Electronic Preprint for Journal of Information Processing Vol.25

Regular Paper

Assistive Typing Application for Older Adults Based on Input Stumble Detection

Toshiyuki Hagiya[1,2,a)] Toshiharu Horiuchi[1,b)] Tomonori Yazaki[1,c)] Tsuneo Kato[3,d)] Tatsuya Kawahara[2,e)]

Received: September 6, 2016, Accepted: February 9, 2017

Abstract: Smartphones offer new opportunities to improve the lives of older adults. Although many older adults are interested in smartphones, most of them face difficulties in self-instruction and need support. Text entry, which is essential for various applications, is one of the most difficult operations to master. Therefore, we propose an assistive typing application that detects input stumbles and provides instructions for typing presented sentences, instead of having human tutors help older adults resolve the input stumbles by themselves. First, we investigated the ways that novice older adults have problems with text entry on smartphones. Next, we confirmed the acceptability of being provided with instructions for text entry by Wizard-of-Oz (WoZ). Then, we constructed an assistive typing application based on the collected data from two user studies. An evaluation with novice older adults (60+) showed that the assistive typing application increased typing speed by 17.2% and reduced input stumble incidence by 59.1% compared with the users' initial performance. Improvement rates were almost the same as those achieved with human tutors.

Keywords: text entry, mobile devices, older adults, tutoring system

1. Introduction

Smartphones offer new opportunities to improve the lives of older adults by providing various creative activities and the ability to communicate with a wider circle. Although these individuals would like to learn about smartphones [1], those who have never used a smartphone may face difficulties because of their lack of experience. Some older adults give up using a smartphone and instead revert to their earlier feature phone. Therefore, support in the initial stage is very important. To make full use of the functions of a smartphone, it is essential to master text entry on a touch screen. However, text entry is one of the operations that novice older adults find most difficult because considerable background is needed, such as knowing which key corresponds to a character and how to select a suggestion. In addition, touch interfaces lack both the mechanical stability and tactile feedback of a keyboard, making it harder for users to accurately select targets [2].

Many researchers have tackled the issues of making text entry easier using a number of approaches, e.g., changing the layout, adjusting the key target areas to suit users, and presenting suitable suggestions [3], [4], [5]. These aids are effective for users who are accustomed to the way a smartphone operates. However, in addition to the problems mentioned above, smartphone novices, especially older adults, tend to have cognitive problems, such as forgetting and losing how/what to type next [6]. According to a prior work by Leung et al. [7], older adults tend to prefer an instruction manual to trial-and-error. They therefore need patient assistance during the initial stages of use. An ideal approach to mastery is to receive assistance from a human tutor whenever required, but this is not always possible.

Accordingly, we have designed a tutoring system that can perform the role of a human tutor who indicates the next action. Although our final goal is to provide a tutoring system for typing free text, we propose an assistive typing application, that is, a tutoring system for entering presented sentences in order to verify the acceptability and the effectiveness of instructions by the system in text entry. The assistive typing application automatically detects input stumbles and provides instructions [8]. We define the term "input stumble" as an occasion when a user makes a mistake or forgets and loses the knowledge of how/what to type next. The targeted user group is older adults who have never used a smartphone but who have owned a feature phone, because most older adults in developed countries own a mobile nowadays.

In this paper, we present the four steps we used to construct an assistive typing application. First, we investigated the ways that novice older adults have problems with text entry on smartphones. Second, we confirmed the acceptability of being provided with instructions for text entry by Wizard-of-Oz (WoZ) [9]. Next, we constructed an assistive typing application, which automatically detects input stumbles and provides instructions on the typing of presented sentences. This was built on the basis of the collected data, including operational stumbles and effective instructions found in the previous two studies. Finally, the assistive

1 KDDI Research, Inc., Fujimino, Saitama 356–8502, Japan
2 Kyoto University, Kyoto 606–8501, Japan
3 Doshisha University, Kyotanabe, Kyoto 610–0394, Japan
a) to-hagiya@kddi-research.jp
b) to-horiuchi@kddi-research.jp
c) yazaki@kddi-research.jp
d) tsukato@mail.doshisha.ac.jp
e) kawahara@i.kyoto-u.ac.jp

typing application was evaluated by comparing its performance with human tutors.

2. Related Work

2.1 Interface Design for Older Adults

Many studies have attributed older adults' difficulties in learning to use technology to a number of user characteristics, such as declines in perceptual performance and lack of relevant technology experience. Docampo et al. [10] noted that older adults have difficulty learning current user interfaces because they generally have less experience with current devices than younger adults, and need to learn different types of user interfaces from former technologies. In addition, user interfaces on mobile devices often have buttons that perform different context-dependent tasks.

To deal with such problems, Fisk et al. [11] provided a list of guidelines for older adults. These guidelines emphasize interface changes that respond to age-related changes in perception and cognition: larger displays, fonts, and buttons, and accessibility improvements in specific contexts. With respect to the touch interface, Jin et al. [12] have investigated optimal target size, spacing and position to derive recommendations and general guidelines for older adults. While these guidelines are very useful, needs can differ among individuals [13]. To respond to older adults' various needs, researchers have considered adaptive solutions based on user profiles such as user behaviors [14], knowledge [15], and skills [16].

2.2 Text Entry for Older Adults

Focusing on text entry, some studies proposed methods for transforming detection areas based on language models and users' touch distribution [3], [4]. For example, Rodrigues et al. [4] analyzed the way to influence the typing behavior of older adults by varying the keyboard, including the color and the width of keys. Gunawardana et al. [3] proposed methods for varying the detection areas based on typing history in the language model. As a different approach, Bi et al. [5] optimized an algorithm for presenting suitable suggestions to make correction and completion easier. However, Kurniawan [17] reported that older adults usually dislike text-prediction features. Therefore, Komninos et al. [18] proposed a keyboard that makes users aware of any errors by highlighting text in the body of the message and using a color bar at the top of the keyboard.

Also, various typing applications have been proposed in the smartphone market. In most typing apps, users are prompted with text showing what they are required to type. When users mistype a character, they are typically notified through auditory and visual feedback, such as beeps and squiggly underlines. [19]. This is effective for highly motivated users with a certain degree of knowledge and skill. However, at least one report [7] indicates that older adults tend to prefer an instruction manual to trial-and-error. Nicolau et al. [6] reported detailed analyses of how older adults learn text entry, and found their most common errors were due to cognitive problems in the initial stage. We focused on support to mitigate cognitive problems rather than addressing physical aspects such as mistyping.

2.3 Tutoring System for Older Adults

Effective ways to help older adults have been well researched. However, that wealth of research has largely focused on designing better instructional resources for learning to use desktop computers. For example, Hichman et al. [20] studied the type of guidance most suitable for older adults. Morrell et al. [21] have studied the quantity of guidance that was most suitable. Rogers et al. [22] investigated the kind of resources most useful in the learning process, and found that step-by-step interactive tutorials were the most effective approach in the learning process for older adults. With respect to using smartphones, Leung et al. [7] surveyed and investigated how older adults learned. According to their report, older adults tend to prefer an instruction manual to trial-and-error. Kelleher et al. [23] proposed stencil-based tutorials that overlay step-by-step instructions on the screen.

Although tutoring systems have been studied in other domains, no tutoring system for text entry has been studied. We considered that providing instructions for the next action is an effective way for novice older adults to learn text entry on smartphones.

3. Preliminary User Study

First, we needed to know the characteristic actions of novice older adults when making text entries on smartphones. Therefore, in this section, we describe a user study to investigate the problems they encounter.

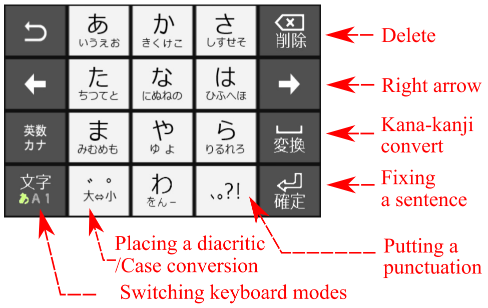
3.1 Participants

Thirty participants, fifteen males and fifteen females between the ages of 60 and 83 (mean 72.1, sd = 8.2), took part in the user study. They were recruited from a local social institution. None of them had any previous experience using a smartphone, but all had owned a feature phone for more than one year, i.e., one with a physical keyboard with a 12-key layout. They were familiar with this kind of keyboard because the same layout has been used in feature phones for a long time. Twenty-six of them had entered text with a feature phone before, while the others had never done so, using the phone only for calling. Ten had used their own PC in the usual ways. None of them had a tremor disorder, eye problems or other health problems.

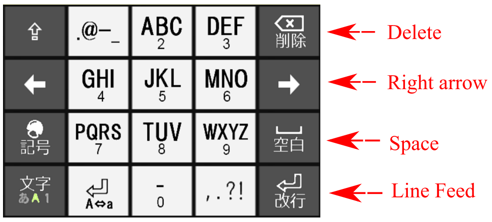
3.2 Apparatus

A Samsung Galaxy S3 running Android 4.1.2 with a 4.8-inch screen having a resolution of 1,280 × 720 pixels (306 ppi) was used in the experiment. The software keyboard was designed for multi-tap input with a 12-key layout for Japanese as shown in **Fig. 1**. The keyboard recorded all touch events using the standard Android API and all linguistic information, e.g., typed keys, displayed characters, and suggestions. During use, an overhead video camera recorded participants' operations as well.

The 12-key layout is common in Japanese smartphones. Kana, that is Japanese syllables, are combinations of a consonant and a vowel as shown in **Table 1**. In this keyboard, a key corresponds to a consonant, and the number of presses corresponds to a vowel. For example, the kana "ku (く)" is input by typing the key of "ka (か)" three times. When typing two characters with the same consonant consecutively, the first character must be fixed by pressing



(a) Kana keyboard mode



(b) Alphabetic keyboard mode

Fig. 1 Key layout.

Table 1 Japanese syllabic alphabet represented by Roman alphabet letters.

	A	I	U	E	O
	a (あ)	i (い)	u (う)	e (え)	o (お)
K	ka (か)	ki (き)	ku (く)	ke (け)	ko (こ)
S	sa (さ)	si (し)	su (す)	se (せ)	so (そ)
T	ta (た)	ti (ち)	tu (つ)	te (て)	to (と)
N	na (な)	ni (に)	nu (ぬ)	ne (ね)	no (の)
H	ha (は)	hi (ひ)	hu (ふ)	he (へ)	ho (ほ)
M	ma (ま)	mi (み)	mu (む)	me (め)	mo (も)
Y	ya (や)	-	yu (ゆ)	-	yo (よ)
R	ra (ら)	ri (り)	ru (る)	re (れ)	ro (ろ)
W	wa (わ)	-	-	-	wo (を)
	nn (ん)	-	-	-	-

the right arrow key before typing the second character. Diacritics, punctuation marks, and other symbols are added by other keys. Kana-kanji conversion is performed with the convert key or by following touching suggestions. A character string being edited is not finalized until either a kana-kanji conversion is complete or the fix key is typed.

3.3 Procedure

First, the participants were given three minutes of explanations and examples of how to operate a smartphone, including touch and swipe operations, by a human tutor. Then, they tried to operate the smartphone for one minute. After that, they were briefly instructed on how to use the software keyboard, including the correspondence between each key and a character and how to correct text or select suggestions by watching a video tutorial on the smartphone. Afterwards, they typed twelve sentences in Japanese by using a typing application as shown in **Fig. 2**. The sentences were selected from an email corpus gathered originally from personal conversations via email, such as greetings and appointments to meet. The corpus contained roughly 30,000 sentences (average character length = 23.6). The experimental period was limited to 75 minutes. Participants had a one-minute rest after each sentence. After the experiment, they took part in an interview.

They operated the smartphone while holding it in their hand or resting it on a desk while sitting on a chair. A human tutor sat next to them. They were instructed to type by themselves if at all possible, with no help from the human tutor. However, they were permitted to ask the human tutor when they lacked the confidence on what to do next. They received operating instructions on request and when the human tutor judged that they needed assistance.

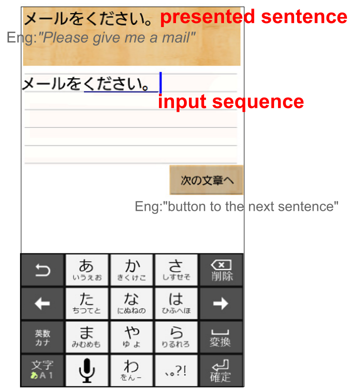


Fig. 2 Screenshot of a typing application.

Table 2 Actual input stumbles, with two frequency metrics.

The operation in which an input stumble has occurred	ISR with respect to chances (%)	ISR with respect to participants (%)
(1) How to type characters allocated to the same key consecutively	27.1	86.7
(2) How to fix a sentence	17.3	63.3
(3) How to select a suggestion	15.2	90.0
(4) How to switch keyboard modes	15.0	50.0
(5) How to insert line feeds	13.0	53.3
(6) How to delete previous character	12.7	26.7
(7) How to convert to a small letter	9.4	20.0
(8) How to enter a symbol	5.0	40.0
(9) How to enter a diacritic	4.6	43.3
(10) How to enter punctuation	4.3	20.0
Average	12.4	49.3

3.4 Observation Results

Twenty-eight participants completed the task. The average completion time was 43.4 minutes (sd = 12.0). The other two participants did not complete the task within 75 minutes. These were the two who had no experience entering text with their own feature phone. A total of 355 sentences were typed, with an average character length of 26.1.

Three annotators independently extracted the pattern of input stumbles from the logs and the videos of the study to maintain inter-rater reliability, and classified the input stumbles into 10 categories based on discussion. **Table 2** shows the input stumbles as categorized. Next, the three annotators labeled input stumbles and possible stumble opportunities according to the logs of each input. We adopted the labels that were applied by more than two annotators. The concordance rate using Fleiss' kappa [24] was 0.82. We classified every category in terms of the two ratios in Table 2. One is the input stumble ratio (ISR) with respect to chances, which is the percentage of stumbles compared with the possible stumbles in each category. The other is the ISR with respect to participants, which is the percentage of participants who stumbled in the category at least once.

These ISRs were higher than average in both metrics: (1) How to type characters allocated to the same key consecutively, (2) How to fix a sentence, (3) How to select a suggestion, (4) How to switch keyboard modes, and (5) How to insert a linefeed. These operations have strong cognitive aspects. The input stumbles on operations (1) and (2) were typical stumbles of omission due to cognitive lapses. The common point is that no character or string changed when these keys were typed. Many participants seemed not to understand that they needed to type these keys, until the human tutor provided instructions several times. In terms of operation (3), many participants found it difficult to judge the timing for selecting the desired suggestion. They focused so much on typing that they forgot to select the target suggestion until they had typed a long sentence. When typing a long sentence, they tended to mistype or they had to decide where the sentence was divided. Therefore, it was useful to recommend selecting the suggestion with each short sentence. Input stumbles of operation (4) occurred frequently because the participants had no previous experience of such an operation on their own feature phone. Operation (5) is confusing because operations (2) and (5) are assigned to the same key and its response changes with the context. Although the concept is the same as that on the keyboard of a feature phone, some participants did not understand the changeable response.

In contrast, there were few stumbles associated with operations (6)–(10) because participants only had to check the final character and the next key position. Once they had received instructions, most of them were able to perform these operations.

In this experiment, we confirmed that older adults needed more repetitive support for complex operations, such as (1)–(5). In addition, we found that instructions that included the purpose of an operation were more effective than instructions that simply said what to do next. However, only a few repetitions of the instructions were required for the simple operations ((6)–(10)), such as the operation to point to the next position.

4. User Study with Wizard-of-Oz

We hypothesized that a tutoring system for text input with a smartphone would be useful for older adults. To confirm the acceptability of being provided with instructions, we first conducted a user study using Wizard-of-Oz (WoZ) [9]. In this study, the tutoring system of WoZ provided instructions when a human tutor judged that participants needed instructions.

4.1 Participants

Five participants took part in the user study, three males and two females ranging in age from 65 to 77 years (mean age 73.2, sd = 5.5) who were recruited from a local social institution and did not participate in the preliminary experiments in this study. They had no previous experience with a smartphone or tablet, but had owned a feature phone with a 12-key physical keyboard for more than one year and had sent emails more than twice a week. None of them had a tremor disorder, eye problems or other relevant health problems.

Fig. 3 Screen shot of a WoZ instruction.

4.2 Apparatus

A Samsung Galaxy S3 running Android 4.1.2 and with a 12-key multi-tap layout was used in this experiment, as in the first study. The keyboard recorded all touch events using the standard Android API and all linguistic information e.g., typed keys, displayed characters and suggestions.

The WoZ system included a video camera, a monitor, and a PC that was connected to the smartphone using Bluetooth. All participant operations were recorded by the video camera and observed by a human tutor via the monitor. The human tutor was able to send an instruction by selecting from a list of instructions in a PC application. The list contained ten instructions appropriate for the input stumbles observed in the first study. The instruction was provided to the participant by a combination of text instruction, key highlighting overlaid on the key, and voice instruction that matched the wording of the text instruction. The highlighted key had a yellow rectangle overlaid on the key. The voice instruction was produced by Text-to-Speech. **Figure 3** shows one of the instructions responding to input stumble (2): “How to fix a sentence.” The text instruction and voice instruction say *“Please fix the sentence by typing bottom-right key,”* with the “fixing” key highlighted.

4.3 Procedure

First, the participants were given explanations and examples of how to operate a smartphone, including touch and swipe operations, by a human tutor. Next, they were briefly instructed on how to use the software keyboard, including the correspondence between each key and a character, and how to correct input text or select suggestions. Then, they typed ten sentences in Japanese by using the same typing application as in the previous study. They operated the smartphone while holding it in their hand or resting it on a desk while sitting on a chair. When the human tutor observing through the monitor judged that the participant needed help with an operation in the instruction list, the human tutor sent the instruction to the smartphone used by the participant. Participants could rest for one minute after writing each sentence. The experimental period was limited to 60 minutes. After the experiment, they filled out a 5-point Likert-scale questionnaire [25] in which five indicated strong agreement, and took part in an interview.

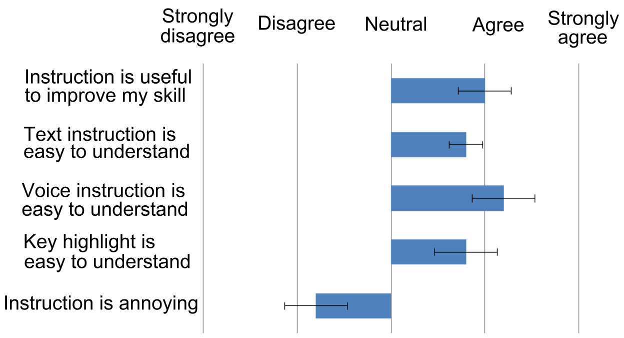


Fig. 4 Questionnaire responses regarding WoZ study.

4.4 Results of WoZ study

All participants completed the task. The average completion time was 40.2 minutes (sd = 12.0). **Figure 4** shows the results of the subjective questionnaires for the tutoring system using WoZ. There was a positive response to the item *"Instruction is useful to improve my skill."* The instruction methods were also positively rated, particularly voice instruction. These results showed that the tutoring system was generally supported by the participants.

On the other hand, the following negative opinions were expressed;

> *"Some instructions were provided even when I didn't make a mistake. Incorrect instructions induce a feeling of anxiety."*

This opinion indicated that precision is more important than recall in designing an automated tutoring system for text entry.

In addition, two participants made the following comment in regard to the manner of instruction;

> *"First, I didn't realize that I could press the overlaid yellow key before the highlight disappeared. It's confusing to me."*

This study showed that the tutoring system was generally supported by the participants. However, we confirmed that the way of instruction should be improved.

5. Assistive Typing Application

To support self-instruction for learning text entry techniques on smartphones, we propose an assistive typing application, that is, a tutoring system for entering presented sentences. The application has two functions: detecting input stumbles and providing instructions. On the basis of the recorded data and instructions from the previous two studies, an input stumble detection model and instruction controls were constructed. **Figure 5** shows the architecture of the application. First, the application observes the input data, such as the typed key and suggestions. Then, the application detects input stumbles based on the input stumbles detection model that was constructed using data from the earlier study. Lastly, instructions for the next action are provided simultaneously by voice instruction, text instruction, and key highlighting based on the instruction model.

5.1 Input Stumble Detection Model

The input stumble detection model included two steps: stumble detection and stumble classification.

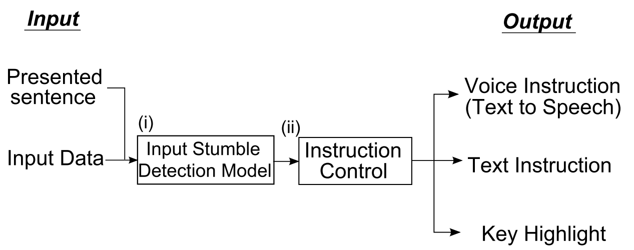


Fig. 5 Architecture of the assistive typing application.

A stumble was detected when the user stopped typing for longer than a certain threshold time ϵ, defined as $\mu_t + 2\sigma_t$. The variables μ_t and σ_t are the mean and the standard deviation of the touch intervals of the thirty participants in the preliminary experiment. 97.8% of the touch intervals labeled input stumbles exceeded the threshold time. The stumble classifier used with the machine learning technique classifies input data into 11 categories: the 10 input stumble categories in Table 2, and a category of no input stumble. To select a proper learning machine, we compared the performances of three models, a support vector machine (SVM) [26], a C4.5 [27] and a Naive Bayes (NB) [28]. Each model was trained by using data labeled 11 categories from the first experiment in Section 3. A total of 44 features was obtained on the basis of input sequences including unfixed string and fixed string, presented sentence, and suggestions. The features were selected from three points of view. One was characteristics of the input string itself, such as the number of phrases in an unfixed sentence and the category ID of the morpheme of the last phrase. Another was the keyboard mode. The third considered whether the partial input string or the suggestion was consistent with the presented sentence. Each model adopted some of the feature amounts so that the accuracy rate by F measure was the highest in the evaluation of each participant in a 30-fold cross validation.

Table 3 shows the performance of each model. The F measures were 0.941 with SVM, 0.942 with C4.5, and 0.938 with NB. Anova (significance level $\alpha = 0.05$) showed no significant differences between the three ($F(2, 58) = 0.74$), so we adopted C4.5, with the highest accuracy rate. The features adopted by C4.5 were the ten shown in **Table 4**. Two features directly related to the input string referred to the type of the last character in an unfixed string and the length of the unfixed string. Two features were relevant to the keyboard mode: the current type of keyboard and the keyboard to be used next. There were six either-or features: whether the fixed string was consistent with the presented sentence, whether the unfixed string was consistent with the beginning of the rest of the presented sentence, whether the word to be input next was in the suggestion list, whether the last character in the unfixed string should include a diacritic, such as a voiced sound, whether the last character in the unfixed string should be converted to a small letter, and whether the character to be input next was allocated to the same key, such as "kaki (かき)."

5.2 Instruction Control

Instruction control, which provides an instruction corresponding to each type of input stumble, was provided manually based on the effective advice observed from human tutors. Concretely,

Table 3 The classification performance of 10 input stumble categories of each model.

Classification model	F-measure (Precision [%])
SVM	0.941 (98.2)
C4.5	0.942 (98.2)
NB	0.938 (98.0)

Table 4 Effective features to classify input stumbles with C4.5.

ID	The feature amounts
1	The type of the last character in U
2	The length of U
3	The type of the current keyboard
4	The type of the keyboard to be used next
5	The binary whether F is consistent with P
6	The binary whether U is consistent with the beginning of the rest of P
7	The binary whether the word to be input next is in the suggestion list
8	The binary whether the last character in U should include a diacritic
9	The binary whether the last character in U should be converted to a small letter
10	The binary whether the character to be input next is allocated to the same key

F: fixed string, U: unfixed string, P: presented sentence

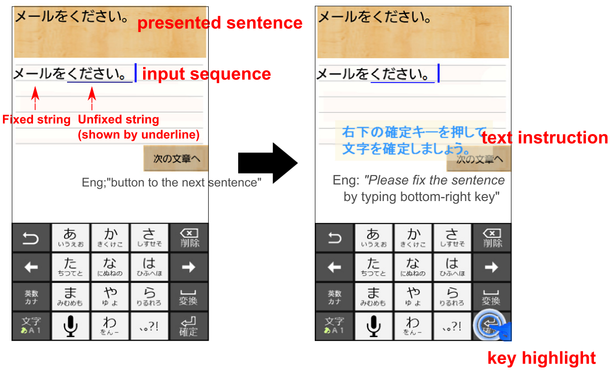


Fig. 6 Screenshot showing the provision of instruction.

the instruction model provides the procedure and the purpose for operations (1)–(5), while providing only the procedure for (6)–(10). For instance, the instruction in operation (1) is *"Please press the right arrow key to move cursor for typing characters allocated to the same key consecutively"* and that in operation (6) is *"The delete key is upper right."* The instruction model produces instructions in the form of voice, text, and a highlighted key with an overlaid finger animation. The procedure for indicating the key position was improved based on the results of the other preliminary experiment because participants in Section 4 pointed out that simply indicating the key with a yellow overlay was confusing.

Finally, the assistive typing application was implemented on the smartphones. **Figure 6** shows that the assistive typing application detects input stumbles of punctuation and provides instructions, such as *"Please fix the sentence by typing bottom-right key"* which is the instruction given in response to input stumble (2) "How to fix a sentence," with the text instruction and the voice instruction given at the same time as the key is indicated by the overlaid finger animation.

6. Evaluation of Assistive Typing Application

In this section, we evaluate the performance of the assistive typing application by comparing it with that of a human tutor.

6.1 Participants

Twenty-four participants, twelve males and twelve females between the ages of 60 and 82 (mean 70.2, sd = 6.5) who did not participate in preliminary experiments in this study, took part in this experiment. They were recruited from a local social institution. None of them had previous experience with a smartphone, but all had owned a feature phone for more than one year. Nineteen of them had entered text using a feature phone before, while the others had never done so, using their feature phones only for calling. Nine of them had used their own PC in the usual ways. None of them had a tremor disorder, eye problems or other health problems.

6.2 Apparatus and Procedure

The tutorial in the experiment on how to use the smartphone and the software keyboard was the same as in the previous experiment. After the tutorial, the participants typed twenty sentences in Japanese (average character length of 16.5) using the typing application within a 75-minute time limit.

When the input sequence matched the presented sentence perfectly, the next sentence was presented. The participants were divided into three groups. In group A, twelve participants used the proposed assistive typing application. In group B, six participants typed the sentences with support from a human tutor, who gave instructions based on the same procedure as the assistive typing application, but only when the participants did not type for a set period. In group C, the other six participants typed the sentences with support from a human tutor who gave instructions without any control. We did not make any comparison with a group typing without instructional support because it was clearly too difficult for them to complete the task without any instruction. After finishing the task, the participants in group A answered a questionnaire about their impressions and the effectiveness of the assistive typing application.

6.3 Experimental Results

All participants completed the task. As in the study in Section 3, three annotators labeled input stumbles and possible stumble opportunities according to the logs of each input. We adopted the labels that were applied by more than two annotators. The concordance rate using Fleiss' kappa was 0.83. We calculated the input stumble ratio (ISR) with respect to chances, i.e., the percentage of possible stumbles that actually occurred.

We evaluated performance by three metrics: the typing speed, the input stumble ratio (ISR) with respect to chances, and the subjective questionnaire.

6.3.1 Typing Speed

We compared typing speeds in characters per minute (CPM) [29] not including the time used by instructions. **Figure 7** shows the measured CPM of all groups for the first five sentences, the middle five (from the 8th to the 12th), and the

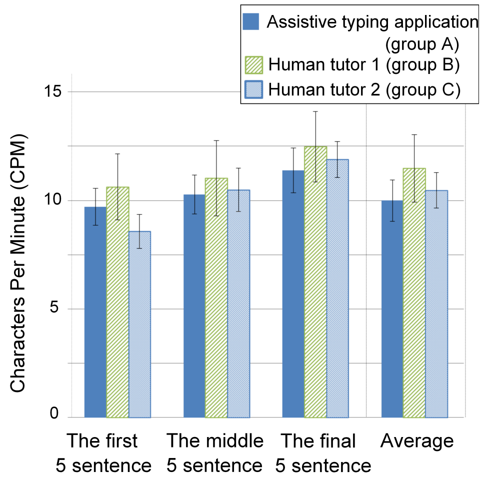


Fig. 7 User progress in CPM for each group.

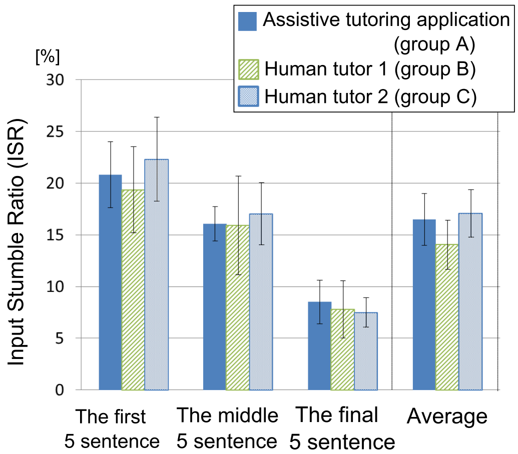


Fig. 8 User progress in the input stumble ratio with respect to the chances for each group.

last five, and the overall average. All participants completed the tasks. The CPM improved over the course of the sentences in all groups. The Anova (significance level $\alpha = 0.05$) did not show any significant differences between three groups in any stage; $F(2, 21)$ was 0.44 in the first five, 0.12 in the middle five, 0.96 in the last five and 1.14 in the overall average. However, focusing on the improvement rate as measured by CPM, the improvement rate of group C was greater than the others: the rates were 17.2% in group A, 17.5% in group B and 35.5% in group C.

6.3.2 Input Stumble Ratio

The average number of input stumbles per participant was 22.1. We compared the input stumble ratios (ISR) with respect to chances. **Figure 8** shows the ISRs of all groups at each step. The ISR for all groups decreased with practice. Anova (significance level $\alpha = 0.05$) showed no significant differences between the three methods in any stage; $F(2, 21)$ was 1.55 in the first five, 1.05 in the middle five, 0.31 in the last five and 1.43 in the overall average. The improvement rate of this measure did not show a noticeable difference from that of the ISR; the rates were 59.1% in group A, 59.8% in group B and 66.4% in group C.

6.3.3 Questionnaire and Interview

Figure 9 shows the results of the subjective questionnaires on the assistive typing application. It shows that the assistive typing application was generally well accepted. The instruction methods

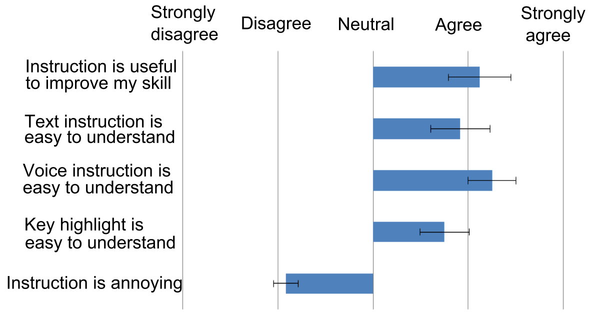


Fig. 9 Questionnaire responses regarding the assistive typing application.

were also positively rated, with voice instruction the same as in the WoZ study in Section 4.

Participants expressed the following positive opinions:

"Multimodal instructions were easy to understand."

"Instructions were helpful to learn smartphone."

On the other hand, there were some negative opinions on the instruction,

"An instruction was provided once when I didn't need it. The instruction made me confused."

The precision for detecting input stumbles was 97.8%. Four participants were provided incorrect instructions at least once. On the other hand, one interviewed participant responded:

"I didn't mind the misdirection because the instruction didn't interrupt my operation."

The comments suggest that it is essential to provide instructions in a way that does not interrupt the user's operations.

7. Discussion

The discussion first considers the results of the evaluation experiment. Anova for CPM and ISR showed no significant differences among the three groups. Additionally, the improvement rates of participants using the assistive typing application were almost the same as those with human tutors. This shows that the assistive typing tutor works effectively and prompts the user's operations well. In an interview, there were similar comments:

"First, I didn't know operations at all. But I could begin to operate gradually as I understood the difference between my feature phone and the smartphone. These kinds of instructions were useful for me."

This comment indicates that the assistive typing application is effective for users who use their feature phones to email frequently but have trouble with self-instruction.

From a different perspective, although Anova did not show any significant difference, the CPM and ISR of participants using the assistive typing application were lower than those of participants given instructions without any control. Participants made not only single stumbles but also complicated stumbles that included multiple components. This was probably because they took more time to resolve these complicated stumbles with the simple operations. The provision of step-by-step instructions for a complicated task is currently difficult to achieve. This difficulty is due to the requirement for a detailed understanding of what stumbles are currently occurring and what the user wants to do next. These challenges form the basis of one of our future research tasks.

On the other hand, a participant made a negative comment in the interview:

> *"Instructions were provided sometimes. But I didn't need the instructions because I preferred trial-and-error."*

He was one of the most skillful participants from the beginning. His CPMs were 13.2 in the first 5 sentences and 19.5 in the final 5 sentences. This comment indicates that the assistive typing application is not needed for users who prefer trial-and-error and have high proficiency. Therefore, providing instructions that consider skill level is essential. Skill estimation is one of our future tasks.

Finally, we note our final goal, which is providing support for free text input. Some participants commented:

> *"I want to learn typing in free text input practically rather than in the typing application."*

It is difficult to estimate what may be typed next with free text input, unlike with the presented sentences in this study. Therefore, we need to further develop the input stumble detection algorithm.

8. Conclusion

We proposed a novel tutoring system, that is, an assistive typing application for text entry that detected input stumbles and provided instructions to help users to resolve stumbles on their own, instead of receiving instructions from a human tutor. First, we collected data including operational input stumbles and effective instructions through a user study that included thirty participants. Next, we confirmed the acceptability being provided with instructions for text entry using Wizard-of-Oz (WoZ). Then, a prototype of the assistive typing application was created based on the previous two studies. An evaluation experiment with novice older adults (60+) showed that the assistive typing application increased characters per minute (CPM) by 17.2% and reduced the stumble ratio with respect to chances for stumbles by 59.1% compared with users' initial rate. Improvement rates were almost the same as those achieved with human tutors. The subjective assessment and the interview showed that the assistive typing application was generally-supported.

References

[1] Beisge, B. and Kraitchman, M.: *Senior Centers: Opportunities For Successful Aging*, Springer Publishing Company (2003).

[2] Hoggan, E., Brewster, S.A. and Johnston, J.: Investigating the effectiveness of tactile feedback for mobile touchscreens, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, *CHI '08*, pp.1573–1582, ACM (2008).

[3] Gunawardana, A., Paek, T. and Meek, C.: Usability Guided Key-Target Resizing for Soft Keyboards, *Proc. 15th International Conference on Intelligent User Interfaces*, *IUI '10*, pp.111–118, ACM (2010).

[4] Rodrigues, É., Carreira, M. and Goncalves, D.: Enhancing typing performance of older adults on tablets, *Universal Access in the Information Society*, pp.1–26 (2014).

[5] Bi, X., Ouyang, T. and Zhai, S.: Both Complete and Correct? Multi-Objective Optimization of Touchscreen Keyboard, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, *CHI '14*, pp.2297–2306, ACM (2014).

[6] Nicolau, H. and Jorge, J.: Elderly text-entry performance on touchscreens, *Proc. 14th International ACM SIGACCESS Conference on Computers and Accessibility*, *ACCESS '12*, pp.127–134, ACM (2012).

[7] Leung, R., Tang, C., Haddad, S., Mcgrenere, J., Graf, P. and Ingriany, V.: How older adults learn to use mobile devices: Survey and field investigations, *ACM Trans. Accessible Computing*, Vol.4, No.3, pp.11:1–11:33 (2012).

[8] Hagiya, T., Horiuchi, T., Yazaki, T. and Kato, T.: Typing Tutor: Automatic Error Detection and Instruction in Text Entry for Elderly People, *Proc. 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, *MobileHCI '15*, pp.696–703, ACM (2015).

[9] Kelley, J.F.: CAL - A Natural Language program developed with the OZ Paradigm: Implications for Supercomputing Systems, *Proc. 1st International Conference on Supercomputing Systems*, pp.238–248 (1985).

[10] Rama, M.D., Ridder, H. and Bouma, H.: Technology generation and Age in using layered user interfaces, *Journals of Gerontechnology*, Vol.1, No.1, pp.25–40 (2001).

[11] Fisk, A.D., Rogers, W.A., Charness, N., Czaja, S.J. and Sharit, J.: *Designing for Older Adults: Principles and Creative Human Factors Approaches*, CRC Press (2009).

[12] Jin, Z.X., Plocher, T. and Kiff, L.: Touch screen user interfaces for older adults: Button size and spacing, *Universal Access in HCI, Part I*, *HCII2007*, *LNCS4554*, pp.933–941 (2007).

[13] Gregor, P., Newell, A.F. and Zajicek, M.: Designing for dynamic diversity: Interfaces for older people, *Proc. 5th International ACM Conference on Assistive Technologies*, pp.151–156, ACM (2002).

[14] Sloan, D., Atkinson, M.T., Machin, C. and Li, Y.: The Potential of Adaptive Interfaces as an Accessibility Aid for Older Web Users, *Proc. 2010 International Cross Disciplinary Conference on Web Accessibility*, Article 35, ACM (2010).

[15] Wobbrockand, J.O., Kane, S.K., Gajos, K.Z., Harada, S. and Froehlich, J.: Ability-Based Design: Concept, Principles and Examples, *ACM Trans. Accessible Computing*, Vol.3, No.3 (2011).

[16] Edlin-White, R., Cobb, S., D'Cruz, M., Floyde, A., Lewthwaite, S. and Riedel, J.: Accessibility for Older Users through Adaptive Interfaces: Opportunities, Challenges and Achievements, *Human-Computer Interaction, Part III*, *HCII2011*, *LNCS6763*, pp.483–489 (2011).

[17] Kurniawan, S.: Older people and mobile phones: A multimethod investigation, *International Journal of Human Computer Studies*, Vol.66, No.12, pp.889–901 (2008).

[18] Komninos, A., Nicol, E. and Dunlop, M.D.: Designed with Older Adults to Support Better Error Correction in SmartPhone Text Entry: The MaxieKeyboard, *Proc. 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*, *MobileHCI '15*, pp.797–802, ACM (2015).

[19] Rudchenko, D., Paek, T. and Badger, E.: Text Text Revolution: A Game That Improves Text Entry on Mobile Touchscreen Keyboards, *Proc. 9th International Conference on Pervasive Computing*, *Pervasive '11*, pp.206–213, ACM (2011).

[20] Hickman, J.M., Rogers, W.A. and Fisk, A.D.: Training older adults to use new technology, *Journals of Gerontology: Series B*, Vol.62, No.1, pp.77–84 (2007).

[21] Morrell, R.W., Park, D.C., Mayhorn, C.B. and Kelley, C.L.: Effects of age and instructions on teaching older adults to use eldercomm, an electronic bulletin board system, *Educational Gerontology*, Vol.26, No.3, pp.221–235 (2000).

[22] Rogers, W.A., Cabrera, E.F., Walker, N., Gilbert, D.K. and Fisk, A.D.: A survey of automatic teller machine usage across the adult lifespan, *Human Factors*, Vol.38, No.1, pp.156–166 (1996).

[23] Kelleher, C. and Pausch, R.: Stencils-based tutorials: Design and evaluation, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, *CHI '05*, pp.541–550, ACM (2005).

[24] Fleiss, J.L.: Measuring nominal scale agreement among many raters, *Psychological Bulletin*, Vol.76, No.5, pp.378–382 (1971).

[25] Likert, R.: A technique for the measurement of attitudes, *Archives of Psychology*, Vol.22, No.140 (1932).

[26] Cortes, C. and Vapnik, V.: Support-Vector Networks, *Machine Learning*, Vol.20, No.3, pp.273–297 (1995).

[27] Quinlan, J.R.: *Programs for Machine Learning*, Morgan Kaufmann Publisher (1993).

[28] Cortes, C. and Vapnik, V.: On the optimality of the simple Bayesian classifier under zero-one loss, *Machine Learning*, Vol.29, No.2-3, pp.103–130 (1997).

[29] Arif, A.S. and Stuerzlinger, W.: Analysis of Text Entry Performance Metrics, *Proc. International Conference on Science and Technology for Humanity*, *TIC-STH '09*, pp.100–105, IEEE (2009).

Toshiyuki Hagiya received his B.E. and M.E. degrees from Kyoto University in 2008 and 2010. He is currently with KDDI Research, Inc. He has been engaged in research and development of spoken dialogue systems and interactive user interface. He is a member of IPSJ, ACM.

Toshiharu Horiuchi received his Bachelor's, Master's and Ph.D. degrees from Nagaoka University of Technology in 1999, 2001 and 2004. From 2002 to 2006, he was a researcher at the Spoken Language Communication Research Laboratories, Advanced Telecommunications Research Institute International (ATR). From 2006 to 2009, he was a research engineer at KDDI R&D Laboratories, Inc. Since 2009, having joined the KDDI Corporation, he has worked as a R&D manager at KDDI Research, Inc. His research interests include acoustic signal processing and adaptive signal processing, and in particular spatial signal processing. He received the 17th Telecommunications Advancement Foundation Award in 2002, the 20th Image Information and Media Future Award in 2012, and the Best Technical Demonstration Runner-up at ACM Multimedia 2012. He is a member of ASJ, IEEE and IEICE.

Tomonori Yazaki is a senior manager in User Interface Laboratory at KDDI Research, Inc. He has been engaged in the research and development of Human Computer Interaction system/application. He received his B.E. and M.E. degrees from The University of Electro-Communications in 1996 and 1998, respectively. He is a member of IEICE, JSSD and JSKE.

Tsuneo Kato received his B.E., M.E. degrees and Ph.D. from The University of Tokyo in 1994, 1996, and 2011, respectively. He joined Kokusai Denshi Denwa Co. Ltd. in 1996. He is currently an associate professor in the department of Intelligent Information Engineering, Doshisha University. He has been engaged in research and development of automatic speech recognition, spoken dialogue systems, and interactive user interface. He received IPSJ Kiyasu Special Industrial Achievement Award in 2011. He is a member of IPSJ, ASJ, IEICE and IEEE.

Tatsuya Kawahara received his B.E. in 1987, his M.E. in 1989, and his Ph.D. in 1995, all in information science, from Kyoto University, Kyoto, Japan. From 1995 to 1996, he was a visiting researcher at Bell Laboratories, Murray Hill, NJ, USA. Currently, he is a professor in the School of Informatics, Kyoto University. He has also been an invited researcher at ATR and NICT. He has published more than 300 technical papers on speech recognition, spoken language processing, and spoken dialogue systems. He has been conducting several speech-related projects in Japan including speech recognition software Julius and the automatic transcription system for the Japanese Parliament (Diet). Dr. Kawahara received the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology (MEXT) in 2012. From 2003 to 2006, he was a member of IEEE SPS Speech Technical Committee. He was a general chair of IEEE Automatic Speech Recognition and Understanding workshop (ASRU 2007). He also served as a Tutorial Chair of INTERSPEECH 2010 and a Local Arrangement Chair of ICASSP 2012. He is an editorial board member of Elsevier Journal of Computer Speech and Language, APSIPA Transactions on Signal and Information Processing, and IEEE/ACM Transactions on Audio, Speech, and Language Processing. He is VP-Publications (BoG member) of APSIPA and a senior member of IEEE.