How Voice Augmentation Supports Elderly Web Users

Daisuke Sato Masatomo Kobayashi Hironobu Takagi Chieko Asakawa Jiro Tanaka†

IBM Research – Tokyo
1623-14, Shimo-tsuruma, Yamato City,
Kanagawa Pref. 242-8502 Japan
+81 46 215 {4793, 4679, 4557, 4633}
{dsato, mstm, takagih, chie}@jp.ibm.com

†University of Tsukuba
1-1-1, Tenno-dai, Tsukuba City,
Ibaraki Pref. 305-8573 Japan
+81 29 853 5343
jiro@cs.tsukuba.ac.jp

ABSTRACT

Online Web applications have become widespread and have made our daily life more convenient. However, older adults often find such applications inaccessible because of age-related changes to their physical and cognitive abilities. Two of the reasons that older adults may shy away from the Web are fears of the unknown and of the consequences of incorrect actions. We are extending a voice-based augmentation technique originally developed for blind users. We want to reduce the cognitive load on older adults by providing contextual support. An experiment was conducted to evaluate how voice augmentation can support elderly users in using Web applications. Ten older adults participated in our study and their subjective evaluations showed how the system gave them confidence in completing Web forms. We believe that voice augmentation may help address the users' concerns arising from their low confidence levels.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces; K.4.2 [**Computers and Society**]: Social Issues – *Assistive Technologies for Persons with Disabilities*.

General Terms

Experimentation, Human Factors

Keywords

Older adults, Web accessibility, voice-based augmentation

1. INTRODUCTION

The United Nations has reported that elderly people (60 or older) constitute 11% of the world's population and 21% of the population of developed regions, and the percentage of older people is continuing to rise [1]. People experience degenerative effects of ageing in their senses of vision and hearing, in their psychomotor abilities, and in their attention and memory [2]. Even people who are happily using the Web now may face problems in the future, because the technologies are evolving and the new technologies will be different [3]. The Web has become an essential tool and online applications have made our daily lives much more convenient. Everyone, including the elderly and people with disabilities, should be equally able to use the Web.

ASSETS'11, October 24–26, 2011, Dundee, Scotland, UK.

Unfortunately there are many webpages that are not friendly for seniors, including Web applications for our daily lives such as online banking and online shopping. Although the elderly may be gradually losing some of their cognitive and physical abilities, they can still learn new technologies. Kolodinsky et al. reported that the most significant problems for elderly Web users are not age-related functional impairments, but fears of the unknown and of the consequence of incorrect actions that inhibit exploration [4]. Many older adults offered such comments as "We can use new applications if we get used to them" or "The problems are in the initial attempts." This same kind of feedback also comes from visually impaired users.

This paper describes how a voice-based augmented interface can make elderly users more confident in completing tasks with online Web applications. This voice augmented interface was originally evaluated for people with visual impairments. Those results showed that the second channel for voice guidance increases blind users' confidence in navigating in Web applications [5]. We assume that although the main channels of the user interfaces are different (voice vs. graphics), the second channel using voice may help older adults complete the tasks in Web applications by increasing their confidence in their operations. Proper support can reduce their cognitive load and help them remember and learn. For example, some new home electrical appliances have function to provide voice guidance about proper usage. Users can operate such an appliance without any manual, and such appliances are increasingly popular with the elderly, though most Web applications still lack corresponding approaches to customer support. This is primarily because the Web tends to be designed by and for younger people, a situation that is steadily changing. That is why we are investigating new ways to apply voice augmentation methods to operations in Web applications.

One of the advantages of voice augmentation approach is that voice support can provide additional content for existing Web applications, extending the lives and utility of those applications. Of course voice augmentation can coexist with other alternative interfaces and just provide the basic components for such groups as older adults or novice users.

This paper describes two experiments after reviewing related work. The first experiment focused on online banking and shopping applications to observe the behavior of older adults and assess the effects of voice augmentation. The other experiment evaluated relative performance in completing Web forms with and without voice augmentation, comparing younger and older adults.

The results and future possibilities are discussed in the concluding section.

2. RELATED WORK

Ageing societies have been a focus for some years now. This section mainly describes studies of older adults using the Web. In addition, we cover studies of user interface agents that augment voice-based user interfaces.

2.1. Studies for Elderly Users

There have been many field studies with older adults. [6,7] cover the differences between younger and older adults in their Web navigation behaviors. Meyer et al. [6] conducted a study with thirteen older and seven younger adults. They reported that the older group needed more steps to find information, but both groups decreased their steps after a hands-on tutorial session. One of the interesting behaviors of older adults reported in this study is returning back to a "home" location if they became disoriented during navigation. Fairweather [7] reports that older adults tended to use the least risky method in navigating. Chadwick-Dias et al. [8] studied how Web experience influenced the behavior of users on the Web. They report Web experience is the same as Web expertise, but older adults take more time to develop their Web expertise. They concluded that older adults need more opportunities for collaborative learning with other people to learn Web navigation techniques.

[9,10,11] presented observations of older adults using existing systems. Sayago and Blat [9] conducted a 3-year study of everyday interactions with the Web and reported that problems with remembering steps, with understanding Web and computer jargon, and with using the mouse are more significant than problems with perceiving visual information, with understanding icons, or with using the keyboard. Akatsu and Miki [10] studied the unexpected behaviors of older Japanese adults using Automated Teller Machines (ATMs). Some users overreacted to certain voice messages or repeated the same error when they couldn't understand the situation. Leitner et al. [11] found that older adults do not show major differences from younger people in their needs and preferences related to an online ticket service.

Hanson et al. [12] evaluated a voice browsing application that provides functions to interact with the browser using speech. Inexperienced users tended to use long commands instead of the brief voice commands that the system could easily recognize.

2.2. Guidelines for Elderly Users

Many studies about ageing Web users have been conducted and many guidelines for Web content have been published to improve accessibility. The Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C) discussed accessibility for older adults in their three-year WAI-AGE project that contributed to the Web Content Accessibility Guidelines (WCAG) 2.0 in 2010. Web accessibility for people with disabilities has gradually improved due to guidelines and new accessibility technologies. In addition, work continues on new assistive technologies for older adults. ISO IEC Guideline 71 [13] also requires considering the needs of the elderly and of persons with disabilities. This guideline affects local standards that implement accessible designs such as JIS X8341, which is also based on the WCAG.

The SPRY Foundation published a Web guideline for older adults in 1999 [14]. This guideline mainly focuses on vision, cognition, hearing, and motor skills. The criteria of this guideline are very similar to WCAG. Newell and Dickinson [15] used a case study approach to the development of a simple Web application for elderly users. Chandwick-Dias et al. [16] studied how older users surf and their problems with the accessibility of Web 2.0 content. They found that each functional impairment of elderly people was relatively smaller than the corresponding impairment of people who need special support. Most of the special needs guidelines are easily applicable for elderly people. In addition, we need to address problems related to memory and learning to support older adults, limitations that have rarely been considered in accessibility technologies to date.

2.3. Assistive Technology for Elderly Users

Although there is no prior study that supports older adults in navigating Web applications by themselves while using audible user interfaces, there are many assistive approaches for older adults. BrookesTalk is an audible Web browser designed for people with visual impairments. A BrookesTalk extension called Voice Help provides guided support for older people with visual impairments [17]. It provides the status of the applications and lists of possible next actions in a way similar to the interface of IVR. Zajicek and Morrissey [18] used BrookesTalk to study the effects of multimodality with older adults. They reported that long instruction messages interfered with the correct operations and the users preferred text instructions rather than mixed text and speech. They also mentioned that older adults found synthesized voices hard to understand.

Hailpern [19] proposed a wizard interface that tracks the current status of elderly users. The system provides a simple interface within a single window and uses a history list to recognize the status of users. Milne et al. [20] proposed a minimal application interface for senior users. Their prototype browser has only five buttons and highly intuitive labels. For example they used "look up" and "look down" for "page up" and "page down". Muta et al. [21] developed a Web browser extension for older adults. It provides functions to read the selected content out loud, to magnify it, or to manage the colors to improve the contrast of the text and background.

Some online shopping applications use online support systems with chat, telephone support, or special applications for remote control by an operator [22,23]. Basically these forms of help are provided by humans, but some systems include intelligent agents combined with frequently asked questions. These applications target novice users (including older adults) who are customers or potential customers.

2.4. User Interface Agents with Voices

Maes [24] talked about the concepts of interface agents to help users reduce their efforts and avoid information overload. Bederson [25] created an automated tour guide prototype that uses audio to guide tourists. Sawhney and Schmandt [26] worked on Nomadic Radio, an agent system to decide how to most effectively present information to the user based on the context, interruption settings, and automatic text understanding. Wagner and Lieberman [27] introduced Woodstein, which predicts and assists the next user action based on analysis of collected sequences of previous actions on the webpages. Roth et al. [28] created an agent to provide audio feedback for the user's cursor location. Yu et al. [29] designed context-aware Web agents to provide audio and haptic feedback for the user's cursor location in a screen reader. Dontcheva et al. [30] created a Web agent that can help record and organize user sessions for comparison and analysis. The authors reduced the users' memory load and

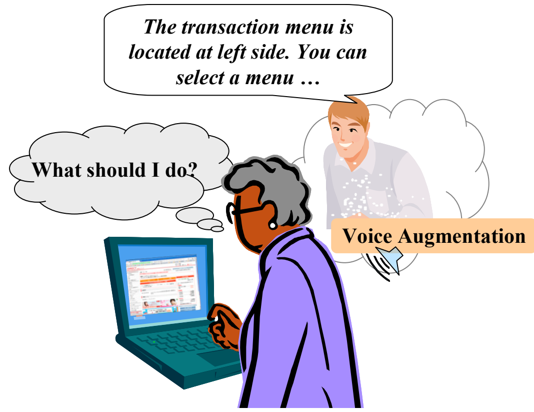

Figure 1. Concept of voice-augmented Web browsing

simplified their tasks. Hartmann et al. [31] described Augur, a context-based smart agent that can highlight, suggest, and automate by analyzing the context data with pre-defined rules.

3. VOICE-AUGMENTED WEB

The voice output we tested is a very simple concept to support operations in Web applications (see Figure 1). The voice can provide instructions for users even when they visit an application for the first time. During the processing of each Web form in the application, the voice repeats the user's input and tells the user about the next action. Users can be notified of errors that are augmented with voices or sounds more easily than by error text alone. Such a voice can be pre-recorded or synthesized. Older adults prefer a pre-recorded voice to a synthesized voice, but that approach is less flexible and more expensive. Although we have developed a voice-based augmented interface for people with visual impairments, a system for older adults requires different type of augmentation, so we investigated the use cases and categorized the augmentations into four types: confirmation, notification, contextualization, and summarization.

Confirmation provides confirmation of a user's input. It reads all types of form fields such as text, radio buttons, and so on, checking whenever the value of the field is changed.

Notification makes users aware of a status change on a webpage, such as errors in a form (e.g., incorrect input) and the progress of content loading.

Contextualization suggests the next action a user should perform in a situation, such as the choices that can be performed and operational tips. For example, "Press the search address button to input the address automatically using the postal-code". The system can also explain the results of some actions.

Summarization explains choices available on a page, summarizes the page structure, or lists the steps to be performed by the user.

4. STUDY 1: BANKING AND SHOPPING

Our first study observed how users interacted with unfamiliar Web applications and how they felt about voice augmentation with a Wizard of Oz implementation [32].

4.1. Participants

Ten older adults participated in this study. Half of them were in their 60s and the other half were in their 70s. They were familiar with computers because they had retired after working for IT companies. Some of them were engineers and others were in sales. Since current employees are quite experienced with computers and the Web, we believe that our participants typify the older population of the future. Table 1 summarizes their experiences with online Web applications, online banking, and online shopping. Most of them were experienced online shoppers but only three had experience with online banking. The participants without online banking experience said they felt it was insecure, they were satisfied with ATMs, and that it was too much trouble to sign up for and learn how to use new services and applications.

Table 1: Experience with online banking and shopping. "*" indicates that the user has experience with the specific application that was used in the task for our research.

User ID	Age-group	Experience with online banking	Experience with online shopping
1	60-64	Yes	Yes
2	60-64	No	Yes *
3	65-69	No	Yes
4	65-69	Yes	Yes *
5	65-69	No	Yes
6	70-74	No	Yes *
7	70-74	No	No
8	70-74	No	Yes
9	70-74	No	Yes *
10	75-79	Yes	Yes

4.2. Procedure

Each user was asked to perform two tasks with Web applications: (1) a fund transfer using an online banking application, followed by (2) a purchase using an online shopping application. Table 1 shows the experience of each user. For each task, the observer first told the user about the task and then the user attempted to perform the task without voice augmentation or human assistance. The observer manually recorded the user's behavior, including the page navigation history, struggles, errors, and so on, as precisely as possible. After the task was finished, the user was asked about the task with reference to the recorded notes about the session.

The three users with experience using online banking had not used the specific application used in this study. The four users with experience in online shopping had used the same application as the study (the *s in Table 1). It was not feasible to register for the online banking application for each user in this study, so the authors provided two online banking accounts and the participants transferred money from one account to the other. This gave the users an authentic feeling of making transactions on the Web. In contrast, the purchases were not executed, but the users were told to stop just before clicking on the last button in the ordering process.

After doing the two tasks without voice augmentation, the user and the observer walked through the tasks again with voice augmentation using a Wizard of Oz protocol and the user was interviewed again. The observer manually used a text-to-speech application to play predefined messages suitable for the user's operations. Typical messages were (Japanese) instructions such as "Please click the red login button on the right side of this page to start online banking", "Please input your account number and the password", or "Please click the Continue button. The transaction will not be executed yet."

4.2.1. Task 1: Online Banking

First the user was given an account card that describes the user ID with a table of random numbers, the password for the account, and the account information for the recipient of the transfer. Next the user was asked to open the webpage of the banking application[1] based on the observer's instructions. Then the user was told to transfer a specified amount of money from the account to the recipient. Here are the required steps for the task.

1. Click the "login" button at the top right of the webpage to open a new window for the transaction. The user must do all of the banking transactions in this new window. The window will initially be 700 pixels wide and 600 pixels high (though users can resize it).
2. Input the user ID and password for the account to get to the account page.
3. Click on the "transfer" menu at the top left of the account page.
4. Select the "new recipient" button after scrolling down approximately one screen (for the initial window size).
5. Select the bank of the recipient and click on the button with the first letter of the branch of the recipient account to navigate to the next page.
6. Select the proper branch from a combobox and select the account type, and input the account number and the amount of money. Then click on the "next" button to confirm the information.
7. Check the information and input two requested random characters from the table on the account card. Finally click on the "execute" button to finish the transfer.

4.2.2. Task 2: Online Shopping

Here are the required steps for this task, starting after the user had opened a product page on the shopping website[2].

1. Put the item into the shopping cart to open the page of the shopping cart.
2. Click the "proceed to checkout" button for the next page.
3. Input the user's name, address, and e-mail address and click the "next" button to open the next form.
4. Select a payment option and a delivery option using radio buttons and click the "next" button to open the last form
5. Confirm the information for the order but stop before clicking the "order" button.

4.3. Observations

Here are some characteristic observations from the sessions. Most of these points are addressable by voice augmentation. Some items confirm findings from earlier studies.

4.3.1. Could not grasp content structure and meaning

Participants struggling with a task tend to read and reread content that was not relevant to the task, and scanned the content sequentially seeking the correct path for the task. A participant might scroll up and down rapidly looking for a target. Such behaviors were described in earlier studies. Participants sometimes lost their partial work on a task because of confusion about their status.

They often failed to select required radio-button options in the shopping application. The form of application was hard to understand because a needed set of radio buttons could not be seen within the initial window because the descriptions of the options were too long. In addition, most of the participants with problems overlooked the error messages that appeared at the top of the page when the incomplete form was submitted. In addition, some participants were confused by the expired-page warning that appeared if the browser "back" button was used within the form.

4.3.2. Did not understand widgets

Participants in their 70s tended to click on non-clickable elements, being misled by bright colors or disabled radio buttons. They also tried clicking on unneeded buttons or links even when they had figures showing how to use the application. They could not understand the meaning of some widgets by looking at them.

One user sometimes clicked (left and right) on some breadcrumb navigation links with a distinctive background to try to input the information. This was because the default-size window was too small and users had to scroll down to complete the transaction.

4.3.3. Did not know the function of the application or understand the general GUI metaphors

About half of participants were not aware of standard functions that are generally used in Web applications, such as a function to search for an address from a postal code. Some participants needed a long time to understand the functions needed for the application, such as how to use the table of randomized numbers. One user quickly found the login button and succeeded in logging in, but after that he returned to the initial page because he accidentally clicked outside of the new window. He said "The window disappeared" in the interview. After that he assumed that he had logged into the application and he searched for "transfer", but got lost in a FAQ page that describes how to transfer funds.

4.3.4. Anxiety Interference

One user, whenever he tried to click a link or button for the next action seemed to nervously confirm the action to himself. Also he said "What?" and struggled with a page for a while when the behavior of the application was different from his expectations.

The top page of the banking application provides a menu with over 20 items, various types of statuses and notifications for the account, and also advertisements for some financial products. A participant said about this page "I feel that [this] important thing (banking) was done as an advertisement leaflet."

5. STUDY 2: WEB FORM

Operations with a Web form are likely to cause errors because users must input or select values that are acceptable to the application. Errors include typos, long or short input, illegal characters for a text field, null selections, and so on. Another reason is that older adults tend to type keys while looking at the keyboard instead of the screen. Our hypotheses were that the

[1] Bank of Tokyo-Mitsubishi UFJ: http://direct.bk.mufg.jp/, the experiment was conducted during April 18-22, 2011.

[2] Rakuten Ichiba: http://www.rakuten.co.jp/, the experiment was conducted during April 18-22, 2011.

Table 2: A translation of the items from the questionnaire

Label	Question
Accurate	Compared to the normal mode, I found voice augmentation to be more accurate for input
Fast	Compared to the normal mode, I could input faster with voice augmentation
Comfortable	Compared to the normal mode, I felt sure that I would finish with voice augmentation
Distracting	Compared to the normal mode, I could not concentrate on the tasks with voice augmentation

voice augmentation could enhance their focus and help support their accurate input, thus reducing errors and increasing the confidence of the users.

The users were asked to fill out several types of forms in two modes, one without voice (normal) and one with voice. The forms are: input a number (task-1), input a user's name in Japanese (task-2), input a bank account type and number (task-3), and choose a valid option (task-4). The voice assistant read aloud each key when the user typed the forms in the first three tasks. After a short delay, the assistant would read all of the input text. For example, a user would input "1000" into a text field and the assistant read "one, zero, zero, zero, (pause), one thousand". For task-4, the voice read the current status. For example if there was an unchecked checkbox by the words "mail notification", then the voice read a phrase such as "mail notification is off". Users had a practice session before the actual experimental session.

The users first did the tasks without the voice, followed by the tasks with the voice for training. Then users were asked to input 8 things in each task. In total, 8 × 4 tasks × 2 modes = 64 actions were to be performed. The order of the tasks was randomized. After finishing all of the tasks, we used a survey with seven-point Likert items from -3/definitely-disagree to +3/definitely-agree to compare the test conditions. Table 2 shows a translation of the items from the questionnaire related to accurate, fast, comfortable, and distracting.

5.1. Participants

Five younger adults and ten older adults participated in this study. The younger adults are all in their 30s and are all familiar with computers and have advanced computer skills. The older adults were the same participants from Study 1.

5.2. Apparatus

For this study, we implemented a simple Web form application with voice augmentation using Eclipse ACTF [33] and a Japanese male synthesized voice. All events from the mouse and keyboard were recorded by the application. The application was running on Windows, in an A4 notebook with a Japanese keyboard. The users could point with the trackpoint on the notebook or with a USB mouse with a scroll wheel.

5.3. Results

Figure 2 compares the average task completion times for each participant group (30s, 60s, and 70s) for each task. The overall average values without the voice augmentation were 6.67 (*SD* = 3.82), 10.17 (*SD* = 4.86), and 13.57 (*SD* = 6.20) seconds for 30s, 60s, and 70s, respectively. The values with voice were 6.72 (*SD* = 3.72), 10.47 (*SD* = 5.04), and 15.77 (*SD* = 9.47) seconds. Obviously the task completion times are increasing with age. In addition, they tend to slightly increase with the voice augmentation for people in their 70s.

Three-way mixed ANOVA showed significant main effects on the task completion times of the age ($F_{2,12} = 43.05$, $p < .001$), the task ($F_{3,924} = 309.48$, $p < .001$), and the mode ($F_{1,924} = 10.52$, $p < .005$). It also showed significant interaction effects of the age and the task ($F_{6,924} = 9.66$, $p < .001$) and the age and the mode ($F_{2,924} = 6.78$, $p < .005$). Only the participants in their 70s were significantly slowed down by the voice augmentation ($F_{1,924} = 23.65$, $p < .001$). A post-hoc analysis found that the participants in their 30s were significantly faster than those in their 60s ($p < .005$) and 70s ($p < .001$) while the 60s were significantly faster than the 70s ($p < .005$). We also found that task-3 took significantly longer than the other three tasks ($p < .001$) while task-4 took significantly less time than the other three tasks ($p < .001$).

Figure 3 compares the error rates of each group on each task and the overall error rates. The overall values without the voice augmentation were 2.5%, 4.4%, and 4.4% for 30s, 60s, and 70s, respectively. The values with voice were 2.5%, 0.6%, and 3.8%. Three-way mixed ANOVA showed a significant main effect of the task on the error rate ($F_{3,924} = 4.43$, $p < .005$). The age and the mode had no significant main effects. A post-hoc analysis found that task-4 caused significantly more errors than task-1 or task-2 ($p < .05$).

Figure 4 shows a comparison of the average scores for the subjective questionnaires. For the question on "accurate", the values were 1.2 (*SD* = 0.98), 1.6 (*SD* = 0.49), and 2.4 (*SD* = 0.8) for 30s, 60s, and 70s, respectively. For "fast", the values were 0.2 (*SD* = 0.75), 1.2 (*SD* = 0.75), and 2 (*SD* = 1.10). For "comfortable", the values were 1 (*SD* = 0.90), 1.4 (*SD* = 0.8), and 2.6 (*SD* = 0.49). For "distracting", the values were -2 (*SD* = 1.10), -0.6 (*SD* = 1.36), and -2.2 (*SD* = 0.4). The respondents in their 70s gave relatively more positive scores for each question. Based on the recorded times, the participants in their 70s took longer with voice than with the normal mode, but they said they could input faster and did not need more time with voice than with the normal mode.

6. DISCUSSION

6.1. Tradeoff of Confidence for Speed

The participants, especially those over 70, reported that the voice augmentation sped up their operations. However, the actual task-completion times increased in spite of their own reports. These results were surprisingly contradictory. We believe this shows that participants had increased confidence in their operations due to the support of the voice augmentation. However the actual time increased, because participants listened to the voice while pausing in their operations which could be observed in logged events. The time seemed shorter because of their higher confidence. It is known that stressful situations lengthen subjective time [34]. Some of the participants' comments support this interpretation, such as "The voice makes us feel relaxed", "It is useful on the first attempt", "I'm sure the input is correct with the voice", and "I could confirm the input without watching the screen". The participants also reported that they could do the tasks more accurately, but there was no significant difference in the actual error rates. Given the relatively small number of errors in the

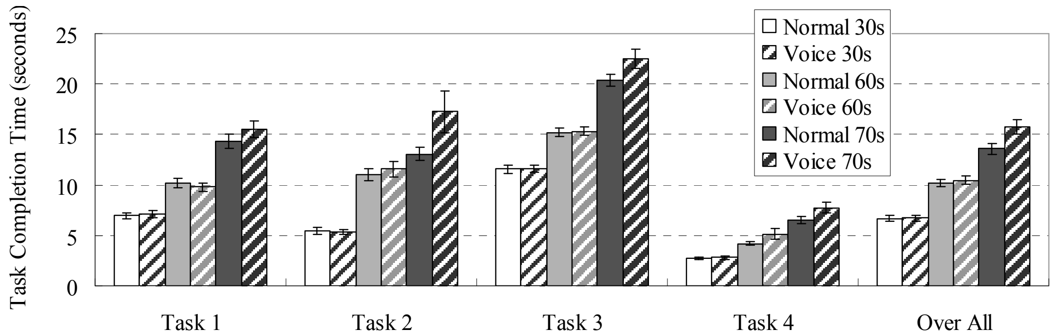

Figure 2: Comparison of task completion times

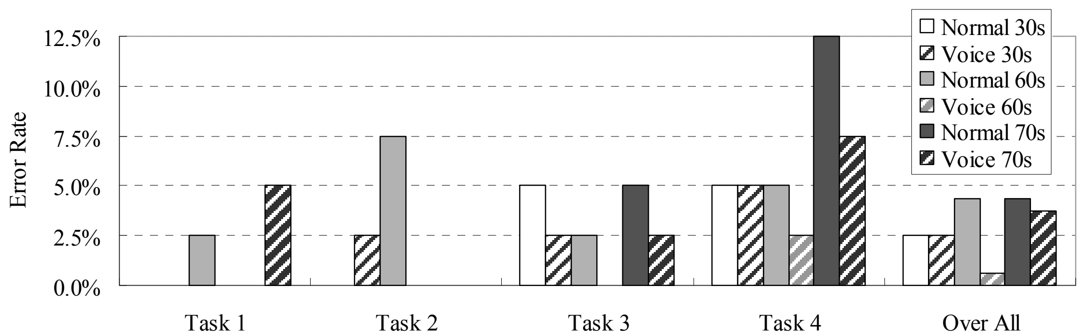

Figure 3: Comparison of error rates of tasks

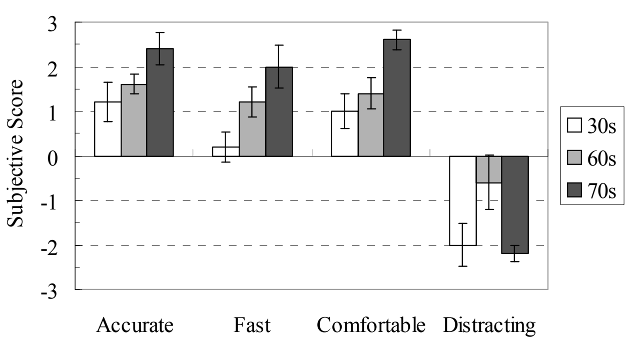

Figure 4: Subjective scores for the questions

experiments, more experiments and observations are required to determine how voice augmentation affects accuracy.

It is known that there are generally small correlations between empirical measurements (speed and accuracy) and subjective evaluations [35]. Since many elderly people tend to resist using technologies due to their fears [4], we believe that subjective factors should be regarded as more important than objective performance in the design and development of senior-friendly applications.

The participants in their 30s and 60s also reported they could input accurately and felt comfortable. Although their subjective scores are relatively lower than those of the participants in their 70s, they were also confident about using voice augmentation compared to the normal condition. They tended to click on the complete button before the confirmation message was finished (which stopped the voice), therefore there was no significant difference in their task completion times. The final message began 1.5 seconds after the final key was pressed for the normal speech rate. Using a faster speech or and shorter pauses for the confirmation messages linked to the user's input speed might change the results.

6.2. Another Way to Gain Confidence

The participants in their 70s tended to first focus on the most appealing content or on the content located at the center of the screen without grasping the structure of the page. Especially with the small window, the buttons the participants must click are located elsewhere and scrolling is needed. This often worried the older participants. They tended to try to read all of the visible content (which was mostly a warning statement about the timing of the transfer). One participant tried to click on an unclickable element without scrolling down.

The next element that should be focused on by the user can be identified for most of the webpages used here. The voice augmentation can say where the next focal element is located. Also visual feedback with highlighting using dynamic HTML technologies would help the users more effectively than voice augmentation alone and would give them more confidence. Participants also commented about such support, with comments such as "The element mentioned by the voice should be highlighted" while the content that should be focused on by the user in that status could be changed according to the user's intention. For example, though the user wants to transfer money, it may be hard for the system to anticipate the user's intention. An instruction for the page structure can help in such a situation, with a message such as "A transaction menu is located at the top left of this page."

6.3. Will They Want to Stop the Assistant?

Many people have bad memories of Microsoft's "Clippit" or "Clippy" (a dolphin in Japanese versions), an intrusive assistant avatar for a user interface agent for GUI applications, and a frequently asked question was "How to disable Clippy?" That strongly indicated that users want to control the assistance shown to them based on the context and their own skills.

Out of four types of voice augmentation, **repetition** and **notification** seem acceptable in many situations for older adults, giving them confidence in completing forms. **Contextualization** and **summarization** mainly support users in constructing mental models of the applications. Therefore they may stop supporting and start interfering with the users as they try to complete the Web forms. More investigation is needed to answer such questions as how older adults learn about applications, how long they remember what they have learned, and how can we assess the mental models of the users from their behaviors. Studying the navigation history in a Web application and the interaction events may help in understanding the users.

Appropriate analysis of a user's behavior and skills may be useful in deciding on the proper presentation for that user. For example a financial Web application may offer a new financial product to a user, based on frequent visits to the application, and lead the user to contact a sales representative.

6.4. Possible Implementations

Sloan et al. [36] reported on the potential of adaptive assistive technology for people whose abilities are gradually declining. In their study, they proposed a new application architecture that provides on-demand assistive technologies for each level of impairment (e.g. vision, hearing, motor, and so on). Instead of providing OS-dependent assistive technologies for certain impairments, users could be supported by multiple forms of assistance running within applications.

The most important component of our system is the voice output component. A prerecorded voice is one solution to provide good voice quality and it is preferred by older adults over a synthesized voice. However, synthesized voices are needed for a voice augmentation system because the Web content is dynamic and fluid and users can input text freely. Synthesized voice is mainly provided by library applications installed in a client system, while some applications provide synthetic voices through the Internet. Client-side voices have advantages in reduced response latency. Server-side voices have disadvantages in latency but the users don't need to install any voice libraries. WebAnywhere [37] is an audible Web browser that provides a server-side synthesized voice through the Internet. This system tries to predict the user's next action to reduce the latency of the speech response by analyzing the keyboard events with a hidden Markov model. Although synthesized voices are disliked by older adults [18], most of the participants could understand the synthesized voice messages. Several of them complained about the quality of voice and could not understand the meaning of the sentence until they heard it a second time. They said "Hmm?" or "What?" to ask for a repetition of the sentence, which could provide feedback to a speech recognition component. [18] also mentioned that a deeper male voice is generally easier for older adults, while some participants suggested a female voice would be better.

Most participants reported the voice did not distract their attention from the input, though two participants in their 60s reported they were distracted during the task. They disliked the timing of the voice presentation, which again indicates that the voice presentation should be optimized to consider the user's typing speed or some other criteria.

7. CONCLUSION

We investigated a voice augmentation system that supports elderly people in online banking transactions and online shopping. Subjective evaluations showed that the system made them feel confident (especially for people in their 70s) when they needed to accurately fill out electronic forms for online banking. This means the voice augmentation succeeded in reducing the mental barriers for using Web forms, giving the users confidence in their accuracy. The results of the experiment showed that the voice augmentation system can to encourage elderly Web users in using Web services even if their abilities are declining. Further exploration is required to clarify how the voice augmentation system might help elderly users navigate and complete forms in heterogeneous applications.

8. ACKNOWLEDGMENTS

We'd like to express appreciation for the participation of retirees of IBM Japan and to Oyagamo-kai.

9. REFERENCES

[1] United Nations. World population ageing 2009. 2010. http://www.un.org/esa/population/publications/WPA2009/WPA2009-report.pdf.

[2] Harper, S. and Yesilada, Y. *Web accessibility: a foundation for research*. Springer, 2008.

[3] Hanson, V.L. Age and Web access: the next generation. *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibility (W4A)*, ACM (2009), 7−15.

[4] Kolodinsky, J., Cranwell, M., and Rowe, E. Bridging the generation gap across the digital divide: Teens teaching Internet skills to senior citizens. *Journal of Extension 5*, 10 (2002), 2.

[5] Sato, D., Zhu, S., Kobayashi, M., Takagi, H., and Asakawa, C. Sasayaki: augmented voice Web browsing experience. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, ACM (2011), 2769−2778.

[6] Meyer, B., Sit, R.A., Spaulding, V.A., Mead, S.E., and Walker, N. Age group differences in World Wide Web navigation. *CHI 1997 Extended Abstracts on Human Factors in Computing Systems: Looking to the Future*, ACM (1997), 295−296.

[7] Fairweather, P.G. How older and younger adults differ in their approach to problem solving on a complex website. *Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM (2008), 67−72.

[8] Chadwick-Dias, A., Tedesco, D., and Tullis, T. Older adults and Web usability: Is Web experience the same as Web expertise? *CHI 2004 Extended Abstracts on Human Factors in Computing Systems*, ACM (2004), 1391−1394.

[9] Sayago, S. and Blat, J. About the relevance of accessibility barriers in the everyday interactions of older people with the web. *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibility (W4A)*, ACM (2009), 104−113.

[10] Akatsu, H. and Miki, H. Usability research for the elderly people. *Oki Technical Review (Special Issue on Human Friendly Technologies) 71*, 3 (2004), 54−57.

[11] Leitner, M., Subasi, Ö., Höller, N., Geven, A., and Tscheligi, M. User requirement analysis for a railway ticketing portal with emphasis on semantic accessibility for older users. *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibility (W4A)*, ACM (2009), 114−122.

[12] Hanson, V.L., Richards, J.T., and Lee, C.C. Web access for older adults: voice browsing? *Proceedings of the 4th International Conference on Universal Access in Human Computer Interaction: Coping with Diversity*, Springer-Verlag (2007), 904−913.

[13] ISO/IEC. ISO/IEC Guide 71:2001 - Guidelines for standards developers to address the needs of older persons and persons with disabilities. 2001. http://www.iso.org/iso/catalogue_detail?csnumber=33987.

[14] SPRY Foundation. SPRY's Work: Education - A Guide for Web Site Creators. 1999. http://www.spry.org/sprys_work/education/web_guide.html.

[15] Newell, A.F., Dickinson, A., Smith, M.J., and Gregor, P. Designing a portal for older users: A case study of an industrial/academic collaboration. *ACM Transaction on Computer-Human Interaction 13*, 3 (2006), 347−375.

[16] Chadwick-Dias, A., Bergel, M., and Tullis, T.S. Senior Surfers 2.0: a re-examination of the older Web user and the dynamic Web. *Proceedings of the 4th International Conference on Universal Access in Human-Computer Interaction: Applications and Services*, Springer Berlin Heidelberg (2007), 868-876.

[17] Gregor, P., Newell, A.F., and Zajicek, M. Designing for dynamic diversity: interfaces for older people. *Proceedings of the Fifth International ACM Conference on Assistive Technologies*, ACM (2002), 151−156.

[18] Zajicek, M. and Morrissey, W. Multimodality and interactional differences in older adults. *Universal Access in the Information Society 2*, 2 (2003), 125-133.

[19] Hailpern, J.M. WISE: a wizard interface supporting enhanced usability. *Proceedings of the 8th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM (2006), 291−292.

[20] Milne, S., Dickinson, A., Gregor, P., Gibson, L., McIver, L., and Sloan, D. Not browsing, but drowning: designing a Web browser for novice older users. *Proceedings of HCI International*, (2005), 7 pages.

[21] Muta, H., Ohko, T., and Yoshinaga, H. An activeX-based accessibility solution for senior citizens. *Proceedings of the Center on Disabilities Technology and Persons with Disabilities Conference 2005*, (2005).

[22] Dell. DellConnect™. http://support.dell.com/support/topics/global.aspx/support/en/dellconnect.

[23] Citrix Online. Remote Support And Remote Support Services | GoToAssist. http://www.gotoassist.com/.

[24] Maes, P. Agents that reduce work and information overload. *Communications of the ACM 37*, 1994, 30−40.

[25] Bederson, B.B. Audio augmented reality: a prototype automated tour guide. *Conference Companion on Human Factors in Computing Systems*, ACM (1995), 210−211.

[26] Sawhney, N. and Schmandt, C. Nomadic Radio: speech and audio interaction for contextual messaging in nomadic environments. *ACM Transaction on Computer-Human Interaction 7*, 3 (2000), 353−383.

[27] Wagner, E.J. and Lieberman, H. Supporting user hypotheses in problem diagnosis. *Proceedings of the 9th International Conference on Intelligent User Interfaces*, ACM (2004), 30−37.

[28] Roth, P., Petrucci, L., Pun, T., and Assimacopoulos, A. Auditory browser for blind and visually impaired users. *CHI 1999 Extended Abstracts on Human Factors in Computing Systems*, ACM (1999), 218−219.

[29] Yu, W., McAllister, G., Strain, P., Kuber, R., and Murphy, E. Improving Web accessibility using content-aware plug-ins. *CHI 2005 Extended Abstracts on Human Factors in Computing Systems*, ACM (2005), 1893−1896.

[30] Dontcheva, M., Drucker, S.M., Wade, G., Salesin, D., and Cohen, M.F. Summarizing personal Web browsing sessions. *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, ACM (2006), 115−124.

[31] Hartmann, M., Schreiber, D., and Mühlhäuser, M. AUGUR: providing context-aware interaction support. *Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, ACM (2009), 123−132.

[32] Dahlbäck, N., Jönsson, A., and Ahrenberg, L. Wizard of Oz studies -- why and how. *Knowledge-Based Systems 6*, 4 (1993), 258-266.

[33] ACTF. Accessibility Tools Framework Project Home. http://www.eclipse.org/actf/.

[34] Droit-Volet, S. and Gil, S. The time−emotion paradox. *Philosophical Transactions of the Royal Society B: Biological Sciences 364*, 1525 (2009), 1943 -1953.

[35] Hornbæk, K. and Law, E.L.-C. Meta-analysis of correlations among usability measures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2007), 617−626.

[36] Sloan, D., Atkinson, M.T., Machin, C., and Li, Y. The potential of adaptive interfaces as an accessibility aid for older Web users. *Proceedings of the 2010 International Cross-Disciplinary Conference on Web Accessibility (W4A)*, ACM (2010), 35:1−35:10.

[37] Bigham, J.P., Prince, C.M., Hahn, S., and Ladner, R.E. WebAnywhere: a screen reading interface for the Web on any computer. *Proceedings of the 2008 International Cross-Disciplinary Conference on Web Accessibility (W4A)*, ACM (2008), 132−133.