Assessment of Writing Text in Mobile Devices

Rodrigo Capa-Arnao [1], Cristina Manresa-Yee [1],

Ramon Mas-Sansó [1], Martin Larrea [2]

[1] Math and Computer Science Department.
Universitat de les Illes Balears. SPAIN

[2] Laboratorio de Investigación y Desarrollo en Visualización y Computación Grafica (VyGLab), Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur, ARGENTINA
correo@derodrigo.com, {cristina.manresa, ramon.mas}@uib.es,
mll@cs.uns.edu.ar

Abstract. The input of text on mobile devices has evolved from physical keyboards to virtually displayed touch keyboards. The aim of the paper is to assess and compare two mobile phone text input methods (mini-QWERTY and on-screen keyboards) to analyze which one is more efficient and effective. We review previous works related to text entry metrics and studies on mini-QWERTY physical and virtual keyboards.

Keywords: HCI, usability, text input, mobile phones, mini-QWERTY

1 Introduction

Mobile phones have evolved to become multiple-featured communicating devices. The increase of the processing power and storage capacity, the introduction of new and more complex operating systems and the use of larger screen sizes allow current mobile devices to perform tasks traditionally devoted to computers.

One of the main restrictions of mobile devices is the input of text, a task that is very easily and comfortably done with personal computers using traditional keyboards. Mobile devices manufacturers have addressed this problem from two different perspectives: to use a mini-QWERTY physical keyboard (i.e. BlackBerry and some high-end Symbian based Nokia phones) or to use a tactile keyboard on the touch screen (i.e. IOS and Android devices).

At first, and if we not take into account the 12-key keypad, mobile devices used either a virtual keyboard or character recognition systems. Virtual keyboards were operated using a stylus because of the resistive nature of the touch screens. Character recognition was offered as an alternative input text method, but high error rates combined with a long learning-training process highly reduced its usability and performance [1, 2, 3, 4].

Current devices introduce new and more sophisticated text input techniques combined with text prediction and correction. For example, Paek et al. [5] and Riadi

[6] compared input techniques considering advances in text prediction and on-screen keyboards of several formats.

Although mini-QWERTY physical keyboards are proven to be faster and less prone to errors they are not always the users' choice [4].

In this paper we review the text-input metrics and we compare the use of touch screen to physical mini-QWERTY keyboards in mobile phones.

2 Previous Work

2.1 Performance measure

The Words per minutes (WPM) value is a measure of writing speed. As a rule-of-thumb a word is considered to be five characters long (in average) [5, 7]. It is important to highlight that WPM does not count keystrokes but only output characters. This means that if the user writes "Alamo" or "alamo" both count as one word, even if the first character in the first word requires two keystrokes (SHIFT+'a') hence requiring more time to write than just a "a". This difference becomes relevant when writing in a language that requires special characters, like Spanish with characters such as 'ñ' or 'á'.

The rate of keystrokes over characters (KSPC) [8] is a measure of the user's errors when writing. If the user does not insert any errors when writing, then the KSPC is equal to one. This means that there was one keystroke for each character. If the text entry has a mistake then there was more than one keystroke involved, then increasing the KSPC. When using traditional writing methods, the KSPC can never be lower than one. But, with predicted writing systems, a word can be introduced with less keystrokes than characters in it. This results in a KSPC lower than one.

The WPM metric does not reflect the user's errors and therefore it cannot be used as a real measure of performance. A way of improving this metric is to weight in the number of errors. For example, by subtracting the number of errors from the WPM.

Another method used to measure performance is the minimum string distance (MSD) [9]. This metric computes the minimum distance between two strings defined in terms of editing primitives. The primitives are *insertion, deletion,* and *substitution*. Given two character strings, the idea is to find the smallest set of primitives that applied to one string produces the other. The number of primitives applied is the MSD. This measurement can be used as a performance metric by using the MSD as an indicator of the number of corrections required to fix a misspelled word. Then, the user's error rate, based on the MSD, can be calculated as the rate between the MSD and the length of the sentence (Equation 1).

$$error\ rate = \frac{MSD\ (A,B)}{\max(|A|,|B|)} \qquad \text{Eq. 1}$$

For example, considering the following sentences:

This phrase is correct
This phrase nis correct

In order to transform the second sentence into the first one, we need to eliminate one character, the 'n'. The MSD for this example is 1. If we rate the MSD with the length of the second sentences, 23, the result is the error rate. In this case is 0.043. This metric is interesting because it considers both the omission as well as the inclusion of characters.

2.2 Performance analysis based on a theoretical model

The analysis of input methods on mobile devices has been done from different point of views. One of them is Fitts' law (1954), which involves the time of movement, the distances travelled and the size of the object to select. There are many studies on writing performance based on Fitts' law but none of them cover all the available devices [10]. In order to include some of these results, we will describe only three of them.

The first one is from Silfverberg et. al. [11], in this work the authors compared the performance of multitap input (one character is obtained after several keystrokes) versus a predicted T9 dictionary on a Nokia 5110 mobile device. They concluded that a user using the T9 dictionary and with an optimal training increase his writing speed between 52% and 109% over the multitap method.

Another study [12] compared user's performance on a QWERTY keyboard versus an OPTI one, both projected onto a touch screen and operated using a stylus. They concluded that the maximum writing speed on the QWERTY keyboard was 43 WPM, and 58 on the OPTI. In both cases, the results are theoretical. The authors then tested several real users using these keyboards and got 40 WPM for the QWERTY and 45 WPM for the OPTI.

Finally, in 2002 MacKenzie and Soukoreff [13] studied a mini-QWERTY keyboard and concluded that the average input speed was between 50 and 60 WPM.

3 Methodology

This section aims at evaluating the current text entry mechanisms in mobile phones. Two systems will be compared by using the Curran et al. [4] experiment in laboratory conditions.

3.1 Apparatus

Two families of mobile devices which count with a tactile screen are globally present: iPhone and Android. iPhones do not count with devices with mini-QWERTY, but there are Android devices which count with mini-QWERTY keyboards plus on-screen tactile keyboards. Therefore, an HTC Desire was used to test both text entry strategies using the same device (see Fig. 1).

The application MyTextSpeed was used to carry out the tasks. In the superior half screen, the phrase to copy is presented and in the low half screen, there is space to write the sentence. The writing duration is automatically registered, starting when the user introduces the first character and finishing when the user introduces the Enter key. In Curran's experiment this last character is not introduced, but for the system to register automatically the duration is the simplest mode. This software does not allow computing the error rate using MSD, so after finishing each subtask, a photograph of the screen was taken to control errors.

Fig. 1. Mobile phone with mini-QWERTY

3.2 Procedure

This experiment involved a within-group study, that is, all users had to try both input strategies (mini-QWERTY and tactile). At the beginning of the session, the experiment conductors described and explained the evaluation objectives to the users and described the two input strategies.

There were two tasks with four subtask each: users had to write four phrases with both input strategies. The phrases were in English, and in order for it not to have an influence on the users, this criterion was taken in to account to select the users. Before writing the phrases, users could read them as many times as desired.

The phrases were:

1. I have never sent a text message before
2. Your flying lesson's cancelled today. Call Andrew from 7:00pm onwards to arrange another lesson.
3. Plane gets in at 10:00pm to Gate 11. Aerlingus flight no. EI 987. Can you meet me? My e-mail address is biggles@hotmail.com!
4. let me no where u r and il pic u up 18r

As Curran et al [4] state, the first phrase is simple and containing no punctuation characters. The second phrase is more complex because it includes punctuation, numbers and lower and upper case letters. Then, the third phrase is more complex and uses more letters, numbers and punctuation characters. Finally the fourth phrase had abbreviations for words and phrases.

The learning bias that could influence the repetition of the task with the second input strategy was reduced by selecting as the first input device the one that the user used usually, in this way, the level of complexity increased gradually. Users before starting the tasks could try the input strategy as long as desired and they usually trained between three and five minutes.

Fatigue or boredom did not appear because the test was of short duration (less than five minutes each task).

At the beginning of the test, users answered a questionnaire compiling demographics information (age, gender), their experience level using Android devices and using both input strategies, kind of tasks frequently carried out and style of writing with mobile phones (use of predictive text, correction of errors, number of fingers and fingers used when writing)

3.3 Participants

Ten users were recruited from the general campus population to test the game. Participant demographics included ages ranging from 18 to 40, and 4 men and 6 women. The criteria followed to select the users were:

- Users had to have an average level of English
- Users had to be Android users
- Users had to know how to write with the QWERTY layout.

4 Results

In Table 1 the results of the speed and error rate are shown. It can be observed that the mini-QWERTY text entry strategy is slightly faster than the on-screen keyboard and with less errors.

Observing the results regarding the users' experience is interesting to highlight those users with experience with mini-QWERTY as they obtained faster results with this system, but the error rate does not vary. Regarding those users without experience, their performance in speed is similar for both systems but inserting more errors when working with the on-screen keyboard. We could observe that users who relied usually on the predicting mechanism for on-screen did not notice errors on words, therefore increasing the error rate (see Table 2 and 3).

Table 1. Average and standard deviations of error rate and WPM.

Input strategy	WPM(mean)	MSD (mean)	WPM (SD)	MSD (SD)
Mini-QWERTY	14,22	1,07	10,67	1,42
On-screen keyboard	13,32	1,48	7,8	2,03

Table 2. Average and standard deviations of WPM regarding miniQWERTY experience.

Experience with mini-QWERTY	WPM mini-QWERTY (m)	WPM on-screen keyboard (m)	WPM mini-QWERTY(SD)	WPM on-screen keyboard (SD)
Yes	16.05	14.32	3,55	2,67
No	11.48	11.85	9,41	14,08

Table 3. Average and standard deviations of error rate regarding miniQWERTY experience.

Experience with mini-QWERTY	MSD mini-QWERTY (m)	MSD on-screen Keyboard (m)	MSD mini-QWERTY(SD)	MSD on-screen Keyboard (SD)
Yes	0,95	0,95	1,95	1,05
No	1,26	2,27	0,93	2,96

If we compare the results of our experiment with the ones obtained by Curran et al [4], the results for the mini-QWERTY strategy are very similar regarding the speed but results for WPM on-screen keyboard differ being much faster in our experiment than in Curran's (see Fig. 2). This can be due to the difference in the device used, as they used a Motorola PDA. On the other hand, error rate is higher in our experiment for both systems; being greater the error rate for on-screen keyboards (see Fig. 3).

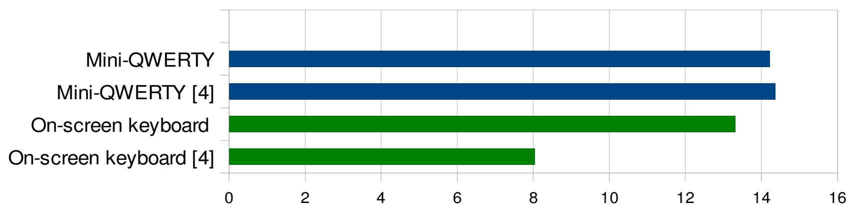

Fig. 2. WPM comparison

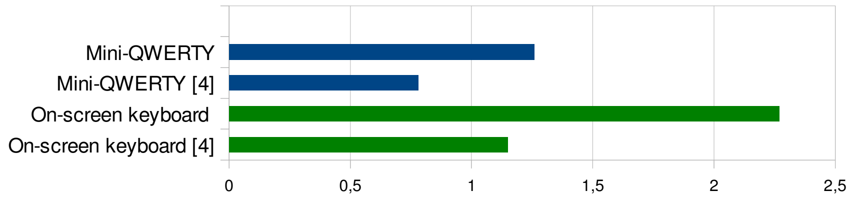

Fig. 3. Error rate comparison

5 Discussion and conclusions

The most efficient input device for text-entry is the mini-QWERTY, as users have achieved better writing speed and fewer errors. However, modern devices tend to include only on-screen keyboards due to aesthetics, weight and size reduction and building costs.

Moreover, as users acquire experience with the on-screen keyboard their typing speed gets closer to the mini-QWERTY, nevertheless inserting more mistakes particularly due to the predictive text.

Nowadays new text input strategies appear to work faster with on-screen keyboards such as sliding your finger through the keyboard instead than pressing each key. Following the evolution, it seems that in a short period the on-screen keyboard with new text input strategies may allow trained users to write faster than with mini-QWERTY, and hopefully with less errors.

Acknowledgments. This work is partially supported by Spanish MAEC- AECID A1/037910/11.

References

1. Pila Suomalainen, L. K., Pääkkönen, R. 2010. A comparision of the usability of a laptop, communicator, and handheld computer. *Journal of Usability Studies 5*, 111–123.
2. Yatani, K., Truong, K. N. 2007. An evaluation of stylus-based text entry methods on handheld devices in stationary and mobile settings. In Proceedings of the 9th international conference on human computer interaction with mobile devices and services. MobileHCI '07. ACM, New York, NY, USA, 487–494
3. Commarford, P. M. 2004. An investigation of text throughput speeds associated with pocket pc input method editors. *Int. J. Hum. Comput. Interaction 17,* 3, 293–308.
4. Curran, K.,Woods, D., Riordan, B. O. 2006. Investigating text input methods for mobile phones. *Telematics and Informatics 23,* 1, 1 –21.
5. Paek, T., Chang, K., Almog, I., Badger, E., Sengupta, T. 2010. A practical examination of multimodal feedback and guidance signals for mobile touchscreen keyboards. In *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. MobileHCI '10. ACM, New York, NY, USA, 365–368.

6. Riadi, D. 2005. An overview and usability measurement of virtual keyboard.
7. Arif, A.S. and Stuerzlinger, W. 2011. Analysis of Text Entry Performance Metrics. In Proc. IEEE TIC-STH 2009. IEEE New York (2009), 100–105
8. MacKenzie, I. S., Soukoreff, R. W. 2002b. Text entry for mobile computing: Models and methods,theory and practice. *Human–Computer Interaction 17,* 2-3, 147–198.
9. Soukoreff, R. W., MacKenzie, I. S. 2001. Measuring errors in text entry tasks: an application of the levenshtein string distance statistic. In CHI '01 extended abstracts on Human factors in computing systems. CHI EA '01. ACM, New York, NY, USA, 319–320.
10. Soukoreff, R. W., MacKenzie, I. S. 2004. Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. International Journal of Human-Computer Studies 61, 6, 751 –789. Fitts' law 50 years later: applications and contributions from human computer interaction.
11. Silfverberg, M., MacKenzie, I. S., Korhonen, P. 2000. Predicting text entry speed on mobile phones. In Proceedings of the SIGCHI conference on Human factors in computing systems. CHI '00. ACM, New York, NY, USA, 9–16.
12. MacKenzie, I. S Zhang, S. X. 1999. The design and evaluation of a high-performance soft keyboard. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*. CHI '99. ACM, New York, NY, USA, 25–31.
13. MacKenzie, I. S., Soukoreff, R. W. 2002a. A Model of Two- Thumb Text Entry. In Proc. Graphics Interface. 117–124.