C

Consiglio Nazionale delle Ricerche

Search engine UIs: remote usability testing with blind persons

P. Andronico, M. Buzzi, C. Castillo, B. Leporini

IIT TR-15/2005

Technical report

Luglio 2005

Istituto di Informatica e Telematica

Index

Abstract

This paper describes a remote usability testing which was the final phase of a research project aimed at improving usability of web search tools for blind users who interact via screen reader and voice synthesizer.

The testing aimed to evaluate a new implementation of Google user interfaces – according to a set of criteria previously proposed specifically for search engine user interfaces - for the simple search and the result exploration. To prepare the environment for the remote testing we needed to re-implement the original Google interfaces, using Google APIs, PERL programming and XSLT transformations.

The results of the testing highlight how Google interfaces, although accessible, may be further improved in order to simplify the interaction for the sightless.

In this article, first an overview of the project is introduced; then we discuss the design and implementation of the UIs. Finally, we describe in detail the usability testing which involved 12 totally blind persons.

Keywords

Usability Testing and Evaluation, User Interface Design, User Research, User studies, User-Centered Design.

Project statement

Due to the growth of the Internet, search engines have become a very common tool for retrieving information; thus a simple and effective interaction is important for all and especially for blind individuals who interact by a screen reader, and encounter more difficulties in web interface navigation. Although legislation in the last few years aims at assuring accessibility for public on-line Web sites and services, digital barriers are very frequent in the private sector [14].

We believe that the user interface design (UIs) is crucial for improving accessibility and usability. Therefore, in a previous phase of our research, based on international standard guidelines and on the personal experience of one of the authors in using screen readers and voice synthesizers, we proposed a set of specific criteria for improving interaction with search engine user interface by the blind [12].

To evaluate the validity and impact of these criteria on UI, we applied them to a popular search engine: Google (http://www.google.com), which offers simple interfaces and effectiveness in search results. Our goal was to demonstrate that it is possible to have a graceful UI look&feel assuring satisfaction and efficiency of use. We decided to maintain the original graphical layout of Google, to avoid any impact on the interaction of sighted users. The goal of our study was to demonstrate how a simple interface can be improved by preserving the same (or very similar) visual layout.

Once the prototype was developed, tested and revised, we decided to measure "concrete" results against our goals. Thus, we set up a user testing in order to collect feedback concerning the original and modified Google interfaces.

In the following we first introduce the project and related work. We then briefly outline the criteria that inspired our design and describe problems encountered during the design and implementation of the new interfaces. Finally we describe how the remote testing with the blind was conducted, and discuss the results.

Background

The project was carried out at the Institute for Informatics and Telematics, which is part of the Italian National Research Council. The project was born as curiosity driven research, i.e. a group of persons spontaneously decided to

join their different expertise to investigate common interest. The Project began in January 2004 and will conclude at the end of 2005.
The team was composed of:

- A Technologist: project Manager,
- A Web Designer: communication and usability practitioner,
- A PhD Researcher: main accessibility investigator,
- Another PhD Researcher who joined us in October 2004 as programming developer.

In different phases of the project each member of the team collaborate in multiple tasks thus a clear separation between tasks and roles did not exist.

In the following points we summarize steps and times of our study:

- We began following a student who was preparing his bachelor's degree thesis (laurea) in Information Engineering at the University of Pisa. His study concerned search tools (search engine, meta-searches and directories), as well as the checking of accessibility with automatic tools and manual verification to compare the most important accessibility problems in the seven different products analysed.
- As an initial step, we conducted a survey aimed at understanding problems that both blind and sighted persons encounter using search tools. The questionnaire preparation, data collecting, analysis and comparison: 2 estimated person months;
- Based on international standards, experience in using screen reader and voice synthesizer (of one of the authors) and usability observations, we suggested a set of criteria for improving accessibility and usability of search engines. Estimated person months: 2;
- To apply the proposed guidelines to Google we needed to redesign the interface structure: i.e, re-arrange the logical position of the page elements, defining new logical sections. This implied rewriting the entire code of the pages, using XHTML and CSS, and adding features for keyboard interaction. This step (which defined the new "static" code of the interfaces) required 3 estimated person months;
- To render the interface code dynamic we designed the interaction with Google. We chose to interact with the Google APIs via web server (as discussed in the following). Thus we needed to implement the prototypes for the simple search and result page both for the "original" and the "modified" Google interfaces. The environment set-up, programming and testing of the code took 3 estimated person months;
- The usability test conducted with 12 blind subjects (discussed in the following) required 3 estimated person months.

Challenge

The main limit of the project was our budget. The cost of the IIT personnel involved in the study was covered by CNR. We had about 12.000,00 Euros, which was a residual amount from a previous project, to cover costs of this research; we used this amount for one contracted staff member (4.5 months/person for a graduate PhD Researcher) and for the visit of a PhD Researcher. The rest of the budget was used to present results in international conferences.

Despite this low budget, members of the team enthusiastically collaborated on the project and this spirit helped us to overcome any difficulties. All project team members had their regular activities, so the time devoted to this project was free of any constraints and whenever possible.

Related work

Search engines are particularly difficult for a blind person to use, since difficulties in Web navigation [5] add to the complexity of the search engine's interface and functions. Specifically, for people using a screen reader actions take longer and tasks are more difficult since additional actions are required. The Manchester Metropolitan University highlighted issues of non-visual access performing experiments on a sample of blind and visually-impaired users who carried out four information-seeking tasks, including the use of search engines [3]. The gap between blind and sighted users' efficiency when performing online search tasks is quantified in [8], where, when executing a set of tasks, blind participants took twice as long as sighted users to explore search results and three times as long to explore the corresponding web pages. In this study authors, aimed at identifying page features that could be presented in results, and the circumstances might help users to decide whether to explore search results or not. In most cases, participants expressed a desire for additional page features, which varied depending on their visual ability and their ability to specify criteria for controlling the order of results (ranking). Authors also suggest various ways the user's search experience to improve.

In recent years, although numerous tools for automating checking of conformance to accessibility guidelines have been developed and publicly made available, practically speaking there was no consequent qualitative improvement, on a large scale [20]. In addition, often accessibility is not enough for a good interaction through special devices: usability issues should be designed to improve the navigation via screen reader. Takagi et al. [20] highlighted the need to spend more attention on practical aspects of usability for improving productivity of the blind. The syntactic checking of Web pages, in fact, does not evaluate "time-oriented" usability factors, such as the speed to reach target content or complete a certain task, the ease of understanding the page structure, and the interface navigability. To better understand and then overcome usability limits of web interfaces, authors designed and implemented a tool for helping developer. The tool in fact allow loading a page and moving the mouse over sections to see the "reaching-time" i.e. the time necessary for the screen reader to announce this content.

Usability evaluation methods include several techniques: heuristics evaluation, cognitive walkthroughs, guideline evaluation and usability testing. Jeffries et al. [11] compared these approaches showing that heuristics evaluation enable one to find the majority of problems at low cost while usability testing are second in effectiveness but bring higher cost, generated by travel expense (to bring users to a central laboratory or to send usability professional to users work setting). In recent years usability testing have been “remotely” executed frequently. Hartson et al. define remote evaluation to be usability evaluation wherein the evaluator, performing observation and analysis, is separated in space/or time from the user [7]. Thus different techniques such as videoconferencing, automatic logging of user paths and tasks, specific tool for usability test have been applied to monitor and analyze data collected by remote testing [19], [6], [7]. Ivory and Hearst grouped automatic usability evaluation methods along their features (method class and type, type of automation and effort level) [9].

Recent studies carried out a comparative analysis showing that remote testing is an effective if not better than traditional testing executed in the laboratory [21].

The Research Process

In our study we focused the analysis on the interaction with search engines user interfaces. According to Nielsen [16], users know very well their expectations in the behavior of a search. In fact, people are looking for at least three main components: a *box*, where performing the search, a *button* "search" to be clicked for carrying out the query, and a *list* of search results - appearing in a new page - that Nielsen calls "the search engine result page" (SERP).

Very often, these three components are not so clear in a search tool UI, especially for a sightless person using assistive technology. Sighted persons may also have difficulty understanding the label on a search button, or even going through the list of the results and all the information presented.

Our design was completely user-centered, thus we took into consideration the needs one might encounter while exploring a search engine UI.

The analysis was restricted to Google interfaces for text retrieval (not for news, image searching, or directory exploration).

In our previous study on search engine user interfaces, we proposed a set of guidelines to improve users' work when performing a query. We focused our work on the main problems encountered by a blind person while exploring a web page with the voice synthesizer of a screen reader. The effort required is very difficult and too much time is spent navigating the interface without reaching the important part.
Our criteria may seem a little too empirical, but in a previous, informal test – also conducted by one of the authors – we found an important improvement. They are mainly concerned grouping the most important part of the interface and repositioning them in a more appropriate way in the code, while maintaining the original "look&feel" given by the visual designer. We worked with Google for several reasons:

- It is the search tool most widely used by blind people in Italy;
- Even if the UIs are quite simple, they are not totally compliant with the W3C WCAG guidelines [22];

We can group the guidelines into two main categories:
- General considerations:
    - **Easy location and labeling of edit field and search options.** Place edit fields, option buttons and any other search element at the top

of the web page; avoid secondary elements (links, texts, banner frames, etc.).

- **Navigating more quickly.** Assign a scale of importance (i.e. by the tab index attribute) so users can reach the most important elements quickly.
- **Alerting by sound.** Different sounds for different events should provide useful information for blind users.

- Navigation in the result page:
  - **Highlighting the search result.** Use a heading level (i.e. <h1> or <h2>…<h6>) at the beginning of the result list; if possible, this heading element should be the first on the page source.
  - **Arranging the results.** Put the list of the result links with their summary, just after the search result notification (nothing else should be located in the middle).
  - **Recognizing sponsored links.** Keep sponsored links separate from the other results.
  - **Adding navigation and help links.** Place the links pointing to result pages at the end of the list (not before).

More details regarding guidelines can be found in [12].

The next step was to implement our guidelines re-designing Google UIs without changing the visual appearance. In particular we wrote the Home Page (simple search page) and the Result Page as described in the following sections of this paper.

Before setting up the environment and the survey for remote user testing, we prepared a scenario to better focalize all the steps a sightless person has to perform when using a screen reader during a search [2].

Solution details

In our project we were faced with two implementation aspects: 1) how to effectively modify the interface code and 2) how to set up interaction with Google to carry out a query. In the following we describe the two corresponding design phases.

Interface code

Restructuring the code implied defining logical sections of the interface by grouping sets of homogenous text and elements together and by giving the user the possibility of jumping rapidly from one part to another. In fact, these logical sections were then structured by heading levels. In addition, some important components were equipped with access keys and could be activated by simply pressing a specific shortcut. Based on this new structure, we also defined an empirical visiting order of links by Tab key. This method – i.e. moving through Tab key - in fact is often used by the blind for positioning more quickly on a relevant part of the page.

We structured the page of query results in seven sections:
- Navigation bar
- Search box and options
- Advanced search and preference links
- Results
- Sponsored links
- Result Pages (previous, next, numbered pages)
- Google links (Google home, advertising programs, etc.).

Thanks to those heading levels, users can get on the fly a kind of “page index” by the Jaws [10] command Insert+f6 (available only in Internet Explorer and Jaws 4.5 or higher). In this way users can see the page content at a glance as well as skip quickly from a section to another.

A finer granularity was applied for the visit via Tab key, according to the following sequence:

1. Search results status (i.e. Results 1 - 10 of about... or no results);
2. First result, Second result, etc. (at this level cached and similar links are skipped. User can access these secondary links of the result explored with the arrow keys);
3. Result pages (Prev, 1, 2,…, Next);
4. Search Tools (i.e. "Search within results" and "Search Tips");
5. Sponsored Links;
6. Searching for (simple search box and options);

7. Advanced Search and preferences;
8. Navigation bar;
9. Google Info and other links;
10. Cached and similar page links.

Practically, an "importance" order is given to the elements (i.e. links) in order to "guide" the user's navigation.

Furthermore, access keys were associated with relevant objects, such as the search edit box, advanced search, "next result" page and "previous result" page links. A navigation help page – reachable by a hidden-link with the shortcut alt+h - was added as support for beginning users on interface features.

Accessory elements such as hidden labels and result numbering were added as well, for further simplifying user interaction and user orientation. And lastly, some sounds were associated to relevant events such as: focus on the search edit field and on the search outcome (two different sounds were defined to indicate "results found" and "results not found").

It is clear that our choices were aimed at optimizing user interaction especially when navigating by keyboard. Search engine companies in fact, should have a different preference for the order in which to visit elements and logical sections, according their needs (for instance sponsored links may be announced by screen reader before those of results). Our point of view took into account user preferences.

As regards the code, we used XHTML and CSS properties for separating the content from its rendering; we replaced tables by using <DIV> elements instead, and the CSS position property for arranging the object in specific areas into the graphical interface.
For details concerning the source code of the modified interfaces and for examples, the reader may refer to [21].

Dynamic generation of the interfaces

There are several alternatives to modifying the layout of a search engine's results before presenting them to the users:

- Capturing the results of the search engine as static HTML pages, and then modifying those pages. This has the disadvantage that the whole experience of querying, browsing, and retrieving is not represented by the experiment.

- Modifying the HTML pages returned by the search engine, by using parsing to locate elements and re-write the page. This is typically discouraged by search engine administrators since it can generate too many requests, and may result in the application being banned from using the search engine. Moreover, the HTML coding of the results is not stable over time.
- Directly accessing a machine-readable version of the search engine's result. This is the option we used.

Programmatic access to the search results by Google is provided by the "Google API", available online at http://www.google.com/apis/. When using this API, messages are exchanged between the application and Google's server using SOAP, a XML-based messaging that uses HTTP for communication. Figure 1 shows the architecture of our implementation.

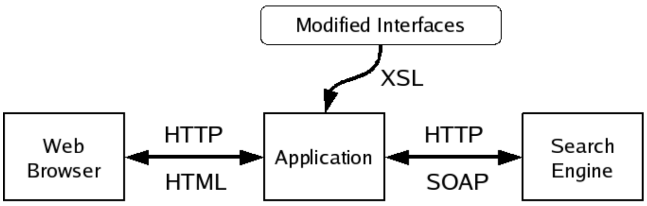

Fig. 1 - Logical architecture of our implementation

SOAP messages are short XML messages describing invocations of remote methods. In the case of the Google API, there is a method for querying, and a data format for returning a response. A typical response can be seen in Figure 2. The "snipped" tag contains a short passage from the text, typically a sentence or paragraph containing the query terms.

```
<?xml version='1.0'?>
<Envelope> <Body>
 <return><resultElements>
  <item>
   <title>Rent a Car in Pisa</title>
   <URL>http://www.auto...co.uk/guides/</URL>
   <snippet>... Each of these ...</URL>
  </item>
  <item>
    <title>City of Pisa</title>
    <URL>http://travel.yahoo.com/pisa</URL>
    <snippet> ... City of Pisa ...</snippet>
  </item>
...
```

Fig. 2 - Short example of a SOAP response from the Google API, re-formatted for clarity

By manual inspection, we found out that the responses from the API are not the exactly the same as the results returned either by google.it or google.com; they may correspond to a slightly older, or smaller, version of the index. This would be a problem if the quality of the responses differed too much from the usual responses, but we found out that this was not the case.

The XML response is not suitable for being shown directly to end-users, and it must be modified. Fortunately, there is a simple language for expressing a transformation of an XML document, called XSLT (eXtensible Stylesheet Language Transformations). Using XSLT, we wrote a small program (a "stylesheet") describing how to format each of the elements on the response (title, URL, and text snippet) to display it according to the guidelines stated above.

The total delays from both the SOAP request and the XSLT transformation were not noticeable, as the SOAP request is sometimes faster than the API when the size of what is transmitted is smaller. For the XSLT transformation we used the Sablotron engine (http://www.gingerall.com/), a fast XML parser and XSLT transformation engine written in C, and the parsing speed was always less than 0.1 seconds, even displaying a complex page.

An interesting point concerning this solution is that modern Web browsers (Mozilla, Firefox, Internet Explorer, Opera) include their own XSLT transformation engines. They are included to be used by applications like this one, in which the Web server transmits the data and the formatting separately, and the Web browser merges both and displays the results. Although at present these technologies are not widely used, we expect that large search engines will start providing pages that are structured differently for different users in the next few years. This study proves that this would be simple to implement, and helpful for final users.

Usability testing

The main goal of the testing was to compare the user experience while performing a query and navigating through the original and the modified Google UIs, in order to understand whether or not our proposals would result more efficient and easy to use by blind persons.

We could have tested the new interfaces with sighted persons using a screen reader and having more expertise in computing and in making queries. According to [13] this method seems to be one of the most effective for testing accessibility problems, whereas blind people in remote testing revealed less. In our opinion in our study we need to test accessibility problems as well as improvement of usability of the interfaces. For this reason we decided to perform a remote test directly with blind subjects, who were free to terminate the questionnaire whenever they wanted.

Why a remote testing

Because of the choice to test only with blind people, the first difficulty we encountered was to find a group of these subjects available near our work location. Another and more important problem is the fact that any sightless person prefers to use his/her own computer with his/her own screen reader. Any set-up or change of version can increase complications during navigation. Users are more comfortable in their own environment.

For this reason as well as the mobility problems due to the blind persons geographical distribution, we decided to opt for remote testing. The choice seemed to be well received also for our group of test volunteers.

Preparing the test

First of all we needed the email addresses of all the participants in the test. After collecting these, we performed a preliminary questionnaire to better contextualize our group of users, with their technical skills, age, educational background, habits in a computer environment, and use of a search tool and a screen reader. This preliminary survey helped us in setting up a code ID for the real test that we then used to cross the two questionnaires for analysis purposes.

Our aim was to identify each participant with an identification number in order to be as anonymous as possible. This fact was really appreciated by the sightless participating in the test, which preferred not to be identified.

The test itself was divided into two phases: some tasks people had to do with the two interfaces (the original and our modified Google UIs), and a questionnaire to fill out afterward concerning difficulties encountered or significant differences between the UIs. The survey, consisting of 22 questions, was divided into three sections: the first requesting information about the subjects experience performing the assigned tasks, the second more about difficulties in making the queries required, and the last concerning navigation issues. The results are discussed in detail in this paper.

Problem with remote testing

There is no doubt that remote testing can present more or different problems than a lab usability test. We contacted several persons interested in helping us with this kind of accessibility evaluation. Unfortunately at the end only 12 persons participated completely in the all phases of the test, maybe due to additional difficulties we did not consider.

Test environment

According Ivory et al. [9] our test may be classified as Remote Testing (the tester and users were not co-located), the automation type was none, and effort level required of users both formal and informal.

The test environment for executing the search tasks was available on-line at a specific URL. The page only contains two links to the original and new interfaces.

Test analysis

Data from the preliminary questionnaire provided us a characterization of our sample of a total of 12 blind subjects.

The sample consisted of 2 women and 10 men; age ranged from 25 – 55+ years, as shown in Fig.3.

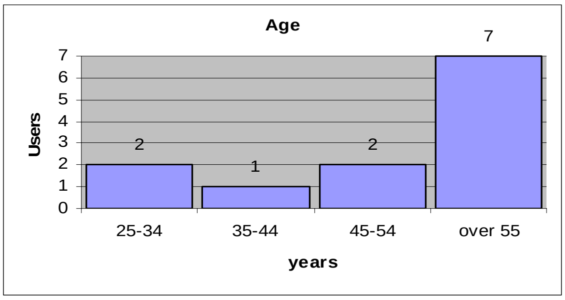

Fig. 3 - Sample Age

The education level of all the participants was high: eight had a secondary school degree and four had a university master or degree.

Nearly all users had used the computer for more than 5 years, only one for 3-5 years and another for 1-2 years. Concerning the use of Internet and web services, the sample has 5 users with basic knowledge, 5 intermediate and 2 advanced. The screen reader used to carry out the test by all users was Jaws on different platforms (at home and at work) and versions, as showed in Figure 4. The minimal experience in using Jaws is 2 years (1 user), from 3-5 years (6 users) and more than 5 years the other participants. It is obvious that different screen reader as well as different versions of the same screen reader may have different features, but their basic behaviors are similar, thus the study can be easily extended.

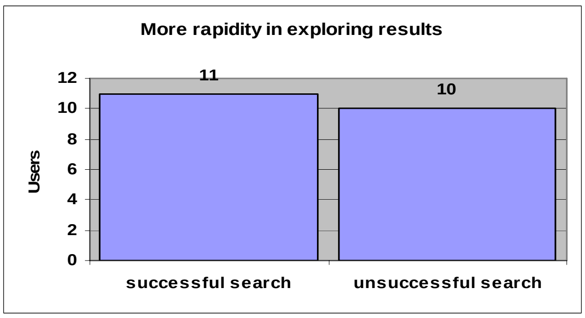

Fig. 4 – Operating system known by the users

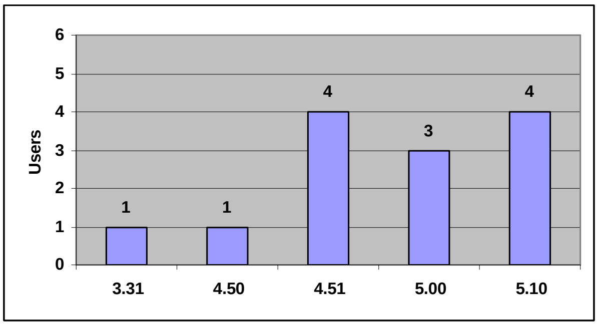

Fig. 5 – Version of Jaws used by users

Only one of the users utilized another http client in addition to IE: the textual browser Lynx on Unix and MS-DOS systems. This person also preferred another screen reader (called Parla) instead of Jaws (although he executed the test in a Windows environment using Jaws). This user also furnished us the feedback for Lynx, as described in the Results section.

Regarding the knowledge of Google (Fig. 6), our results were at first, biased by the language we used in the test. We asked about the subjects' degree of knowledge of Google interfaces, but the term "interfaces" proved to be ambiguous so we were surprised to receive the answer "we do not know". A similar problem occurred for another question in the second questionnaire where we ask the subjects to specify any feature that makes the "navigation" of result page easier: one user reported any differences when clicking the result link (i.e regarding the content of the first/second/… result).

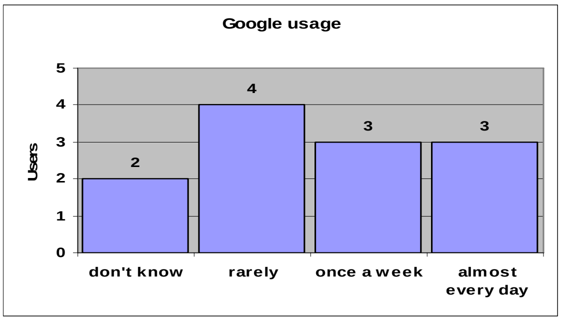

Fig. 6 – Usage of Google

We decided to call those persons who gave some ambiguous answers to try to clarify the misunderstanding.

Measuring results against project goals

Due to the introduction of structural and accessory interface changes we expected to gain improvements both in rapidity to reach the search box and in simplicity while exploring the results. For instance, numbering the results - feature invisible in the visual interface and perceptible only by the screen reader - improved user orientation in result exploration.

Results

As we already mentioned we performed the test with 12 totally blind users, using the Jaws screen reader in a Windows environment with the IE browser. All the 12 questionnaires reported very important user feedbacks. The user testing was executed by users following our written instructions, in complete freedom in time and methods.

All users appreciated the simplified interaction and especially the immediate positioning on the search box and results. Specifically, the participant declared that the modified home page interface simplified the search set-up compared to the original, and 11 out 12 of them also thought the result interface was clearer and easier to use, as shown in Fig. 7.

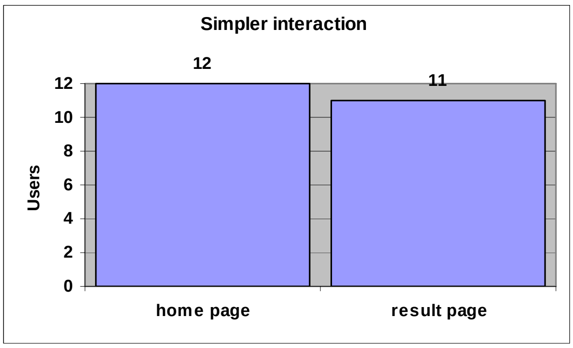

Fig. 7 – Modified interfaces vs the original ones: simplification of user interaction

Concerning the speed of accomplishing the assigned search tasks, 11 out 12 acknowledge they feel the simplified interaction and the major clarity in result exploration reduced the time required to carry out the search (Fig. 8). The most skilled user considered that time to reach desired results was reduced by 20-30% compared to time needed with the original Google interface. In contrast, it is interesting to observe that the less skilled user of the sample, who did not use access keys nor Jaws special commands for exploration (but only arrow and tab keys) stated he/she felt unable to evaluate the interface changes; however, his/her feedback although incomplete was in perfect agreement with the trend of the other answers.

He/she in fact arrived in one step to the first result and understood immediately if the query gave no results.

The sound associated with the unsuccessful search (no result) was appreciated by 10 users (see Fig.8); one was indifferent and one, as previously mentioned, did not answer most of questions.

Regarding evaluation of specific features, participants judged important not only sounds, access keys, and tab keys but also hidden labels and numbering of results which assured more clarity and aided orientation in result exploration.

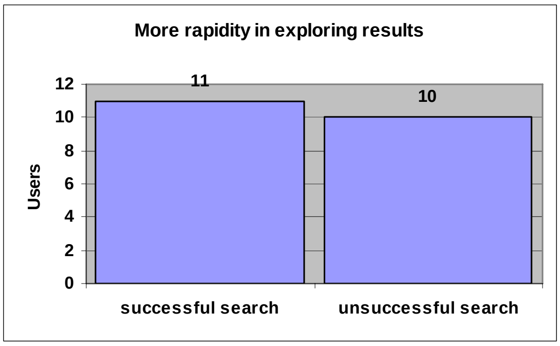

Fig. 8 - Modified vs original interfaces: exploration of results in successful and unsuccessful search

All users found differences between the original and the modified interfaces and suggested that Google adopt all (9 users) or some (2 users) of the proposed changes. Again the same person felt not to be able to give any suggestions.

Other opinions concerned the use of sounds: 10 users found the sounds useful and 6 of them gave some suggestions for their usage. For instance one of the participant observed that the sound chosen for the focus on the search box was too similar to those of MS Outlook Express and thus may create some confusion.

In addition to this evaluation for interaction in a Windows environment, we received other useful feedback. One user acknowledged that the modified version of Google UIs run very well with Lynx (better than the original one) although the textual browser does not take advantage of shortcuts, sounds, and Tab keys. He reported that the arrangement of the information is rational and favours both input specification and results navigation. The main advantage was to have moved all the secondary links to the bottom of the page, in order to have the cursor immediately positioned on the useful information.

Lesson learned

We conducted very few usability tests; thus this experience gave us the keys for understanding how to effectively set up a remote test. For the qualitative and quantitative answers we defined open questions in order to obtain as many observations and suggestions as possible from participants but obviously this required more time for elaborating the collected data with respect to multiple choices surveys.

Conclusion

In this paper we described a usability testing performed with 12 blind users in order to evaluate the real impact on a search engine UI of the usability criteria we proposed in a previous work. To this end, we modified Google user interfaces in order to build a case study to be evaluated.

After introducing the whole project, outlining the criteria, and summarizing the implementation phases needed for setting up the test environment, results collected from the remote testing are shown and discussed in detail.

The feedback clearly showed that users appreciated the new Google UI, giving positive comments and detecting the main differences and features added to the modified Google interfaces.

In the future, we plan to study issues of interaction with user interfaces considering the changes in society due to the globalization and the aging population.

References

[1] Andronico P., Buzzi M., Castillo, C., Leporini B. Improving Search Engine Interfaces for Blind Users: a Case Study. To appear in International Journal Universal Access in the Information Society, Springer.

[2] Andronico P., Buzzi M., Leporini B. Increasing Usability of Search Engine Interfaces for the Blind. Proceedings of HCI International 2005 (Las Vegas, July 2005), Lea.

[3] Craven, J., Brophy, P. Non-visual access to the digital library: the use of digital library interfaces by blind and visually impaired people. Technical report, Manchester: Centre for Research in Library and Information Management. http://www.cerlim.ac.uk/pubs/index.php.

[4] Damery E. JAWS for Windows version 4.51. New features and accessing the Internet. Proceedings of CSUN 2003 (Los Angeles, 2003).

[5] Goble C., Harper S., Stevens R. The Travails of Visually impaired Web Travellers. Proceedings of Hypertext 2000 (San Antonio, June 2000), ACM, 1-10.

[6] Hartson, H. R., Castillo, J. C., Kelso, J, Neale, W. C., Remote Evaluation: The Network as an Extension of the Usability Laboratory. Proceedings of CHI 1996 (Vancouver, April 1996), ACM, 228-235.

[7] Hartson, H. R., Castillo, J. C. Remote evaluation for post-deployment usability improvement. Proceedings of CHI 98, (Los Angeles, April 1998), ACM, 22-29.

[8] Ivory, M. Y., Yu, S., Gronemyer, K. Search result exploration: a preliminary study of blind and sighted users' decision making and performance. Extended abstracts of CHI 2004 (Vienna, April 2004), ACM, 1453-1456.

[9] Ivory, M. Y., Hearst, M. A. The state of the art in automating usability evaluation of user interfaces. Computing Surveys, Vol. 2001, N.4, (December 2004), ACM, 470-516.

[10] Freedom Scientific. Jaws for Windows. Available at http://www.freedomscientific.com/

[11] Jeffries, R., Miller, J. R., Wharton, C., Uyeda K. M. User interface evaluation in the real world: A comparison of four techniques. Proceedings of SIGCHI conference on Human factors, 1991.

[12] Leporini, B., Andronico P., Buzzi, M. Designing Search Engine User Interfaces for the visually impaired. Proceedings of International Cross-Disciplinary Workshop on Web Accessibility (NY, May 2004) ACM, 57-66.

[13] Mankoff J., Fait H., Tran T. Is your Web Page Accessible? A Comparative Study of Methods for Assessing Web Page Accessibility for the Blind. Proceedings of CHI2005 (Portland, April 2005), 899-908.

[14] Milliman R. E. Website accessibility and the private sector: disability stakeholders cannot tolerate 2% access! http://www.rit.edu/~easi/itd/itdv08n2/milliman.htm

[15] Nielsen J., Alertbox January 21, 2001. Usability Metrics. Available at http://www.useit.com/alertbox/20010121.html.

[16] Nielsen J., Alertbox May 9, 2005. Mental models for search are getting firmer. http://www.useit.com/alertbox/20050509.html.

[17] Nielsen J., Usability Engineering, Academic Press, Inc., Chestnut Hill, MA, 1993.

[18] Pearrow M., Web Site Usability Handbook, Charles River Media Inc., Rockland, Massachusetts, 2000.

[19] Scholtz, J. Adaptation of Traditional Usability Testing Methods for Remote Testing. Proceedings of HICSS 2001 (Hawaii, January 2001), IEEE, 1-9.

[20] Takagi, H., Asakawa, C., Fukuda, K., Maeda, J. Accessibility designer: visualizing usability for the blind. Proceedings of SIGACCESS 2004 (Atlanta, October 2004), ACM, 177-184.

[21] Thompson, K. E., Rozanski, E. P., Haake, A. R. Here, There, Anywhere: Remote Usability Testing That Works. Proceedings of SIGITE 2004 (Salt Lake City, October 2004), 132-137.

[22] W3C. Web Accessibility Initiative. http://www.w3.org/WAI/.

Acknowledgements

We thank all the participants in our user testing who made an enthusiastic effort to help us in our research.