Voice-Controlled Search User Interfaces for Young Users

Tatiana Gossen, Michael Kotzyba, Sebastian Stober, Andreas Nürnberger
Data & Knowledge Engineering Group, Faculty of Computer Science
Otto-von-Guericke University, Germany
http://www.dke.ovgu.de/

ABSTRACT
Children need advanced support during web search or related interactions with computer systems. At this point, a voice-controlled search engine offers different benefits. Children who have difficulties in writing will not make spelling errors using a voice control. Voice control is a natural input method and is supposed to be easier to use for children than a keyboard or mouse. To integrate a suitable voice control into search engines, it is necessary to understand the children's behavior. Therefore, we investigate children's speech patterns and interaction tactics during a web search using a voice-controlled search engine. A user study in form of a *Wizard-of-Oz-Experiment* was conducted and we found out that children are motivated to use voice-controlled search engines. However, voice control in combination with touch interactions should be possible as well. Furthermore, the analysis of the speech patterns suggests that it is possible to build a speech recognition program. The results of this study can serve as fundamentals to develop voice-controlled search dialogues for young users.

Author Keywords
Web Search Interface; Children; Voice Control; User Study.

ACM Classification Keywords
H.3.3. Information Storage and Retrieval: Search Process; H.5.2. User Interfaces: Voice I/O

General Terms
Human Factors.

INTRODUCTION

During a (web) search, children often need to be more supported and motivated than common users like adults. Children can get easily frustrated if they do not find relevant results, do not understand the search engine output or if a failure emerges [2]. The fact that children also have difficulties to evaluate the relevance of retrieved documents to their information needs [6] aggravates this. Furthermore, most children have difficulties with typing [3]. They are not able to type commands without looking at the keyboard (touch-typing). Instead they typically hunt-and-peck on the keyboard for correct keys. By looking at the keyboard while typing, children often do not spot spelling mistakes. In addition, some interaction techniques like scrolling or drag-and-drop are difficult for young users [3]. Therefore, young users would benefit from support mechanisms to formulate their information need and interact with a search user interface.

In order to tackle these problems, a voice-controlled search interface can be used that allows a voice interaction in both directions, i.e., voice input and output. By using speech recognition, the user does not need to be good in writing. Besides, the interaction with a voice-controlled system can be more intuitive and hence more motivating for children as they do not have to learn the cumbersome interaction with mouse and keyboard. These advantages motivate us to investigate voice-controlled search user interfaces (SUI) for young users.

Voice control offers further advantages. There are methods to extract emotions from the users speech (e.g. [7]). Hence a speech interaction can provide us with the necessary information about the emotional and dispositional state of the user. Thus, if a child feels unhappy, e.g., caused by a search failure, his or her emotional condition could be recognized using the voice information and the system can initiate countermeasures. User emotions and dispositions were considered in information retrieval before, however mostly for relevance assessment [1] or for the enrichment of an object descriptive feature space, e.g., in collaborative filtering [9, 8]. Our long term goal is to take the emotional states of young users into account in order to provide a better support from the SUI side. To our knowledge, voice control has not been studied before in the context of search engines for children (see [5]).

In order to develop a well-functioned voice-controlled search engine for young users, it is necessary to analyze users' acceptance towards a system of this kind and how they would use it. One essential part of this is the user behavior. Therefore, in this paper we investigate children's speech patterns and interaction tactics that are used by operating a voice-controlled search engine. For this, a user study was conducted in form of a *Wizard-of-Oz-Experiment*. During this experiment a user interacts with a program that seems to be autonomous, but is remotely controlled through a hidden person (the wizard) instead. This method allows to study users' behavior without any limitations and even appropriately react to unexpected user actions.

HCIR '13, Vancouver, BC, Canada, Oct 3-4, 2013

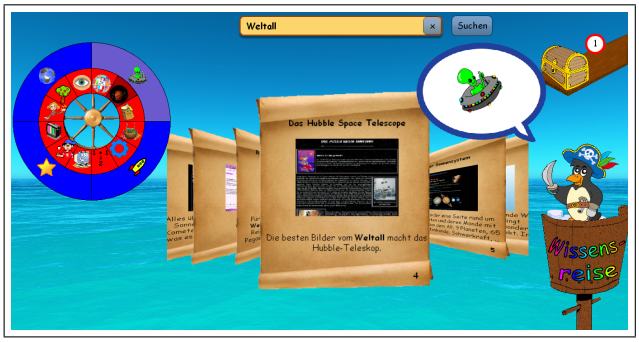

Figure 1. Screenshot of the *Knowledge Journey* user interface: a guidance figure and a treasure chest on the right hand side, query input elements on the top, a navigation menu on the left hand side and a coverflow with search results in the middle.

Figure 2. Screenshot of the user interface: journey journal with favorite web pages.

In order to illustrate the work and its results, the paper is structured as follows: In Sect. *Preliminaries* we briefly describe the child-centered search engine we have used for this study. Sect. *User Study* contains information about the user study procedure and its results. The last section provides a summary and outlines of the future work.

PRELIMINARIES

For this study, we used the child-centered search engine *Knowledge Journey* [4]. The search engine uses the metaphor of a treasure hunt, where a user takes a journey to gather relevant search results. In Figure 1 the interface of the *Knowledge Journey* is illustrated. The interface consists of five groups of elements: a guidance figure (here a penguin pirate), a treasure chest, a coverflow visualization of results, elements for keyword search and a pie-menu for navigation.

Using the query input element, a standard keyword search can be done. The navigation menu helps children who have difficulties in query formulation. This menu is illustrated as a pie menu placed on a steering wheel. It contains different categories and subcategories like sports, nature, persons or entertainment and by choosing a category its name is automatically spoken by the system. The third element is a coverflow visualization of search results. Coverflow view of results offers an attractive browsing animation. Its central element is clear separated from the rest and a child can concentrate on one item a time, thus resulting in a small cognitive load. Using the last element, a treasure chest, the user can store relevant search results. The treasure chest contains a journey journal where for each stored webpage its thumbnail, a textual summary and a title are shown (Figure 2). Furthermore, the search engine contains multimedia elements, large fonts and buttons in the user interface design to make it more attractive for children. It is also easy to operate for children with click-and-point interactions. The guidance figure supports children's search process in order to avoid frustration, e.g. by providing a spelling correction after a misspelled query is submitted.

Figure 3 depicts the possible user interactions with the Knowledge Journey. Overall, the SUI offers many ways for users to interact with it. In the following, we present our user study which was conducted in order to study children's speech patterns and interaction tactics during a web search using the Knowledge Journey via voice interaction.

USER STUDY

In this section we describe the design, results and implications of the user study we conducted.

Study Design

The goal of the user study was to examine how children would interact with a voice-controlled version of the search engine *Knowledge Journey* and what voice commands and interaction patterns they would use. Therefore, a study in form of a so-called "Wizard-of-Oz-Experiment" was designed. Here the user assumes, that he or she interacts with a fully autonomous application but actually it is controlled by a hidden operator. In general, two investigators were involved in the study. One was the "wizard", whereas another one conducted the user study with participants by interviewing them, giving instructions etc. The study procedure consisted of the following four steps:

Pre-interview: We used a pre-interview to gather the user's demographic information and their experience with computer systems and the Internet. Besides, we asked the participants about their experience with different input methods, i.e., keyboard, mouse, touch and voice control.

Introduction: In the next step, we gave the participants an introduction to the Knowledge Journey and how to use it. We briefly explained what the different elements are and what purpose they have. Intentionally, we gave no information about how to use the interface per voice.

Search Experiment: In this step, the actual search was done. Children could execute a free, explorative search where they were able to look for everything they liked and use the elements how they want, but only using voice commands. If a young user had no idea for what or how he or she should use the system, an investigator gave some assistance, e.g., "Currently its Christmas time and there are a lot of things one can do during this time. Maybe you can search for these things?". If a child used an ambiguous command or the system (the Wizard) could not understand the command, a prepared audio message "I cannot understand you." was triggered.

Post-interview: The last step was a post-interview to evaluate the users' attitude towards the system and if they would use

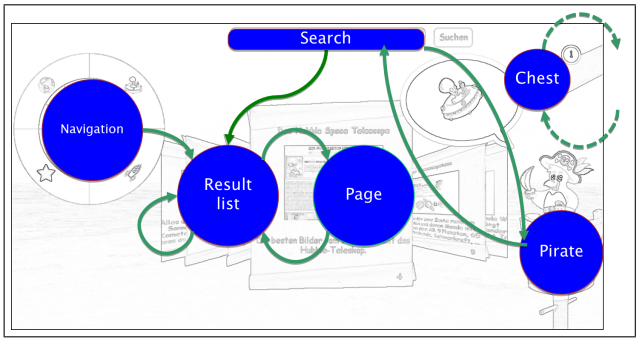

Figure 3. Interaction graph of the *Knowledge Journey*. A user can search for results, navigate using the pie menu, interact with the coverflow, interact with a result page, interact with a guidance figure and operate the treasure chest.

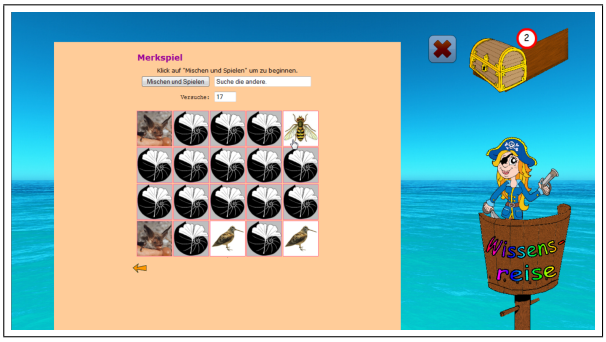

Figure 4. Screenshot of an opened webpage with a memory game: For children, it is difficult to name a certain card and its position.

voice-controlled SUI in the future. We also asked for recommendations to improve our search engine. Each session with these four steps took about 30 minutes.

Study Results

The user study was conducted in December 2012 at the trilingual international elementary school in Magdeburg, Germany.[1]

Participants: Our ten participants were of age eight to ten (average 8.8 years), seven female and three male. They were mostly third (8 children) and fourth (2 children) grade pupils. All children had experience with computer systems and the Internet. Eight children could easily and one child had some difficulties to handle mouse and keyboard, one child mainly used touch devices. The frequency distribution of the Internet usage is: everyday (1 pupil), two-four times a week (3 pupils), once a week (1 pupil), once a month (4 pupils), less than once a month (1 pupil). We noticed no significant correlation between the frequency of usage and age or school grade. Four participants use the Internet without any supervision, three participants together with relatives and three participants do both from time to time. The children use the Internet mostly to play online games and watch videos on Youtube, but also to search information for school. These activities were mentioned by almost all participants. Some of them also write messages or look at Amazon or Ebay. In order to search for information, eight participants use Google and five of them also use web search engines for children.

Patterns and Tactics: In the following we describe the command patterns, tactics and the response of the young users to the interaction per voice control. Furthermore, we address the emerged difficulties. All children enjoyed to interact per voice control. This is a good sign as their motivation to use the search engine increases. They perceived as beneficial, that queries do not need to be written using a keyboard. However, voice control was perceived as unusual by 90% of the children. Especially at the beginning of the study, the young users were overstrained a little bit because they did not know what they can actually say to the system. The researcher had to emphasize that a child could tell anything and that there are no restrictions.

Nearly half of the children interacted with the system in a very polite way, e.g., “Please enlarge” or “Could you open the treasure chest please?”. But with time they changed the interaction tactic and used only pragmatic, relatively short voice-commands. Eight children used almost exclusively *elliptical constructions* for commands where a word or phrase implied by context is omitted from a sentence, e.g., “close”, “next page” or “go back”.

If an explicit keyword input per voice was done or if a category in the navigation menu was selected per voice, then half of the children also continuously used the descriptive terms for the UI element, e.g., “I’m searching for ...” or ”Select calculations on the steering wheel.”. Users who actively used the storage function, also used the specific terms “store” and/or “(in) treasure chest”. This behavior is very fortunate in order to detect possible voice commands automatically. Different children also used different words, i.e., synonyms, for the same controls. For example, to flip through the search results, six children used relative descriptions like “continue” or “next page” and the others used an absolute description like “second page”. Hence, the system also has to support alternative voice commands for each interaction.

The commands used during the search were very often ambiguous. For example, a child previously used the search bar or the navigation menu and than he or she used “space” for the next command. It was unclear, whether the pupil intended to search for it or wanted to select a corresponding menu category. For such commands a context, i.e., the previously used element, is important to carry out the commands without errors. The system has to take the context and the user interaction sequence into account to interpret the user correctly.

Nevertheless, some children’s interaction tactics were not always conform with the original interaction graph of the *Knowledge Journey*. For example, while being within the “treasure chest” participants wished to start a new search skipping the command of closing it first. For some of the children, it was extremely difficult to name “clickable” elements that only contained pictures within web pages or on the SUI surface. For example, Figure 4 shows a website with a memory game that was opened by one of the pupils. It was

[1] All the relevant agreements from the caretakers had been obtained. The Parents agreed in advance, that their children can participate in the study.

problematic for this child to name a certain card and its position, in particular the cards in the center presented the greatest difficulty. While the card at the borders could be named like "upper left" or "second row right", the cards in the center, which require numbers for both rows and columns, were difficult to explain. Also the thumbnails for stored elements in the treasure chest (see Figure 2 at the bottom) were difficult to select. Children often named an object within a certain thumbnail to access a website. In future, this problem could be solved by labelling those elements with numbers or symbols. Another solution is to combine the voice-controlled search engine with touch input. Thereto, 70% of the young participants explicitly told us that they would rather use touch to accomplish certain tasks directly on the user interface. But, they would like to use voice control for keyword search as it is faster than using a keyboard.

In order to get a deeper insight into the children's commands, some examples of the used voice controls are listed in Table 1.

Interaction	Voice Command
Search for	- I'm searching for Lord of the Rings movies - I would like to ahem ... search for ... YouTube - I would like to ahem ... I'm searching for a ... for ahem information about - put animals in above - hmm horses - the stars
Menu navigation	- I'd like er to go to the bag - at culture er at history I'd like ahem to the steering wheel at the volcano - I would like ... (click) on st at in steering wheel to ahem the calculation task - click nose below - steering wheel tree - once again to the tree - that where the man is running
Next page	- I would like to see the second page please - I would like to see the second page - page five - the next
Open a search result	- open the (web) page - I want that you open it - click on it - please enlarge - show me

Table 1. Examples for young user's voice commands.

CONCLUSION AND OUTLOOK

In this paper, we described a user study that was conducted to investigate children's speech patterns and interaction tactics during a web search using a voice-controlled search engine. The results of this study indicate that a voice controlled search user interface in combination with touch can increase the usability of web search engines for children. However, this combination has to be investigated in future work as the participants in this study were only able to use a voice control and no touch interaction was possible. In order to develop completely voice-controlled search dialogues, further conceptual adaptations are necessary, in particular to provide young users with suggestions of possible voice interactions. Our long term goal is to take the emotional states of young users into account in order to support search processes, e.g. by providing users with motivation in case of failure.

ACKNOWLEDGMENTS

We would like to thank the participants of this user study and the employers of the trilingual international elementary school in Magdeburg. This work was supported by the Transregional Collaborative Research Centre SFB/TRR 62 "Companion-Technology for Cognitive Technical Systems" funded by the German Research Foundation (DFG).

REFERENCES

1. Arapakis, I., Athanasakos, K., and Jose, J. M. A comparison of general vs personalised affective models for the prediction of topical relevance. In *Proc. of the 33rd international ACM SIGIR conf. on Research and development in IR*, ACM (2010), 371–378.
2. Bilal, D., and Kirby, J. Differences and similarities in information seeking: children and adults as Web users. *Inform. Proc. & Management 38*, 5 (2002), 649–670.
3. Budiu, R., and Nielsen, J. *Usability of Websites for Children: Design Guidelines for Targeting Users Aged 3–12 Years, 2nd edition*. Nielsen Norman Group Report, 2010.
4. Gossen, T., Nitsche, M., and Nürnberger, A. Knowledge journey: A web search interface for young users. In *Proc. of the Sixth Symposium on Human-Computer Interaction and Information Retrieval (HCIR 2012)*, ACM (2012).
5. Gossen, T., and Nürnberger, A. Specifics of information retrieval for young users: A survey. *Information Processing & Management 49*, 4 (2013), 739–756.
6. Jochmann-Mannak, H., Huibers, T., Lentz, L., and Sanders, T. Children searching information on the Internet: Performance on children's interfaces compared to Google. *SIGIR'10 Workshop on accessible search systems* (July 2010), 27–35.
7. Krell, G., Glodek, M., Panning, A., Siegert, I., Michaelis, B., Wendemuth, A., and Schwenker, F. Fusion of fragmentary classifier decisions for affective state recognition. In *Multimodal Pattern Recognition of Social Signals in Human-Computer-Interaction*. Springer, 2013, 116–130.
8. Moshfeghi, Y. *Role of emotion in information retrieval*. PhD thesis, University of Glasgow, School of Computing Science, 2012.
9. Moshfeghi, Y., Piwowarski, B., and Jose, J. M. Handling data sparsity in collaborative filtering using emotion and semantic based features. In *Proc. of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, ACM (2011), 625–634.