Proceedings of the Second IASTED International Conference
HUMAN-COMPUTER INTERACTION
March 14–16, 2007, Chamonix, France
ISBN Hardcopy: 978-0-88986-654-6 / CD: 978-0-88986-655-3

USING CATALOGUE BROWSING FOR SPEECH-BASED INTERFACE TO A DIGITAL LIBRARY

Yael Dubinsky, Tiziana Catarci, and Stephen Kimani
University of Rome "La Sapienza"
DIS (Dipartimento di Informatica e Sistemistica)
Via Salaria 113, 00198 Rome, Italy
{dubinsky,catarci,kimani}@dis.uniroma1.it

ABSTRACT
Catalogue browsing is a well known activity in the library world. We are all familiar with the box of cards that is used by the librarian who browsed it over in order to find the card for the book we were searching for. The Catalogue Browsing Access Paradigm (CBAP) that is suggested in this paper aims at providing the users with the same feeling of the actual catalogue browsing activity while using the benefits of the digital world. In particular, relationships between the physical and the digital realms are investigated. We illustrate the CBAP concept using a prototype presentation of a speech-based mobile interface to a digital library (DL). A physical to digital relationship is examined when using vocal commands to control the search activity. A digital to physical relation is examined when using digital instructions from a positioning system in order to navigate in a physical library. We use agile software development and user centered design as the methodologies in the development of CBAP, and present user evaluation data that emerged from this process.

KEY WORDS
Catalogue browsing, speech-based mobile interface to digital libraries, agile software development, and user centered design

1. Introduction

The design of modern digital libraries (DLs) is a complex task that usually requires the participation of a number of different parties, such as end-users, librarians, publishers and developers, as well as the formation of a multi-disciplinary team including design and human factors experts [14,5]. A major challenge involved in such a process is to acquire and formulate the functional and non-functional requirements that drive appropriately the design of the DL in question.
In previous work, functional and non-functional DL requirements were identified and a model is suggested for DL usage lifecycle [6]. In addition, interaction paradigms are explored based on the DL requirements study. This paper focuses on one of the interaction paradigms named the Catalogue Browsing Access Paradigm (CBAP) that has been defined together with specific recommendations to design effective catalogue in the DL case.
Catalogue browsing is a well known activity in the library world. Browsing the box of cards the librarian finds the card for the book being searched for. This card contains the information that characterizes the specific book e.g., the name of the book, the name of the authors, the publisher name, the year of publication, and most important its physical location. The CBAP concept is to provide the users with the same feeling of the actual catalogue browsing activity while tapping into the benefits of the digital world. The relationships between the physical and the digital realms are of interest.
In Section 2 we describe the research that is conducted as part of the DELOS Network of Excellence on Digital Libraries[1]. Specifically, we present the concept, design, and implementation of CBAP by presenting a prototype of a speech-based mobile interface to DL. In Section 3 we describe the way we merge agile software development[2] and user centered design methodologies, and present evaluation data with respect to the speech aspect of CBAP. We conclude in Section 4.

2. The Catalogue Browsing Project

In this part we present the research that is conducted as part of the "User Interfaces and Visualisation" Work Package (#4) of the DELOS project (Section 2.1), and delve into the details of CBAP (Section 2.2). We note that this part is based on DELOS internal reports [6,3]. The CBAP prototype is detailed in Section 2.3.

2.1 Research Description

Traditional information access systems consist of interfaces that are mainly based on "search" and "search-refinement" activities. In order to use them, the information need of the user is always assumed to be

[1] See DELOS site at http://www.delos.info/.
[2] See the agile manifesto in http://agilemanifesto.org/.

well-defined and clear, and ready to be articulated into specific queries for the system. However there are various situations in which this assumption does not hold, i.e., user information need is not always clear and precisely formalized, hence making such query-based interfaces hard to use. These scenarios include cases like novice user, new collection, vague information need [12], vocabulary problem [8,9], and exploratory learning. The foregoing scenarios show that alternative access paradigms should be explored.

In recent years we have experienced rapid technological evolution like the success of the World Wide Web, the diffusion of various kinds of interactive applications, and the availability of different end-user devices. Still, Digital Library interfaces are still based mainly upon "search" and "search refinement" mechanisms. In previous work we identify functional and non-functional requirements of DLs, with the aim to establish an empirical basis for user interface design for DLs and to define taxonomy and visualisation paradigms [6]. Additionally, a preliminary DL usage lifecycle model has been elaborated, targeted to facilitate further analysis of user requirements [3].

The results of such an empirical study call for the investigation of the potential effectiveness and benefits to the user stemming from a full adoption of alternative interaction paradigms, and especially of novel techniques for navigation such as browsing by catalogues, semantic linking, information visualisation, interactive maps, social navigation, etc., which are seldom and occasionally employed in current DLs. Our specific task aimed at systematic investigation of non-conventional interaction paradigms, and the correlation of such paradigms with different usage phases of DLs.

2.2 The Catalogue Browsing Access Paradigm

The Catalogue Browsing Access Paradigm (CBAP) is a non-conventional access paradigm for DLs that may be effectively used in DLs to meet more sophisticated (and often neglected) user needs, going beyond traditional query-based interfaces. A specific effort is devoted to the investigation of CBAP, bringing expertise and design knowledge common in other domains (such as e-commerce, e-learning) to the lifecycle of DL development.

Motivation. As a basic requirement, an online catalogue should support the primary functions of a card catalogue: finding and collocation functions [18]. The online catalogue can help to better identify library entities "in terms of their nature, scope and orientation through different data fields such as intellectual level, document type, genre, language code, geographic area code and additional notes" [7]. The design of online catalogues should explore the new possibilities offered by technology to better match emerging needs and requirements of online behavior. In fact it has been argued before that online catalogues are still hard to use because they often are designed without sufficient understanding of searching behavior [4]. An assessment of the effectiveness of online catalogue design should not be based on its success in matching queries but rather by its success in answering questions. All these processes may help to disambiguate or take into account the context of a user's information need, thus enabling her/him to find appropriate answers to a need and acquire a better understanding of knowledge structures in a certain domain. It should also be pointed out that the design of online catalogue systems has often failed to consider the social collaborative dimension of searching behavior that can be easily observed in physical libraries. An online catalogue system should provide assistance to the searcher where necessary, and follow the idea that bibliographic records serve as information "seeds" to fertilize subsequent searching [10].

In light of this, the following should be addressed:

- Complexity of searching material that has been classified according to librarians' taxonomies;
- Dynamic evolving information needs;
- Time constraints, channel or device constraints and digital convergence.

High-level Design. We suggest a DL solution that is based on CBAP to provide library access and seamless interaction with physical and digital entities [2]. The effort relies on mobile computing [13] and CBAP in order to facilitate the library artefact access and seamless interaction, since it proposes a more natural link between the card catalogue artefact (as used by visitors of a physical library) and online catalogue interfaces (as part of the services provided by DL). The relationships between the physical and digital realms are part of the design and are described in Section 2.3 as part of the user requirements.

Cognitive Aspects. DL systems are hardly equipped to adapt to *natural user behaviour* and help the users circumvent the problems that arise [11]. The challenge in DL design is to better understand users' *information seeking behaviour*. One aspect of the users' information seeking behaviour is their adoption of different *Cognitive Interaction Strategies*. A cognitive interaction strategy is a category of cognitive processes which take place while a user interacts with a system in order to make one of the information retrieval decisions e.g., analysis of need situation, goal evaluation or planning. A user's cognitive strategy is a particular way of thinking about the relationships of the information that is processed.

A cognitive interaction strategy can be characterized as follows:

- It is based on a certain kind of mental representation usually called a mental model of the environment.
- It is based on a particular interpretation of the observations of the environment.
- It serves to help in information retrieval decisions.
- It requires a particular set of resources to be successful.

CBAP imports browsing and exploration paradigms traditionally used in other application context to the domain and needs of DL users and designers. We suggest

that the cognitive interaction strategy that is used by users of CBAP is the interaction strategy by analogy.
The mental model of this strategy is represented by a metaphor or an analogy [15] that provides a mapping between two worlds of concept. With respect to CBAP they are two main mappings. The first is the library metaphor that maps between a physical library and a digital one. The second is the catalogue metaphor that maps between the physical catalogue box and the appropriate digital interface. These mappings assist the users in the process of information interpretation. The main resource requirement is the awareness of users to these mappings that are conveyed by the interface. Such awareness enables users to gradually adopt the digital world and the relationships involved.

2.3 The CBAP Prototype

The main goal of the CBAP prototype development project is to provide a seamless interaction between the physical and the digital realms in accessing library artefacts based on the concept of CBAP. The development process adopts an integration of agile principles [1] and the user-centred design methodology [17]. The interface that is developed is a speech-based mobile interface to a DL. Beyond the search activity, two additional features were defined. The first feature is enabling vocal commands for artefact searching and localization. The second feature is enabling artefact localization in the physical library using a digital positioning system. Speech input is enabled for navigating the application, and speech output is enabled for the positioning instructions. Table 1 summarizes the set of requirements in details as were presented and prioritized by the customer at the beginning of the first release.

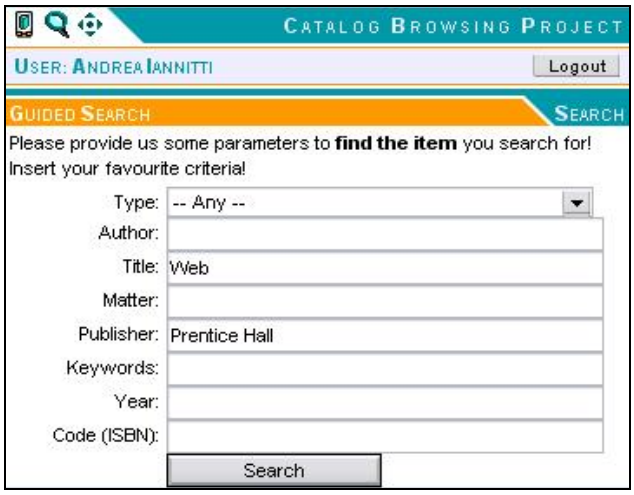
Table 1. Customer stories for CBAP prototype

Customer story	priority
1. Mobility - a) The application is web-based and able to run on mobile devices b) The user can move while using the application (a librarian with a bunch of books in one hand and a PDA in the other)	I II
2. CBAP - The user interface should be inspired by the typical catalog cards box.	I
3. Navigating between physical and digital realms – location and speech a) The user should be able to search in order to find a book of his/her interest. The information can be filtered by queries on topics and/or authors. It should be useful to readers and librarians. [digital□:digital] b) The user should be able to receive instructions about the physical location of the book that he/she searches for (as a reader) or wants to put back on the shelf (as a librarian). That is, the application has to provide information about the path to follow inside the library, in order to be able to physically hold the book or put it back on the correct shelf. [digital□:physical] c) The application has to show the path (see 3.b) with an output speech interface. In general, the system has to provide a speech I/O infrastructure. Example: the librarian is walking or has a bunch of books on his/her hands, so he/she prefers voice over looking at the PDA screen. Real-world example: GPS. [digital□:physical] d) The application must be able to understand and execute commands given by a vocal input. Example: the librarian has a bunch of books on his/her hands so cannot easily press any PDA button. [physical□:digital]	I II
4. Artifacts for navigating between physical and digital realms – barcodes a) The application should provide a barcode reading infrastructure. [physical□:digital] b) The application should demonstrate the identification of physical items and related digital data by reading their barcodes. [physical□:digital]	I II
5. Related search - The application should suggest, given an item, similar-related ones. Example: once he/she reads a book, the reader is particularly satisfied and curious, so wants to know if there is something else written by the same author, or other same arguments,... Real world example: web interfaces that present "Who bought this, also bought...".	II
6. Clustering - The application should be able to cluster related items when showing them. Example: the librarian wants to retrieve information about every thing produced by the same author that the library has.	II
7. New items disposal - If new products arrive, the application should be able to suggest to the librarian where to put them, considering places congenial to the library's shelves management. A rules interface is to be provided by the system for properly disposing into the library.	II
8. Top list - The application should be able to highlight the "top-"products as the "most popular" or "latest" charts.	II
9. Advertisement - In case of new artifacts, the application should send an email to inform users that new products are available.	II
10. Pictures - The application should show pictures of the items, to let a user recognize them through a visual approach.	II
11. Digital Libraries - The application should be based on a new Digital Library to be developed for the project sake.	II

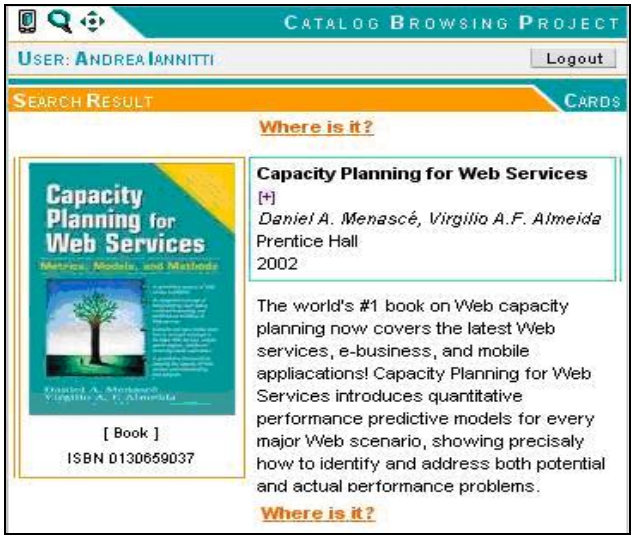
The customer further emphasized that nowadays libraries involve several kinds of artefacts like DVDs, audio and video CDs, etc. This note is significant when analyzed using the library and catalogue metaphor since as perceived by the users, the traditional setting of the library includes mainly books.

The first release of CBAP was performed by two developers during 4 months (from the middle of May till the beginning of September 2006) and was composed of four iterations. Customer collaboration and evaluation by users were emphasized during the process. Measures were taken to control the progress.

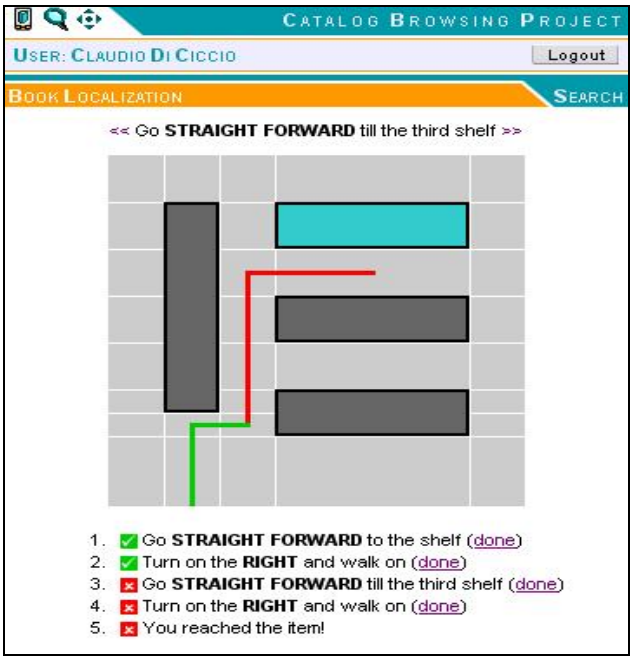
The implementation of the CBAP prototype is performed using the Opera 8.5 browser[3] (W3C® compliant) that supports VoiceXML and XHTML for Microsoft Windows XP systems. In addition, it provides a small screen view that enables the development for mobile applications. Figure 2 presents CBAP screenshots of the *Guided Search* interface (2a), *Search Result* interface (2b) and the *Book Localization* interface (2c).



(a)



(b)



(c)

Figure 2. Screenshots from the CBAP prototype

3. Evaluating the Speech Aspect of CBAP

The evaluation process of CBAP is composed of evaluation iterations that each examines the artefacts of the previous development iteration and results in design changes for the current or next development iteration. The 1st development iteration provides its artefacts. During the 2nd development iteration, the 1st evaluation iteration took place to evaluate and reflect on the artefacts produced in the 1st development iteration and further to decide upon changes that should be introduced. During the 3rd development iteration, the 2nd evaluation iteration took place to evaluate and reflect on the artefacts produced in the 2nd development iteration, and so on. Each iteration is of 3-5 weeks and as aforementioned CBAP first release was composed of 4 such iterations. Following this process of combining the agility concepts with on-going user evaluation that contributes to the design and is performed by the team members, we join the call for the HCI community that is made recently by Norman [16] to embrace rapid and iterative methods and be part of the team for the sake of continuous HCI design.

In the first two iterations the user groups were identified to include librarians and readers, and questionnaires and semi-structured interviews were prepared in order to better understand user needs. In the third iteration a cooperative evaluation was performed with two users in order to learn about users' behaviour with the system and encounter major problems. After the forth iteration ended, meaning the first release was over, we planned and conducted a controlled experiment for the purpose of the evaluation of the speech aspect. The main goal is to

[3] See http://www.opera.com/ about the Opera browser.

provide guidelines for evaluation of speech-based user interfaces that ensures better design of these interfaces.

We conducted a *within experiment* with six participants who are computer science students in different levels, 3 male and 3 female. The experiment task includes login to the system, search activities and book localization activity. The task can be performed using speech (S) or without speech (non-S). Each of the participants performed the task in both modes S and non-S, while 3 participants follow S and then non-S and 3 follow non-S and then S. Further, before starting the experiment, each participant filled an attitude questionnaire and received ten-minute training on how to use CBAP. After the experiment each of the participants filled a questionnaire to reflect on his/her activities.

In what follows we present the experiment qualitative and quantitative data. Table 2 presents the answers of the participants to some initial attitude questions with respect to speech aspect, where SD means that the participant Strongly Disagrees with the statement, D means disagree, A means agree, and SA means strongly agree.

Table 2. Participants attitude to speech interfaces

Statement	SD	D	A	SA
I like interfaces with speech features	1	1	3	1
I have experience with speech interfaces	1		5	
I use speech interface when I can	1	3	2	
People whom I know do not like speech interfaces		4	1	
Speech interfaces are slow	1	2	1	1
I feel uncomfortable with speech interfaces		2	4	
Speech interfaces are fun	1		4	1
Speech interfaces are annoying		4	2	
I expect to use more speech interfaces in the future			3	3
I prefer interfaces that do not include speech		3	3	

We note that when for a specific statement the sum of answers is less than 6, it means that some participants did not answer on this one. As can be observed, the attitudes with respect to speech interfaces are mixed and do not follow a consistent approach. Though speech interfaces are fun they are also annoying, and though participants like them, they do not always prefer them.

The same questionnaire had some open questions asking the participants to provide features that they consider important to be included in speech interfaces, advantages and disadvantages of speech interfaces, and a personal scenario that happened to them when using such interfaces. Following are some of the expressions of participants answering those questions:

- "[consider important] using realistic voices"
- "…I had to provide some information to an automatic operator – it was boring waiting for its answers"
- "[disadvantage] it can take several minutes to interact with speech interfaces"
- "[advantage] they can become friendly"

Examining the answers, two main categories are observed which are user interaction and user friendly. One phenomenon that was found is that participants see speech interfaces as both friendly and not friendly, or as both fast and slow. For example one participant answers the following in two consecutive rows, "[advantage] faster than normal interfaces"; "[disadvantage] a user may wait too long before achieving [his/her] purpose".

After filling the questionnaire, the participants receive one-page users` guide and when completed to read with no more questions, they received the task page according to their appropriate experiment order of S and non-S. An automatic time measure, which was developed as part of the system, provides us with the time stamps of the login / logout and with the time stamps of each search start /end. Table 3 presents the averaged time in minutes that was invested on the two search activities by both experiment groups together with its division per mode.

Table 3. Averaged search time (in minutes)

Group	Averaged search duration	Averaged Non-S search duration	Averaged S search duration
Non-S → S	54.66	28	81.33
S → Non-S	26.58	14	39.16

As can be observed, the S→Non-S group performed the entire task almost twice faster than the Non-S→S group. When looking into the data of speech and non-speech per each group, we see that the participants in both groups performed the speech task slower then the non-speech task. This implies that although the speech task required more time from the participants, they learned better the system when first using it with the speech option.

After completing the CBAP task, participants were asked to fill a questionnaire to reflect on their own activities. Table 4 presents their level of agreement to some statements.

Table 4. Participants reflect after using CBAP

Statement	SD	D	A	SA
I like searching using speech commands		4	2	
I have experience with vocal GPS	1	2	3	
I prefer to work with the silence mode		2	4	
People will feel that the speech mode is too slow		2	4	
I feel uncomfortable with the system I use	1	1	3	1
It was fun			4	2
It was annoying	1	4	1	
I expect I will see such systems in the future			6	

As can be observed, most participants find it hard to use CBAP in its current stage, though it was fun and they expect to such interfaces in the future.
The same questionnaire had some open questions asking the participants to describe what they liked with CBAP, what are the problems they have encountered, their severity ranking between 1– not so important and 5– very important, and to recommend on how to deal with the specific problem. Following are some of the expressions of participants answering those questions:

- "I expect the system to vocally recognize also the value I want to search"
- "It was easy to use; Funny to use"
- "[rank 4] Instructions too fast"
- "It's been a new experience to me"
- "[rank 3] too sensitive to pronunciation"
- "[rank 5] unstable"
- "[I like] the GPS system"
- "I like activating commands by voice"
- "[rank 5] sometimes it doesn't understand what I say"
- "[rank 5] I have to repeat"
- "[recommendation] try to translate to Italian; it should be more flexible with pronunciation"
- "[I like] moving the cursor by speaking"

Examining the answers we learned that we should focus on some improvements that concerns with implementing speech for all interface features and improving the on-line usage information. This is based on our observation that when users are introduced to a speech-based interface they expect it to be fully speech-based meaning no using of keyboard at all. Further, they expect to receive vocal on-line help to assist them in the process of using the application.

4. Conclusion

In this paper we present the concept of the Catalogue Browsing Access Paradigm (CBAP) and the development process of a prototype to illustrate this concept. Agile software development and user centred design are used as the methodologies, thus providing rapid and iterative development that is ongoing evolved based on user evaluation and customer feedback.
We found that with respect to the methodologies used, the user evaluation is fostered by the process agility and the product development benefits from keeping the design updated according to the evaluation outcomes. In addition, a set of evaluation tools including automated features is built and refined during the process and can be further used.
Concerning with the catalogue browsing, we found that users accept this access paradigm in a natural manner as if they used to browse over a catalogue. They enjoy giving vocal instructions to the system and receiving vocal localization guidance from the system. Still, they expect the system to understand as much as possible natural language utterances and be more sensitive to languages and pronunciation.

Acknowledgements

This research is supported by the DELOS Network of Excellence on Digital Libraries (http://www.delos.info/). In addition, we would like to thank Claudio Di Ciccio and Andrea Iannitti for their significant work in the prototype development.

References

[1] Beck, K. with Andres, C.: Extreme Programming Explained, 2nd edition, Addison-Wesley (2005).
[2] Bertini, E., Catarci, T., & Kimani, S.: Library and Ubicomp: Supporting Seamless Interaction with Physical and Digital Entities. In Proceedings of the Workshop on Interaction Design for CSCL in Ubiquitous Computing at Mobile HCI'04 (2004).
[3] Bolchini, D., Colazzo, S.: D1.a – Advanced Accessibility Design (UNISI input), T4.7 DELOS progress report, (2005).
[4] Borgman, Christine L.: Why are Online Catalogs Still Hard to Use? Journal of the American Society of Information Science 47(7), (1996) 493-503.
[5] Borgman, C.L.: What are digital libraries? Competing visions, Information Processing & Management, 35, (1999) 227-243.
[6] Catarci, T. (Ed.): Report on functional and non-functional digital library requirements. DELOS Deliverable D4.1.1 (2004).
[7] Fattahi, R.: The Relevance Of Cataloguing Principles To The Online Environment: An Historical and Analytical Study. PhD Thesis, University Of New South Wales (1997).
[8] Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T.: The vocabulary problem in human-system communication. Communications of the ACM, 30(11), (1987)964–971.
[9] Heaney, M.: Object-Oriented Cataloging. Information Technology and Libraries, 14(3), (1995)135-153.
[10] Hildreth, C. R.: The GUI OPAC: Approach with Caution. The Public-Access Computer Systems Review 6, Number 5, (1995) 6-18.
[11] Hong, W., Thong, J. Y. L., Wong, W. M., & Tam, K. Y.: Determinants of user acceptance of digital libraries: An empirical examination of individual differences and system characteristics. Journal of Management Information Systems, 18(3), (2001) 97-124.
[12] Ingwersen,P.: Users in Context, Lecture Notes in Computer Science, Vol 1980, (2001) 157.
[13] Kimani, S. & Bertini, E.: Exploiting Ubiquitous Computing to Support Digital Library Tasks. Proceedings of HCI International, Las Vegas US (2005).
[14] Kling, R. & Elliott, M.: Digital library design for usability. In Proceedings of Conference on Theory and Practice of Digital Libraries, College Station, TX, (1994).
[15] Lakoff, G. and Johnson M.: Metaphors we live by, Chicago : University of Chicago Press (1980).
[16] Norman D.: Why Doing User Observations First Is Wrong. ACM Interactions, July-August, 2006.
[17] Norman, D. and Draper, S.: User Centered System Design; New Perspectives on Human-Computer Interaction, Lawrence Erlbaum Associates Inc, NJ (1986).
[18] Wilson, P.: The Second Objective. In Elaine Svenonius (Ed.), The Conceptual Foundations of Descriptive Cataloging. Academic Press, San Diego, Calif., (1989) 5-16.