keyboard are bigger than those from the AC keyboard. This could be explained as follows: small keys on the AQ keyboard are difficult to accurately touch, which leads to greater variances.

Horizontal offsets. Horizontal offset is defined as $x_p - x_e$, where x_p is the x coordinate of a touch point and x_e is the x coordinate of the geometric center of the expected key. The positive/negative horizontal offset indicates the actual touch point located to the right/left of the expected key center.

Figure 3(c) and Figure 4(c) show the mean values of the horizontal offsets of the keys’ touch points. The mean horizontal offset of AC is -11.48 pixels (SD=78.06) while that of the AQ is -4.81 pixels (SD=88.00). There was a main effect of keyboard type on horizontal offset ($F_{1,50} = 6.95, p < 0.01$). For single left-handed input (left-handed keyboard), it is easier to touch the left half keyboard than the right half. We can observe the trends from Figure 4(c): the left half key touches have a trend towards right, because it is slightly harder to touch the keys near the palm, while the right half has an obvious trend towards left, because fingers have a reachable limitation. Especially for the rightmost keys, their leftward trend is greater than other positions.

Vertical offsets. Vertical offset is defined as $y_p - y_e$, where y_p is the y coordinate of a touch point and y_e is the y coordinate of the geometric center of the expected key. The positive/negative vertical offset indicates the actual touch point located to the top/bottom of the expected key center.

Figure 3(d) and Figure 4(d) show the vertical offsets of all the keys’ touch points. The mean vertical offset of AC is -11.48 pixels (SD=26.78) while the AQ is -25.68 pixels (SD=134.07). There was a main effect of keyboard type on vertical offset ($F_{1,50} = 32.49, p < 0.0001$). We can clearly observe the downward trend across both keyboards, especially the AQ. This finding is also consistent with previously reported finding [2, 12, 14, 34], where a similar trend is observed. In other words, there is a touch points bias towards the hand.

For the AQ layout from Figure 3(d), we can see it bias towards the hand more when the keys are close to the hand. The decisive factor of this bias is not the distance, but the occlusion. When using the small keyboard (i.e., AQ), the finger is prone to keep the whole layout visible. Therefore, when pressing the keys in the bottom row, it is downward. But when pressing the keys in the top row, it must occlude several keys, in order to make touches accurately, the finger will try to touch the middle of the keys.

Subjective Feedback. We interviewed the participants when they finished the first study. Most of them (9 out of 10) thought the AC keyboard was better than the AQ keyboard. They felt that the bigger keys of the AC keyboard greatly boosted their typing. The common concern they had is, because of the deformed layout, they need to spend slightly longer time to find the locations of some keys. So, they generally liked the practice section at the start of the

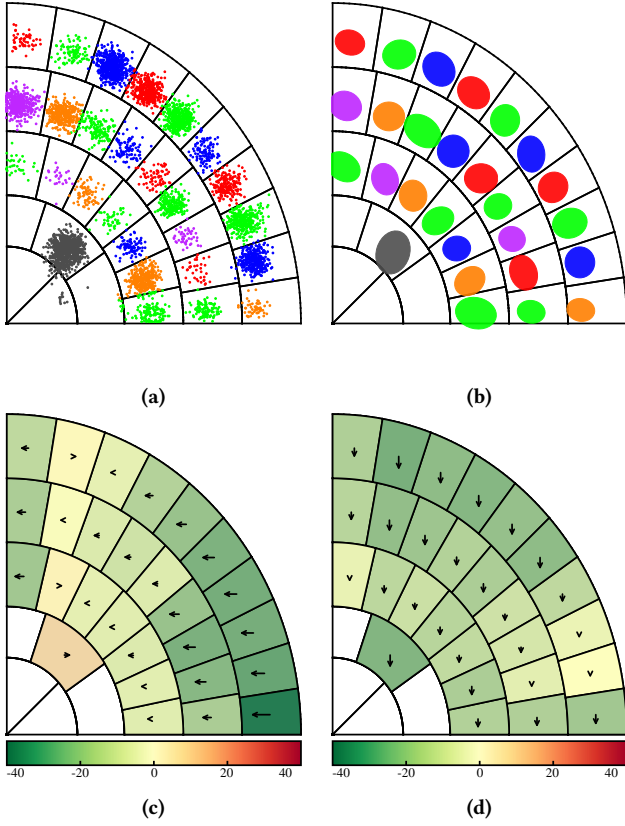


Figure 4: Data extracted from the user-adaptable Curved (AC) keyboard. (a) Touch points. (b) 95% confidence ellipses. (c) Horizontal offsets. (d) Vertical offsets

study. They also mentioned that the practice section was not long enough to get familiar with the AC keyboard.

4.2 Adopting Spatial Model and Language Model for Single-handed Keyboards

Inspired by previous research [5, 34], we also designed keyboard hint functions to predict the user’s input based on the spatial model (or called the touch distribution model) and the language model (i.e., the most frequently used word lexicon). When the user touches keys, the hint algorithm combines the probabilities from the two models to obtain the highest likelihood of the word. The hint is word-based, not sentence-based. The following part gives the details of the hint computation algorithm.

Language Model. The language model comes from a 100K most-frequently-used word lexicon; the assigned weight of each word is the reverse of its frequency-based rank, for example, the first ranked word has a weight of 100K while the last-ranked word has a weight of 1.

Spatial Model. Each key touch has a series of probabilities associated with nearby keys (e.g., a key touch within the key boundary “s” may originally expect to press “s”, “w”, “a”, or “q”, each of them has a probability). From the above study #1, we already acquired

the spatial model; thus, the associated keys’ probabilities from each touch can be calculated directly.

As reported in previous research [2], touch points for tapping a key follow a bivariate Gaussian distribution. Therefore, assuming the coordinates of a touch point c_i are (x_i, y_i) , and the intended key is k , the probability $p(c_i|k)$ can be calculated as:

$$p(c_i|k) = \frac{1}{2\pi\sigma_{i_x}\sigma_{i_y}\sqrt{1-\rho_i^2}} \exp\left[-\frac{z}{2(1-\rho_i^2)}\right], \quad (1)$$

where

$$z \equiv \frac{(x_i - \mu_{i_x})^2}{\sigma_{i_x}^2} - \frac{2\rho(x_i - \mu_{i_x})(y_i - \mu_{i_y})}{\sigma_{i_x}\sigma_{i_y}} + \frac{(y_i - \mu_{i_y})^2}{\sigma_{i_y}^2}. \quad (2)$$

Here (μ_{i_x}, μ_{i_y}) is the center of the touch point of k , σ_{i_x} and σ_{i_y} are standard deviations, and ρ_i is the correlation.

A word may consist of multiple characters. Assuming that a word W is comprised of n characters: k_1, k_2, \dots, k_n , C has n touch points: c_1, c_2, \dots, c_n , and touches are independent of each other, we obtain:

$$P(C | W) = \prod_{i=1}^n P(c_i|k_i) \quad (3)$$

Statistical Decoding for Hints. The following part is used for hint selection. Given a set of touch points $S = \{s_1, s_2, \dots, s_n\}$ on the keyboard layout, the algorithm is to find a word W in a lexicon L that satisfies:

$$W = \arg \max_{W \in L} P(W | S). \quad (4)$$

According to the Bayes’ theorem,

$$P(W | S) = \frac{P(S | W)P(W)}{P(S)}. \quad (5)$$

As $P(S)$ is an invariant across words, Equation 4 becomes:

$$W = \arg \max_{W \in L} P(S | W)P(W). \quad (6)$$

Here $P(W)$ is the language model, and $P(S | W)$ is the spatial model.

We used the same language model for all the keyboards, the language model is a 70K frequency dictionary. In light of the layout similarity, we used the same spatial model for both AC and CK. AQ used its own spatial model (refer to the above study # 1).

4.3 Study #2: Evaluate the single-handed keyboards with spatial and language models

The purpose of the study #2 is to measure and compare the performances of the three single-handed keyboards: AQ, CK and AC. Based on the acquired data from the first study, we implemented the spatial and language models for the three studied keyboards: AQ, AC, and CK, following the methodology in [5, 34]. Then, we conducted the following study #2 to compare their performances. The AQ is the baseline in this study.

Experiment Setup. For each of the three studied keyboards, participants are instructed to type texts on the experimental smartphone using the left hand at their best speeds, given some reference phrases one by one. Whenever one phrase was completed, the participants can press the Enter key to continue the typing of the next phrase. At the beginning of the experiment, all the participants need to draw a quarter circle as shown in Figure 2(a). Two

user-adaptable-sized keyboards (i.e., the AC and AQ keyboards) are constructed accordingly. They can redo the keyboard customization process if they want.

Participants. 18 participants (6 female, 12 male; ages from 23 to 45) participated in this study. All of them have at least 5 years experience of typing on touchscreen smartphones. One of them is left-handed, but the participant used both hands to type normally.

Phrase number. 20 different phrases.

User experience questionnaire (UEQ). UEQ [29] is a widely-used, direct measurement of user experience. The questionnaire is designed for use as a part of a normal usability test. The used UEQs contain 26 items grouped into six aspects (scales) :

- **Attractiveness:** Do users like or dislike it? Is it attractive, enjoyable, or pleasing? 6 items: annoying/enjoyable, good/bad, unlikable/pleasing, unpleasant/pleasant, attractive/unattractive, and friendly/unfriendly.
- **Perspicuity:** Is it easy to get familiar with the keyboard? Is it easy to learn? Is the keyboard easy to understand and clear? 4 items: not understandable/ understandable, easy to learn/ difficult to learn, complicated/ easy, and clear/ confusing.
- **Efficiency:** Can users solve their tasks without unnecessary effort? Is the interaction efficient and fast? Does the keyboard react fast to user input? 4 items: fast/ slow, inefficient/ efficient, impractical/ practical, and organized/ cluttered.
- **Dependability:** Do users feel in control of the interaction? Can they predict the system behavior? Do users feel safe when working with the keyboard? 4 items: unpredictable/ predictable, obstructive/ supportive, secure/ not secure, and meet expectations/ does not meet expectations.
- **Stimulation:** Is it exciting and motivating to use the keyboard? Is it fun to use? 4 items: valuable/ inferior, boring/ exciting, not interesting/ interesting, and motivating/ demotivating.
- **Novelty:** Is the keyboard innovative and creative? Does it capture users' attention? 4 items: creative/ dull, inventive/ conventional, usual/ leading-edge, and conservative/ innovative.

The above measures are not assumed to be independent. In fact, the user's general impression is captured by the Attractiveness scale, which could be influenced by the values of the other 5 scales.

Results. In the second study, we collected a total of more than 14,664 valid touch points. Each keyboard has nearly 5 thousands touch points on average. Other than the objective key touch data, the participants took 3 UEQs (see above) when they finished the task. In each UEQ, they were asked to rate a keyboard. In the following, we analyze the collected data.

Error Rate. We use the refined KSPC (keystrokes per character) in [30] as the error rate, which better fits for smartphones. It is calculated as follows:

$$KSPC \approx \frac{C + INF + IF + F}{C + INF}, \quad (7)$$

where C stands for the number of correct characters, INF stands for the number of incorrect and unfixed characters, F stands for the number of fixes, and IF stands for the number of incorrect but fixed characters.

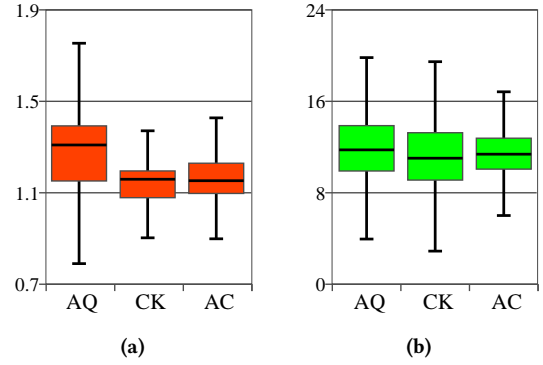


Figure 5: Boxplot of the performance comparisons of the three keyboards with objective performance data. The vertical lines represent the ranges. (a) KSPC Error Rates. (b) WPM Entry Rates

The KSPC error rate reflects how many trials are needed to correctly enter a character. Due to our bi-variable Gaussian hint prediction algorithm, a higher error rate may not greatly hinder the performance of a keyboard. To ensure a fair comparison among the keyboards, we did not implement a user library in the second study so that all the keyboards have the same hints given the same input sequence. This may somehow weaken the performance of the hint function, but the results of our comparison study would be fair.

The means of the KSPC error rates are 1.26 (SD=0.020) for AQ, 1.14 (SD=0.006) for CK, and 1.16 (SD=0.007) for AC. The ANOVA analysis shows a significant main effect of keyboard type ($F_{2,51} = 6.33, p = 0.0035$). Pairwise comparison with Bonferroni adjustment shows the difference is significant for both AQ vs. AC and AQ vs. CK. The error rate of AQ is significantly higher than the other two layouts due to the smaller size of its keys and thus the inaccuracy of typing, while CK and AC have similar average KSPC error rates (1.14 vs. 1.16).

Input Speed. The input speed WPM (words per minute) is calculated as:

$$WPM = \frac{|S - 1|}{T} \times \frac{1}{5}, \quad (8)$$

where S is the length of the given phrase in characters, and T is the time elapsed in minutes from the first touch event to the last touch event of the phrase. Five characters are counted as one word in our analysis.

The mean WPM is 11.80 (SD=10.20) for AQ, 11.03 (SD=7.42) for CK, and 11.07 (SD=4.61) for AC. The ANOVA analysis does not show a significant main effect of keyboard type with regard to WPM. Although the mean WPM of AQ is slightly higher (but fewer than 7%) than the other two layouts, the input speed advantage may come from the different degree of familiarity, because none of the participants claimed to be familiar with the two curved layouts, while all of them are proficient with the Qwerty keyboard.

Subjective Comparison Results. We analyzed the collected UEQ data in the second study, and summarized the calculated means and variances in Table 1. As shown in the column 1 in Table 1, when the CK keyboard is evaluated, all the 6 scales are positive numbers.

So, from this perspective, all the participants generally approved the CK keyboard for its way of deforming the keyboard. The highest achieved score is "Novelty", which means that the participants thought the curved-shape is innovative, compared to widely-known rectangle-shaped keyboards. The scores of "Attractiveness" and "Stimulation" are also reasonably high, which shows the interest levels of the participants on the CK keyboard. By contrast, the relatively low (still positive) scores are "Perspicuity", "Efficiency", and "Dependability". The above relatively low scores could be due to the fact that the CK keyboard is generally new to the participants, and thus their typing could slow down when they cannot easily find some keys in their habitual way.

The column 2 in Table 1 shows the UEQ result of the AC keyboard. The relatively high scores of the AC keyboard are in the scales of "Attractiveness", "Stimulation", and "Novelty". But the "Dependability" and "Efficiency" scores are lower than the corresponding ones in the CK group (see column 1 in Table 1). From the participants' perspectives, the change from the CK keyboard to the AC keyboard is not drastic. This indicates, the participants generally approved AC was useful for single-handed typing. On the other hand, the "Dependability" score is the lowest among the three groups (CK, AC, and AQ), which means the AC has a similar dependability as the users' previous experiences.

The column 3 in Table 1 shows the UEQ outcome when the AQ keyboard was evaluated. The highest score is "Perspicuity". The reason is that since all the participants are already familiar with the Qwerty layout, it is easy for them to be adept at the new "adaptable-size" Qwerty keyboard. In light of the fact that the AQ keyboard allows all the participants to type with only one hand. The scores in "Novelty", "Stimulation", and "Attractiveness" are relatively low compared to the CK and AC keyboards. This could be explained as follows: the changes from the regular sized Gboard to the AQ keyboard is relatively small, and thus it is hardly considered as an innovation.

4.4 Study #3: Evaluate The Learning Curves of Single-handed Keyboards

The purpose of the study #3 is to acquire the learning curves of the single-handed keyboards. By performing experiments in 3 consecutive days, participants are expected to have a similar familiarity with all the single-handed keyboards and thus we can compare their performances over time.

Experiment Setup. 9 participants (1 female, 8 males; ages from 29 to 45) participated in this study. All the participants have at least 5 years experience of typing on the touchscreen keyboard on mobile devices. 15 different phrases are used.

Results. More than 14,788 valid touch points were collected each day. Each keyboard has nearly 5 thousands touch points per day. We analyzed the error rates and entry speeds, and obtained the learning curves by comparing the performance data in 3 days, illustrated in Figure 6.

Day 1. The obtained data from day 1 is very similar to that in the study #2.

Error Rate. The means of the KSPC error rates are 1.23 (SD=0.023) for AQ, 1.12 (SD=0.003) for CK, and 1.11 (SD=0.002) for AC. The ANOVA analysis shows a significant main effect of keyboard type

Attribute	CK		AC		AQ	
	Mean	Vari.	Mean	Vari.	Mean	Vari.
Attractiveness	1.306	0.99	1.608	1.39	0.991	2.59
Perspicuity	1.000	1.72	1.042	1.35	1.278	2.29
Efficiency	1.111	1.94	1.049	2.08	1.028	2.65
Dependability	0.986	0.58	0.740	1.23	0.889	1.88
Stimulation	1.389	0.53	1.697	1.39	0.764	1.87
Novelty	1.569	0.64	1.961	1.01	0.833	1.35

Table 1: The calculated UEQ scales (the means and variances) of the three keyboard layouts in the experiment #2. The range of the mean value is from -3 to 3: the positive value means participants prone new keyboard shows more given attribute. On the contrary, the negative value means participants prone previous experienced keyboard shows more given attribute. Variances reflect the consistency of participants' opinion, the less, the more consistency.

($F_{2,24} = 3.822, p = 0.03$). Pairwise comparison with Bonferroni adjustment shows the difference is significant for both AQ vs. AC and AQ vs. CK. The error rate of AQ is significantly higher than those of the other two.

Input Speed. The means of WPM are 11.66 (SD=11.97) for AQ, 10.99 (SD=6.00) for CK, and 11.42 (SD=7.36) for AC. The ANOVA analysis does not show a significant main effect of keyboard type with WPM.

Day 2. From the input speed data at day 2, we observe a clear speed-up compared to the data at day 1. This suggests that the one-day practice did take some effect. The error rates show little difference, compared to the rates in the first day.

Error Rate. The means of the KSPC error rates are 1.26 (SD=0.016) for AQ, 1.15 (SD=0.003) for CK, and 1.13 (SD=0.007) for AC. The ANOVA analysis shows a significant main effect of keyboard type ($F_{2,24} = 5.48, p = 0.01$). Pairwise comparison with Bonferroni adjustment shows the difference is significant for both AQ vs. AC and AQ vs. CK. The error rate of AQ is significantly higher than those of the other two.

Input Speed. The means of WPM are 12.51 (SD=9.06) for AQ, 12.44 (SD=7.92) for CK, and 13.41 (SD=9.28) for AC. The ANOVA analysis does not show a significant main effect of keyboard type with WPM. Up to day 2, the AC shows the best performance of all the keyboards, and its advantage is about 8% of the means. This is a substantial improvement (17.5%) compared to the first day's data. This also indicates that among the studied keyboards, AC is the easiest to learn for users.

Day 3. The input speed data in the third day has a speed up, compared to those in day 2 and day 1. But still, the error rates do not show much change.

Error Rate. The means of the KSPC error rates are 1.23 (SD=0.014) for AQ, 1.12 (SD=0.003) for CK, and 1.11 (SD=0.003) for AC. The ANOVA analysis shows a significant main effect of keyboard type ($F_{2,24} = 5.81, p = 0.009$). Pairwise comparison with Bonferroni adjustment shows the difference is significant for both AQ vs. AC and AQ vs. CK. The error rate of AQ is significantly higher than those of the other two.

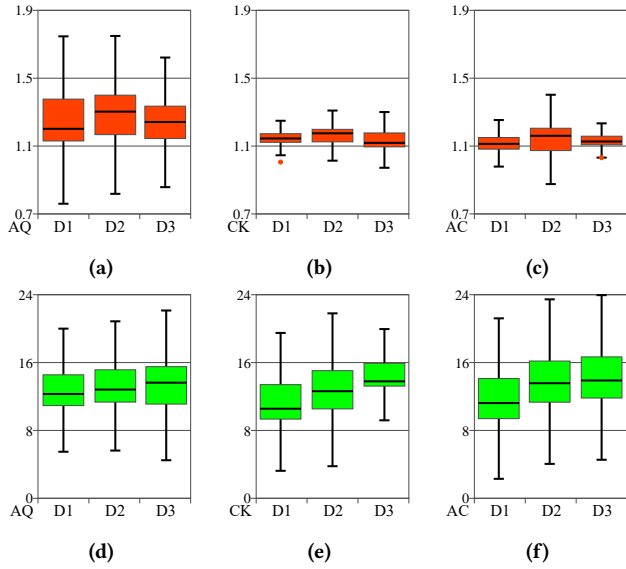


Figure 6: (a) - (c) are the error rate data in three days, (d) - (f) are the input speed data in three days. (a) AQ’s KSPC in 3 days. (b) CK’s KSPC in 3 days. (c) AC’s KSPC in 3 days. (d) AQ’s WPM in 3 days. (e) CK’s WPM in 3 days. (f) AC’s WPM in 3 days.

Input Speed. The means of WPM are 13.12 (SD=10.46) for AQ, 13.72 (SD=5.66) for CK, and 13.85 (SD=9.09) for AC. The ANOVA analysis does not show a significant main effect of keyboard type with WPM. Up to day 3, the AC still shows the best performance among all the studied keyboards. The CK at day 3 shows a substantial improvement (10.3%), compared to day 2.

Prediction of WPM over time. Following the prediction analysis in previous research [21], we found the WPMs of CK and AQ reached the crossover point after the second day: $R_{AQ}^2 = 0.9985$, $R_{CK}^2 = 0.9932$, and $R_{AC}^2 = 0.9538$. The prediction of the upper-bound text entry rate can be plotted in Figure 7. It shows that AQ outperformed the two curved keyboards (AC and CK) at day 1; then, AC outperformed AQ at day 2 and CK achieved a close performance to AQ at day 2. At day 3, both the curved keyboards (AC and CK) clearly outperformed AQ. The prediction curves show that the advantage of curved keyboards will become more obvious after the practice of several more days. But because of the similarity of the two curved keyboards (AC and CK), the prediction curves also show that both of them would have similar WPM performances over time.

Subjective Feedback. After the 3-days user study, each participant was asked to rate the keyboards using a 5-level scale: 1 (Very dislike) to 5 (Very like). The means of subjective ratings are 1.2 for AQ (SD = 0.15), 3.06 for CK (SD = 0.53), 4.28 for AC (SD = 0.69). The self-reported results are consistent with the trend observed in Figure 7, where AC is considered the best while the AQ is perceived to be the worst among the three over time.

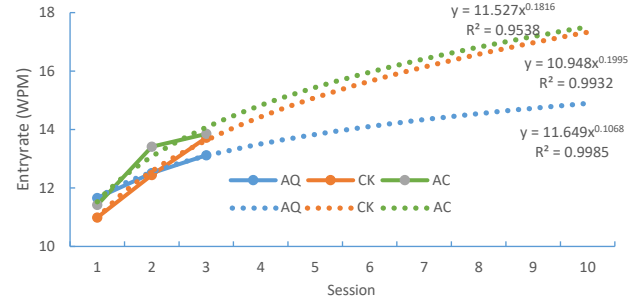


Figure 7: The prediction of the WPM upper-bounds of the AQ, CK and AC keyboards.

5 DISCUSSION

First, the subjective feedback from both the studies #2 and #3 support the finding that the AC keyboard is more preferred than the other keyboards by participants. Second, as shown in Figure 6, the KSPC rates of all the single-handed keyboards are consistently low. This indicates that the spatial and language models are reasonably effective at handling input noise from single-hand typing, although it is well recognized that single-hand typing has substantially larger noise than two-hands typing. Third, in terms of the KSPC error rate and input speed WPM, the AC keyboard performed best among the studied keyboards. This can be more clearly observed after users get more familiar with the keyboards (e.g., more sessions of practice). Fourth, the input speed WPM can be measurably improved over a study period of 3 days, as clearly observed in Figure 6 and Figure 7. By contrast, we cannot observe consistent improvements of the KSPC error rate for the three studied keyboards, which indicates that the error rate of single-handed text entry cannot be easily reduced through practices in several days but could largely depend on the keyboard layout design.

6 CONCLUSION

In this work we study single-handed typing behaviors on several representative keyboards on mobile devices. We found that, (i) the AC performs best among all the studied keyboards; (ii) users’ familiarity with the Qwerty layout plays a significant role at the beginning, but after several sessions of practice, the AC can have the best learning curve and performs best, which is supported by both the collected objective performance data and subjective feedback data. It also fits the well-known Fitts movement theory [13] and its refined version [1]; (iii) generally the statistical decoding algorithms via spatial and language models can well handle the input noise from single-handed typing. Our current study has certain limitations. For example, in our user studies, we asked participants to sit on a chair in a stationary environment while doing the user studies. But for a daily use, users may sometimes walk or stand to do typing on mobile devices.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Xiaojun Bi for helpful discussion and feedback. This work was in part supported by NSF Grant IIS-2005430.

REFERENCES

- [1] Johnny Accot and Shumin Zhai. 2003. Refining Fitts' Law Models for Bivariate Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (CHI '03). ACM, New York, NY, USA, 193–200.
- [2] Shiri Azenkot and Shumin Zhai. 2012. Touch Behavior with Different Postures on Soft Smartphone Keyboards. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services* (San Francisco, California, USA) (*MobileHCI '12*). ACM, New York, NY, USA, 251–260.
- [3] Joanna Bergstrom-Lehtovirta and Antti Oulasvirta. 2014. Modeling the Functional Area of the Thumb on Mobile Touchscreen Surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI '14). ACM, New York, NY, USA, 1991–2000.
- [4] Xiaojun Bi, Yang Li, and Shumin Zhai. 2013. FFitts law: modeling finger touch with fitts' law. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1363–1372.
- [5] Xiaojun Bi and Shumin Zhai. 2013. Bayesian Touch: A Statistical Criterion of Target Selection with Finger Touch. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) (UIST '13). ACM, New York, NY, USA, 51–60.
- [6] Youli Chang, Sehi L'Yi, Kyle Koh, and Jinwook Seo. 2015. Understanding Users' Touch Behavior on Large Mobile Touch-Screens and Assisted Targeting by Tilting Gesture. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). ACM, New York, NY, USA, 1499–1508.
- [7] Youli Chang, Sehi L'Yi, Kyle Koh, and Jinwook Seo. 2015. Understanding users' touch behavior on large mobile touch-screens and assisted targeting by tilting gesture. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1499–1508.
- [8] Mauro Dell'Amico, José Carlos Díaz Díaz, Manuel Iori, and Roberto Montanari. 2009. The single-finger keyboard layout problem. *Computers & Operations Research* 36, 11 (2009), 3002–3012.
- [9] Mark Dunlop and John Levine. 2012. Multidimensional pareto optimization of touchscreen keyboards for speed, familiarity and improved spell checking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2669–2678.
- [10] Rachel Eardley, Anne Roudaut, Steve Gill, and Stephen J. Thompson. 2017. Understanding Grip Shifts: How Form Factors Impact Hand Movements on Mobile Phones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). ACM, New York, NY, USA, 4680–4691.
- [11] Leah Findlater and Jacob Wobbrock. 2012. Personalized Input: Improving Ten-finger Touchscreen Typing Through Automatic Adaptation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) (CHI '12). ACM, New York, NY, USA, 815–824.
- [12] Leah Findlater, Jacob O. Wobbrock, and Daniel Wigdor. 2011. Typing on Flat Glass: Examining Ten-finger Expert Typing Patterns on Touch Surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) (CHI '11). ACM, New York, NY, USA, 2453–2462.
- [13] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.
- [14] Mayank Goel, Alex Jansen, Travis Mandel, Shwetak N. Patel, and Jacob O. Wobbrock. 2013. ContextType: Using Hand Posture Information to Improve Mobile Touch Screen Text Entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI '13). ACM, New York, NY, USA, 2795–2798.
- [15] Sunghyuk Kwon, Donghun Lee, and Min K Chung. 2009. Effect of key size and activation area on the performance of a regional error correction method in a touch-screen QWERTY keyboard. *International Journal of Industrial Ergonomics* 39, 5 (2009), 888–893.
- [16] Huy Viet Le, Sven Mayer, Patrick Bader, and Niels Henze. 2018. Fingers' Range and Comfortable Area for One-Handed Smartphone Interaction Beyond the Touchscreen. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). ACM, New York, NY, USA, Article 31, 12 pages.
- [17] Luis A. Leiva, Alireza Sahami, Alejandro Catala, Niels Henze, and Albrecht Schmidt. 2015. Text Entry on Tiny QWERTY Soft Keyboards. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (CHI '15). ACM, New York, NY, USA, 669–678.
- [18] Lawrence Leung and Parham Aarabi. 2014. Mobile circular keyboards. In *Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on*. IEEE, 1–4.
- [19] Yanzhi Li, Lijuan Chen, and Ravindra S Goonetilleke. 2006. A heuristic-based approach to optimize keyboard design for single-finger keying applications. *International Journal of Industrial Ergonomics* 36, 8 (2006), 695–704.
- [20] 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* (Ft. Lauderdale, Florida, USA) (CHI EA '03). ACM, New York, NY, USA, 754–755.
- [21] I Scott MacKenzie and Shawn X Zhang. 1999. The design and evaluation of a high-performance soft keyboard. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 25–31.
- [22] Microsoft.Com. 2018. Word Flow Keyboard. Retrieved May 13, 2018 from <https://www.microsoft.com/en-us/garage/profiles/word-flow-keyboard/>
- [23] Anna Ostberg, Mohamed Sheik-Nainar, and Nada Matic. 2016. Using a Mobile Device Fingerprint Sensor As a Gestural Input Device. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) (CHI EA '16). ACM, New York, NY, USA, 2625–2631.
- [24] Antti Oulasvirta, Anna Reichel, Wenbin Li, Yan Zhang, Myroslav Bachynskiy, Keith Vertanen, and Per Ola Kristensson. 2013. Improving Two-thumb Text Entry on Touchscreen Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (CHI '13). ACM, New York, NY, USA, 2765–2774.
- [25] Philip Quinn and Shumin Zhai. 2016. A cost-benefit study of text entry suggestion interaction. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, 83–88.
- [26] Hanae Rateau, Yosra Rekik, Edward Lank, and Keiko Katsuragawa. 2021. Edge-Mark Menu: A 1D Menu for Curved Edge Display Smartphones. In *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*. 1–9.
- [27] Taebeum Ryu, Jihyoun Lim, Joobong Song, Myung Hwan Yun, and Moonsoo Shin. 2013. Performance comparison between the preferred right and preferred left hands in text entry using Qwerty touch keyboard smartphones. *International Journal of Industrial Ergonomics* 43, 5 (2013), 400–405.
- [28] Oliver Saylor. 2018. Handy - The One-Hand Keyboard for iPhone 6 and 6 Plus Quick & Easy 1 Handed Typing. Retrieved May 13, 2018 from <https://itunes.apple.com/us/app/handy-one-hand-keyboard-for-iphone-6-6-plus-quick-easy/id951349222?mt=8>
- [29] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *IJIMAI* 4, 4 (2017), 40–44.
- [30] R. William Soukoreff and I. Scott MacKenzie. 2003. Metrics for Text Entry Research: An Evaluation of MSD and KSPC, and a New Unified Error Metric. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) (CHI '03). ACM, New York, NY, USA, 113–120.
- [31] Hui-Shyong Yeo, Xiao-Shen Phang, Steven J. Castellucci, Per Ola Kristensson, and Aaron Quigley. 2017. Investigating Tilt-based Gesture Keyboard Entry for Single-Handed Text Entry on Large Devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (CHI '17). ACM, New York, NY, USA, 4194–4202.
- [32] Hui-Shyong Yeo, Xiao-Shen Phang, Steven J Castellucci, Per Ola Kristensson, and Aaron Quigley. 2017. Investigating tilt-based gesture keyboard entry for single-handed text entry on large devices. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 4194–4202.
- [33] Ying Yin, Tom Yu Ouyang, Kurt Partridge, and Shumin Zhai. 2013. Making touchscreen keyboards adaptive to keys, hand postures, and individuals: a hierarchical spatial backoff model approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2775–2784.
- [34] Suwen Zhu, Tianyao Luo, Xiaojun Bi, and Shumin Zhai. 2018. Typing on an Invisible Keyboard. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). ACM, New York, NY, USA, Article 439, 13 pages.