How Unfamiliar Words in Smartphone Manuals Affect Senior Citizens

Tatsuya Ishihara, Masatomo Kobayashi, Hironobu Takagi, and Chieko Asakawa

IBM Research – Tokyo, NBF Toyosu Canal Front, Toyosu 5-6-52, Koto, Tokyo, Japan
{tisihara,mstm,takagih,chie}@jp.ibm.com

Abstract. Elderly people are motivated to continue working, but may have difficulties working in full-time jobs and need flexible working styles to compensate for their declining physical abilities. ICT can help support flexible working styles by enhancing communication between people in distant places. Smartphones offer various features for communication and information gathering, thus creating more opportunities to work. However, smartphone adoption has been slow for the elderly. One of the reasons is that elderly people have lower familiarity with computer terminology and therefore find the manuals difficult to understand. In this study, we investigated factors that make smartphone manuals hard to understand. We first asked elderly people about their familiarity with words found in smartphone manuals. Our second survey asked about sentences extracted from the smartphone manuals. By analyzing these results, we found that the comprehension was highly correlated with their familiarity with the specialized vocabulary.

Keywords: Word familiarity, text readability, ageing, smartphone.

1 Introduction

Elderly people want to work and contribute to their societies [1], but often have difficulties with full-time employment because of their declining physical abilities. They need more flexible work [2] where they can participate in part-time jobs or from remote offices. Flexible working styles can be supported by ICT (Information-Communication Technology). Smartphones are rapidly gaining popularity because they offer various useful functions for communication and information gathering such as e-mail, Web browsing, and GPS navigation. In addition to these functions, they offer various accessibility features such as zoom and screen reading functions. Many of these features were originally designed for computer-use by people with disabilities, but are also helpful for senior citizens whose physical abilities are gradually declining. If the elderly can fully utilize smartphones, they will find more opportunities to work and easily communicate with their friends and families and be more involved in their local communities.

However, smartphones are not yet popular among elderly people. In the US, 45% of all people over 18 years old own a smartphone, but only 12% of people over 65

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2013, Part III, LNCS 8011, pp. 636–642, 2013.

have one [3]. One of the reasons for their limited popularity is that elderly people who have weaker computer skills find it difficult to adopt new computer-like devices.

Nicholas et al. [4] showed that elderly people with less Internet experience had difficulties in using touch screen kiosks because of their lower literacy as regards computer-related terms. Leung et al. [5] found that elderly people preferred to study instruction manuals rather than learn by trial and error. Elderly people often felt frustration when the manuals were inadequate or difficult to understand [5]. Creating easy-to-understand manuals is essential to help them adopt the new technologies. Though many of elderly people experience declines in their cognitive skills and short-term working memory, it is known that the language skills of elderly people are good [6]. Therefore, we hypothesized that elderly people may have special difficulties with these manuals because they are unfamiliar with the computer-related terms.

In our two-part experiment, we first asked elderly people about their familiarity with various technical words extracted from smartphone manuals. We then asked about their comprehension of sentences extracted from smartphone manuals. By analyzing both of these results, we investigated factors that make smartphone manuals hard to understand for elderly people.

2 Related Work

Text readability research involves metrics to relate the reading difficulty to educational levels. In English, several metrics such as FOG, SMOG, and Flesch-Kincaid are used to estimate the appropriate educational level of a text [7]. All of these metrics uses sentence structures, such as the number of words per sentence, average number of syllables per word.

Collins-Thompson et al. [8] showed that a statistical language model can help assess text readability more accurately. They used a unigram language model to predict the grade level of text. Pitler et al. [9] showed a language model can help estimate the text readability for various purposes. They used a unigram language model that was calculated from newspaper articles, and estimated the text quality for educated adults. Kanungo et al. [10] assessed the text readability for the summaries of webpages that were created by search engines.

Although the Japanese language has a special problem in that the words are not separated by spaces, several methods have been devised to predict the readability of a text by analyzing the sentence at the character level [11, 12].

3 Word Familiarity and Sentence Difficulty

In contrast to the methods for calculating text readability, we focused on the vocabulary of the elderly to recognize the kinds of sentences that elderly people may find difficult. Word familiarity describes how familiar people are with specific words. In Japanese, there is a major database for word familiarity [13], but the participants were young people in their twenties, and only general vocabulary was tested. Because

elderly people have higher language skills and are already familiar with this general vocabulary, we focused only on computer-related terms.

To assess the senior citizens' familiarity with computer terms, we used a similar approach to that used in [13]. To collect the computer terms, we analyzed five smartphone manuals and extracted some frequently used terms. In our study, all of the participants were senior citizens over 60 years old. Each participant indicated how familiar they were with each of the words by selecting an answer from three candidate answers: "I know this word" (scored as 3 points), "I recognize this word" (2 points), and "I do not know the word" (1 point). In this paper, we defined word familiarity by averaging the numeric values for all of the participants, so each word had a familiarity value between 1 and 3.

We defined sentence difficulty in a similar way. We selected certain sentences from the smartphone manuals, where each sentence should contain more than one word included in the word familiarity survey. For some of the sentences, the words from the word familiarity survey were replaced with other words that have similar meanings. For each sentence, the participants answered by selecting from three candidate answers: "The sentence is clear" (scored as 3 points), "The sentence is partly unclear" (2 points), "I do not understand the sentence" (1 point). We defined sentence difficulty as the average value for all of the participants, so each sentence also had a difficulty value between 1 and 3.

4 Experiments

4.1 Experimental Settings

When collecting the word familiarity data, we prepared both paper and online versions of our surveys. A total of 341 words were studied. For the online survey, the questions appeared at the side of the top page of our SNS. Each question asked about the user's familiarity with one word. When participants answered one question, another question was automatically displayed. A total of 52 elderly people participated in this online survey. For the paper-based survey, 24 elderly people participated. All of the participants live near our lab, and all of them are over 60 years old (mean=68.8, SD=5.3, two people who did not provide their accurate age were excluded from the analysis). All of them were native Japanese speakers, and the experiment was conducted in Japanese. In total, 76 people participated and the average for the number of responses from each participant was 143.84 (SD=100.45). The most active participant responded to 341 words, while the least active participant answered for 1 word. The average number of participants for each word was 32.15 (SD=10.5).

For sentence difficulty, we used only an online survey with 105 experimental sentences. Each question asked about the sentence difficulty of one sentence. At least 30 participants answered one or more questions. All 30 of the sentence-level participants had also answered word familiarity questions. All of the sentence-level participants were over 60 years old (mean=66.0, SD=8.7). The experiment was conducted in Japanese. The average number of questions answered about sentences was 28.9

(SD=36.39). The most active participants evaluated 105 sentences, while the least active answered for only 1 sentence. The average number of responses for each sentence was 8.09 (SD=1.3).

4.2 Results for Word Familiarity

We categorized the words in the word familiarity survey into three categories: (A) Words that are specific to smartphones, cell phones, or touchscreen devices, (B) Words that are specific to a vendor or a product, and (C) Other computer-related terms. Table 1 shows the five most difficult and five easiest words for each category. The number of words classified in categories (A), (B), and (C) were 100, 29, and 211, respectively. The average word familiarity for the categories (A), (B), and (C) were 2.25, 1.86, and 2.5. Analysis of variance showed a significant main effect for the category (F_2=103.9, p<.001). A post-hoc analysis found that the word familiarity for all three categories was significantly different from each other (p<.001).

Table 1. Five most difficult and five easiest words in each category: (A) Words that are specific for smartphones, cell phones, or touchscreen devices, (B) Words that are specific to vendors or products, and (C) Other computer-related terms. The numeric value in parentheses for each word is the word familiarity. The Japanese word in brackets for each word is the word that was used in our questions. The words that do not have Japanese equivalents in parentheses were asked using English.

Category	Aver. Word Familiarity (SD)	Five most difficult words	Five easiest words
A	2.25 (0.87)	Pinch Open (1.15, ﾋﾟﾝﾁｵｰﾌﾟﾝ), Pinch Close (1.15, ﾋﾟﾝﾁｸﾛｰｽﾞ), Pinch Out (1.23, ﾋﾟﾝﾁｱｳﾄ), Scrub (1.24, ｽｸﾗﾌﾞ), Scrub Bar (1.29, ｽｸﾗﾌﾞﾊﾞｰ)	Vibrator (2.95, ﾊﾞｲﾌﾞﾚｰﾀ), Redial (3, ﾘﾀﾞｲﾔﾙ), Emoji (3, 絵文字), International Call (3, 国際電話), Received Calls (3, 着信履歴)
B	1.86 (0.91)	i-Concier (1.14, iｺﾝｼｪﾙ®), Siri® (1.15), mi-niUIM (1.16), VoiceOver (1.17), ToruCa (1.23, ﾄﾙｶ®)	YouTube™ (2.57), Google Maps™ (2.71, Googleﾏｯﾌﾟ), iPad® (2.77), Twitter (2.85), iPhone® (2.93)
C	2.5 (0.79)	CalDav (1.03), DiXiM (1.16), Widget (1.21, ｳｨｼﾞｪｯﾄ), HDR Picture (1.27, HDR写真), WPA2 (1.27)	Reboot (3, 再起動), Device (3, 端末), Press and Hold (3, 長押し), Transfer (3, 転送), Junk Mail (3, 迷惑ﾒｰﾙ)

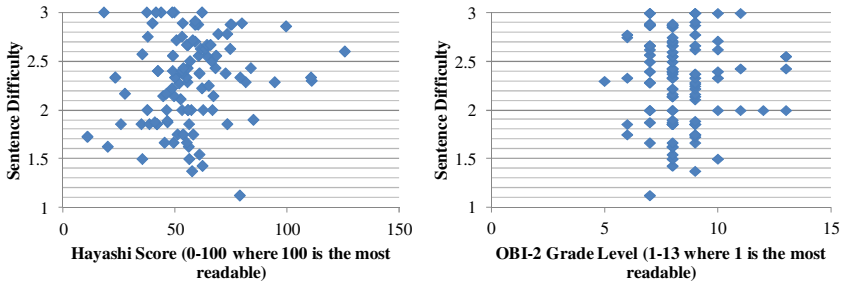

Fig. 1. Relationships of Japanese text readability and sentence difficulty for the elderly

4.3 Results for Sentence Difficulty

For this survey, we first evaluated the relationship between the sentence difficulty as reported by our participants and two text readability metrics for Japanese text, the Hayashi Score [11] and the OBI-2 Grade Level [12] rating. The cb's Japanese Text Analysis Tool[1] was used to calculate the text readability scores. The left-side scatterplot in Fig. 1 shows the relationship of sentence difficulty and Hayashi score for all of the tested sentences, while the right-side shows the relationship of sentence difficulty and OBI-2 Grade Level. The vertical axis shows the sentence difficulty and the horizontal axis shows the text readability score. Each dot shows the results for one of the tested sentences. The sentence difficulty and the Hayashi score are not highly correlated (R^2=.02). The correlation is even weaker for the OBI-2 Grade Level (R^2=1.0 $\times$ 10^{-4}).

We then evaluated the relationship between reported sentence difficulty and word familiarity. The left-side scatterplot in Fig. 2 shows the relationship of sentence difficulty and average word familiarity for each sentence. The average word familiarity for each sentence was calculated by averaging the word familiarity for the words that were tested in word familiarity survey. The right-side of Fig. 2 shows the relationship of sentence difficulty and the least familiar word of each sentence. By comparing this with the text readability, the sentence difficulty and the average word familiarity shows a meaningful correlation (R^2=.35), and the sentence difficulty and the smallest word familiarity is even more strongly correlated (R^2=.49).

From these results, it appears that the existing methods for measuring text readability are not appropriate for predicting the sentence difficulty for elderly people. It appears that it is more effective to estimate the sentence difficulty for elderly people based on the least familiar words in the sentences, which means that the key is their familiarity with the most difficult words.

[1] https://code.google.com/p/japanese-text-analysis-tool/

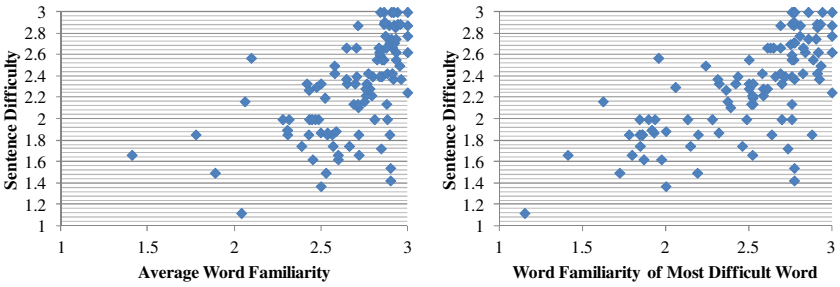

Fig. 2. Relationships of word familiarity and sentence difficulty for the elderly

5 Discussion

As shown in Table 1, the words that are specific for vendors or products are the least familiar words for senior citizens. The words that are specific for smartphones, cell phones, or touchscreen devices are more familiar, but less familiar compared with general computer-related terms. The writers of manuals should be careful when using these words. Additional explanations for these unfamiliar words may help to make the manuals easier to understand.

As shown in Fig. 1, the text readability has a weaker correlation with the sentence difficulty for elderly people, and the word familiarity for the most difficult words in each sentence has a stronger correlation. From these results, word familiarity appears to be a suitable metric for an automatic checking tool that will detect the sentences that elderly people will find hard to understand. To create such tools, we need to collect more data for statistical reliability and test more terms. One of the limitations of this study was that some of the same participants answered both in the word familiarity survey and the sentence difficulty survey. To evaluate the quality of an automatic checking tool, we need to test these two categories separately.

In this study, we focused on problems with manuals. Creating natural and intuitive menus is also important to improve usability. We are planning to investigate if word familiarity will also help to detect menu items that are difficult for senior citizens to understand and use.

6 Conclusion

In this study, we conducted a survey to investigate the word familiarity of computer-related terms that were extracted from smartphone manuals. Our second survey studied the difficulty of sentences that were also extracted from these smartphone manuals. Our results indicate that the word familiarity is the dominant factor when estimating the sentence difficulty, and showed that existing text readability metrics are not fully suitable for estimating the sentence difficulty for senior citizens who have good language skills but limited computer literacy. This research is part of our preliminary work for automatic checking tools that will help us assess manuals to be

used by elderly people. Such tools will help to create easy-to-understand manuals, and thus help elderly people adopt new technologies.

Acknowledgement. This research was partially supported by the Japan Science and Technology Agency (JST) under the Strategic Promotion of Innovative Research and Development Program.

References

1. United Nations University, Active Ageing, http://wisdom.unu.edu/en/active-aging (accessed February 20, 2013)
2. Leibold, M., Voelpel, S.: Managing the Aging Workforce. John Wiley & Sons (2006)
3. Pew Internet: Mobile, http://pewinternet.org/Commentary/2012/February/Pew-Internet-Mobile.aspx (accessed February 20, 2013)
4. Nicholas, D., Huntington, P., Williams, P.: Delivering Consumer Health Information Digitally: A Comparison Between the Web and Touchscreen Kiosk. J. Med. Syst. 27(1), 13–34 (2003)
5. Leung, R., Tang, C., Haddad, S., McGrenere, J., Graf, P., Ingriany, V.: How Older Adults Learn to Use Mobile Devices: Survey and Field Investigations. TACCESS 4(3), 11 (2012)
6. Ouchi, Y., Akiyama, H. (eds.): Gerontology – Overview and Perspectives, 3rd edn. Univ. of Tokyo Press (2010) (in Japanese)
7. Feng, L., Jansche, M., Huenerfauth, M., Elhadad, N.: A comparison of features for automatic readability assessment. In: Proc. COLING, pp. 276–284 (2010)
8. Collins-Thompson, K., Callan, J.: A language modeling approach to predicting reading difficulty. In: Proc. HLT/NAACL (2004)
9. Pitler, E., Nenkova, A.: Revisiting readability: a unified framework for predicting text quality. In: Proc. EMNLP, pp. 186–195 (2008)
10. Kanungo, T., Orr, D.: Predicting the readability of short web summaries. In: Proc. WSDM, pp. 202–211 (2009)
11. Hayashi, Y.: A three level revision model for improving Japanese bad styled expressions. In: Proc. COLING, pp. 665–671 (1992)
12. Sato, S., Matsuyoshi, S., Kondoh, Y.: Automatic Assessment of Japanese Text Readability Based on a Textbook Corpus. In: Proc. LREC, pp. 28–30 (2008)
13. Amano, S., Kondo, T.: Nihongo no Goi tokusei (Lexical properties of Japanese), vol. 1-6, pp. 1–6. Sansei do, Tokyo (1999) (in Japanese)