Analysis of Interface Vocabulary and Users' understanding in Integrated Iranian Library Software

Dr Mohsen Nowkarizi

LIS Dept., University of Birjand, Birjand, IRAN
mnowkarizi@birjand.ac.ir, nowkarzi@yahoo.com

Abstract

The purpose was to survey vocabulary readability of Simorgh (SIM) and Pars Azerakhsh (PAZ) OPACs in two university libraries of Iran. The population included undergraduate and post graduate (except PhD.) students in Ferdowsi (to examine the SIM OPAC) and Shiraz (to examine the PAZ OPAC) universities. Data gathered by a questionnaire from 380 persons of whom 275 (145 of 200 from Ferdowsi and 130 of 180 from Shiraz Universities) returned the questionnaire. Findings showed understanding rate of interface vocabulary was %35.1 for SIM OPAC and %33.9 for PAZ one. Any significant difference was not observed in readability of interface vocabulary between various degree students. Understanding rate of vocabulary used in SIM OPAC interface did not show any significant difference with the same rate in PAZ OPAC. In SIM OPAC and PAZ, a significant difference was observed between vocabulary readability of first page and other pages (for other pages). The impact of users' experience in SIM and PAZ OPACs on their understanding rate of interface vocabulary did not show any significant difference among correct answers obtained from students of different school years. The impact of study areas (humanities, basic sciences and engineering) on vocabulary readability of SIM and PAZ OPACs was examined and results showed a significant difference between means gained from respondents' scores of the three areas in Ferdowsi University (SIM OPAC). In Shiraz University, that impact was not significant.

Research limitations/Implications- This study was limited to a certain population in two Iranian universities that cannot be easily considered as an representative of the heterogeneous population of academic environments. In addition, the data collected was based on a researcher made questionnaire rather than a standardized one.

Originality/value- Most studies in this area are dealing with user satisfaction and emphasize on how much OPAC interfaces satisfy users. However, this paper addresses an aspect of them that has received little attention up to now. The paper explores users' comprehension of OPAC vocabulary, and the elements that affect it.

1. Introduction

The evolution of library catalog from traditional card catalog to Online Public Access Catalog (OPAC) requires special interaction between human and computer. In this interaction, on the one hand, there is a user who does a simple search to find a given document or book. On the other hand, s/he has a vague need and tries to fulfill it in the OPAC. In the process of interaction, that need must be clarified and evaluated which makes the process more complex. The user prospects to find a reply or a document which may contain some information to his or her need. The OPAC finds one or more documents which are stored in the database. These may be relevant or irrelevant to that need. These differences may defect the communication and interaction and have some negative effects for the user. It is necessary, therefore, to understand the interface as clear and easy as possible.

Lack of understanding the interface or the difficulty of its understanding may be considered from various aspects. One of the important aspects of the interface not already investigated enough is linguistic one: i.e. the considerable language factors and components in the interface and their roles in understanding the environment better. Kukulska-Hulme points that words in the screen may impede communication, and users who use the documentary files usually do not come to the convenient result. In addition, the users often do not use the help files, as the language used in them may not be understandable (1999, p.3).

OPACs have emerged since 1980s and have developed fast. The issue of their interfaces was soon represented as a new research area, especially in the

USA, the UK and Canada. This research area is now receiving a great significance in most countries including Iran. In the following, a selective and analytical literature review is presented.

The OPACs are investigated in a variety of aspects. Various studies are undertaken using QUIS. All these investigations also have surveyed the vocabulary used in the OPACs interface or other databases in part. Chin and others (1988), Fox and others (1993), Notess and Swan (2003), Sittig et al. (1999), Slaughter et al. (1995), Tomasek and others (1998) among others, can be cited.

Mathews (1984) investigated the problem of vocabulary selection in the screen as part of OPACs interface and concluded that selecting terms used for the screen tags were as complex as selecting the convenient format selection. He claimed that the tags used for bibliographical display had generally a meaning beyond the current meaning of this term. For example, title has broader meaning compared to the document title, as the users can observe title area, physical description, as well as imprint under the name of this tag.

Before computer terms and Internet Jargon became widespread, Naismith and Stein (1989) conducted a research on the students' comprehension level of technical language used by librarians. They found that more than 50% of the students did not know what "search expression" meant. One third of them could not even define an OPAC. 65% of the participants did not know what "citation" meant. In general, there was communication problem between the librarians and the patrons.

Using a method similar to Naismith and Stein's research, Chaudhry and Choo (2001) investigated users' comprehension of technical language used by librarians in reference interviews. The results were more positive than Naismith and Stein's and were indicating that librarians had to provide technical terms regarding readers' needs, use less jargon and make sure that the readers have the same understanding of the terms used in the reference as themselves. Kalin (1991), Spivey (2000), Klein (2001), and Hutcherson (2004), among others, have conducted some studies to find out the impact of jargon on users' comprehension.

Using QUIS, Zerehsaz (2005) investigated students' satisfaction degree of SIM and found out that their overall satisfaction was about 50%. In part of his research, he investigated users' comprehension of the vocabulary used in the SIM interface. The results showed that 48.3% of respondents did not understand some words and   expressions used in the interface.

The above mentioned studies reviewed mostly issues of designing or kind of interface design and contained guidelines for designing OPAC interface.

None of these studies have been directly investigated the role of language and linguistic elements in OPACs' interfaces. The present study, therefore, aims to meet this goal.

In general, the interface plays an important role in communicating between the users and the system. In the interface, the language role in a convenient communication between the two mentioned components is irrefutable. In our country where there are only a few OPAC softwares, it is very important for the users know how to access library resources through them. Since language is one of the main factors to communicate in these environments, its role on communication requires further investigation.. The main purpose of this research was to find out the vocabulary readability of Simorgh (SIM) and Pars Azerakhsh (PAZ) OPACs in the Iranian university libraries and the factors affecting it. In order to achieve this goal, some objectives are considered such as to survey readability of each software separately, the students' degree of study, experience and study area on the understandability of the mentioned softwares.

Since most Iranian university libraries use those two softwares, it is not clear whether they have good interfaces from the vocabulary point of view, the users easily understand them, or there are any differences between the vocabulary used in their various pages regarding simplicity and readability. This research aimed at finding some answers to these questions.

1.1. Hypotheses

1. There is a significant difference between the users' degree of understanding the interface vocabulary of Persian library softwares and the degree of desired understanding (%50 point).

2. There is a significant difference between the students with different educational level with regard to understanding the interface vocabulary of Persian library softwares.

3. There is a significant difference between the users' degree of understanding with regard to the interface vocabulary of SIM and PAZ Persian softwares.

4. There is a significant difference between the users' degree of understanding with regard to the first page interface vocabulary of Persian softwares and the following pages.

1.2. Questions

1. How much does the students' experience with the OPACs affect the interface vocabulary understanding?

2. How much does the students' study area affect the understanding of the OPACs interface vocabulary?

2. Methodology

2.1. The Participants

This study was initiated to examine the degree of users' comprehension of the interface vocabulary in SIM and PAZ library softwares. The participants were the students of undergraduate and post-graduate (only master) degrees as SIM and PAZ users in Ferdowsi and Shiraz Universities. Data gathered by a questionnaire from 380 persons of whom 275 (145 of 200 from Ferdowsi and 130 of 180 from Shiraz Universities) returned the questionnaire. The response rate was %72.4. The sampling was done by simple (proportional) stratified method. The participants were chosen from various fields of study.

2.2. Procedure

This survey had a quantified approach to the topic. The data were collected by a researcher-made questionnaire. The participants were required to fill out a questionnaire involving 3 parts: part one (demographic information), part 2 (main questions), and part 3 (including vocabulary suggestions for the OPACs interfaces).

In the part 2, the respondents were asked multi-choice questions on their comprehension of the words and expressions used in the interface search and display pages. To facilitate answering, the related interface pages were printed out to help the respondents. These questions were designed to answer the hypotheses as well as research questions.

In the third part, some general concepts used in the two OPACs were offered and they were required to select the best alternative among the three or to write down a desired one. This part was mainly designed to provide designers with convenient words or expressions which might be included in the OPACs and were offered by the users.

To answer research question 1, the users' experience is measured by the respondents' educational years of study (Freshman & Sophomore, Junior, Senior, and postgraduate).

As the freshmen number was few, they were considered with the sophomores as one group as well as the master students were considered as an independent group.

3. Findings

To test the first hypothesis, all the words and expressions used in the various search and display pages of the SIM including 30 ones were primarily extracted and offered to the users in the second part of questionnaire to express their comprehension regarding each word and expression. The lowest rate of comprehension related to the expression "reserve ad" was by 14% (n=136) and the highest one was by 65% (n=143) of respondents. If the 50% point is considered as a base for the users' comprehension, it is observed that 24 out of 30 (80%) words and expressions, were understood by less than 50% of respondents, and only 20% of them by more than 50% of the users, i.e. the users had some problems with the comprehension of 80% of the SIM interface vocabulary.

The users' overall comprehension of SIM interface vocabulary was 35.1%. Binomial test results showed a significant difference (∞=0.01and p-value= 0.000) between this amount and the average comprehension (50% point).

The same process was also used for the PAZ. The words and expressions used in the various search and display pages were 21. "Simple Search" and "Type of Document" had the lowest (3.1% [n=130]) and the highest (84.4% [n=130]) comprehension rate respectively. If the 50% point is considered as a base for the users' comprehension, it is observed that 17 out of the total 21(about 81%) words and expressions were understood by less than 50% of respondents, i.e. the users has some problems with the comprehension of 81% of the PAZ interface vocabulary.

The users' overall comprehension of PAZ interface vocabulary was 33.9%. Binomial test results showed a significant difference (∞= 0.01p = 0.000) between this amount and the average comprehension (50% of users).

To identify the significant difference between the averages gained of the undergraduate and post-graduate students' comprehension rate of the SIM interface vocabulary based on the data collected in the part 2 of questionnaire, two sample t-test was used. The results did not show any significant difference between the two degrees respondents' comprehension rate (at α= 0.01, p = 0.547). In testing the same hypothesis in PAZ, it was observed that there was not any significant difference between the means of the two degrees students' overall comprehension rate (at α=0.01, p = 0.795).

To be informed whether there is any significant difference between the comprehension rate of the vocabulary used in the two OPAC interfaces, using a two sample t-test, the data collected in the two questionnaires were compared. The results showed that there was not any significant difference between the

comprehension rate of the SIM and PAZ interface vocabularies at ($\infty = 0.01$, p = 0.366).

The respondents' scores means to the first page words and expressions (16 items) and the other pages ones of SIM were primarily calculated. The same process was conducted in PAZ in which the first page words and expressions and the other pages ones were respectively 14 and 7 items. A two sample t-test was performed on the means of the data acquired and the results showed that there was a significant difference between the two groups of words and expressions on the SIM respondents' views (behalf on the other pages ones) (at $\alpha = 0.01$, p = 0.003). The results of PAZ in this case showed a significant difference between the two variables (at $\alpha = 0.01$, p = 0.000) just as the other one.

To answer research question 1, the data gained of the part 1 and 2 of the questionnaire were used. The users experience impact on the vocabulary comprehension was evaluated by this question.

The SIM and PAZ respondents total scores based on their educational years (here considered as their experience) compared together. One-way ANOVA is used to analyze the data, and the findings showed that there was not any significant difference (at = 0.1) between the various educational year students' score means of the SIM (Table 1) and PAZ (table 2) OPACs (respectively p = 0.604; p = 0.409).

Table1. Ferdowsi students' experience impact on total vocabulary understanding ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	22.255	3	7.418	.619	.604
Within Groups	1689.786	141	11.984		
Total	1712.041	144			

Table2. Shiraz students' experience impact on total vocabulary understanding ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	16.153	3	5.384	.970	.409
Within Groups	699.116	126	5.549		
Total	715.269	129			

To answer research question 2, the data acquired of the vocabulary comprehension rate of three groups (the students of Humanities, Pure Sciences, and Engineering study areas) were compared to identify whether the students' study area nature influence their vocabulary comprehension rate or not.

The data of each study area using One-Way ANOVA were first analyzed for SIM (Table 3) and showed that there were significant differences (at 0.01) among the means acquired from the respondents vocabulary meaning scores in the three study areas, i.e. Humanities, Pure Sciences, and Engineering. To identify which two study areas were significantly different, Duncan Test for multi-comparisons were used (Table 4). The results suggested that there was a significant difference at 0.05 between Pure Science(area 2) and Engineering (area 3) students score means, but any difference between Humanities and Pure Science ones (p = 0.149), and between Humanities and Engineering ones (p = 0.062) were not observed.

Table3. Ferdowsi Students' Study Areas Impact on Their Vocabulary Understanding ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12.302	2	61.151	5.462	.005
Within Groups	1589.739	142	11.195		
Total	1712.041	144			

Table4. Post Hoc Multiple Comparisons on Study Areas of Respondents Duncan

AREA	N	Subset for alpha=.05	
		1	2
2	45	9.42	
1	51	10.41	10.41
3	49		11.69
Sig		.149	.062

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 48.202.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

The same test was performed for PAZ respondents (Table 5) and the results did not show any significant difference among the means gained of the respondents' vocabulary meaning scores in the three investigated study areas as mentioned.

Table5. Shiraz students' study areas impact on their vocabulary understanding ANOVA

	Sum of Squares	dif	Mean Square	F	Sig
Between Groups	16.153	3	5.384	.970	.409
Within Groups	669.116	126	5.549		
Total	715.269	129			

At the end of the two questionnaires a part 3 was included which aimed to provide a list of appropriate

words and expressions in the respondents' view for some concepts used in the two OPACs interfaces. They have chosen some of the words and expressions which are used in some of library softwares which had represented by the researcher, but in some cases, they suggested some new ones which is applicable in the future versions of the OPACs, if the designers will use them. These are in Persian and are not represented here.

4. Discussion

The users' overall comprehension of SIM and PAZ interface vocabulary showed a significant difference between users' comprehension of the two OPACs and the average comprehension (50% of users). These results support Naismith & Stein's. Investigating the students' comprehension rate of library jargon, they found that more than half of them did not know what, for example, did mean "Search Expression". Some other researchers, such as Yee (1991), Shires & Ulszak (1992), Shneiderman (1998), suggest that in interface designing, one must use a simple, expressive, and comprehensive vocabulary. Graham's (1995) study results differ considerably with these ones. In part of her study, she reported that 73.4% of the respondents found the interface vocabulary clear and simple. They thought the vocabulary was comprehensive, and rejected jargon.

It was observed that there was not any significant difference between the means of the two degrees students' overall comprehension rate in the two OPACs, Park (1997) in his study claimed that the students' degree of study affected the kind of selected OPAC. But he did not conduct this study on the comprehension rate and its relation on educational degrees.

The results did not show any significant difference between the comprehension rate of the SIM and PAZ interface, while Green and Head (1998), comparing two OPACs, emphasized on their interface vocabularies as a distinguishing characteristic.

It is often expected that the designer uses simpler, more fluent and comprehensible vocabulary in the first page of the software interface than in the other pages, that s/he considers the first page for the simple and fast applications and the other pages for the more specified ones. But in these two OPACs, this is not true and is vice versa.

The findings showed that there was not any significant difference between the students' comprehension rate and their experience with the two investigated OPACs.

These results support Novotny (2004) who pointed out that each of two students groups, inexperienced and experienced, did not specialize using OPAC and had not a complete comprehension. But these two studies results did not support Hildreth study (2001) where he reported that juniors and seniors announced more ease of use in comparison with freshmen and sophomores in using OPAC.

The results of SIM OPAC showed significant difference between the students' comprehension rate of various areas investigated, but the PAZ results was vice versa.

In this case, Park (1997) showed that various fields of study did not affect on the students' selected OPAC type. One element that is important for users to select upon a certain OPAC, may be the vocabulary of it.

5. Conclusion

Both of the two OPACs are comprehensible less than the 50% point. To increase the comprehension rate, the designers must interact more with their users or try to know them more and more. They ought to benefit of the research outcomes such as ones of this study. They can also support some research projects to promote their softwares. One of the factors that increases the vigorous interface vocabulary of OPACs is that these, compared with the traditional and manual systems (such as Card Catalog), have some new words and expressions, which are added to the interface by two groups: the words and expressions that librarians have proposed to express some concepts and bibliographic element functions in the records, and in the overall search and retrieval environment; and the words and technical terms which computer specialists and software designers have offered to express the software functions in the search and retrieval environment. It is clear that these new words and terms were not available in the card catalog environment, and catalog cards were self-evident. But OPAC environment is a new one, thus it may be vague and unfamiliar for some users. Consequently, some Jorgen of library and computer science (such as ISSN, ISBN, Form, all Database Types, Reset, and Brief Form) are observed in the OPACs. Even some words with a broad scopes for librarians, such as Author, are appeared there which may mean for the users more than usual meaning, in that case as a writer, compared with this concept in Library and Information Science which means beyond that meaning as a writer, such as editor, translator, compiler and so on. This issue affects low comprehension of vocabulary. Users' low experience may also be a factor of low comprehension.

Regarding what are said, it is recommended that, conducting surveys among their users, software designers increase as far as possible the readability and comprehensibility of their software vocabulary. Librarians are recommended to consider the words and expressions comprehensibility by users when proposing library terms needed to represent library concepts to designers and propose the words and expressions based on users' public understanding.

The two frequent-used OPACs in the Iranian universities seem have some problem in terms of designing the interface vocabulary. The designers need, therefore, primarily interact more and more with librarians to know their points of view and to use convenient vocabulary since librarians interact closely with the users and know their views and language to somewhat.

One may expect that educational level affects the comprehension rate of users because of the masters' more knowledge depth, especially in abstract things such as vocabulary comprehension, but what was showed in the study does not confirm this statement. This is a problem in the two OPACs that must be considered and resolve.

As experience may affect comprehension rate of vocabulary, encouraging users by librarians to use OPAC more and more in various ways (such as training, help them to find their real information need, express orally the high characteristics and possibilities of OPACs or through publishing brochures and so on) is important.

It seems the interface vocabulary comprehension rate among various study areas is different. In some areas such as Engineering, it is expected this rate increases because of the students' greater familiarity with computer terms. In some other, such as Humanities area in which the students deal more with the abstract concepts and especially in some of its fields (i.e. Language, Psychology, Education, and Library and Information Science) linguistic concepts are investigated, it is expected that the users' comprehension rate of the vocabulary be more. What the designers must remember always is that the interface is designed so that all users can understand its vocabulary not the users of some special fields. As the results of the two investigated OPACs were different in this case, more research is necessary.

Since most of their users may be amateur with little experience, these OPACs must have a deductive structure where the general, simple applications primarily appear in the first page, and complex and more specialized applications come in the other ones. But this study showed vice versa.

Acknowledgment

References

[1] Chaudhry, A. S., Choo, M. (2001).Understanding of Library Jargon in the Information-seeking Process, *Journal of Information Science*, 27, 343 349

[2] Chin, J. P., Diehl, V. A., and Norman, K. (1988), "Development of an instrument measuring user satisfaction of the human-computer interface", *In CHI '88. Conference Proceedings on Human factors in computing systems* (pp. 213-218), New York: ACM Press. Retrieved November 21, 2006, from http://www.lap.umd.edu/QUIS/index.html

[3] Fox, E. A., Hix, D., Nowell, L. T., Brueni, D. J., Rao, D., Wake, W. C., Heath, L. S. (1993). Users, user interfaces, and objects: Envision a digital library. *Journal of the American Society for Information Science*, 44(8), 480-491.

[4] Graham, M. E. (1995). Evaluation of the British Library network OPAC in selected UK academic libraries. *Newcastle: Department of Information and Library Management University of Nothumria*. Unpublished

[5] Green, E., Head, A. J. (1998). Web-based catalogs: is their design language anything to talk about? Online, 22(4), 98-104. Retrieved December 10, 2006, from http://web1.epnet.com

[6] Hildreth, C. R. (2001). Accounting for users' inflated assessments of online catalogue search performance and usefulness: an experimental study. Information Research, 6(2), Retrieved December 5, 2006, from http://informationR.net/ir/6-2/paper101.html

[7] Hutcherson, N. B. (2004). Library jargon: student recognition of terms and concepts commonly used by librarians in the classroom. College & Research Libraries, 65(4), 349-354.

[8] Kalin, S. W. (1991). Support services for remote users of online public access catalogs. RQ, 31, 197-213. Retrieved October 19, 2006, from http://www.ala.org/ala/rusa/rusapubs/rusq/specialfeatures/rsp awardwinning/1994/1994.htm

[9] Klein, L. R. (2001). The web is not your library. *Library Journal Net Connect*, 4, 36 37.

[10] Kukulska-Hulme, A. (1999). Language and Communication: essential concepts for user interface and documentation design. *New York: Oxford University Press*.

[11] Matthews, J. R. (1984). Screen displays for public online systems. In P. E. Peters (Ed.), *Command Language and Screen Displays for Public Online Systems*, pp.25-36, Washington D.C: Council on Library Resources

[12] Naismith, R., Stein, J. (1980). Library Jargon: Student Comprehension of Technical Language Used by Librarians. College & Research Libraries, 50, 543-552.

[13] Notess, M. & Swan, M.B. (2003). Predicting user satisfaction from subject satisfaction, *In Proceedings of the conference on Human factors in computing systems*, Ft. Lauderdale, Florida, USA, (pp. 738-739), New York, NY: ACM Press

[14] Novotny, E. (2004). I don't think I click: a protocol analysis study of use of a library online catalog in the Internet age. College & Research Libraries, 65(6), 525-537.

[15] Park, I. J. (1997). A comparative study of Major OPACs in selected Academic libraries for developing countries: users study and subjective user evaluation. *International Information & Library Review*, 29(1), 67-83.

[16] Shires, N. L., Ulszak, L. P. (1992). What our screen should look like: an introduction to effective OPAC screen. RQ, 31(3), 357-369.

[17] Shneiderman, B. (1998). *Designing user interface: strategies for effective human computer interaction* (3rd ed.). Reading, MA: Addison Wesley Longman.

[18] Sittig, D. F., Kuperman, G. J., Kiskio, J. (1999). Evaluating physician satisfaction regarding user interactions with an electronic medical record system, *Proceedings of AMIA of AMIA Annual Symposium*, (pp.400-404), Chestnut Hill, MA: Partners Healthcare System. Retrieved October 25, 2006, from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retriev e&db=PubMed&list_ uids=10566389&dopt=Abstract

[19] Slaughter, L., Norman, K. L., & Shneiderman, B. (1995). Assessing users' subjective satisfaction with the Information System for Youth Services (ISYS), *Proceedings of the Third Annual Mid-Atlantic Human Factors Conference*, (pp.164-170), Retrieved August 2, 2006, from ftp://ftp.cs.umd.edu/pub/hcil/ Reports-Abstracts-Bibliography/95-09html/95-09.php

[20] Spivey, M. A. (2000). The vocabulary of library Home Pages: an influence on diverse and remote-users. Information Technology & Libraries, 19, 151-156

[21] Tomasek BA, Sagoo N, & Xiao Y. Telemedicine: Pilot evaluation of user satisfaction survey tool. Proceedings of 1998 American Medical Informatics *Association (AMIA) Annual Fall Symposium*, 1085, 1998, Retrieved January 7, 2007, from http://www.amia.org/pubs/symposia/D005055.PDF

[22] Yee, M. M. (1991). System design and cataloging meet the user: user interfaces to online public access catalogs. *Journal of the American Society for Information Science*, 42(2): 78-98.

[23] Zerhsaz, M. (2005). Elements, Characteristics Investigation and Analysis Presented on Simorgh software Interface and Identifying Satisfaction degree of Interaction of Education and Psychology Students of Ferdowsi University. Thesis of Library and Information Science, Ferdowsi University

[24] Zipf, G.K. (1949). *Human Behavior and the Principle of Least Effort*, Cambridge, MA.: Addison-Wesley.