

# Voice Browsing Approach to E-Business Access: A Blind's Perspective

Muhammad Nabeel Talib (corresponding author)

School of Management, Huazhong University of Science and Technology

Wuhan, 430074, China

E-mail: nabeeltalib@hotmail.com

Shuqin Cai

School of Management, Huazhong University of Science and Technology

Wuhan, 430074, China

E-mail: caishuqin@sina.com

Muhai Hu

School of Management, Huazhong University of Science and Technology

Wuhan, 430074, China

E-mail: humuhai@163.com

## Abstract

Information accessibility on World Wide Web (WWW) still remains to be a complex issue for blind users as majority of websites are invaded by contents, both non-visual (audio) and visual (video & images). Accessibility measures in terms of a usable and blind-friendly website should be made available to blind users. A particular area that can be improved on is the e-Business segment for blinds. Easily-accessible, e-Business platforms could further promote entrepreneurship practices through the development of blind-friendly features such as voice. In current study, we investigated the existing difficulties (such as inconsistent webpage structure, incompetent voice recognition engines and use of long sentences as commands, etc) to access the web-content and suggested a different (numeric hotlist) approach, in order to make the web more accessible, exclusively for blinds.

**Keywords:** Voice browser, E-business, Web accessibility, Vocabulary, Numeric hotlist, Dictation mode

## 1. Introduction

Most web applications attract a majority of users but still the efforts to make them blind-friendly are not enough. The core issues of insufficiency of these applications are: a) Accessibility standards (provided by Web Content Accessibility Guidelines (WCAC), under Web Accessibility Initiative (WAI)) are not followed b) Inefficient speech recognition engines in recognizing voice commands accurately. Waibel & Lee (1990) pioneered the voice technology systems which enabled tasks such as bank balance inquiries, flight schedules and phone call transfers. Such voice technology-related advancements were especially useful in the field of medicine, acknowledged later on by Bajgoric (2006). Investigated by Carter & Markel (2001), following are the major organizations that address web-accessibility issues.

- W3C's (World Wide Web Consortium) Web Accessibility Initiative is the mother lode of web accessibility information.
- The Center for Applied Special Technology (CAST) sponsors Bobby, a program to check the accessibility level of a website.
- The Trace Center is a research organization at the University of Wisconsin-Madison.
- National Center for Accessible Media (NCAM), researches, develops and tests methods of integrating web access technologies.
- WebABLE encourages web accessibility.
- AWARE (Accessible Web Authoring Resources and Education), educates developers to focus on web accessibility.
- Microsoft's Enable also encourages developers to manufacture accessible products.

A consistent structure by using headings and lists would really facilitate to state the browsing issues as we

propose the idea of using Numeric Hotlist for navigating a webpage through voice, which would assist blind users to interact with the system more naturally. The impacts of voice enabled systems have yet to be revealed but speech-enabled mobile phones, navigational systems, and automobiles might make a distinct contribution towards blind people's life. Although future would certainly bring a lot of technological advancements but the pace of development, in this particular area (voice enabled systems), remains to be slow. The paper is structured as follows; Section 2 provides a brief literature review on existing voice technology and its weaknesses. An overview of voice browsing is given in section 3. Section 4 presents analysis and suggestions to enhance the existing models with the help of voice technologies (such as Microsoft Speech Software Development Kit). And section 5 gives a future outlook for the web semantics with special emphasis on blinds.

## 2. Literature Review

### 2.1 Problem Definition

The term web accessibility refers to easy access by all, however, Web Accessibility Initiative, a group managed by World Wide Web Consortium (W3C), describes web accessibility as to create and develop the web in such a way that allows blind and older people to not only access but contribute to the web same as any normal person Talib, Shuqin, Abrar & Shafiq (2009). WAI has defined certain standards for web accessibility to keep the web evolving in a single direction. These standards involve every type of disability where access to the web is affected either visually, aurally, physically, cognitively, neurologically or by speech.

### 2.2 Existing Web

Web Content Accessibility Guidelines (WCAG 2.0) provides a foundation for developing web applications by offering necessary web accessibility support but unfortunately the practical use of these guidelines is very inadequate and insufficient. Most of the websites do not meet even the basic standards provided by WCAG 2.0, making the websites inaccessible for blind users. WCAG 2.0 listed seven common barriers to web accessibility for blinds:

- images without alternative text
- image map hot spots without alternative text
- misleading use of structural elements on a webpage
- uncaptioned audio or un-described video
- lack of alternative information for users who cannot access frames or scripts
- tables that are difficult to decipher when linearized
- sites with poor color contrast

The web developers should understand the practical difficulties faced by blind users, as Asakawa (2005) and Fukuda, Saito, Takagi & Asakawa (2005) believed that web developers try to comply with the web accessibility guidelines without understanding the pragmatic requirements of blind users. The structure of the web pages are complex, and simple accessibility rules are not followed which results in web contents not being accessible by blind and elderly users Rana, Reynolds, Cirstea & Entecott (2007). Only in China, approximately 12.33 million people of the population have some form of visual impairment. About 314 million people are visually impaired worldwide; 45 million of them are blind World Health Organization (2009). These statistics make it clear how fast the numbers of blind users in the world are growing.

In our opinion, web accessibility is not only important for blind or elderly people; it is also a concern for those people who are naïve to the computer or web. Discussing on this issue, it was disclosed that almost 97% of the websites do not provide the basic accessibility standards and this statement was based after accessibility agency Nomensa (2006), tested the leading websites in five different sectors (travel, retail, banking, government and media), across 20 countries. This news is not only shocking but shows the lack of responsibility taken by web developers who completely ignore the needs of a whole range of (blind) people and hence contradict to the idea of web access for all.

### 2.3 Technology

WWW is becoming the ultimate source of information, so access to web gets more vital as for web developers; the most complex development task is to address the blinds. Some solutions were proposed to overcome this difficulty, out of which one was, Screen Readers Charles, Hemphill & Thrift (1995) but unfortunately, screen readers can only provide the text of the applications and its contents; however it is difficult to deliver the same interaction by audio (by converting text into speech). Screen readers and voice browsers (another proposed

solution) usually read the document from top left to the bottom right corner which makes it difficult for blind users to access a specific point on the page especially when they are at the bottom. It is because of the inconsistent structure of the document which always varies from site to site as the guidelines are not followed Matlay (2004).

Furthermore, Buzzi, C, Buzzi, M, Leporini & Akhter (2009) described that interaction via screen reader involves several issues such as:

- Information overloading, hence time consuming
- Repetition of information to reach the required one
- Poor structure leads to user annoyance in finding the relevant information
- Reading row wise from tables makes the information out of order
- Links, buttons and content's titles should be self explanatory
- Finding it difficult to work with form control elements
- Alternative descriptions for visual and non-visual contents should be provided.

### 3. Voice Browsing

A browser is a software application typically used to bring the contents of (an HTML) webpage into display, which allows quick access to the links, buttons, text and pictures. Asakawa & Itoh (1998) emphasized that blind people can easily access the web by using non-visual browsers such as a voice browser. Furthermore, Christian, Kules, Shneiderman & Youssef (2000) described that a voice browser is at least capable of rendering web pages in audio format or it can interpret speech input for navigation. Supported with technology, an ideal web browser is capable of two-way communication with the user through speech or voice. It not only allows the user to listen to what is on the screen, but also allows the uttering of specific commands (e.g. next page, skip, active links on page and submit forms), all by voice.

Figure 1 demonstrates how the Voice Browser works; the input voice passes through the voice recognition process and would invoke an action in the browser only if it matches with one of the voice commands/vocabulary available in the voice recognition engine. However, integrated use of scripts like JavaScript, and visual interfaces like flash, makes it even more difficult for voice browsers to audibly present the web page to blind users. Gupta and Kaiser (2005) observed that ads, non-description link text e.g. 'click here', 'see more' and inaccessible forms on the web pages seems to be a big problem which restrict accessibility developers to develop a generic approach to extract the useful contents from a webpage. Theofanos and Redish (2003) blamed excessive use of graphics and image-links preventing a solid solution for blinds. Moreover, Rudall & Mann (2006) found that long input/command is more complicated and prone to errors. But inconsistency of web structure remains a restriction for developers to come up with a universal solution; as it is impossible to change the whole web.

#### 3.1 Web Architecture

The heading tags such as <H1>, <H2> are not used by web developers and to synchronize the contents, navigational aids are used. However, navigation through these aids is not a trouble-free task for blind users. Relating to this, the structure of the web can be built with a semantic approach, interpreting the visual content (by scanning) into an appropriate, sequenced audio (Optical Character Recognition technology can be used to convert the image if the image represents some text, more accurately printed text) which will not only assist blind users to navigate properly but issues like time constraint, user annoyance and informal structure, will also be addressed as well Kouroupetroglou, Salamasis & Manitsaris (2006).

#### 3.2 Related Work

Charles, Hemphill & Thrift (1995) offered the idea of flexible vocabulary and dynamic grammar interface by designing a speech user Agent. The user can interact with the system by commands or speak-able hotlist, to control the browsing experience like scroll down, back as shown in the figure 2. The user can use the speak-able hotlist in which a grammar can be associated with a URL, for example, 'Weather in Wuhan' should open the <http://www.bbc.com/weather/wuhan>. The inspiration of smart pages was introduced to implement this in which, grammar is defined and for a given link many alternative grammars can be introduced. So, an idea which started to make the existing web useful ended in a list dependent approach.

Ramakrishnan, Stent & Yang (2004) realized the actual problem with voice browsers (and screen readers), which are now fully capable to read the text on the screen (or alternative text for images), but are unable to convey the

logical structure and semantics of the content in a web document. In our opinion a dynamic conversion approach, is required instead of attempting to change the whole existing web, which is not possible in anyway. So, a universal solution is required using existing browsers (or using a new standard browser), which can convert the existing web into a useable format (for blinds).

#### 4. Analysis and Findings

In order to improve voice accessibility, we need to have a voice recognition engine and in current scenario we opt to use Microsoft's Speech Software Development Kit (MSSDK) v5.1, which provides a fine opportunity to develop high quality speech applications, offering enormous voice functions. However, MSSDK does not accept a long list of voice commands at any one time, which is around 400-500 commands for the numeric list. For "natural text commands" (or word commands), this capacity is almost less than 70 commands. The programs can be set in to Dictation Mode on user request, where the user can add and/or alter the entered commands word by word, as described in the table 1. Reasonable speech input option should be offered, where the user can comfortably input the data through speech and navigates between contents. Voice output is also one of the biggest setbacks as sometimes the voice quality is poor and hard to understand but the good news is that there are 'Voices' available in the market which are very similar to the normal speech output for Voice Browser.

From the above mentioned information, we can understand that a command should not be longer than 3 words to be recognized successfully, i.e. if each command consists of a single word then nearly 60-85 commands can be offered otherwise we would have to offer only 30-50 set of commands (see table 1). This indicates that only limited set of commands can be offered for recognition at any one time.

The proper and efficient use of numeric commands (hotlist) in Dictation Mode is quite helpful but is a huge task to be achieved. Dynamic grammar and improved voice recognition in dictation mode are some of the open options to be explored in this field. Web Accessibility through speech can be greatly improved, if all the applications, algorithms and web structure developed in compliance with the WCAG guidelines. There is a need of such tools, which can help developers to build web pages and generate contents according to the guidelines making it more convenient for blinds to access web. In an attempt to enhance the accessibility features for the web, Voice Browser also provides prospects to increase the accessibility for windows, its applications and commands. Opening new horizons to pursue the voice access for operating system and web based technologies is indeed a huge task, which requires more research, corroboration and support from all the scientific community.

#### 5. Conclusion

Voice accessibility, has a great potential as a marketing tool, and might not only increase the financial worth, by providing opportunities for the blind to become a part of online business activities but also the company's repute for its social responsibilities. Improvement in speech recognition abilities would definitely help to enhance the capabilities of voice browsers, so that more generic and effective solutions to escalate the formal e-Business access could be obtained, explicitly for blinds. The optimum focus of current effort can be summarized as: introducing numeric hierarchies (same as HTML heading tags) by including dynamic voice commands; introducing the whole range of document object model elements to offer comprehensive voice accessible browsing solutions; reading data from pictures (and graphics) using Optical Character Recognition (OCR) techniques and offering it through speech; reading webpage data and associating related data with each other by introducing semantic relations; reading data from objects and scripts e.g. Flash, JavaScript, VB Script and converting it to understandable (speech) format; extracting and converting pop-ups, alerts and banners into useful data; searching specific text within elements; and finally moving ahead by handling windows through voice commands. Furthermore, we also need to take account of noise level, microphone being used and computer hardware to get the fine quality of voice recognition. Use of help command at any segment of navigation between web content, can significantly reduce the height of complexity for blind users. Global expansion in e-Business volume will open new prospects for the blind to be a part of it; either as a consumer or entrepreneur, which will reveal niche markets creating new opportunities with an enormous potential in previously an untargeted e-Business area.

#### References

- Asakawa, C. & Itoh, T. (1998). User interface of a home page reader. *Proceedings of ACM ASSETS, ACM SIGACCESS Conference on Assistive Technologies*, pp. 149-156
- Asakawa, C. (2005). What's the web like if you can't see it. *ACM International Conference Proceeding Series*, 88, 1-8
- Bajgoric, N. (2006). Information systems for e-business continuance: a systems approach. *Kybernetes*, 35(5),

628-648

Buzzi, M. C., Buzzi, M., Leporini, B. & Akhter, F. (2009). User Trust of ecommerce Services: Perception via Screen Reader. *Proceeding of NISS*, pp. 1166-1171 in press.

Carter, J. & Markel, M. (2001). Web accessibility for people with disabilities: an introduction for Web developers. *IEEE transactions on professional communication*, 44(4), 225-233

Charles, T., Hemphill, C. T. & Thrift, P. R. (1995). Surfing the web by voice. *Proceedings of ACM International conference on Multimedia*, pp. 215-222

Christian, K., Kules, B., Shneiderman B. & Youssef, A. (2000). Comparison of voice controlled and mouse controlled web browsing. *ACM SIGACCESS Conference on Assistive Technologies*, pp. 72-76

Fukuda, K., Saito, S., Takagi H. & Asakawa, C. (2005). Proposing new metrics to evaluate web usability for the blind. *Conference on Human Factors in Computing Systems*, CHI, pp. 1387-1390

Gupta, S. & Kaiser, G. (2005). Extracting content from accessible web pages. *ACM International Conference Proceeding Series*, 88, 26-30

Harper, S., Goble, C. & Stevens, R. (2000). A pilot study to examine the mobility problems of visually impaired users travelling the web. *ACM SIGCAPH Computers and the Physically Handicapped*, 68, 10-19

Horton, S. (2006). Designing beneath the surface of the web. *W4A: Proceedings of international cross-disciplinary workshop on Web accessibility*, pp. 1-5

Kouroupetroglou, C., Salampasis, M. & Manitsaris, A. (2006). A semantic-web based framework for developing applications to improve accessibility in the WWW. *ACM Proceedings of international cross-disciplinary workshop on Web accessibility (W4A): Building the mobile web: rediscovering accessibility*, 134, 98-108

Matlay, H. (2004). E-entrepreneurship and small e-business development: towards a comparative research agenda. *Journal of Small Business and Enterprise Development*, 11(3), 408-414

Nomensa (2006). Accessibility Surveys. [online]. Available at: <http://www.informationweek.com/news/internet/showArticle.jhtml?articleID=196601676> (August 24, 2010)

Ramakrishnan, I. V., Stent, A. & Yang, G. (2004). Usability and accessibility- hearsay-enabling audio browsing on hypertext content. *Proceedings of the 13th international conference on World Wide Web*, pp. 80-90

Rana, M. M., Reynolds, T., Cirstea, M. & Entecott, A. (2007). Check points to design an accessible web application in Microsoft dot net framework for visually impaired people. *2nd International conference on Internet technologies and applications*, UK, pp. 173-180

Rudall, B. H. & Mann, C. J. H. (2006). Advances in the development of semantic e-business. *Kybernetes*, 35(5), 613-615

Talib, M. N., Shuqin, C., Abrar, M. & Shafiq, M. S. (2009). E-business access for blinds: A semantic approach. *Proceedings of International Conference on E-Business and Information System Security, Wuhan*, 1, 214-217

Theofanos, M. F. & Redish, G. J. (2003). Guidelines for accessible and usable web sites: Observing users who work with screen readers. *Interactions*, 10(6), 36-51

Theofanos, M. F. & Redish, J. G. (2003). Bridging the gap: between accessibility and usability. *ACM Interactions*, 10, 36-51

WAI Group, W3C. Introduction to web accessibility. [online]. Available at: <http://www.w3.org/WAI/intro/wcag20.php> (July 11, 2010)

Waibel, A. & Lee, K.F. (1990). *Readings in speech recognition*. San Mateo, California: Morgan Kaufmann, pp. 1-3

World Health Organization (WHO) (2009). Visual Impairment and Blindness. [online]. Available at: <http://www.who.int/mediacentre/factsheets/fs282/en/> (August 24, 2010)

Table 1. Command list in dictation mode for voice browser

<b>Command List in Dictation Mode</b>			
<i>Sr</i>	<i>Type</i>	<i>Command</i>	<i>Purpose</i>
1	Text Entry	Enable Text Entry	To start a Text input
2		Disable Text Entry	To finish a Text input
3	Custom: Link Entry	Enable Link Entry	To start a Link or URL Entry input where space is not inserted
4		Disable Link Entry	To Finish a Link or URL Entry input
5	Advanced Custom: Command	Enable Command	To Enable Command Processing Mode
6		Disable Command	To Disable Command Processing Mode
7	Print	Print Page	Automatically sends the print of the current webpage to the default printer
8	Refresh	Refresh Page	Refresh the current webpage
9	Home	Home Page	Takes to the Home webpage (preset in Internet Explorer)
10	Forward	Forward Page	Takes to the next page in memory(if visited and available in memory)
11	Backward	Backward Page	Takes to the previous page in memory
Windows Commands			
12	Log off	Log off	Will force the Windows to log off
13	Restart	Restart	Will force the Windows to Restart
14	Shutdown	Shutdown	Will force the Windows to Shutdown

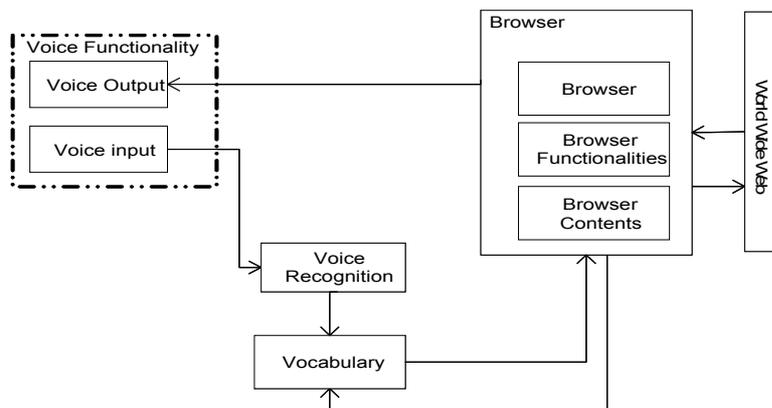


Figure 1. Conceptual components and their interaction diagram

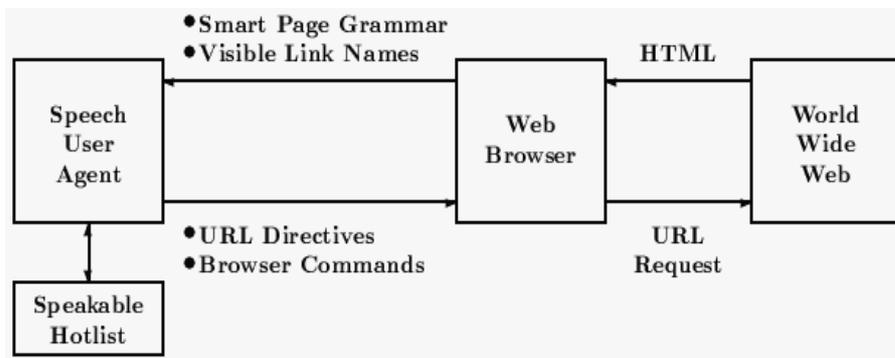


Figure 2. Adopted from Surfing the web by voice, (Charles et al, 1995)