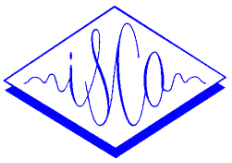

6th International Conference on Spoken Language Processing (ICSLP 2000)
Beijing, China
October 16-20, 2000

THE CONSTRUCTION OF SPEECH OUTPUT TO SUPPORT ELDERLY VISUALLY IMPAIRED USERS STARTING TO USE THE INTERNET

Mary Zajicek

The Speech Project, School of Computing and Mathematical Sciences,

Oxford Brookes University

ABSTRACT

This paper is concerned with the construction of dialogues for speech output with large text, in a Web browser for elderly visually impaired people. Studies made in our research group, show that elderly visually impaired users experience great difficulty in getting going on the Internet when they have no previous experience in computing [6]. Age Associated Memory Impairment (AAMI) is a major problem for these users [2]. This coupled with their visual impairment demands a new dialogue interface modality.

1. INTRODUCTION

This paper explores interface design issues for speech dialogues, which can affect the uptake of Information Technology by elderly users with impaired memory. It is particularly concerned with issues concerning the complexity of functionality presented to a first time user of the Internet and supporting strategy building at the interface.

The paper describes a special enhancement to BrookesTalk for elderly users with impaired memory, which in essence, tells them where they are in their interaction and what they can do next. The content of synthetic speech sentences designed to help users in their interaction and the dialogue design, is discussed.

2. STANDARD BROOKESTALK

BrookesTalk is a stand-alone Web browser for visually impaired people developed as part of the Speech Project at Oxford Brookes University [1]. It can be used in speech output only mode or with a large text window, which displays the words as they are spoken. The size of text can be adjusted to suit the user. BrookesTalk also incorporates a standard visual browser so that visually impaired people can work with sighted colleagues. Consequently a range of visual impairment can be accommodated.

The innovative feature of BrookesTalk is the way it enables partially sighted users to move quickly around the Web searching for information in the same way as sighted users. BrookesTalk gives information on the page from different view points, headings, links, keywords and provides an abstract [7] so that blind and visually impaired users do not have to listen to the whole page in order to decide whether it is useful to them. Our solution was inspired by experiments with sighted users, which showed that in assessing the usefulness of a page they looked first at images then links and then heading, before reading the text.

Abstracting is designed to imitate the visual scanning of pages that sighted users do. BrookesTalk also has a special search facility, which retains the real results when a search has been submitted, separate from all the other information on a page. These facilities are designed to compensate for the lack of visual clues and the requirement of using synthetic speech which is slow and often tedious to listen to.

Keys used in BrookesTalk are:

F1: load a URL
F2: perform a search
F3: hear the headings in the current page
F4: Hear the links in the current page
F5: hear the jumps in the current page
F7: change browser settings
F8: hear more details about the options available
F9: hear the keywords in the current page
F11: hear summary of the current page
F12: hear abstract of the current page
Shift F: select from results of the last search
Shift F4: select a link from the current page
Shift F2: select search results from the last search
Right arrow: read out the page in the mode selected
Back arrow: go back in the page in the mode selected
Up and down arrow: cycle through word mode, sentence mode, document mode
B: move backwards through pages
F: move forwards through pages

3. ELDERLY MEMORY IMPAIRED USERS

The problems experienced by elderly visually impaired users were demonstrated initially during trials on BrookesTalk. BrookesTalk was distributed free to over 200 blind and visually impaired users and an evaluation of how users were interacting with the browser was performed using on-line questionnaire and follow-up telephone interview [5].

Browser uptake by elderly visually impaired first time users was very disappointing, 82% of this group were unable to get up and running on the Web. Analysis of their interaction showed that they were unable to build useful conceptual models of the functionality of BrookesTalk or of the workings of the Web. Their confidence in making the decisions needed for the

construction of conceptual models was low and they became confused and frustrated.

Some subject users were unsure as to the functionality of a link. Sighted users are able to see the link, how it is placed in the page, and how it relates to other text on the page. They can also follow the link to reinforce their concepts and easily return to their original position. It is not as easy for visually impaired users to try a link out, see what happens, and learn from the experience. In addition, elderly users find difficulty in remembering sequences of actions they have previously performed [4].

These users also found difficulty in understanding the way a computer application works. Some 'borrowed' the model of a video recorder and expected one press of a button to make everything 'happen'. They were afraid that they would 'break' the software if they did something wrong. The concept of dialogue and learning to use a language at the interface through trial-and-error was very new to them. Impaired memory as described below seriously interferes with exploratory activity that involves remembering many combinations of actions and outcomes.

Poorly developed conceptual models of the Web, as distinct from the browser, also form a major impediment to successful Web use, for blind and visually impaired users. Sighted users rely on complex and contextual, conceptual models and many visual clues to help them find information on the Web [7].

Four important inter-relating conceptual models were identified for Web search.

- The model of the workings of the search engine.
- The model of the results page of a given search engine.
- The model of the Web site being visited.
- The model of the page.

4. ENHANCED BROOKESTALK

It is apparent that many factors affect the take up of the Internet by elderly visually impaired users [6]. However memory loss and visual impairment in are the two factors addressed in the enhanced BrookesTalk system. Users' perception of the Internet and their levels of confidence and technology acceptance [5] are important factors, but not easy to control at the interaction level. It was also assumed that the subject users have a reasonably high level of technology acceptance as they, or their family, had originally requested a trial copy of BrookesTalk.

At the interaction level difficulties can be attributed to two interrelated factors which interfere with conceptual model development, age associated memory impairment and visual impairment both of which reduce the user's ability to benefit from visual clues and contexts. To accommodate memory loss and visual impairment, a speaking front end was built onto BrookesTalk. The idea is to support the user in their construction of conceptual models by 'talking' them through their interaction.

For each possible state of BrookesTalk an optional spoken output is provided. The user is informed as to where they are in the interaction and which actions are possible at this point. Optional further details are also available to describe the consequences of each action. After listening to the message the user chooses an option, presses the appropriate function key and then receives another message describing the new state of the system.

For example the spoken output for those who have just started up BrookesTalk would be:

'Welcome to BrookesTalk your speaking web browser. There is currently no page loaded. Would you like to:
Enter the URL of a page, press F1
Start an Internet search, press F2
Change the settings of the browser, press F7
Hear more details about options available to you, press F3
Repeat the options, press return'

Under the 'Hear more details ...' it was possible for the user to access further expanded details which describe the options in more detail and in a language further removed from computer terms. An example is given below.

' You have just started the browser and have no page loaded. You might want to first change the setting for the browser such as the speed of the voice, the size of the text and other things which affect how the browser works, or you might want to load a page.
Which of the three would you like to do: ……'

With these messages reinforcing the users' knowledge of the state of the system and explaining to them what they can do next, it was hoped that the development of conceptual models will be supported through repetition and that the user will no longer need to rely on memory. The user can function initially with virtually no conceptual models at all, by using the system in a similar way to a telephone answering system and simply responding to questions. The aim of the speaking front end was to familiarize the user with the steps needed to achieve Web interaction goals so that eventually the spoken instructions would be superfluous and the user would 'know' which function key to press for the required result.

5. ENHANCED BROOKESTALK AND THE ELDERLY MEMORY IMPAIRED USER

The system described above was used in a small pilot study to collect first impressions of the effect of the further explanation and to gain better insight into the way elderly visually impaired users reacted to the dialogues.

5.1 Concurrent Synthetic Speech and Large Text Output

The aim in providing synthetic speech output concurrently with a large text version of the speech output, was that they would complement each other. However it has become apparent that the two senses, sight and hearing, are competing. Elderly users said they were either looking at the text or listening to the synthesized speech but not attending to both. Most users preferred to read the text if their sight was good enough.

The issue of competing sensory channels is an important consideration in the design of multi-modal interfaces. The ability to multi-task is diminished with age [3] and might mean that elderly users should be provided with the means to select only one sensory channel for output. Adjusting to synthesized speech was also a major problem for elderly users. Anomalies in the synthetic speech such as acronyms being spoken as words caused great confusion and irritation.

5.2 Information Overload With Synthesized Speech Messages

The idea behind the talking front end is that users who do not know how to map their task in hand onto the tools available at the interface in terms of actions behind function keys will be told what they can do and how to do it. The hope is that familiarity with the sets of actions needed to operate the browser, will eventually be retained in chrystalline memory i.e. the form memory that is retained in later life [2].

The system is simply providing an explanation of the current state and what can be done next. However it was found that elderly visually impaired users were unable to absorb descriptions such as the one below and then make a choice.

'Welcome to BrookesTalk your speaking web browser. There is currently no page loaded. Would you like to:
Enter the URL of a page, press F1
Start an Internet search, press F2
Change the settings of the browser, press F7
Hear more details about options available to you, press F3
Repeat the options, press return'

There was too much information and they could not remember all the instructions. They were unable to complete a conceptual model as had been hoped and commonly simply chose the last option. The further details option, which describes the options in more detail and in a language further removed from computer terms, as shown in Section 4 failed to provide more useful information or clarify the choice of options. It simply added to the confusion and created anxiety. This demonstrates further that these facilities, which might be of benefit to younger users in helping them to work out what is happening at the interface, are not accessible to elderly users as they rely on the absorption of large amounts of information and good memory to remember the messages.

5.3 How Can We Measure Memory?

A free recall memory test [was applied to the subject users to see if there was a link between, memory levels as shown by the test and their performance at the interface. The observed phenomena, of users' remembering the first options or the last options most frequently, primacy and recency, were seen to hold. We also found a link, in our small sample of eight people, between the score on the memory test and users' performance at the interface.

This means that it may be possible to use this test as a component of a user modeling framework. The ultimate aim of our work is to adapt dialogues to user' s memory levels and other indicators [2] It is not only crucial that some users are provided with a low functionality easy-to-use interface, but those who can manage larger selection sets must be provided with them otherwise they will find the interface intolerably pedestrian.

6. DIALOGUE DESIGN ISSUES

The aim of the speaking front end to BrookesTalk was to support elderly users towards the use of full standard interfaces. This happened with some users [6], but others were not able to absorb the information provided and use it to learn how to interact with a computer interface. It is clear from observation of some users struggling to recall long synthetic speech messages, that for these people, messages must be simpler and shorter since these elderly users cannot absorb or remember large amounts of information.

BrookesTalk as described in Section 2, is operated using twelve function keys. In the enhanced system used for experimentation, we divided the function keys into two conceptual groups. The first group contained functions that were involved with page retrieval i.e. F1 ' Enter a URL' and F2 ' perform a search' . The second group contained functions that were concerned with different views of the page once it was loaded i.e. F3 ' hear the headings in the current page' , F4 ' hear the links in the current page' , F5 ' hear the jumps in the current page' , F11 ' hear summary of the current page' , F12 ' hear abstract of the current page' . F7 change settings was made available at logon only.

6.1 Conceptual Groupings

Grouping functions provides for a smaller set of function keys to select between at any one time and therefore shorter messages and less to remember. However conceptual groupings of options in speech output interfaces rely on the user understanding the concepts behind the groupings in order to know where the functions may be found. This knowledge cannot be assumed with first time Internet users and is not required of sighted users making selections on a Graphical User Interface as they can see the selection and refer back to it at any time.

It was observed that a smaller number of selections in any one message makes operation easier. However a smaller numbers of selections means a larger number of groupings with all the attendant conceptual problems. The ideal choice is a selection between two functions with the most commonly required option appearing last.

6.2 Reducing Functionality

It has been decided that as reduction in selection is paramount, the functionality of the system must also be reduced. The advantage gained by presenting several different views of the page is lost when the user cannot even load a page. There is little gain in providing a wide range of views of the Web page to elderly users if they cannot visualize them.

7. MODIFIED DIALOGUES

The reduced functionality version of BrookesTalk has been stripped down to,

F1: load a URL
F2: perform a search
F4: hear the links in the current page

Users are now unable to change settings and there is only one way of reading the page (document mode)

Unfortunately the majority of the innovative aspects of BrookesTalk have been removed. However it must be remembered that our aim is to enable users to start using the Internet and build conceptual models that will sustain their interaction. It is hoped that users would progress from the reduced functionality version of the browser to full functionality when they had serviceable conceptual models in place.

Interaction under the modified dialogue is much simpler. The speech output that corresponds to the full functionality version in Section 4 becomes:

'Welcome to BrookesTalk your speaking web browser. There is currently no page loaded. Would you like to:
Enter the address of a page, press F1
Start an Internet search, press F2
Repeat the options, press return

For the page loaded state the output would be:

You have a page loaded would you like to:
Read the page, press the down arrow key
Read out the links on the page, press F4
Load another page, press A
Repeat the options, press return

8. CONCLUSIONS

The uptake of elderly visually impaired users with poor memory to Information Technology is challenging as they have a short supply of precisely those skills, which are necessary for the use of standard interfaces. Visual impairment usually means that the user must use speech output dialogues which rely far more on memory than graphical interfaces.

This paper has presented an introduction to interface design considerations in speech systems for those with poor memories. The author's work indicates that the best solution for these users is to start on a system with reduced functionality in order to get up and running with a view to moving onto a higher functionality system when strategies are retained in chrystalline memory. The current strategy of the Speech Project is to pilot different dialogue designs on a system with reduced functionality and evaluate the results.

9. REFERENCES

1. http://www.brookes.ac.uk/speech

2. Park, D.C, (1998) Ageing and memory: mechanisms underlying age differences in performance. Australian Journal on Ageing. 17(1), 69-72

3. Stuart-Hamilton, I. (1999) Intellectual changes in late life. In Ed Woods, R.T. Psychological Problems of Ageing Chichester, Wiley

4. Wilkniss, S.M, Jones, M.G., Korol, D.L., Gold, P.E. and Manning C.A. (1997) Age-related differences in an ecologically based study of route learning. Psychology and Ageing 12(2) 372-375

5. Zajicek M., Arnold A., (1999), The 'Technology Push' and The User Tailored Information Environment, *5th European Research Consortium for Informatics and Mathematics, Workshop on 'User interfaces for all'*, Dagstuhl, Germany

6. Zajicek M., Hall S., Solutions for elderly visually impaired people using the Internet, Proceedings of HCI 2000, Sunderland 2000

7. Zajicek M., Powell C. and Reeves C. (1998). A Web Navigation Tool for the Blind, *In Proc. 3rd ACM/SIGAPH on Assistive Technologies*, Los Angeles, USA.