Xamobile: Usability Evaluation of Text Input Methods on Mobile Devices for Historical African Languages

Sunkanmi Olaleye and Hussein Suleman

Department of Computer Science, University of Cape Town
Private Bag, Rondebosch, 7701, South Africa
{solaleye,hussein}@cs.uct.ac.za

Abstract. Customized text input editors on mobile devices for languages with no standard language models, such as some African languages, are vital to allow text input tasks to be crowdsourced and thus enable quick and precise participation. We investigated 4 different mobile input techniques for complex language scripts like |Xam and collected accuracy data from experiments with the Xwerty, T9, Pinyin script and hierarchical entry methods for mobile devices and also usability data from the participants. Our results on usability testing show that Xwerty methods offer substantial benefits to the majority of users in terms of speed for |Xam text entry and ease of use.

Keywords: Error correction, Human factors, Input devices, Text entry, Text entry metrics

1 Introduction

The Bleek and Lloyd collection of handwritten notebooks document the language and culture of some Khoi-San people in South Africa. All the pages of this collection have been scanned but are yet to be completely transcribed. It is made up of about 20000 pages of text in the |Xam and !Kun languages. A Web crowdsourcing platform designed for the transcription of the |Xam text requires desktop computer systems and Internet access, which has limited the potential of this tool [1]. However, volunteer workers of this Transcribe Bleek and Lloyd project indicated a preference for a mobile platform for the transcription tasks. This paper presents the outcome of our users' interaction and testing study on how best transcription can be done with low cost mobile devices using four different input methods (Xwerty, T9, Pinyin script and hierarchical) with |Xam text. Our research focuses on a mobile transcription input editor for |Xam called Xamobile. It currently has 4 text entry techniques. It is designed for Android OS mobile devices and tested on small touch screen phones.

2 Keyboard Design

We carried out a usability experiment that evaluates the text entry rate and ease of using the four input methods shown in Fig. 1 (a-d). We tested these input methods using text entry methods implemented on a Samsung GT-S5301.

(a) Xwerty

(b) T9

(c) Pinyin Script

(d) Hierarchical

Fig 1. (a) – (d) Xamobile Input methods

Fig. 2 shows the characteristics of |Xam text used in this experiment. The substring in a red rectangle is made up of a single base character ‘a’ with diacritic marks above and below, and combined base characters ‘nn’ with diacritic marks spanning across them.

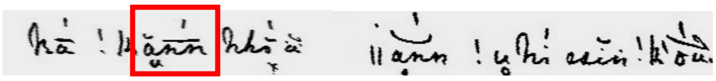
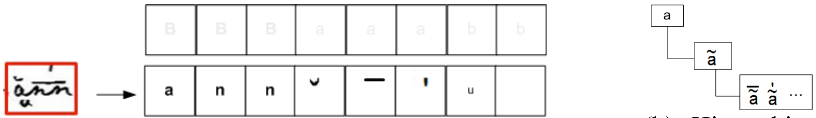


Fig 2. |Xam text

Each of the input methods has some unique feature as an advantage over the other. Xwerty is based on the existing qwerty touchscreen keypad on mobile devices, with popup keypad templates for keys that are polymorphic. T9 uses the 2 key multi-tap feature, with the number of keys reduced compared to Xwerty. It has a word list choice suggestion based on the key pressed [2]. Pinyin script is a qwerty based technique based on Mandarin Chinese text entry techniques but differing in implementation. This does not make use of a keyboard popup for the representation of diacritical based characters, but has a Pinyin Renderer (PR) for the rendering of complex characters with diacritics using the proposed model in Fig 3a. It makes use of the |Xam character model, with 3 columns for base characters, 3 columns for diacritics above and 2 columns for diacritics below. It has 2 frames above the soft keyboard [3].



(a) Pinyin Script Model.

(b) Hierarchical key structure.

Fig 3. Proposed |Xam text entry models

The topmost frame is for PR, where character and diacritics are entered in sequence and PR converts it to suitable |Xam text and displays suggested representations on the second frame below it (Fig 3a). The hierarchical technique reuses its keypad layout for the rendering of the children keys of its immediate parent. It renders the children key text as boldface to highlight them (see Fig 3b). It uses the first keypad layout repeatedly, unlike the Xwerty and T9 popup keypad template. We developed the four input methods for Xam and the text entry evaluation prototype using Java with Eclipse IDE, Android ADT and the Android SDK. The prototype is called Xamobile. Fig 4 shows the entry methods for our evaluation.

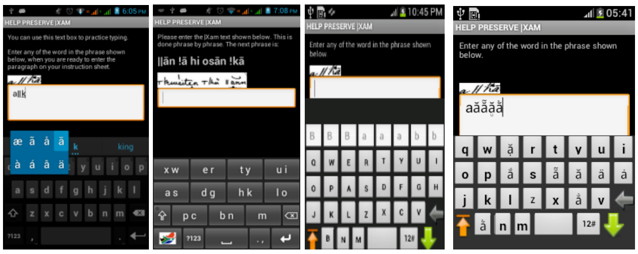


Fig 4. Xamobile text entry methods

3 Evaluation

15 subjects (5 males, 10 females) were recruited to participate in a pilot study. All were recruited from the university campus, aged from 22 to 45, with an average of 1 year of Android QWERTY soft keyboard usage. The criteria for selecting subjects for the study was that they must know how to identity Latin and non-Latin characters from the presented text on the soft keyboard. All participants received their primary education in English.

Subjects were given pre- and post-experiment questionnaires to complete. Pre-questionnaires captured background information while the post-questionnaire captured their rating of the different input methods used in terms of how fast and complex it is for entry of the |Xam phrases on a 5-point Likert scale, where 1 was strongly disagree and 5 was strongly agree. A session consisted of a practice session where participants would familiarize themselves with the chosen text entry method for that session and, after correct entry of the practice presented |Xam phrase, a real session started with 20 different presented |Xam phrases. The experimental design was a within-subjects one-factor analysis of variance. The single factor was 'text entry interface' with four levels: X for Xwerty, T for T9, P for Pinyin script and H for Hierarchical. Table 1 shows the experiment order - T1, T2, T3 and T4, which was governed by a balanced Latin Square. We randomly assigned participants to an experiment order and each participant performed 30 minutes of text entry for each input technique. After each participant completed the experiments using the four input methods in the experiment order assigned, the post-questionnaires were filled and submitted.

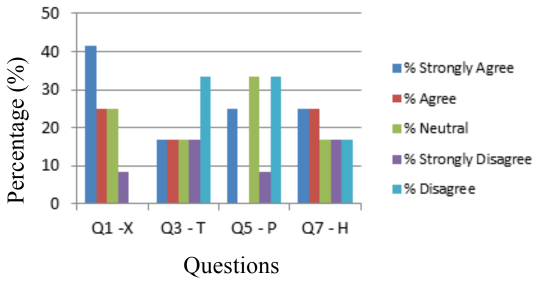
Table 1. Balanced Design Model

T1	X	T	P	H
T2	T	P	H	X
T3	P	H	X	T
T4	H	X	T	P

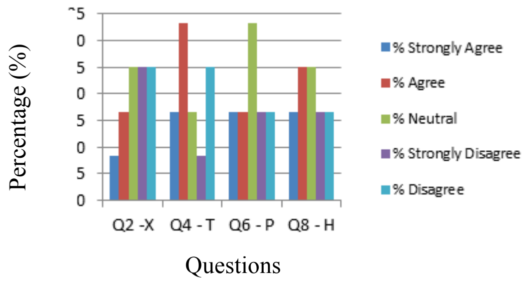
Our experiment results for the highest and lowest typing speeds for the study were 14.88 and 2.92 words per minute (WPM), respectively. An ANOVA analysis established that there was a significant effect of entry techniques on the text input speed ($F_{3,56} = 5.32$, $p < .005$). This shows that Xwerty is the fastest text entry method with the highest word per minute (WPM) for |Xam text. The average WPM for the Xwerty, T9, Pinyin Script and Hierarchical techniques were 7.45 (SD = 1.76), 5.33 (SD = 1.31), 4.43 (SD = 1.18) and 4.85 (SD = 1.42), respectively. The post-survey asked users to comments on speed (Q1 [xwerty], Q3 [t9], Q5 [Pinyin], Q7 [hierarchical]) of text entry. Fig 5a shows the results of the survey regarding the speed of use and Xwerty is clearly the fastest method, followed by Pinyin and then the other methods. Likewise, the ANOVA analysis established that there was a significant difference in the error rates among the entry techniques ($F_{3,52} = 6.66$, $p < .001$). Our results show that hierarchical is the most accurate, with the least error rate for |Xam text. The average MSD for the Xwerty, T9, Pinyin Script and Hierarchical techniques were 17.38 (SD = 13.41), 15.72 (SD = 10.50), 27.68 (SD = 10.62) and 10.80 (SD = 4.52), respectively. There was no significant difference in error rates between Xwerty and T9, but these were significantly more accurate than the Pinyin Script entry technique. Also, our result from the post-survey asked on complexity of use (Q2 [Xwerty], Q4 [T9], Q6 [Pinyin], Q8 [hierarchical]) of text entry as displayed on Fig 5b. It shows that the Pinyin and hierarchical methods are noted as being marginally more complex than the other techniques, with Xwerty being the least complex. One of the comments from our users about Pinyin was that;

"Pinyin was very difficult. It was like old computing 'punch in card system'. What worked well was its ability to make many complex symbols/characters that one could not be easily found in the other systems. If Pinyin could be combined with Xwerty system, it will get better results because it is familiar and it should be able to create complex characters"

This comment further supports both our qualitative and quantitative findings from the experiment and survey in terms of speed of use. Contrarily, results from our usability survey on accuracy of text entry disagree with that of the computation from extracted data for accuracy but the majority of our users after the survey confirmed hierarchical is easy to use, which supports our quantitative findings that the hierarchical technique is the most accurate.



(a) Users comment on speed



(b) Users comment on complexity

Fig 5. Questions and opinion of subjects based on Likert scale

4 Related Work

Digitization of local languages and text processing activities like text translation and transcription are typical applications of crowdsourcing. In recent times, low cost and small devices were used for crowdsourcing. Wismer et al. [4] evaluated input methods using text and image based CAPTCHAs on mobile devices by investigating five different touch and voice input techniques. The outcome of this survey shows that users prefer touch based CAPTCHAs or the voice based CAPTCHAs. Ilinkin and Kim [5] made use of 3 Korean text input methods, namely Chon-ji-in, EZ-Hangul, and SK on mobile for their evaluation. This study is similar in its implementation but differs due to the models of languages used.

Korean is a standard language with known character sets but |Xam is a resource-scarce language.

5 Conclusions

Our research findings show that users prefer the Xwerty text entry technique. Our computed results from device generated data indicate Pinyin Script as the worst of the four tested techniques but contrarily, results from our usability testing indicates Pinyin script was the next to Xwerty in terms of text entry rate and ease of use. Further research is required to distinguish among Pinyin Script, Xwerty-T9 and hierarchical techniques and also to investigate the possibility of adding the Pinyin script model with Xwerty as suggested by a user. Results from device generated data show that there is no significant difference in the speed of entry between these three in terms of speed and accuracy in order to determine the most efficient technique for |Xam text entry. It is obvious with our usability testing that T9 is the worst entry technique for |Xam text due to the lowest percentage in subjects' opinion on speed and the highest percentage in subjects' opinion on the complexity of usage. Ultimately, the fastest and most accurate keyboard will support further research, such as crowdsourcing of transcriptions that require users to type in text in extinct languages such as |Xam on a mobile device.

Acknowledgements. This research was partially funded by the National Research Foundation of South Africa (Grant numbers: 85470 and 83998), University of Cape Town and Telkom SA Ltd. The authors acknowledge that opinions, findings and conclusions or recommendations expressed in this publication are that of the authors, and that the NRF accepts no liability whatsoever in this regard.

References

1. Munyaradzi N., Suleman, H.: Quality Assessment in Crowdsourced Indigenous Language Transcription. Research and Advanced Technology for Digital Libraries. pp. 13-22 (2013)
2. Soukoreff R.W., MacKenzie, I.S.: Metrics for Text Entry Research: An Evaluation of MSD and KSPC, and a New Unified Error Metric. In: ACM Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1-3 (2003)
3. Liu, Y.: Chinese Text Entry with Mobile Devices, Ph.D., University of Tampere (2010)
4. Wismer, A.J., Madathil, K.C., Koikkara, R., Juang, K.A., Greenstein, J.S.: Evaluating the Usability of CAPTCHAs on a Mobile Device with Voice and Touch Input. In: Proceedings of the Human Factors and Ergonomics Society 2012 Annual Meeting. pp. 1228-1232 (2012)
5. Ilinkin, I., Kim, S.: Evaluation of Text Entry Methods for Korean Mobile Phones, a User Study. In CHI(2010), 2023-2026