



MASTER 2 RESEARCH – INTERACTION SPECIALTY

Bi-manual access to two-dimensional information for the blind

Author:
Kalliopi STRATIGI

Supervisor:
Yacine BELLIK
Mallika AUVRAY

Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur
Group Architectures and Models for Interaction
LIMSI - CNRS
Université Paris Sud
91403 Orsay Cedex, France

Secrétariat - tél : 01 69 15 66 36 Fax : 01 69 15 42 72
Courrier électronique : Murielle.Benard@u-psud.fr
02 April – 31 August 2013

Table of contents

Summary	5
1 Introduction.....	6
2 Content and Objective	7
2.1 Context of the internship.....	7
2.1.1 Presentation of LIMSI - CNRS.....	7
2.1.2 Presentation of AMI	7
2.2 Objective of the internship	7
2.3 General Approach	7
3 Related Work.....	8
3.1 Human - Computer Interaction	8
3.1.1 Terminologies.....	8
3.1.2 Assistive Technologies	8
3.2 Haptic Interaction	10
3.2.1 Haptic Displays.....	10
3.2.2 Tactograph	11
4 Analysis and Design	13
4.1 Project Analysis	13
4.2 Documents' Structure	13
4.3 Method 1- Speech Synthesizer	16
4.4 Method 2- Mouse and Speech Synthesizer.....	16
4.5 Method 3- Concept Keyboard.....	17
4.6 Conclusion.....	17
5 Implementation.....	18
5.1 Technologies	18
5.2 Algorithm's Analysis	18
5.3 Method 1 – Speech Synthesis.....	19
5.4 Method 2 – Mouse and Speech.....	20
5.5 Method 3 – Concept Keyboard.....	20
6 Evaluation.....	21
6.1 Evaluation Protocol – Methodology	21
6.1.1 Hypothesis.....	21
6.1.2 Participants.....	21
6.1.3 Tasks.....	22
6.1.4 Experimental Procedure	22
6.2 Analysis Results	24
6.2.1 Analysis Task Results	24
6.2.2 Questionnaire Analysis.....	25
6.2.3 Analysis the user's behavior	28
6.3 Conclusion - Summary	29
7 Conclusion and Perspectives – Future Work.....	30
8 Bibliography.....	32
Acknowledgments.....	35
Appendix A	37
Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI-CNRS)37	
<i>Agreement Document</i>	37
<i>Agreement Document</i>	38
Appendix B	39
Results Document: Subject n°.....	39

Appendix C 45
 Evaluation Questions: Subject n° 45
Appendix D 48
 Pdf Structure: 48
Appendix E 49
 Table: 49
Appendix F 50
 Results Analysis: 50
Appendix G 51
 Figures 6: 51

Summary

It is widely known that life for blind people is too difficult and complicated. Compared to sighted people, they need more effort to finish a task in their daily life. The access to different kinds of documents would be a start for an improved life, giving them the right to a better quality of life. Nowadays, the blinds have access to digital documents with the use of screen readers but unfortunately without perceiving the spatial structure of the document. In this project, we have implemented a spatial structure of four different types of documents: power point, pdf, excel and html. We have used a series of symbols to present the different kinds of information (for instance bullet symbols). The interaction with the document was developed using three different methods. First, we implemented method 1 – speech synthesis, using only a speech synthesizer for the access to the documents. The next method consists of using a mouse and a speech synthesizer. Finally, we used a concept keyboard to accomplish the third method.

Furthermore, this project would not be completed without the evaluation part. Ten subjects participated to the evaluation of the usability and effectiveness of the methods, in addition to the accessibility to the structures of the documents. At the end of the evaluation, we were able to get results regarding the best method and the design of the structure for all the different kinds of documents.

Keywords:

Haptic interaction, Tactile, Audio Feedback, Blind, Document Structure, Speech Synthesis

1 Introduction

In everyday life, almost everything is designed for sighted people, without taking into consideration the fact that 285 million people are visually impaired worldwide. 39 million from them are blind and 246 million have low vision. It is interesting to mention that 90% of blind people live in developing countries [1]. Deprivation of vision leads to many forms of disability, whether in daily life or professional life. Therefore, it is absolutely necessary to use technologies that could improve the life of visually impaired people or the blinds.

A low-tech aid such as the white cane, the guide dog, Braille writing and embossed pictures are some of the aids that blind people often use in their daily life. It is commonly accepted that the interaction between haptic sense and the environment is important. The white cane and the guide dog help the blind to move and inform the user about many features of the environment. Blind people can read text using their Braille terminals or embossed papers. Braille are coded with a six-point (or sometimes eight-point) rectangular matrix. In all these cases, the blind get the information through the environment or a printed Braille paper or terminal but a way to perceive general digital information remains to be found. [9]

The visually-impaired people's life is affected by the lack of access to information. The information is becoming increasingly digital, influencing the effective access through computers or mobile devices which demands excellent visual abilities because of the use of screens or monitors. It is necessary to handle this serious problem encountered by visually- impaired people. The access to digital information present on the Internet is not enough (for instance surfing the web or trying to read emails). There is also a need to find a solution to perceive the structure of the information, besides the information itself for every kind of (digital) document.

Nowadays, the visually- impaired people can access to every digital document using a specific designed software called "screen reader". Screen readers have the ability to read in a linear way the content of a document or the content of a web site but, in fact, it is not so easy for the blinds to explore every digital document or every web site. These web sites have to be designed and implemented according to W3C requirements for the accessibility [2]. Unfortunately, these requirements are not applied by the majority of designers. There is only a small proportion of web sites which meet the above criteria. The same problem exists for different kind of document, such as power point, pdf, excel, etc.

It is assumed that the life of the blinds would be improved if all digital information was designed correctly based on W3C requirements, although it would still not be the perfect solution to this problem. Many blind people have made complaints about this method of access to digital information because of the use of screen readers. The use of this method needs a basic education and considerable time to familiarize with this software. Especially the blind people who lost their vision gradually in their lives will be too tired to get involved in learning the use of this software.

Furthermore, in some cases, it is too difficult to explore a document and perceive the meaning of this document, as in the case of an excel document or an array. Understanding all the content in an array using a screen reader is a slow process because of the linear reproduction of the content, which makes the comprehension of the document more complicated.

This linearization can complicate the assimilation of the information because websites or arrays have been designed to be explored spatially on a screen. The dissatisfaction regarding the use of the screen readers has been mentioned in several studies [3]. It is absolutely necessary to provide the same access and almost the same structure for visually impaired people. This is the main idea of this thesis using several technologies and methods which help to provide a solution to this problem.

This master thesis is structured into six separate chapters. The first chapter presents the main purpose of this thesis. The second chapter presents the research work which was used to understand the needs of blind people. Then it is important to present the analysis and the design part of this work. Thereafter, the implementation of a prototype is presented. In addition, it was necessary to test its usability with real users. The description of the evaluation protocol and the results will be presented in a next chapter. Finally, some ideas for future work will conclude this report.

2 Content and Objective

2.1 Context of the internship

This project was accomplished at the computer science laboratory LIMSI (Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur - Laboratory for Mechanics and Engineering Sciences) within the AMI group (Architectures and Models for Interaction).

2.1.1 Presentation of LIMSI - CNRS

The Laboratory for Mechanics and Engineering Sciences (www.limsi.fr) belongs to CNRS (Centre National de la Recherche Scientifique), and is associated with the University Paris-Sud 11. The LIMSI laboratory consists in two separated departments, the department of CHM (Human Computer Interaction) and the department of MECA (Mécanique-Energétique) which are themselves divided into groups of research. The groups comprise several research topics.

2.1.2 Presentation of AMI

This work was accomplished within the AMI group. Some members of the AMI group study and investigate the use of different modalities such as touch and speech to improve information access for the blind.

2.2 Objective of the internship

The topic of my internship is “bi-manual access to two-dimensional information for the blind”. The main purpose of this subject is the creation of a structure for several documents containing two-dimensional information using embossed papers and to find a solution to make them easily accessible by the blinds.

The main purpose of this research is to find the best method that can help blind people to perceive not only the content of a document but also its spatial structure. It's important for visually impaired people to get the same information as sighted people. This kind of method should be a promised solution solving problems about spatial distribution of information. In this thesis, we propose and we evaluate three different methods aiming to improve the understanding of the spatial structure of different kind of document.

The two-dimensional structure of the documents was developed using three different methods which are analyzed in the following sections. The first one called “Speech Synthesize” uses only speech to help the blind to understand the meaning of the document. The second method uses Speech Synthesize and a mouse. Finally, the last method uses speech synthesize and an external device called “Concept Keyboard”. In the section “Analysis and Design” we will analyze all these methods separately.

2.3 General Approach

The first part of my internship consisted in doing a bibliographical research to identify existing projects, techniques, tools and methods that were used at the past in similar research topics. The next step was the comprehension of the basic needs of the blind. Thereafter, it was important to analyze and design my project according to the blinds' needs. The next part was the most difficult part of my project: to implement all the different methods according to the design part. Finally, I conducted an evaluation study where 10 users took part in.

3.1 Human - Computer Interaction

3.1.1 Terminologies

This master thesis focuses on tactile interaction for visually impaired people using speech, mouse, and concept keyboard device. It aims at providing methods and tools allowing access to spatial structure of several types of documents for the blind.

First, it is necessary to mention some terminologies related to this work. Usability refers to the extent to which a product can be used by specific users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.

Accessibility means that there is a sequence of accessible input actions and associated feedback for successful accomplishment. A system or an application will be accessible, if it is designed to provide facilities according to the characteristics of the environment or taking into account the user's disabilities or limitations (permanent or temporary).

Multimodal interaction refers to the use of different communication modalities. A system can be called multimodal system, if it can use multiple sensory/motor modalities such as vision, speech, touch, etc.

3.1.2 Assistive Technologies

Assistive or adaptive technology commonly refers to products, devices or equipment that are used to maintain, increase or improve the functional capabilities of individuals with disabilities. Assistive technology is a generic term that includes assistive and adaptive devices and the process used in selecting, locating and using them. These technologies promote greater independence for people with disabilities by enabling them to complete tasks that they were formerly unable to accomplish or with great difficulty. The use of assistive technologies is necessary for people with disabilities to make their life easier and independent.

For blind users, the most popular assistive technologies are the screen readers, speech synthesizers and Braille displays. For users with low vision it is screen magnifiers, text prediction systems, and speech synthesizers.

Some of the assistive technologies have been used to help blind people accessing graphical information. One of these technologies is called "Tactile Mouse". This low-cost device has four main components: a Braille cell which consists of a 2x4 pin array representing the tactile interface, the electronics to drive it, an RF transmitter directly underneath the pin array to keep track of absolute position with a graphics tablet and a mouse casing (Figure 3.1). In this project, they implemented and evaluated three methods of presenting Braille with graphics using textures. According to their evaluation, Braille and graphics at the same amplitude level was the most effective method [6].

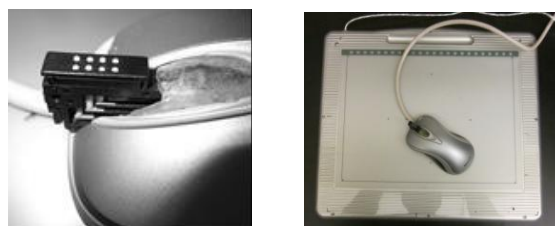


Figure 3.1: Tactile Mouse for tactile graphics display
Extracted from [6]

There is also another commercial device, which is called Logitech Wingman mouse (Figure 3.2). This device is used to browse windows. When the user moves the mouse over an object, a figure or a hyperlink the user can feel a force feedback.



Figure 3.2: Logitech Wingman Mouse¹



Figure 3.3: SensAble PHANTOM²

SensAble PHANTOM is the most cost-effective haptic device used in tactile / kinesthetic interaction. When the user touches the stylus, pen-tool or thimble, the device can return force feedback. It can present a virtual 3D object using haptic interaction. [7] The users can move their finger and the PHANTOM tracks their movements. Users get the feedback from this device using the stylus. They feel a force on their hand creating with this way an illusion of interaction with the solid physical objects. Apart from the stylus, users could use a thimble cursor and get output from their finger about virtual interfaces.

Brewster and Yu evaluated two multimodal interfaces designed for visually impaired people [5]. The output modalities of these interfaces were audio and haptics. They compared the usability of the Logitech Wingman Mouse and the SensAble PHANTOM. They created a project to present graphical information for visually impaired people. They used the tactile mouse in application for browsing graphical data such as bar charts. An example of the evaluated tasks is to find the highest and lowest value or to describe the overall trend of the data using haptic and audio feedback. The results of the evaluation showed that there are no significant differences between these technologies. However, PHANTOM had better performance in the haptic interaction. Also, 2D graphical interfaces, plots, were evaluated with great effects using the Logitech Wingman Mouse.

The PHANTOM device was also used for the implementation of several games for blind children [23]. Multimodal applications for learning were designed for 6 to 10 year old visually impaired children. The children could easily use the system making use of haptic, auditory and visual modalities (Figure 3.4).



Figure 3.4: children use the PHANTOM,
Extracted from [23]



Figure 3.5: children play a haptic memory game
Extracted from [23]

The Tactos tool [10] consists of a tablet PC and a touch cell on the side. The user uses one hand to hold the stylus and move it on the shelf, and the other hand to feel the tactile feedback of the cell. A similar system was used for Optacon [27], a tool that was very used during the last 30 years to facilitate the reading of books and newspapers that are not translated into Braille (Figure 3.6) Indeed, the user can move a mini-camera on the text he wants to read with one hand and get tactile feedback on the other hand. A touch panel display made up of pins arranged in the form of a 24 x 6 matrix is placed in the other hand. The camera recognizes the different characters and the user can get the output from the display.

The "Cat" (Figure 3.7) is a small device that is normally used only for Braille. It is composed of 40 cells, each containing 16 tactile pins arranged in the form of a 2 x 8 matrix. However, this device can be used to present tactile graphics [11]. But it should be noted that this type of tool is limited regarding the tactile rendering, as Braille cells are not very precise, and pins can have only two different behaviors: retracted or extended.

¹ Extracted from <http://www.hardware.fr/articles/159-1/souris-logitech-wingman-force-feedback-mouse.html>

² Extracted from <http://www.senztech.cc/showpros.aspx?proid=36>

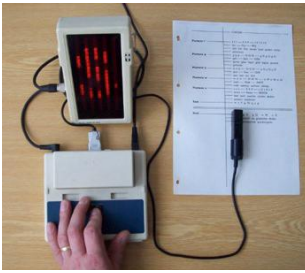


Figure 3.6: Optacon³



Figure 3.7: The “Cat”
Extracted from [28]



Figure 3.8: RDB Braille interface⁴

The RDB Braille interface generates Braille characters using pins. The output from this device depends on the state of the pins. Braille characters can be composed using the top or down position of the pins (Figure 3.8) [26].

The VT Player mouse (figure 3.9) provides also tactile output using a dots matrix 4x4. This mouse has two dot matrixes 4x4 on both sides of the device. The users put their fingers on these matrixes and get the feedback from a series of patterns. In this project, the VT Player mouse was used as an authentication entry to the system. [8]

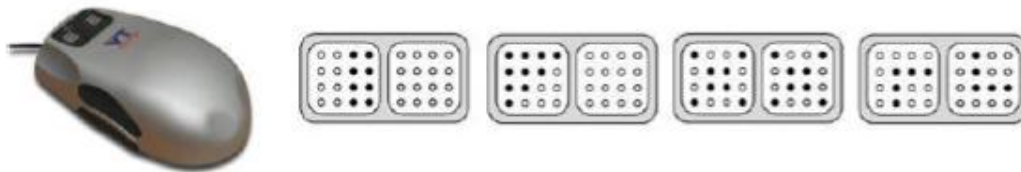


Figure 3.9: VT Player Mouse (at the left) and the different patterns (at the right)
Extracted from [8]

3.2 Haptic Interaction

In some cases haptic interaction is more necessary than vision. For instance, haptic is significant in case to approximate the weight of objects or to feel the hardness of a surface or a texture. On the other hand, it's very important to mention that haptic is not able to provide an overview of a scene [23]. So it is necessary to take into account the natural functioning of haptics, when making research about the usefulness of haptic devices [9].

3.2.1 Haptic Displays

Haptic displays help to provide information about graphical information or virtual objects. They provide force feedback through a “collision” from a stylus, a thimble, or a similar tool in a user's hand, and the surface of a virtual object. In addition to the shape of objects, surfaces properties such as hardness/ softness, texture and friction can be rendered in 3D for haptic exploration.

It is possible to present graphical data information with textures using haptic displays. Textures using haptic displays can be perceived by visually impaired people. However, a general problem concerning this kind of devices is that they may have shown their potentials in laboratory contexts, but the usefulness in real situations may be more problematic. The users may find them difficult to use, possibly because they are not sufficiently adapted to the way haptics work. Furthermore, it's important to mention that guidelines for tactile and haptic interaction between users and relevant devices are under development by ISO (International Standards Organization).

In some cases visual and auditory displays are used together in haptic displays. In case that these displays are used by visually impaired people, they just don't perceive the visual part of the displays and they work without this visual information. The lack of this visual information causes more difficulties for the blind to perceive the meaning of the information. However, haptic displays are the only way to get direct information for some contents [9].

³ Extracted from http://www.tyflokabinet-cb.cz/y_gal_braille.htm

⁴ Extracted from <http://www.dotlessbraille.org/tactiledisplays.htm>

According to the HyperBraille⁵ project, the planar tactile display “BrailleDis9000” is used as input and output device for blind people (Figure 3.10). This tactile device is touch sensitive and consists of an equidistant pin-matrix of 120 columns and 60 rows. The display allows a refreshable two-dimensional tactile output of information and a gestural input [18].

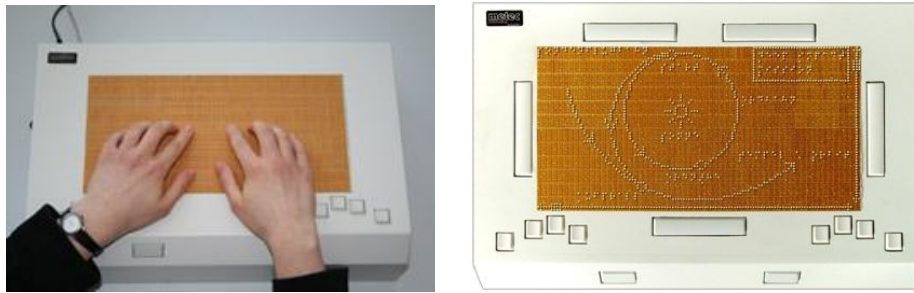


Figure 3.10 : BrailleDis9000^{6 7}

3.2.2 Tactograph

The TactoWeb project which uses the Tactograph device (Figure 3.11) was developed in Haptics Laboratory at McGill University in Montreal. In this project a spatial structure of Web pages with tactile and audio feedback using the Tactograph device has been developed. It consists in a tactile device STReSS² (Stimulator of Tactile Receptors by Skin Stretch), mounted on a 2D haptic device for exploration. The Tactograph is different from other types of touch devices because it stimulates the tip of the finger laterally stretching the skin. The STReSS² has a tactile cell rendering of 10.8 x 12 mm, which corresponds approximately to the skin surface at the end of the index [20]. To deform laterally the skin, 60 actuators (6 x 10) can move from right to left to produce sensations of vibration, ripple, etc. The distance between each actuator is 1.8 x 1.2 mm and each of them can bend from 0.1 mm to the left or right. The STReSS² is integrated into protective black foam, and then mounted on a trolley to explore an area of 11.3 x 6 cm. With this device, it is possible to explore an area, while having a tactile rendering adapted to where it is on the surface.

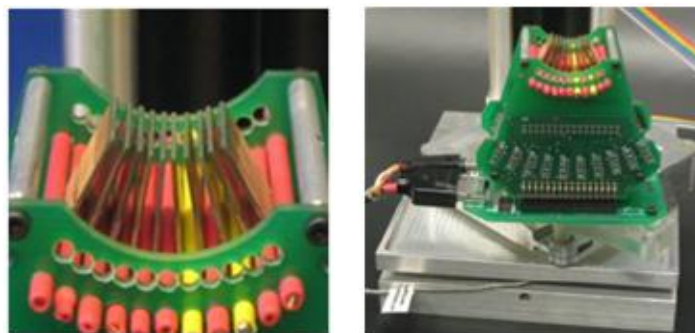


Figure (3.11): Tactograph
Extracted from [28]

In the Tactoweb project, the Tactograph device is used to produce undulations and vibrations depending on where the cursor is in the Web page. The project recreates connections between the different elements composing a Web page that could have disappeared during the linearization of the information. In this project, rectangles are used to present the information of a web page (Figure 3.12). [28]

⁵ www.hyperbraille.com

⁶ <http://www.cs.uni-potsdam.de/aghb/lehre/2009ss.html>

⁷ http://www.inf.tu-dresden.de/index.php?node_id=3245&ln=en

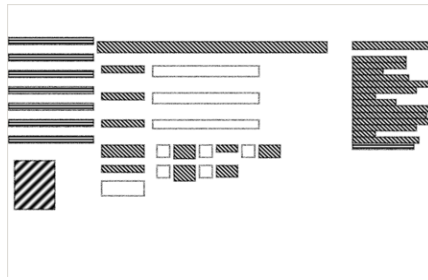


Figure 3.12: The spatial structure of a web page
Extracted from [28]

The main advantage of this device is its touch cell. Indeed, it allows a much finer touch than a set of Braille cells. It is also more accurate than a three-dimensional (3-D) tactile display device [21]. It is possible to have different textures. With this tool, some applications such as a memory game [22] have been created. On the other hand, the disadvantage of this device is the lack of use of multiple fingers. The user can use only one finger, at the end of index, and not all the fingers of the hand or both hands to perceive easily the spatial structure of the web page.

4.1 Project Analysis

The purpose of this project is the creation of a spatial structure for several types of documents using a series of symbols which can help the blind people to perceive the meaning of the document not only the content but also the whole structure of the document. This structure will be created by three different methods. At the end of this project, we will be able to know which method had the best results about the usability and accessibility of the different kinds of document as also, if the structure of all these documents was understanding and easily accessible for the blinds. It's necessary to get results about the meaning of the different symbols that we used, about the difficulty of the structure, and the best method (easier method with high effectiveness to the blind accessibility).

In this project, we present three different methods to help blind people to perceive the structure of four different kinds of documents. The first method called "Speech" using a speech synthesizer, the second method is the "Mouse and Speech" using a mouse and the speech synthesizer. The last method called "Concept keyboard" and it consists of the Universal Plus Concept keyboard, as also a speech synthesizer.

All these methods were evaluated using the structure of four different kinds of document, the first document was a power point, the second document was the structure of a paper (article) at Pdf format, the third document was an excel document and finally an html document presenting a site. All these documents are printed in two-dimensional papers, so, the blind people have to explore these tactile papers and they can easily understand the rectangles.

4.2 Documents' Structure

According to the research of Denise Prescher etc [18], and a second research of Maria Schiewe [19], the blind user can explore the information of the documents at different presentation modes views and levels. These views define the presentation way of the information. They provide four predefined views that cover the majority of the users' need. These views are called: layout, outline, symbol, and operating view.

The outline view provides a general overview of the document presenting the content of the document as abstract rectangles. The outline view maintains spatial relations using simple geometrical shapes, like lines and rectangles that indicates group of objects. (Figure 4.1)

The operating view presents a completely text-based enabling the user to read the content in Braille. It is optimized for a fast work flow and it is too close to screen readers' structured mode. (Figure 4.1)

In the symbol view, preserves the relative position of the elements, but text appears in Braille. It provides spatial relations using also Braille for the text. (Figure 4.1)

The layout view preserves the information of the document but it provides the content of the document in low-resolution tactile screens. Text does not appear as Braille but rather in a tactile version of print offering users without Braille skills, new possibilities in tactile reading. (Figure 4.1)

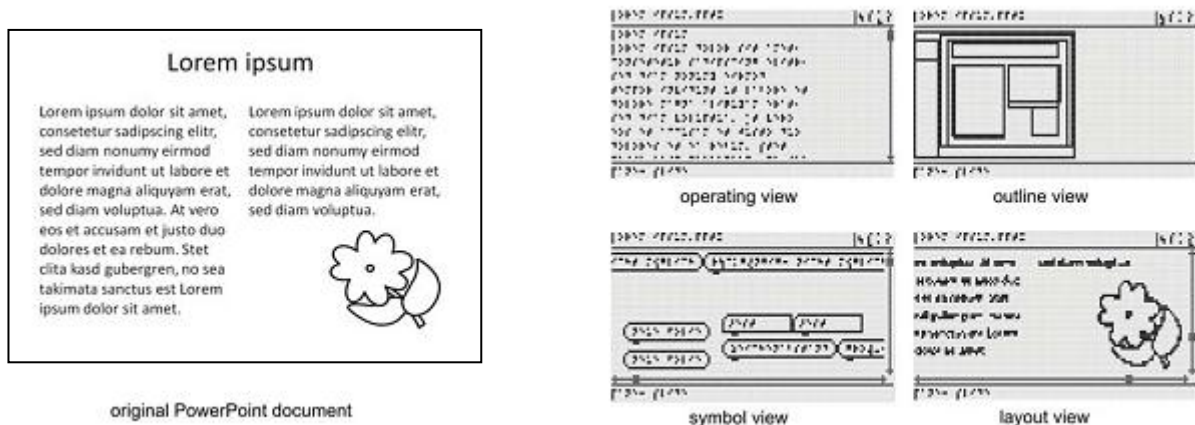


Figure 4.1: Power point document represented in four different views
 Extracted from [18]

Taking into account the analysis of the previous views, we created this project according to the outline view, presenting the general structure of the document as a composition of rectangles and other symbols that were used to present specific information and get an overview feedback.

It's important to mention the different symbols that we used to present different information in the document. In general approach, we used simple rectangles to present a paragraph or a text area, figure 4.2 (a). We used different symbol to present the paragraph with bullet or list, at this way, it's easier for the user to understand the meaning of this area in the document from the first time, if only the user touch the rectangle and feel the differences. This symbol consists of the rectangle and a full cycle at the left of the rectangle, as shown in Figure 4.2 (b). The subsection of a paragraph presented as a rectangle with one vertical line, if the subsection indicates the first level, for instance, a paragraph for "4.2 Documents' Structure", Figure 4.2 (c). The subsection in two levels was presented by a rectangle with two vertical lines, for example a paragraph for "4.2.1 Example", Figure 4.2 (d). It's necessary to mention that the figures and subfigures of the document were presented, also, as a simple rectangle. The difference between the figures and text area determined from the audio feedback.

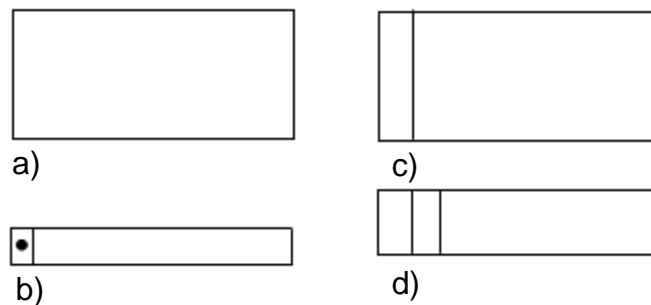


Figure 4.2: The symbols of the different areas, a) rectangle – text, b) rectangle with cycle – bullet, c) rectangle with one line – subsection level one, d) rectangle with two lines – subsection level two

At this point, it would be interesting to present some examples to understand the structure of the document. At the figures 4.3 and 4.4, presented the structure of two different documents using the symbols which determined below.

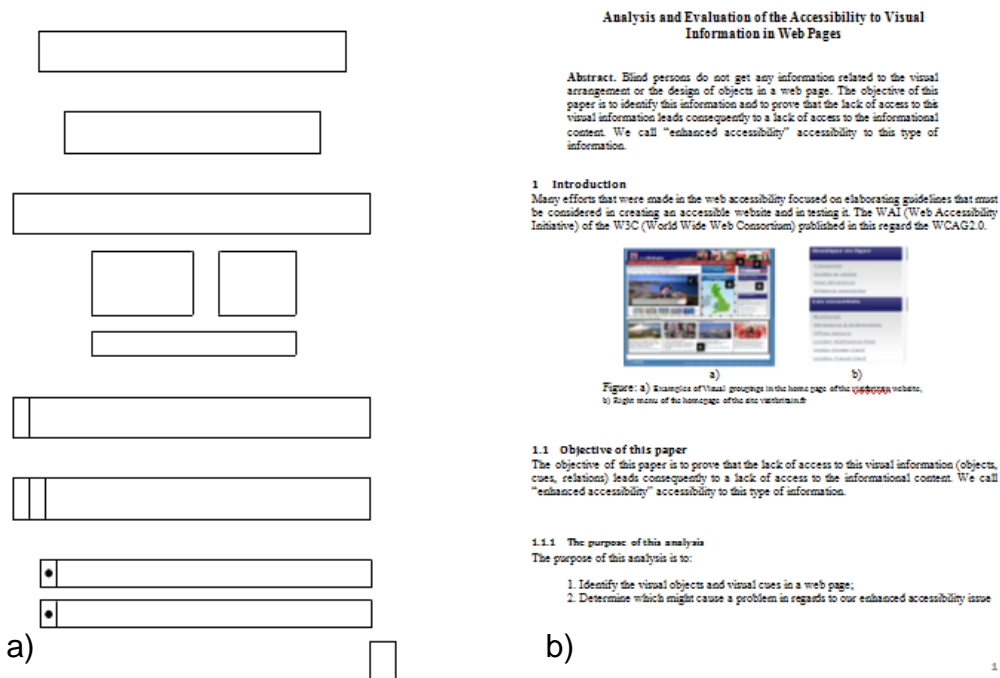


Figure 4.3: The training document during the evaluation, a) the structure of the document using the symbols, b) the original structure of the document

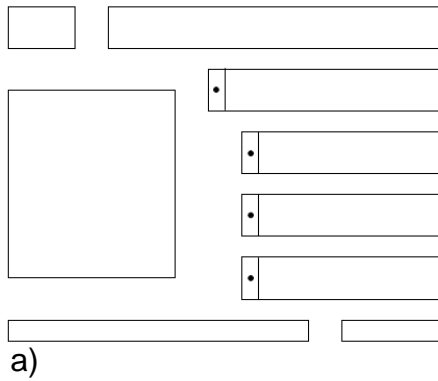
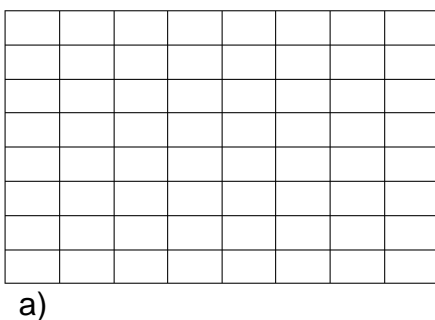


Figure 4.4: The power point document, a) the structure of the power point document using the symbols, b) the original power point structure



Car Sales per month	January	February	March	April	May	June	Total Sales
Citroen	300	600	500	450	250	400	2500
Renault	550	550	200	210	190	300	2000
Toyota	300	400	800	590	300	210	2600
Peugeot	200	550	500	600	300	250	2400
Mercedes	500	300	440	600	560	300	2700
smart	550	500	640	550	310	450	3000
fiat	280	300	270	250	300	300	1700

Figure 4.5: The excel document, a) the structure of the excel document in tactile surface, b) the original excel document

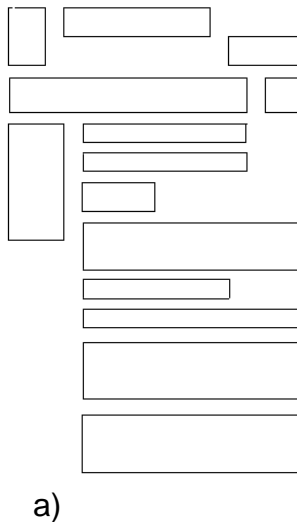


Figure 4.6: The html document, the site of the M2R interaction, a) the structure of the html document, b) the original html document

Taking into account the structure of the power point in Figure 4.4, the last three rectangles with the cycle present the list, and they have a little shift at the right. The reason of this shift is the division in separate areas. To be more specific, the three last rectangles present a list of three objects that is part of the previous rectangle-bullet. In this kind of subsection, the rectangles of the structure will not follow the same alignment with the previous rectangle but it will be shifted at the right side. This symbol (shifted rectangle at the right) indicates the sub-division of a paragraph in separated paragraphs – objects.

It's important to give more declarations about the figures and sub-figures. As already indicated in previous paragraph, the figures and sub-figures have exactly the same rectangle with text area. The difference between them came from the audio feedback. The blind user will listen from the speech synthesizer, a first description of the rectangle as image (a), or part (a) of figure 1, part (b) of figure 1, that indicate in images. In addition, the blind user will listen from speech synthesizer the legend of the image as "Figure 1: part (a) example 1 part (b) example 2". The connection of the image and legend of the image came from the description of speech synthesizer.

4.3 Method 1- Speech Synthesizer

According to this method, the blind users have to listen very carefully the audio feedback. To be more specific, this method consists of a speech synthesizer, a camera and the two-dimensional papers with the structure of the documents.

First of all, the blind users should use a blue sticker on their finger, because this method works with the detection to blue color using a web camera. With the detection to blue color, the system is able to understand the position of the finger when the users move around their hand.

At the beginning, the blind user will explore / touch the papers to feel and understand the different rectangles. When the user's finger is located into the borders of a rectangle, the speech synthesizer will reproduce the description of this area, for example, if the finger of the user is inside the borders of the rectangle with the title (at the first rectangle in figure 4.3 (a)), the speech synthesizer will reproduce the text – description of this rectangle which is "Title". The user touch inside the limits of a rectangle, and he will receive an audio feedback about a general description of this rectangle, for instance, "Title", "Introduction", "References" etc

When the blind user listen the description through the speech synthesizer, they can wait for 4 seconds and then they will listen the full content of the area or they can remove the finger in another rectangle to listen another description.

In summary, the blind users explore the two-dimensional document. When their finger is inside the borders of a rectangle, they will hear a general description of this area, for example "Title". If the users search about the title of the document, they just wait for 4 seconds and they will hear the full content of the title. But, if they are looking for the paragraph "Introduction", they will remove their finger in another rectangle until get an audio feedback through the speech synthesizer as "Introduction".

Using this method, the blind users will have direct manipulation to the system, because they have to do only one thing, first, to hear the description, then, wait for a few seconds, and the speech synthesizer will start the reproduction of the content for this area.

4.4 Method 2- Mouse and Speech Synthesizer

This method consists of a web camera, a mouse, the two-dimensional documents and speech synthesizer. The users should use a blue sticker also, at their finger to detect the blue color, helping with this way to track the finger.

First of all, the users explore the document using the blue color for the finger tracking. They hear the description of an area – rectangle, when their finger are detected inside the borders of the rectangle. The difference between this method and the method 1- speech synthesizer, is the way that users listen the full content of the rectangles. At this method, users have to make a right click with the mouse and then, they will hear the content of the paragraph.

For example, the blind user explores the document and he try to find the page number of this document. During the exploration of the document, the user will hear the description of the rectangles, when the user hears "page number" through the speech synthesizer, it means that he find the position of the page. At this time, the user has to press the right button on the mouse to receive the full content of this area, to hear the number of the page. If the user removes his finger in another rectangle, the speech synthesizer will reproduce the text for the description for the next area – rectangle.

At this method, the users have indirect manipulation because they have to use external devices (the mouse) to perceive the full content of an area.

4.5 Method 3- Concept Keyboard

It's necessary to mention that this method consists of a web camera, the two-dimensional documents, the speech synthesizer and the concept keyboard. A concept keyboard is a flat board that contains a grid of buttons with dimension 16x16. Each button can be programmed according to the needs of the project. (Figure 4.7)

The users use also, a blue sticker to detect the finger position. The two-dimensional documents placed on the concept keyboard. When their finger is detected inside the borders of a rectangle, the speech synthesizer reproduces the description of this area, for example the user will hear a description like "Title". After that, the user can touch/ press inside the rectangle and the speech of the full content of this rectangle will be reproduced automatically.

In summary, the user explores the two-dimensional document using the blue sticker to follow the finger tracking. When the user's finger is detected inside a rectangle, the speech synthesizer will reproduce the text of the description of this rectangle. If the user presses inside the rectangle on the concept keyboard, he will hear the full content of the rectangle. If the user moves his finger to another rectangle, the speech will start automatically for the description of the last rectangle.

This method called direct manipulation, because the user uses only one device, the concept keyboard, to interact with the system, and to perceive the meaning of the documents.



Figure 4.7: Concept Keyboard⁸

4.6 Conclusion

In concluding, it would be necessary to mention some useful terms about these methods. In this project, it's used passive modality because of the finger tracking. In the case of mouse method, it's used active modality because of the use of mouse, which is an external device.

This project meets the requirements of the output multimodality, because of the use of tactile interfaces and the speech, audio feedback. To be more specific, the user uses the tactile and audio feedback, the speech synthesizer automatically. This means that users use more than one modality at the same time.

At general approach, this project was designed to meet the requirements for output multimodality because of the use of tactile and audio feedback at the same time. It gets also input passive multimodality because of the finger tracking.

The purpose of this project is:

- To help the blind to understand the structure of the document easily and effectively
- To perceive a spatial structure of the document
- To find the best method which can help the blind to understand the structure of the documents

⁸ Extracted from http://acorn.chriswhy.co.uk/8bit_Upgrades/ConceptKbd_A4-128.html

Matching the corners of the image recording from the camera, with the corners of the digital image/ document will be created the following system:

$$\begin{aligned} (Xc1, Yc1) &\rightarrow (0, 0) \\ (Xc2, Yc2) &\rightarrow (Xmd, Ymd) \end{aligned}$$

$$\begin{aligned} XDF &= a_x * XcF + b_x \\ YDF &= a_y * YcF + b_y \end{aligned}$$

$$\begin{aligned} &\Rightarrow \begin{aligned} 0 &= a_x * Xc1 + b_x \\ 0 &= a_y * Yc1 + b_y \end{aligned} \Rightarrow b_x = -a_x * Xc1 \end{aligned}$$

$$\begin{aligned} XmD &= a_x * Xc2 + b_x \\ YmD &= a_y * Yc2 + b_y \end{aligned} \Rightarrow \begin{aligned} XmD &= a_x * Xc2 - a_x * Xc1 \\ YmD &= a_y * Yc2 - a_y * Yc1 \end{aligned} \Rightarrow \begin{aligned} XmD &= a_x (Xc2 - Xc1) \\ YmD &= a_y (Yc2 - Yc1) \end{aligned} \Rightarrow \boxed{a_x = XmD / (Xc2 - Xc1)}$$

$$b_x = -XmD / (Xc2 - Xc1) * Xc1 \Rightarrow \boxed{b_x = (XmD * Xc1) / (Xc1 - Xc2)}$$

$$b_y = -a_y * Yc1, \quad YmD = a_y * Yc2 - a_y * Yc1 \Rightarrow YmD = a_y (Yc2 - Yc1) \Rightarrow \boxed{a_y = YmD / (Yc2 - Yc1)}$$

$$b_y = -(YmD / (Yc2 - Yc1)) * Yc1 \Rightarrow \boxed{b_y = (YmD * Yc1) / (Yc1 - Yc2)}$$

Solving this system, it's known the values for a_x , b_x , a_y and b_y , as also, the finger position using these values. The coordinated for the (XcF, YcF) is known from the OpenCV library, with the previous solution of the system, it's known also the coordinates of the (XDF, YDF) which correspond to the original document.

Starting the project, the calibration is the first and the most important thing. The calibration of this project is completed when we picked the two corners of the original documents calling the function `mouseEvent`. It's necessary to pick the first corner at top left side of the document, and the second corner at the bottom right side of the document. Finishing the calibration, it will be created an invisible array 16 x 16, which is used to determine the borders/ limits of the useful rectangles of the document.

According to the array 16 x 16 which is created after the calibration, it's designed the borders of the rectangles. We calculated the empty cells of this array and in which cells there are rectangles, and we made a new file in Visual Studio to identify the limits of every rectangle. Finishing this step, the limits of the rectangles were completed and the structures of the documents were ready.

The last important thing for the implementation was the speech synthesizer. Including the speech library at the project, as also, at the system of the project, it will be prepared to call the function "Speak" or "SpeakAsync" (in case that it's necessary to run the project in parallel with the speech output, we have to use the function "SpeakAsync") to start the speech synthesis of the text. At the beginning of the project, the speech synthesis will be used only for the description. Then, it depends from the method, if the speech output will continue to the content or not.

5.3 Method 1 – Speech Synthesis

At the beginning of this method, it's necessary to finish the part of the calibration, to pick the corners of the documents as described below. After the calibration, it's already called the function about the detection of the blue color, the finger position and the speech output.

When the user, put the finger on the document, it will start automatically the speech synthesizer to reproduce

the description of the areas/ rectangles. At the same time, a timer will start calculate the seconds in which the user's finger are remind at the same area.

If the user doesn't move the finger from a rectangle for more than 4 seconds, the content of the rectangle will reproduce automatically. If the user remove his finger before the end of the 4 seconds, the content of the rectangle will never be reproduced, and the description of another rectangle that explore will start be produced by speech synthesizer.

5.4 Method 2 – Mouse and Speech

This method is too close to the previous method, with the use of the speech. The process for the speech output, the calibration, the finger position and the detection of the color is exactly the same.

The difference between the two methods is the way to get feedback for the content of the rectangles. When the finger tracking is inside the desirable rectangle, the user will press the right button of the mouse. This means that a function `mouseEvent` is called and when this event happens, the speech synthesizer starts the production of the content for this rectangle.

When the user removes the finger before the end of the speech, the speech synthesizer will stop and start another event, for example, the speech synthesis of the description for another rectangle.

5.5 Method 3 – Concept Keyboard

The process for the calibration, finger tracking, color detection, audio feedback is exactly the same as the previous method.

The difference in this method is also, the way to get the output for the full content of a rectangle. When the finger tracking is inside the desirable rectangle, the user will press inside the rectangle and the speech output will start.

This method was implemented taking into account the name of the cells in the concept keyboard. The borders of the rectangles were created using the values of the cells. For example, from the cell number 0 until 7 will be the title and if I press inside this area, I will hear the full content of the title.

The concept keyboard consists of array 16 x16 cells, the number of the first cell was the 0, the numbers of the cells are continued serially, the first line consists of the cells number 0 – 15, the second line starts from 16 – 31 and the last cell is the number 255.

6 Evaluation

The purpose of this evaluation is to find the best method which can meet the requirements for the blind people to explore the structure of the document, as described at the previous section. The goal of this project is to find a better way to give spatial information about a series of documents, to understand the different symbols that was used, and perceive or “read” the content of the document easily, effectively, and quickly.

This chapter are separated in the following subsection, firstly, at the section 6.1 evaluation protocol, will be presented the methodology that was used during the evaluation, the numbers of the users and their abilities, the tasks which was demanded to success, and the experimental procedure which was followed during the evaluation part. Then, there is a chapter about the data collection and, finally, a completely analysis of the results, as also a summary from the evaluation of this project.

6.1 Evaluation Protocol – Methodology

6.1.1 Hypothesis

At the beginning of the evaluation, we considered that the concept keyboard is the faster method and provides less error than the two other methods. The users can easily use this device and understand this technology.

6.1.2 Participants

For the needs of the evaluation, ten participants took place to complete a prototype evaluation for this project. At the follow table are presented the profile of the users. It’s shown their age, their informatics level as also the gender of the users.

Participants	Age	Gender	Informatics Level
1	25	Male	Expert
2	29	Female	Expert
3	26	Female	Expert
4	25	Female	Beginner
5	27	Male	Expert
6	33	Male	Expert
7	26	Male	Expert
8	25	Male	Expert
9	24	Female	Intermediate
10	27	Male	Expert

Table 1: The profile of the participants of the evaluation

It wasn’t necessary to find users with experience in such technologies, or to be experts for the tactile surfaces. The users were not obligated to have knowledge about the technologies that was used, or to be convenient with two-dimensional papers for the documents. Unfortunately, it was too difficult to find real blind people for the evaluation, for this reason, the participants was sighted people but they closed their eyes with the help of a specific mask during all the evaluation part.

I explained orally the three methods, the technologies, and the usage of each method as also a general idea about the two-dimensional documents. I explained the goal of this evaluation without a lot of details. After that, users should wear the mask, and they started to use the training part of the documents for all the methods.

6.1.3 Tasks

We have used a several number of tasks during the evaluation, to get better results about the effectiveness and the usability of the three methods, as also to evaluate the spatial structure of the documents.

Specifically, the purpose of the tasks that was used based on the understanding of the documents' structure. Exploring the document, the users should perceive easily the position of each separated area – rectangle, to get in their mind an overview of the documents.

Furthermore, it's completely necessary to evaluate the symbols that were used. Through the tactile, it's easy for the users to perceive the differences between the simple rectangles, the rectangles with lines and the rectangles with the cycle. Even though, they can understand the difference between the symbols, it's interesting to know if they are able to perceive the deeper meaning of this symbol. For instance, the rectangle with the cycle present the paragraph with bullet or a list, the rectangle with line presents the subsection of the paragraph.

In the task document are contained questions like, “where is the page number” and “where is the title”. The purpose of these questions is to evaluate the symbols. (In the [Appendix B](#), there is a task document which was used during the evaluation with all the questions for every document.)

There is also a question in the tasks, like “how many references were used”, to give the opportunity to search the symbols of the references, to calculate the amount of the rectangles with the same symbol in a specific area on the document.

We asked from the users, to read the full content of the paragraph or find a specific paragraph. In this way, the user is obligated to use correct the technologies of each method separately and find an answer of these questions. For example, we asked the users to tell the first word of a specific paragraph, or to tell the number of the page (more similar tasks are found at [Appendix B](#)). Users obligate to use the technologies and the system of each method correct and answer in these questions.

With this kind of questions, it's evaluated the method, the technologies which was used and it was easy to get an answer about the most easier, effective and quickly way to perceive the meaning of the document, and get the spatial structure of each document.

6.1.4 Experimental Procedure

First of all, the participants should sign in an agreement document which contains their approval for the perception in this evaluation. Signing this agreement document, the users are informed about their privacy rights, and they agree to get place at this evaluation. Users agreed to record video during the evaluation. Inside the agreement document, is described a summary of the project, using three different methods, under the approval of my supervisor.

The agreement document are found at [Appendix A](#), containing the content of the agreement document as also the date of the evaluation, the name and the signature of the participant as also the name and the signature of the experimenter.

The second part of the experiment was information about the different methods, explaining in detail the use of the different methods as also the technologies.

At this point of view, it's absolutely necessary to mention that during the evaluation were used three different documents for each format. To be clearer, we used three different structures for power point document, three different structures for pdf, three different structures for excel and three different structures for html document, it was presented three different sites. In the [Appendix D](#), you can see a structure of pdf document.

We used three different kinds of structures for one type of document because it was necessary to get significant results for this evaluation. To be more specific, we used three different structures of power point or pdf, one structure for each method. In this way, the user didn't know the structure of the documents, because we used other structure for pdf or excel in three methods.

It's important to declare that the series of the method as also the series of the different structures were not always the same. We didn't keep the same sequence of the methods or the same series of the structures. Many researchers have shown that keeping the same series of the methods, the users will improve their abilities until the end of the evaluation, and the results will not represent the reality. At the beginning, the users will face more difficulties, but until the end of the evaluation, the users familiarize with the technology and the way to explore the documents, so, at the end it seems to

improve their skills.

To avoid this problem, the series of the methods was not the same as also the type of the documents. I changed all the time, not only the sequence of the methods but also the structures of the same type of documents for each method. In the following table, it's shown the sequence of the methods that was used as also the series of the structures.

Participant	Method 1 – Speech Synthesis	Method 2 – Mouse and Speech	Method 3 – Concept Keyboard
1	1 st	2 nd	3 rd
2	3 rd	2 nd	1 st
3	2 nd	1 st	3 rd
4	2 nd	3 rd	1 st
5	3 rd	1 st	2 nd
6	1 st	3 rd	2 nd
7	1 st	2 nd	3 rd
8	2 nd	3 rd	1 st
9	3 rd	1 st	2 nd
10	2 nd	1 st	3 rd

Table 2: The series of the methods which were used in the evaluation

The table with the series of the different structure of the documents is found at [Appendix E](#).

After the agreement of the users and their approval for this evaluation, they put the mask on their eyes. At the beginning, there was the training part for the methods and the two-dimensional documents. Firstly, they explored the first training document with my help to understand the rectangles, the different symbols, the meaning of the symbols, as also to familiarize with the technologies and generally with the methods which were used. During the exploration of the training document, the users could ask everything about the document, the method, the symbols and they occupied with the training document as long as they needed. They used the training document with all the methods to be sure that they understand and to make practice with all the methods. After the first training document, there was a second training document to explore the structure of the document without my help. When the training part was completed with success and the users felt comfortable with the methods and technologies, and they were prepared, we started the evaluation with the real documents asking them the tasks.

During the evaluation, I followed the series of the table 2, and the table on Appendix E for the series of the document. I would remind that I changed the sequence of the methods and the type of documents for all the users. At the beginning of the evaluation for every different structure, the users had 30 - 40 seconds to explore alone the document and after that I asked the first task according to the task table (see [Appendix B](#)).

When I finished the question about a task, I started to calculate the time that they need until their answer using a chronometer. I kept all these times in the task table, as also their answers. I kept the errors and the right answers of the user and the total time that they needed to finish a task. It's important to mention that I didn't answer any question during the evaluation as also I didn't answer them about their correct or wrong answers.

At the end of the evaluation, when the users completed all the tasks for all the methods, they had to answer a questionnaire about the difficulty of the methods, and the structures. The purpose of the questionnaire is also, to understand if they perceive the meaning of the symbols as also, if it was easy to realize the different symbols. The completely questionnaire of the evaluation are found at [Appendix C](#).

Finally, I made a short interview / discussion with the users to understand their general estimate of the project, to ask about more suggestions and get an overview opinion about the usage of the methods and the structure of the documents, keeping notes with their answers.

6.2 Analysis Results

6.2.1 Analysis Task Results

During the questions of the tasks ([Appendix B](#)), I wrote the answers as also the time that was used to complete a task from the users. First of all, it would be necessary to refer the total errors from all the participants for every method separated. The results of these errors are presented to the next table (Table 3). This table shows the effectiveness from the users for the three methods. The majority of the errors were occurred at the concept keyboard, because the users touched the lines of the rectangles, but using the concept keyboard, the line of the rectangles is the border of the concept keyboard cells. The users pressed on the line, so, they pressed on the borders of the cells as a resulting to not get a feedback. On the other hand, when they used the mouse a simple detection on the borders of the rectangles was enough to get an output.

A simple solution to this problem would be the use of full rectangles for the rectangles with content, to be clear the position of the rectangle without too much effort. There are more tables about the errors for each method comparing the different types of document at the [Table F.1](#), [Table F.2](#), and [Table F.3](#) for the speech method, the mouse and speech method and the concept keyboard respectively.

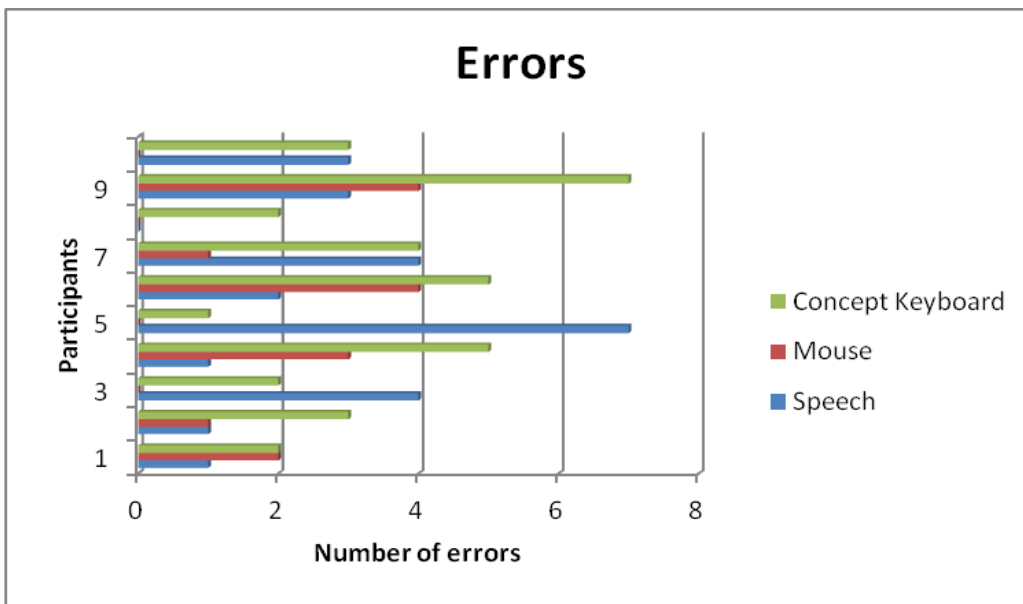


Table 3: Total errors for all the methods from the users.

Participant number	Speech Synthesis	Mouse & Speech	Concept Keyboard
1	0:08:48	0:04:24	0:08:13
2	0:07:55	0:07:16	0:14:03
3	0:04:55	0:04:20	0:09:14
4	0:10:34	0:03:35	0:09:42
5	0:06:30	0:08:28	0:09:50
6	0:10:49	0:08:06	0:07:58
7	0:08:43	0:13:48	0:12:54
8	0:10:21	0:05:00	0:08:38
9	0:04:11	0:04:20	0:08:40
10	0:05:02	0:09:41	0:05:59
Average	0:07:47	0:06:54	0:09:31

Table 4: Total times for all the methods from the users.

It was deserved that the first method which was evaluated kept longer than the next two methods. Users seems to improve their abilities and to be familiarized with the methods or the symbols. At the first method, they needed more time to find a specific area or to give an answer but, at the end of the evaluation, they used fluently the technology as also they understood the meaning of the rectangles emmediately.

6.2.2 Questionnaire Analysis

A questionnaire was created in order to evaluate the usability and effectiveness of our project. The effectiveness and the usability of the project separated in several questions. First question was based on method 1-speech, second on method 2-mouse and speech and third on concept keyboard. Each different type of document, html, excel, power point and pdf, were evaluated by putting a grade ranging from 1-5 and also grades like very good, good, average, bad and very bad. Our results showed that our system provides usability and effectiveness.

The graph (Figure 6.1 in Appendix G) shows the results from the answers about the different types of documents. The results from this graph indicates that the structure of the pdf documents was graded as average, which means that the user find some difficulties to use the mouse for the exploration of the pdf but not so much to be impossible the access of this structure. The majority of the users believe that the excel documents was the most easy and effectively structures to get access using the speech method, because they graded the excel document as “very good”.

The graph (Figure 6.2 in Appendix G) shows the same results as the previous but using the method - mouse. This graph indicates that almost all the users believe that the structure of the excel document was the best, because they graded this structure as “very good”. They choose also that the html documents was “bad” which means that they found difficulties through the exploration of these documents.

The same grades from the last method – concept keyboard are showed in the graph (Figure 6.3 in Appendix G). It’s too clear that the use of this method is too easy for excel documents as also, the access to the html document using the concept keyboard was more effective than the two previous methods. In general, the majority of the users believe that the structure of the documents are more easy accessible using the concept keyboard than using the method with the mouse or the speech synthesis.

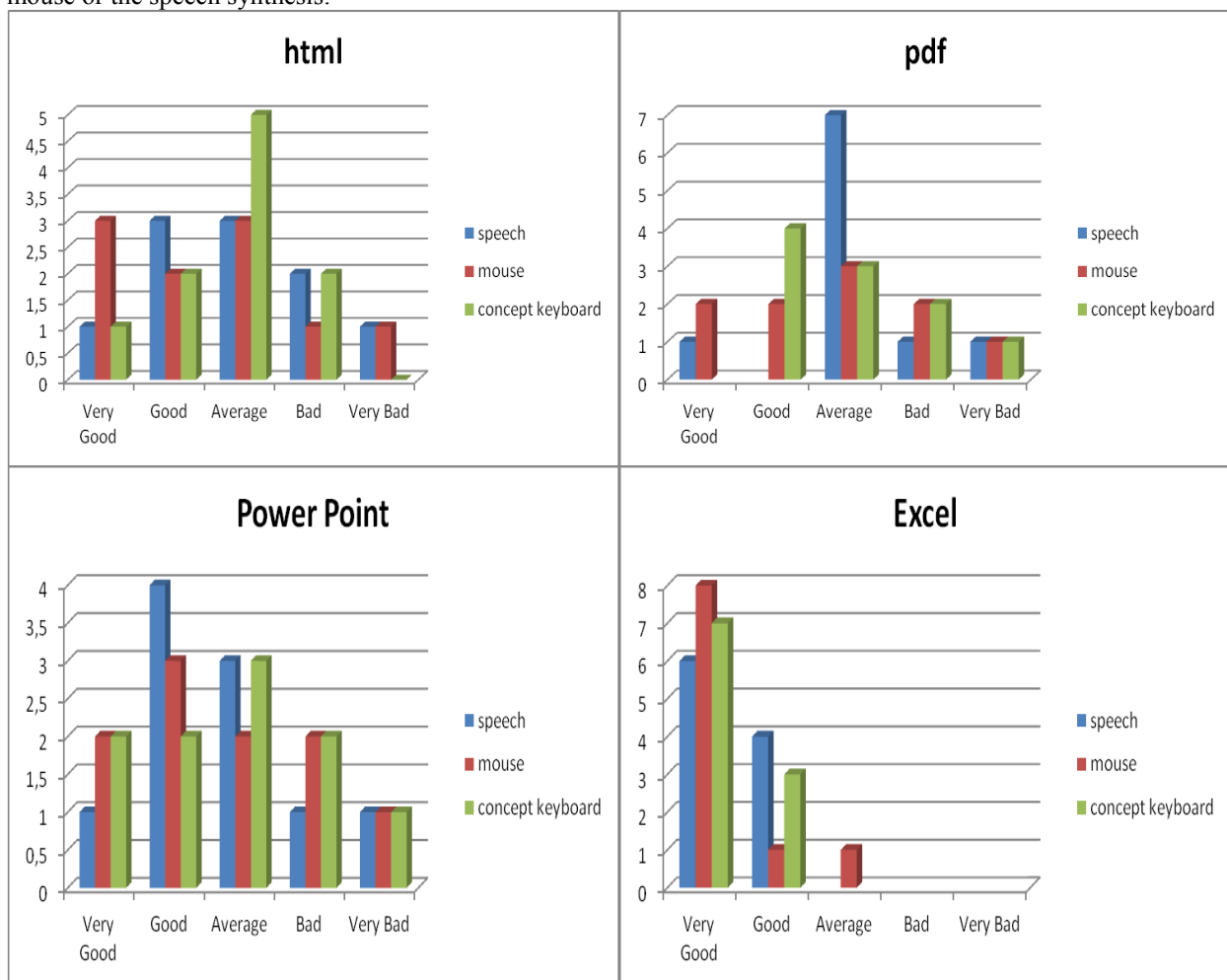


Figure 6.4: Results comparing the methods for each type of document

The figure 6.4 shows the results for every document. We compared the three methods with every type of document. The majority of the users evaluated the html document as average when they used the concept keyboard. The majority of the users believe that the speech method was the average method for pdf documents. The majority of the users believe that the speech method was a good method for the power point documents. Finally, 8 users believe that the mouse method was the best method for the excel document, after the concept keyboard and then the speech method.

At this point, it's interesting to compare the three methods asking from the users to grade as very good, good, average, bad and very bad (Figure 6.5). The results showed that the majority of users believe that the concept keyboard was a good method to use the structure of the documents because; they graded the method – concept keyboard as “good”. The next method that they choose called method 2- mouse and speech synthesis. Four users believe that the method – mouse was very good. The last method seems to be the speech method with the majority of the users grades this method as “good”.

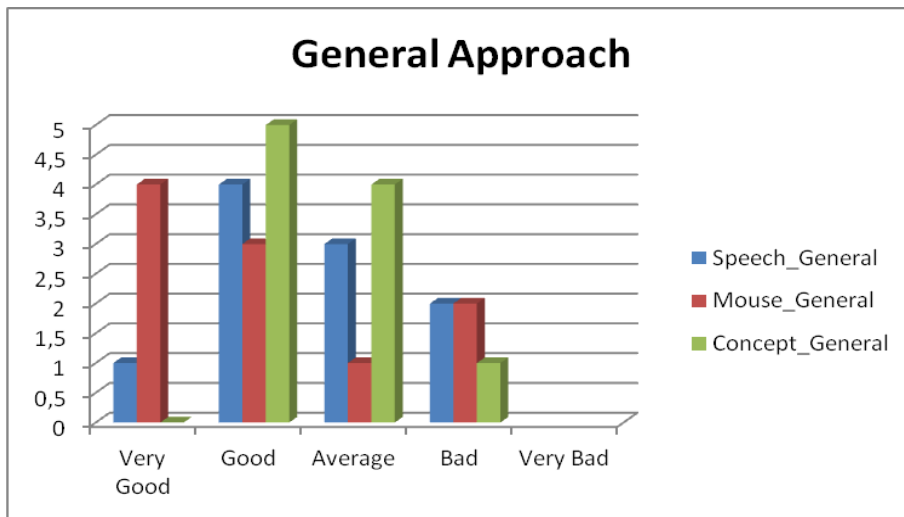


Figure 6.5: Results from the comparison of all three methods

It would be more interesting to mention the results of the last question in the questionnaire. We asked the users, to classify the methods according to their own opinion as “best”, “average” and “worst”. The results of this question are showed at the figure 6.6. The majority of the users believe that the method 1 – speech synthesis are graded as average. The next bigger amount of the users believe that the concept keyboard are the best method and immediately following from the method 2 – mouse and speech which was graded also as the best method but we have to take into account the users who evaluated the method 2 – mouse as the worst method. 40% of the users believe that the method 2- mouse as also the method 1 – speech were the worst methods.

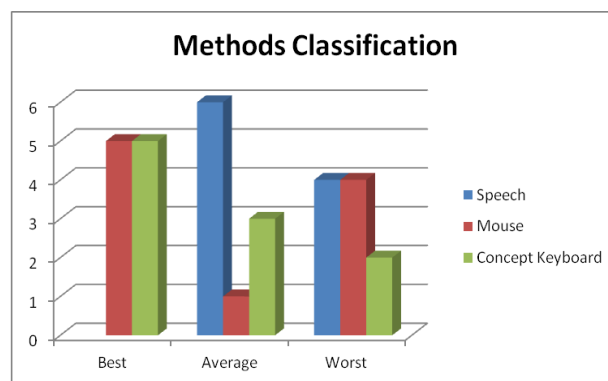


Figure 6.6: Results from the classification of all three methods

According to the questionnaire, 50% of the users believe that the relation between a figure and a simple rectangle are comprehensible. On the other hand, 60% of the participants didn't understand the relation between the figure and the sub-figures. To be more specific, the users didn't receive the connection between an image and the legend of this image. The amount of 75% of the users perceives easily the global structure of the documents. (Figure 6.7)

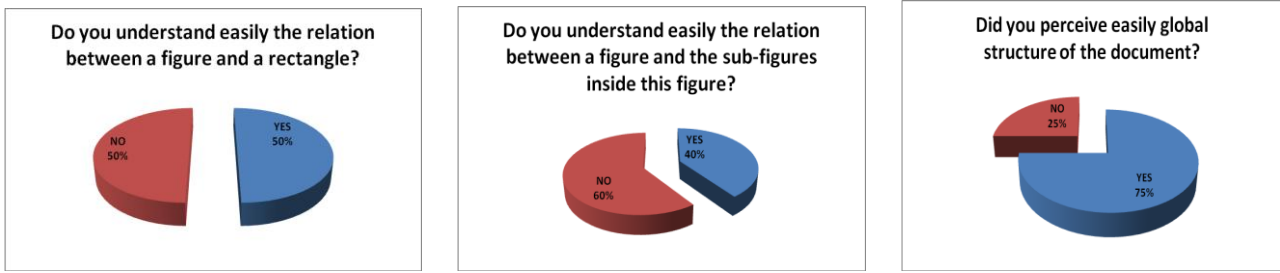


Figure 6.7: Percentages from the users who answer the previous questions

The next graph shows the results about the symbols that were used in the documents. All the users could understand the rectangle and perceive the content of the rectangles with easy. On the other hand, the half of the users didn't understand the differences with the rectangle with vertical lines for subsection. The majority of the users found easy the meaning of the symbol with the cycle which presents the bullets or lists, and they could understand the difference through the tactile. Almost all the users could understand the meaning of the bullets/ list and text as a rectangle with a cycle and a simple rectangle respectively. Some of the users found difficulties for the presentation of the subsection; they didn't recognize the meaning of symbol as also a serious difference with the simple rectangle. (Figure 6.8)

The most important thing is the meaning of the figures and sub-figures and the connection with the rectangles. The users could not understand the connection between the figures and sub-figures as also the meaning of the rectangle which presents a figure. (Figure 6.8)

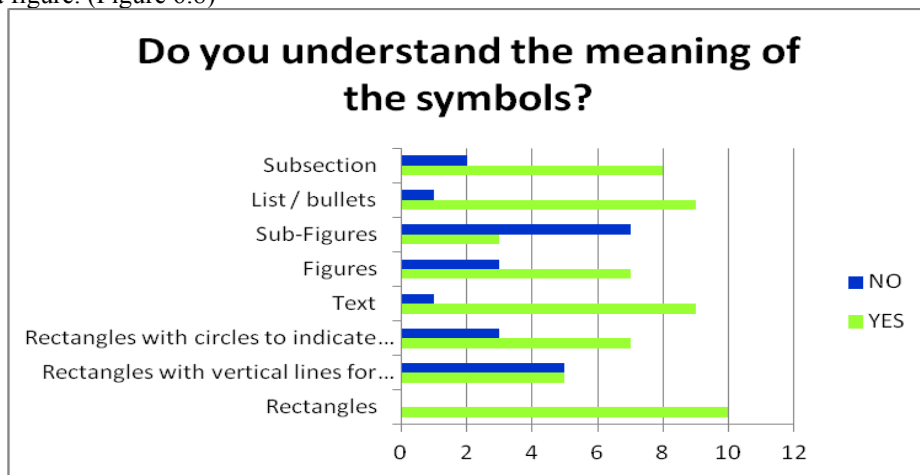


Figure 6.8: The answers of the users for the symbols

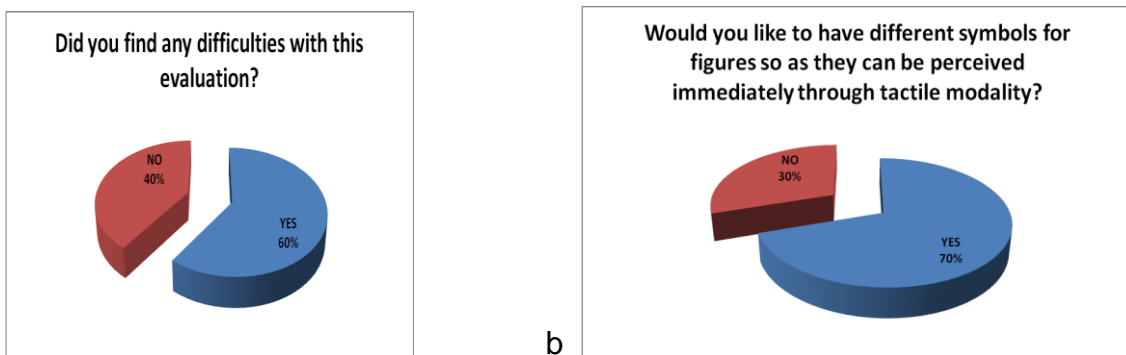


Figure 6.9: a)The answers of the users for the evaluation difficulties, b)the preference of the users to change the symbol of the figures

70% of the users argue that the symbols for the figures should be completely different from all the others symbols, so as they could be perceived immediately through the tactile modality. Many of the users focused their attention to this problem, the connection between the figure and sub-figures as also the symbols for these areas. They proposed to be completely different from the other rectangles, using a symbol that could indicate immediately that this rectangle contains the content of an image. Some of the suggestions were rectangles with a big cycle or a cross at the center, or rectangles with different width for the lines, or to use different high of the rectangles for figures.

Finally, the amount of 60 % of the users found difficulties during the evaluation. They mentioned that the difficulties were focused on html documents because of the high complexity, the html documents consist of many rectangles and it was difficult to find the area that they searched. A user referred that the difficulties were only at the beginning until familiarize with the document and the method. Also, they found difficulties to identify the type of the document. It was interesting the answer of a user that he wanted sometimes to use both hands. At the end, some users mentioned that the voice of the speech synthesizer was not too clear all the time.

6.2.3 Analysis the user's behavior

During the evaluation, we observed the movements from the users' finger as also the user's behavior for every task. We have to mention some necessary factors about the methods, the symbols and the users' reaction for the tasks.

First of all, some of the users explored the document following only the lines, the borders of the rectangle and not all the rectangle, to understand the width of this area. In this manner, they didn't receive the whole structure of the document. They touched only the lines of the rectangle and they got the output from the speech synthesizer, but in case of html documents, which comprised of many areas, many different rectangles, they received wrong output. They touched only the lines, they got the output without understand the position of the text area and the empty area. In some cases, they touched the line with their finger out of the borders, believing that the rectangle with the text was there, but in reality this area was empty.

In many cases, the users were pressing an empty area waiting for an answer, because they didn't recognize the area with the content and an empty area. The reason for this situation was the lack of the full comprehension of the rectangles and generally for the structure. Many participants didn't make the movement to discover the perimeter of the rectangle as a result the lack of the full comprehension of the structure. The users touched the lines and they received the output from speech synthesizer but they were not able to know on which side was the content and where was empty area. This happened because either the users moved their hand for a while as a result to touch an empty area or they thought that the rectangle was found at this position (they didn't try to find the limits of the rectangles).

Furthermore, users explored the structures making only a horizontal or vertical movement with their hand. They crossed the document horizontally or vertically without thinking to explore the spatial position of the rectangles. They heard the audio feedback when they touch incidentally inside a rectangle, without discover the dimensions of the rectangle.

Most of the times, users were hasty in finding an answer, as a result to make quickly movements crossing the documents from the top to the bottom or the inverse, and in some cases, they didn't wait to listen the full description of a rectangle.

The most significant problem was the connection between the figures and sub-figures or the images and the legend of the image. The majority of the users pressed inside the image – figure, and they waited to hear the legend of the image. They heard the text “part (a) of figure 1” which means that this is the first part of the legend but they didn't comprehend this connection. Almost all the users pressed inside the image to hear the legend of the image, the minority of the users understands, after a while, the connection between them and they found the legend but the others didn't give an answer or they gave a wrong answer. Comparing the figures of the pdf structure with all the others figures, almost all the users found the legend of the images. The reason was that the legend in one of the pdf documents is in the same image - rectangle (the legend of the image is in the same rectangle and the sub-figure contains the general overview of the image).

The next important issue for the evaluation came from the excel documents. The users loosed the line or the columns that they tried to follow as a result to change line without understand this movement. They should follow a vertical line but suddenly, they changed line without understand this change. The point is that they gave a wrong answer, they told the value from the next line and not from the desired line even though, they heard the audio feedback with a different description.

Finally, it's necessary to mention that some users made a movement with their fingers hiding the blue sticker as a result to lose the detection of the blue color. In their effort to explore the document, they hid their finger with the sticker with the middle finger or they bend their finger losing also the detection.

In conclusion, some users were using the whole hand to explore the document not only one finger or they used both hands even though they heard the description from the rectangles only from the finger of one hand. In some cases, they put the one hand (usually the left hand) at the top of the document and they used the right hand to explore the structure. They used the left hand just to indentify the top of the document.

Taking into account all the previous observations, it's necessary to suggest solutions about these problems. First of all, it would be necessary to be used full rectangles and not only the perimeter of rectangles, with this way, user could

identify the position of the area with content in contrast the empty area.

In case with the excel documents, it could be used vibration following the lines or the columns depending from the structure of the document. The vibration could provide a manual for the blind to indentify the series of the lines or the borders of the rectangles.

It would be necessary also, to use different symbol for the figures or sub-figures to be understanding through the tactile modality the content of this area not only from the speech synthesizer.

6.3 Conclusion - Summary

According to the previous analysis, the half of the users believes that the best method was the concept keyboard and the others believe that the best method was the mouse method. On the other hand, there was a large amount of the users who underpinned that the mouse was the worst the method. The majority of the users support that the speech method was in average.

Taking into account the previous percentages of the users willing, the best method was the concept keyboard including the times and the type of the errors in concept keyboard and the mouse method. However, the fact that some users pressed on the lines of the rectangles giving wrong answers, requires immediately correction and system optimization.

In concluding, the symbol of the rectangle was completely understanding for all the users but it's absolutely necessary to create a different symbol for the figures according to the users' answers to the questionnaire. Furthermore, the connection between figures and sub-figures requires immediate improvement.

7 Conclusion and Perspectives – Future Work

The implementation of this project showed basic results about the structure of the documents, as also the effectiveness of the three different methods. Finally, the results from the evaluation presented the users' needs and the most necessary technologies.

This is an important step to continue a promising project which could meet the blinds' requirements about the accessibility in spatial information of a series documents.

At the future, there are a series of interesting searching areas that could be implemented. First of all, it's necessary to use full shapes instead of empty rectangles.

Furthermore, we should compare the best method with the classical approach using screen readers to find out the easier way for blind to perceive the spatial structure of documents.

It's necessary also, to use Braille terminal "Bi-manual access" to present the text of the content into the rectangles.

In real life, it would be impossible to use embossed papers to present the structure of the documents. However, the use of dynamic dot-matrix display is necessary to present automatically the structure in this device using HyperBraille.

It would be absolutely necessary to develop a software about automatically inversion of the documents from digital type to spatial – structures documents.

There is also the need to create more complicated structures, using more symbols about different kind of information. In this way, it would be easy for blind people to perceive an overview of the document through the tactile modality.

Specifically, in automatically conversion of the documents from digital to two-dimensional structured document using the necessary technology, there are some points that need more attention. For instance, it will be interesting to find a solution in automatic conversion of the documents when it's used two images in the document. In automatic conversion, it will be difficult for the system to "understand" the connection between the image and the legend of this image automatically, without the human help. The solution of this problem would be a challenge for the researcher to continue this project to the next level.

It would be interesting also to create structures using image inside to text area. This structure is presented in figure 7.1. This schema would be difficult and complicated for the blind users; they should provide too much effort to perceive this meaning of the structure, so it would be necessary to be found a way for easily access on such structures.

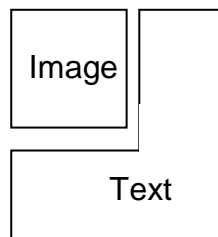


Figure 7.1: A structure of a document in future work

8 Bibliography

- [1] World Health Organization, Visual Impairment and Blind
<http://www.who.int/mediacentre/factsheets/fs282/en/index.html>
- [2] W3C <http://www.w3.org/standards/webdesign/accessibility>
- [3] Lazar, J., Allen, A., Kleinman, J., & Malarkey, C. What Frustrates Screen Reader Users on the Web: A Study of 100 Blind Users. (2007). *International Journal of Human-Computer Interaction*, 22(3), 247 - 269.
- [4] The Universal Access Handbook, Constantine Stefanidis edition, 2009
- [5] Wai Yu, Stephen Brewster, S., Comparing two haptic interfaces for multimodal graph rendering, Haptic Proceedings 10th Symposium on Interfaces for Virtual Environment and Teleoperator Systems, 2002.
- [6] Victoria E. Hribar, Laura G. Deal, Dianne T.V. Pawluk, Displaying Braille and graphics with a “Tactile Mouse”, ASSETS ’12, October 22-24, 2012
- [7] Thomas H. Massie and J. K. Salisbury, The PHANTOM Haptic Interface: A Device for Probing Virtual Objects, Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Chicago, IL, Nov. 1994.
- [8] Ravi Kuber and Wai Yu, Tactile vs Graphical Authentication, 2010, In proceedings of Eurohaptics (Lecture Notes in Computer Science 6191), Amsterdam, Netherlands, 314-319
- [9] Haptic Interaction, Gunnar Jansson and Roope Raisamo, The Universal Access Handbook, 2009, Constantine Stefanidhs (ed)
- [10] Gapenne, O., Rovira, K., Ali Ammar, A., & Leney, C., *Tactos: Special computer interface for the reading and writing of 2D forms in blind people*, (2003) Stephanidis, C. *Universal Access in HCI, Inclusive Design in the Information Society*, Heraklion, Greece, 1270-1274.
- [11] Ebina, T., Igi, S., Miyake, T., & Takahashi, H., *Graph Access System for the Visually Impaired*, (1998). *Proceedings of the 3rd Asian Pacific Computer and Human Interaction*, Shonan Village Center, Japan, 254-259.
- [12] Wang, Q., & Hayward, V., *Compact, Portable, Modular, High-performance, Distributed Tactile Transducer Device Based on Lateral Skin Deformation*. (2006). *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'06)*, Arlington, VA, USA, 67-72.
- [13] Shinohara, M., Shimizu, Y., & Mochizuki, A., Three-dimensional tactile display for the blind. (1998). *IEEE Transactions on Rehabilitation Engineering*, 6(3), 249-256.

- [14] Wang, Q., Lévesque, V., Pasquero, J., & Hayward, V., *A haptic memory game using the STRESS² tactile display*. (2006). Olson, G. *Proceedings of CHI EA '06 Extended abstracts on Human factors in computing systems*, Montréal, Québec, Canada, 271-274.
- [15] J Kildal & SA Brewster, Explore the Matrix: Browsing Numerical Data Tables Using Sound, Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display, Limerick, Ireland, July 6-9, 2005
- [16] Kildal, J. and Brewster, S.A., *VibroTactile External Memory Aids in Non-Visual Browsing of Tabular Data*. In Vol II Proceedings of First International Workshop on Haptic and Audio Interaction Design 2006 (Glasgow, UK), pg 40-43
- [17] S Wall, S Brewster, Providing external memory aids in haptic visualisations for blind computer users, Proceedings of 5th Intl Conference on Disability, Virtual Reality & Assoc. Tech., Oxford, UK, 2004
- [18] Denise Prescher, Gerhard Weber and Martin Spindler, A tactile windowing system for blind users, *ASSETS'10*, October 25–27, 2010, Orlando, Florida, USA.
- [19] Maria Schiewe, Wiebke Kohlmann, Olivier Nadig, and Gerhard Weber, What you feel is what you get: mapping GUIs on planar tactile displays, C. Stephanidis (Ed.): Universal Access in HCI, Part II, HCII 2009, LNCS 5615, pp. 564–573, 2009
- [20] Wang, Q., & Hayward, V. *Compact, Portable, Modular, High-performance, Distributed Tactile Transducer Device Based on Lateral Skin Deformation*. (2006). *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (HAPTICS'06)*, Arlington, VA, USA, 67-72.
- [21] Shinohara, M., Shimizu, Y., & Mochizuki, A. Three-dimensional tactile display for the blind. (1998). *IEEE Transactions on Rehabilitation Engineering*, 6(3), 249-256.
- [22] Wang, Q., Lévesque, V., Pasquero, J., & Hayward, V. (2006). *A haptic memory game using the STRESS² tactile display*. Olson, G. *Proceedings of CHI EA '06 Extended abstracts on Human factors in computing systems*, Montréal, Québec, Canada, 271-274.
- [23] Gunnar Jansson and Roope Raisamo, Haptic Interaction, The Universal Access Handbook, 2009, Constantine Stephanidis (ed.)
- [24] OpenCV library and documentation, <http://docs.opencv.org/>
- [25] Microsoft Speech Synthesizer
<http://msdn.microsoft.com/en-us/library/system.speech.synthesis.speechsynthesizer.aspx>
- [26] Mohamed benali-Khoudja, Jean-Marc Alexandre, and Abderrahmane Kheddar, Tactile interfaces: a state of the art survey, 35th International Symposium on Robotics, Paris France, 2004.
- [27] Goldish, L. H., & Taylor, H. E. The Optacon: A Valuable Device for Blind Persons. (1974). *New Outlook for the Blind*, 68, 49-56.

- [28] G. Petit, "Conception, prototypage et evaluation d'une système pour l'exploration audio-tactile et spatiale de pages web par des utilisateurs non-voyants", Thèse de doctorat, Ecole polytechnique de Montréal, 28 Août 2013.

Acknowledgments

First of all, I would like to thank my supervisors Yacine Bellik and Malika Auvray for giving me the opportunity to work with them and to be a part of the Limsi AMI team. The last 6 months of my life, I have been more productive and pressed ever in my life. Without the support, help, guidance of my supervisors, my master thesis could not be the same. Mr Bellik gave me the opportunity to work with him learning the needs of blind people which I really care about them and I ignored their needs until now.

I would like to thank my professor Constantine Stefanidhs who teach me and encourage me by his lessons to work in Human Computer Interaction laboratory.

I would like to thank also, the users who took place at the evaluation of our system and they helped to complete the most important part of my project.

I would like also to thank the University of Crete, the University Paris – Sud 11 and Mr. Yannis Manousakis who gave me the opportunity to be a part of the Greek-French Postgraduate program.

I am grateful to the secretary in the Computer Science department of University of Crete, for helping me in many different ways. Mrs. Rena Kalairzaki deserves special mention.

I would like also to convey thanks to the Faculty of FORTH – ICS and CNRS – Limsi for providing the financial means and laboratory facilities.

Last but not least, I wish to thank my family. They bore me, raised me, supported me, taught me, and love me. I dedicate this thesis to them.

Kalliopi STRATIGI



Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI-CNRS)

BP 133, 91403 Orsay Cedex, France, <http://www.limsi.fr>

Agreement Document

In this project, we do research about the “Bi-manual access to two-dimensional information for the blind”. This study is conducted under the responsibility of Mr. Yacine Bellik, Lecturer at the University Paris-Sud XI, Orsay.

This study consists in comparing three different methods to help blind people better perceive the structure different types of document. The main purpose of this evaluation is to find the best method that helps the blinds to better understand the content of the document effectively.

Agreement Document

Title of the experiment: Bi-manual access to two-dimensional information for the blind

I certify that I have given my consent to participate in this study organized by LIMSI-CNRS and coordinated by Mr. Yacine Bellik. I voluntarily agree to participate in this study and I understand that my participation is voluntary and I can stop my participation at any time without having to explain myself any liability. My consent does not relieve the organizers of the search of their responsibilities and I keep all my rights guaranteed by law.

During this experience, I agree on data collection of my answers in video recording format and I give all my rights on this recording. I understand that the collected information is strictly confidential and for exclusive use of the involved investigators.

I was informed that my identity will not appear in any report or publication and all information concerning me will be treated confidentially. I agree that data recorded during this study may be stored in a database and used by LIMSI-CNRS.

Date:

Name of the volunteer:

Signature of volunteer (preceded by the words "read and approved"):

Name of the experimenter:

Signature of experimenter:

Results Document: Subject n°.....

Questions on the experience:

1. *You have performed a series of tasks using speech synthesis method, speech synthesis - mouse, and concept keyboard.*

Method

Task_Ppt.1:	Success/ Errors	TIME
Task 1. What's the first word of the title		
Task 2. Where is the logo-image		
Task 3. What is the 1 st word of the paragraph with bullet		
Task 4. What is the 1 st word in 3 rd subsection		
Task 5. What's the page number		

Method

Task_Ppt.2:	Success/ Errors	TIME
Task 1. What's the first word of the title		
Task 2. Where is the image for "smart home"		
Task 3. What is the 1 st word in 2 nd subsection		
Task 4. What is the 1 st word in 1 st subsection		
Task 5. What is the name of lesson		

Method

Task_Ppt.3:	Success/ Errors	TIME
Task 1. Where is the logo		
Task 2. Where is the image		
Task 3. What is the 1 st word in 1 st subsection		
Task 4. What is the content in 2 nd subsection		
Task 5. What's the page number		

Method

Task_Pdf.1:	Success/ Errors	TIME
Task 1. What's the first word of the paragraph "future investigations"		
Task 2. What's the first word of paragraph 4.1.1		
Task 3. What's the legend of sub-figure (a) of figure 1		
Task 4. What's the name of the 1 st author of the 3 rd reference		
Task 5. What's the page number		

Method

Task_ Pdf.2:	Success/ Errors	TIME
Task 1. What's the first word of the paragraph “image categorization”		
Task 2. Where is the image with cycle		
Task 3. What's the first word of the paragraph “discussion and conclusion”		
Task 4. What's the name of the 1 st author of the 2 nd reference		
Task 5. What is the number of references?		

Method

Task_ Pdf.3:	Success/ Errors	TIME
Task 1 What's the first word of the title		
Task 2. What is the 1 st “author keyword”		
Task 3. What's the legend of sub-figure (b) of figure 1		
Task 4. Where is the first image(Prototype tactile system)		
Task 5. What is the 1 st word in paragraph 1.1.1		

Method

Task_Excel Articles:	Success/ Errors	TIME
Task 1. Where is the article "Sport News"		
Task 2. What are the sales of the article "Times" at March		
Task 3 What are the sales of the article "national geographic" at February		
Task 4. What are the total sales for the article "Travel Guide"		
Task 5. What are the total sales for the article "the guardian"		

Method

Task_Excel Cars:	Success/ Errors	TIME
Task 1. Where is the car "Renault"		
Task 2. What are the sales for Renault at April		
Task 3. What are sales for Ferrari at February		
Task 4. What are the total sales for Rover		
Task 5. What are the total sales for Honda		

Method

Task_Excel Grades:	Success/ Errors	TIME
Task 1.What is the name of student Gabriel		
Task 2. What is the date of birth for the student Smith		
Task 3. What is the grade in Physics for the student Aston		
Task 4. What is student's average for the student "Travolta"		
Task 5. What is course's average for the lesson Computer science		

Method

Task_Html.1 (LIMSI):	Success/ Errors	TIME
Task 1. Where is the search area		
Task 2. What is the number of different languages		
Task 3. What is the 1 st word of the title		
Task 4. What is the 1 st word of the content- main area		
Task 5. Where is the main menu area		

Method

Task_Html.2 (FORTH):	Success/ Errors	TIME
Task 1. What is the 1 st Institute in “research institutes”		
Task 2. Where is the “home” link		
Task 3. Where is the area for “More research news”		
Task 4. What is the title		
Task 5. What is the 1 st word at 2 nd new in the content		

Method

Task_Html.3 (M2R_Interaction):	Success/ Errors	TIME
Task 1. Where is the menu		
Task 2. Where is the area to log in		
Task 3. Where is the path of this site		
Task 4. What is the 2 nd choice in menu area		
Task 5. What is the title		

Evaluation Questions: Subject n°.....

User Profile

Date:

Age:

Gender:

Informatics Level:

Expert

Intermediate

Beginner

Braille Level (Completed only by Braille readers):

Expert

Intermediate

Beginner

1. Which kind of technology / method do you prefer to interact with the system? (One presentation way from questions - evaluation)

Method 1- Only Speech Synthesizer:



	5 Very Good	4 Good	3 Average	2 Bad	1 Very bad
HTML					
PDF					
Power Point					
Excel					
In general Or "across all type of documents"					

Method 2- Mouse and Speech:

	5 Very Good	4 Good	3 Average	2 Bad	1 Very bad
HTML					
PDF					
Power Point					
Excel					
In general Or "across all type of documents"					

Method 3- Concept Keyboard:

	5 Very Good	4 Good	3 Average	2 Bad	1 Very bad
HTML					
PDF					
Power Point					
Excel					
In general Or "across all type of documents"					

Answer the questions with "YES" or "NO"

	YES	NO
2. Do you understand easily the relation between a figure and a rectangle?		
3. Do you understand easily the relation between a figure and the sub-figures inside this figure?		
4. -Would you prefer to use only one rectangle for the whole figure or to use a rectangle for each sub-figure?		
5. Did you perceive easily global structure of the document?		

6.

- *Do you understand the meaning of the symbols? (Rectangles, rectangles with lines, number of lines etc). Do you think you have completely understood the meaning of the symbols?*

	YES	NO	Ideas from users
Rectangles			
Rectangles with vertical lines to indicate sub-sections			
Rectangles with circles to indicate lists			
Text			
Figures			
Sub-Figures			
List / bullets			
Subsection			

Do you have suggestions about other symbols to use?

.....
.....

7. *Did you find any difficulties with this evaluation? If yes, what part of the evaluation was the more difficult for you?*

YES **NO**

Notes:
.....

8. *Would you like to have different symbols for figures so as they can be perceived immediately through tactile modality? If yes do you have a suggestion about the symbol to use?*

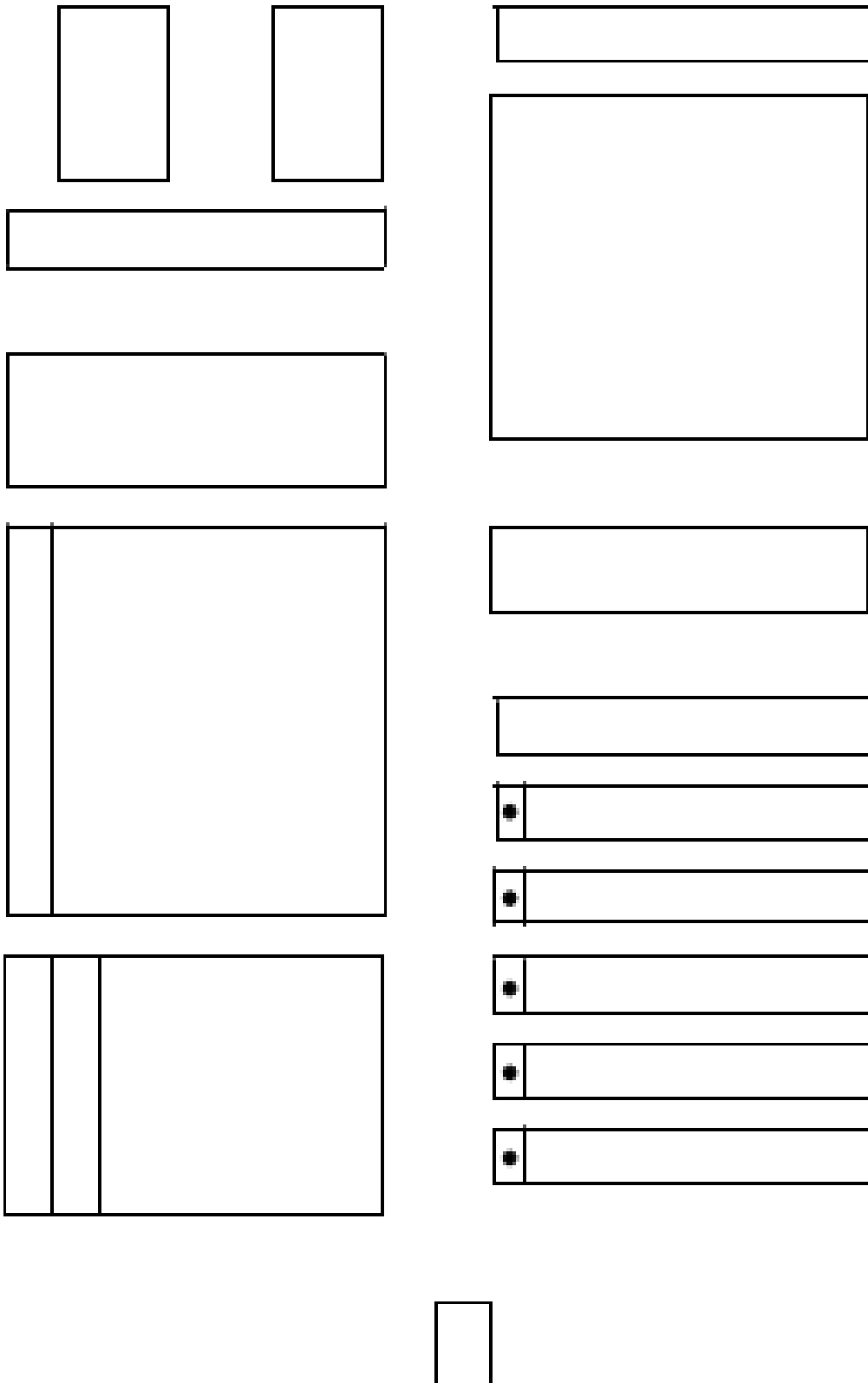
YES **NO**

Notes:.....
.....
.....

9. *Classify the 3 methods: Best: Average..... Worst:*

10. Notes:

Pdf Structure:



Appendix E

Table:

			Ppt1	Ppt2	Ppt3	Pdf1	Pdf2	Pdf3	Excel 1	Excel 2	Excel 3	Html 1	Html 2	Html 3
User1	1	Speech			✓			✓			✓			✓
	2	Mouse	✓			✓			✓			✓		
	3	Concept Keyboard		✓			✓			✓			✓	
User2	3	Speech		✓			✓			✓			✓	
	2	Mouse	✓			✓			✓			✓		
	1	Concept Keyboard			✓			✓			✓			✓
User3	2	Speech			✓			✓			✓			✓
	1	Mouse	✓			✓			✓			✓		
	3	Concept Keyboard		✓			✓			✓			✓	
User4	2	Speech	✓			✓			✓			✓		
	3	Mouse		✓			✓			✓			✓	
	1	Concept Keyboard			✓			✓			✓			✓
User5	3	Speech	✓			✓			✓			✓		
	1	Mouse			✓			✓			✓			✓
	2	Concept Keyboard		✓			✓			✓			✓	

			Ppt1	Ppt2	Ppt3	Pdf1	Pdf2	Pdf3	Excel 1	Excel 2	Excel 3	Html 1	Html 2	Html 3
User6	1	Speech			✓			✓			✓			✓
	3	Mouse	✓			✓			✓			✓		
	2	Concept Keyboard		✓			✓			✓			✓	
User7	1	Speech			✓			✓			✓			✓
	2	Mouse	✓			✓			✓			✓		
	3	Concept Keyboard		✓			✓			✓			✓	
User8	2	Speech			✓			✓			✓			✓
	3	Mouse	✓			✓			✓			✓		
	1	Concept Keyboard		✓			✓			✓			✓	
User9	3	Speech	✓			✓			✓			✓		
	1	Mouse		✓			✓			✓			✓	
	2	Concept Keyboard			✓			✓			✓			✓
User10	2	Speech	✓			✓			✓			✓		
	1	Mouse			✓			✓			✓			✓
	3	Concept Keyboard		✓			✓			✓			✓	

Results Analysis:

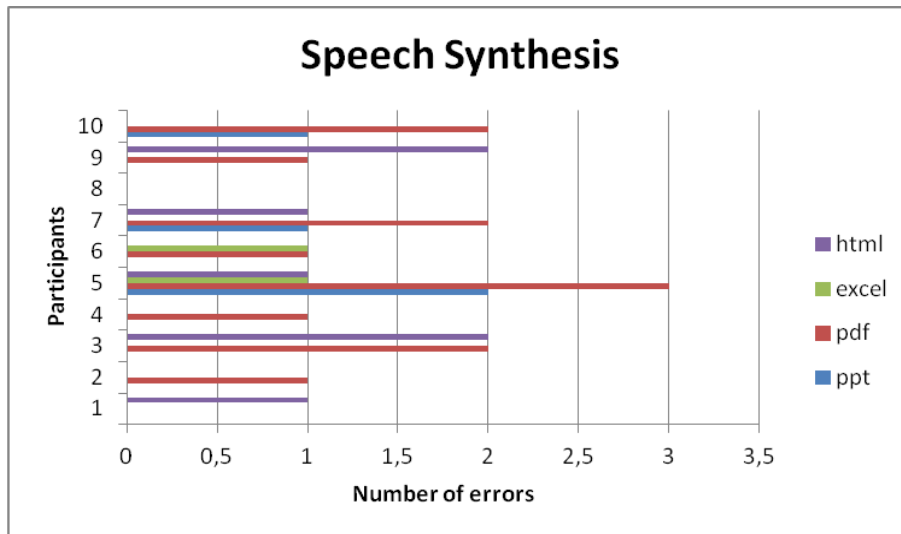


Table F.1: Number of errors from all the users for each type of document for the speech method

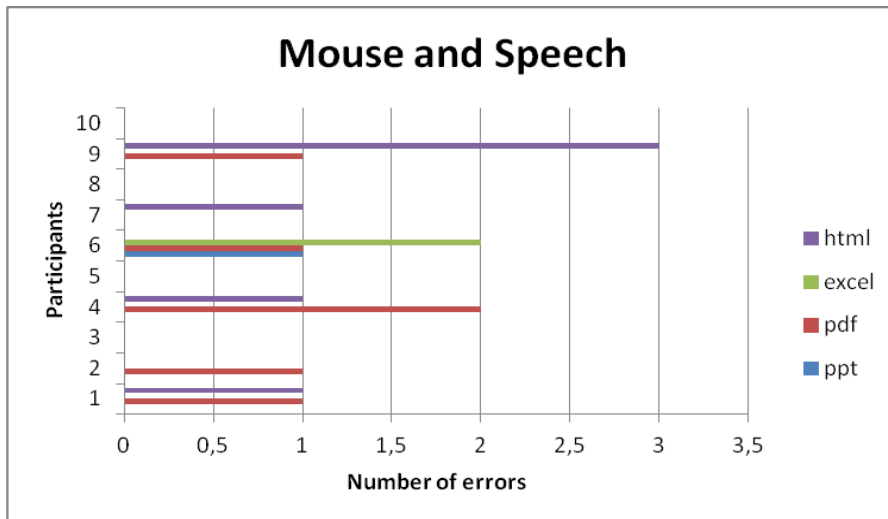


Table F.2: Number of errors from all the users for each type of document for the mouse and speech method

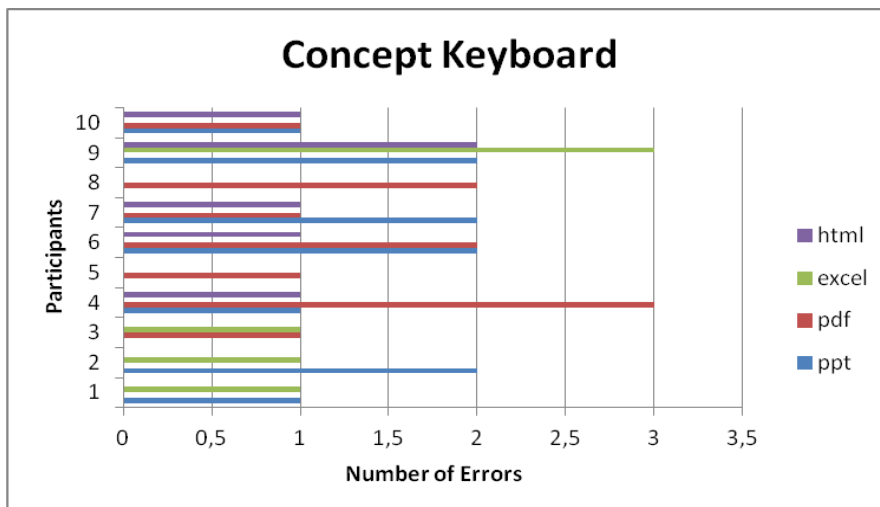


Table F.3: Number of errors from all the users for each type of document for the Concept keyboard

Figures 6:

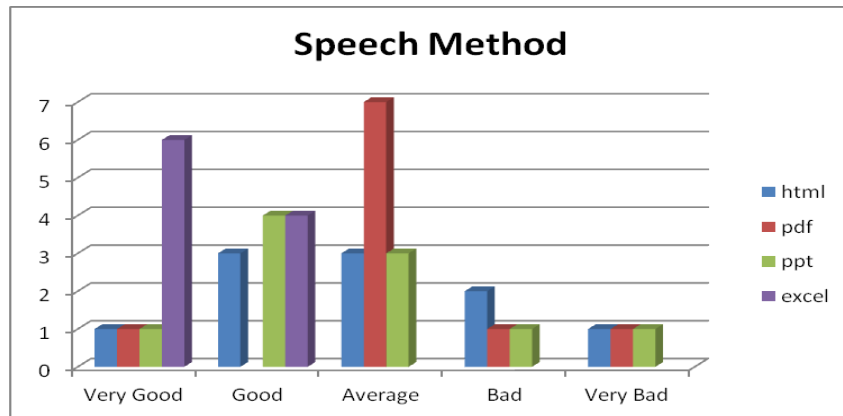


Figure 6.1: Results for the speech method according to all the kinds of document

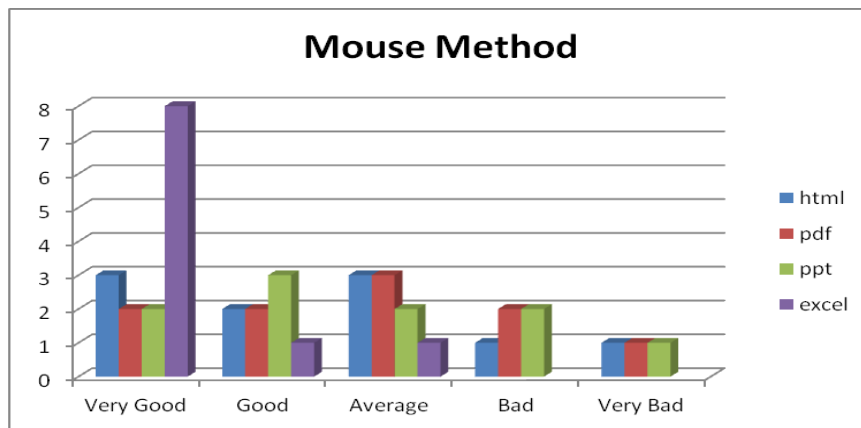


Figure 6.2: Results for the mouse method according to all the kinds of document

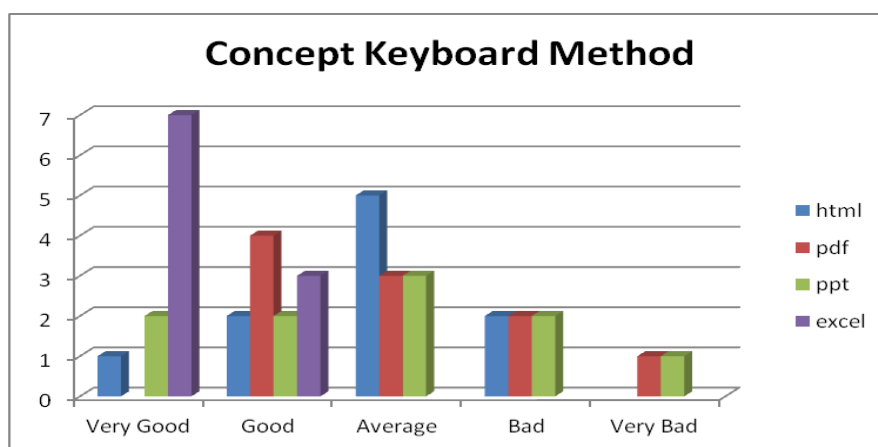


Figure 6.3: Results for the concept keyboard method according to all the kinds of document

