

Using the Web for Critical Learning of Al-Qur'an's Language: an Exploratory Case Study

Matin Saad Abdullah (Corresponding author)

Department of Computer Science, KICT, International Islamic University Malaysia (IIUM),
Gombak, 53100, Kuala Lumpur, Malaysia

Tel: +603-61964000 Ext. 5653 E-mail: matin.abdullah@gmail.com

Al-Sakib Khan Pathan

Department of Computer Science, KICT, International Islamic University Malaysia (IIUM),
Gombak, 53100, Kuala Lumpur, Malaysia

Tel: +603-61964000 Ext. 5653 E-mail: sakib@iium.edu.my

Muhammad Abdur Rahman

Department of Computer Science and Engineering, BRAC University, Dhaka, Bangladesh

Tel: +880-1712147603 E-mail: anon83@gmail.com

Abstract

About 80 percent of the world's Muslim populations are non-native speakers of Arabic language. Since it is obligatory for all Muslims to recite the Qur'an in Arabic during regular prayers, an extraordinary social phenomenon has taken place in some parts of the Muslim world: in schools, children are only taught the complex phonetic rules of the Arabic language in the context of the Qur'an. This has given rise to a demographic segment of adult learners who are interested in a Language for Specific Purposes (LSP) curriculum that would help them learn a closed set of syntactic rules and vocabularies in the context of the Qur'an, so that they can recall an idiomatic translation in their native language while they recite or listen to the Qur'an. Little research work exists on the task modeling and user modeling for this purpose. This research work explores the possibilities of using user stereotypes in the creation of task models to be used in development of a comprehensive Computer Assisted Language Learning (CALL) module. In this paper, firstly, the design and initial prototype of a ubiquitous web based language learning software is presented and some results of the user modeling are shared. Secondly, changes made to the initial implementation based on the user modeling are presented. And finally, merits and drawbacks of the new implementation are shared.

Keywords: Qur'an , E-learning, Computer Assisted Language learning, ICT, Internet

Introduction

About 80 percent of the world's Muslim populations are non-native speakers of Arabic ("Arab people", 2012). Since it is obligatory for Muslims to read and listen to the Qur'an in Arabic during Prayers, an extraordinary social phenomenon has taken place in some parts of the Muslim world: Muslims, men and women, are taught the complex phonological rules of the Arabic language in the context of the Qur'an. They learn these rules and recite the "sounds" of the Qur'an understanding very little of what they are reciting. Similarly when they listen to the Imam reciting the Qur'an in prayer they barely understand what they are listening to (Moore, 2006). This has given rise to a demographic segment who are consumers of Arabic language classes, language learning books and software to overcome this language barrier. Some resources have been developed to address this particular challenge; these include books, courses, and lexical resources focusing primarily on teaching the Arabic language in the context of the Qur'an. Since the main learning goal of the target demography is recalling a closed set of syntactic rules and vocabularies in the context of the Qur'an, so that they can reconstruct a meaning in their native-language (L1) while reading or listening to the verses of the Qur'an, the content falls under the category of Arabic for Specific Purposes (ASP).

Despite the availability of resources for this purpose, to the best of our knowledge, no empirical research has yet identified the learning environments and the unique learning requirements of the target demography and explored the possibilities of emerging E-learning technologies to address

this challenge. This research work explores the possibilities of using user stereotypes in the creation of task models to be used in the development of a comprehensive Computer Assisted Language Learning (CALL) module. In this paper, the design and initial prototype of a ubiquitous web-based language learning software founded on the LSP principles is presented and some initial results of need analysis of the user data are shared. The preliminary results of this study indicate the need for such a CALL module and highlight the areas where improvement and more research are required.

Theoretical Background

Second Language Learning

In the field of second language learning four basic language skills are distinguished. These are ordered along two dimensions **Error! Reference source not found.** (Bransford & Johnson, 1973):

- *Modality*, which is the difference between the auditory language mode versus the visual mode, and
- *Processing activity*, which is the process of either encoding or decoding.

Together these two binary dimensions define the so-called four skills, as indicated (Table I):

Table 1: Language processing activity and modality

		Processing activity	
		<i>Encoding</i>	<i>decoding</i>
Modality	<i>Auditory</i>	Speaking	listening
	<i>Visual</i>	Writing	reading

The learning outcome of the target demography of this project is recalling a closed set of syntactic rules and vocabularies in the context of the Qur'an to reconstruct a meaning in the learners' native-language (L1), while reading or listening to the verses of the Qur'an. So the tasks that the learners have to perform are decoding of the visual and auditory modality - in other words, reading and listening comprehension. The learner comprehends a new lexical item by assigning meaning to it, thus establishing a new form-meaning connection. Comprehension problems are often caused by a form-meaning mismatch, sometimes caused by faulty word-boundary identification or by insufficient lexical knowledge **Error! Reference source not found.**(Byrnes, 1984). Even if a learner identifies the form of a word correctly, if s/he does not have enough knowledge of its meaning, the recognition process will also fail. The listening comprehension is a bit more challenging than the reading comprehension because a characteristic of continuous speech is that speech contains no clear auditory equivalent of the inter-word white spaces that we find in written text (Buck, 1995). The lack of explicit word boundary markers and the distortions of word sounds due to assimilation and reduction are the reasons why words that are known by the language user when presented visually, are often not recognized when they are part of continuous speech.

Because of the narrow domain of the learning outcome –i.e. listening and reading comprehension of only the text of the Qur'an - this project can be categorized as teaching Language for Special Purpose (LSP) (Dita, 2007). In the specific purposes language curriculum, the language is not seen as the main goal in the process of learning, but rather a vehicle for its acquirement. Students study a special purpose language curriculum not because they are interested in the language as such but because they have to perform a task in that language. The focus is placed on specific skills as opposed to general language development. Some of the assumptions made about the learners are:

- Target learners are primarily adults.
- Basic knowledge of the language system is expected.

Computer Assisted Language Learning

Computer-assisted language learning (CALL) is succinctly defined in a seminal work by Levy as "the search for and study of applications of the computer in language teaching and learning" **Error!**

Reference source not found.(Levy, 1997). CALL grew out of the field of Computer Assisted Instruction (CAI) and draws on other related fields such as Educational Psychology, Artificial Intelligence (AI), computational linguistics, instructional design, Human Computer Interaction (HCI) and SLA (Second Language Acquisition) **Error! Reference source not found.** ("Computer-assisted language learning", 2012). The current philosophy of CALL puts a strong emphasis on student-centered materials that allow learners to work on their own. Such materials may be structured or unstructured, but they normally embody two important features: interactive learning and individualized learning. There is a large body of literature concerning how the use of CALL affects the development of language learners' four skills (listening, speaking, reading and writing). Most studies report significant gains in reading and listening and most CALL programs are geared toward these receptive skills because of the current state of computer technology (Levy, 1997).

Related Resources

There is a wide range of material available for addressing this target demography. Since the L1 of this project is English we are going to focus on the English material available for this purpose. The material can be mainly divided into three categories:

- Lexical resources
- Instructional resources
- Interactive software to learn the language of the Qur'an.

Lexical resources give the user the ability to do lexical lookups needed to study the Arabic of the Qur'an. The main category of lexical resources comprises interlinear translations of the Qur'an (Shaikh & Khatri, 2007; Baig, 2002; Mujahid, 1999; Ali, 2003). These books include the interlinear translation of every verse along with the idiomatic translation. Some of them contain deeper lexical information like roots and morphological rules as footnotes (Ali, 2003). The second category of lexical resources available consists of dictionaries of the Qur'anic words (Nadwi, 1996; Penrice, 2005; Azmi, 2003; Parekh, 2003). These are different from the first category because they provide the list of the Qur'anic words in alphabetical order. Some of them include all the morphological forms of a word used in the Qur'an instead of only the canonical form of the words. Both of these resources are available both as books and soft copies (Baig, 2002) over the internet. There are online (Dukes, Atwell & Habash, 2012; "Project Root List", 2012) and offline (Dror, Shaharabani, Talmon & Wintner, 2004; Talmon & Wintner, 2003) applications that serve the need for these lexical lookups. Foremost among these is the Qur'an Corpus (Dukes et al., 2012) project which is an annotated linguistic resource which shows the Arabic grammar, syntax and morphology for each word in the Qur'an. The corpus provides three levels of analysis: morphological annotation, a syntactic Treebank and a semantic ontology. A similar project was developed by the Project Root List ("Project Root List", 2012) whose goals were: to list all the roots used in the Qur'an, including possible meanings (from several classical Arabic dictionaries/lexicons); give all examples of usage (allowing cross-reference); and provide detailed grammar information for every word used. There are also some offline resources developed for similar purposes. Instructional resources focus on teaching the grammar of classical Arabic focused based on the Qur'an. These include books (Nadwi, 1997; Jones, 2005; Thackston, 2000) and videos (Raheem, 2012; Surty, 2012), some of which are available through the web. Lastly, there are some commercial efforts to create interactive CALL software for this purpose (Magnicode, 2012).

Although there are many fragmented efforts like these, there is no comprehensive CALL software that addresses the unique needs of the target demography. The research goal is to examine both the learning environment and the unique learning requirements of the target demography by analyzing existing best practices in the field of CALL and eLearning. That information will then be used to develop and modify an interactive learning management system that would help the user of the system to achieve their learning goals.

Design and Implementation

The goal of this initial prototype was to conduct an initial needs analysis of software that focuses on teaching the Arabic of the Qur'an and gather insight about the user base and the system requirements. The initial implementation of the software was designed for accompanying the UnderstandQuran.com website maintained by Abdulazeez Abdulraheem (Raheem, 2012). The goal of UnderstandQuran.com is to teach the language of the Qur'an by mainly focusing on vocabulary

and morphology of words, and some syntax, in the context of the Qur'an. Abdulazeez Abdurraheem has produced lexical resources like "Dictionary of the Qur'an" and a word list which accounts for 80% of the words in the Qur'an. Our initial prototype utilizes a subset of this wordlist to develop interactive software for learning the vocabulary of the Qur'an.

There were three phases to the initial implementation. In the first phase an accompanying web site called 80percentwords.com was set up. The goal was to teach the different forms of the verbs and their meanings by using an online vocabulary game. While answering a question the user could access an example of that word in the context of the Qur'an where s/he could learn the other meaning of the other individual words in the context promoting incidental vocabulary learning (Nation, 2001). This was developed using Adobe Flex so there was no installation footprint. Adobe Flex ("Apache Flex", 2012) was chosen at the time of development because of Adobe Flash player's 99 percent browser coverage in desktop environments. Social media integration with Facebook, Twitter and StumbleUpon was implemented to broaden the user-base. An email subscription system and a Facebook page was set up to collect user information and to notify the user of new content when it became available. An Youtube video describing the goal and interaction of the software was published as a how-to guide for the website.

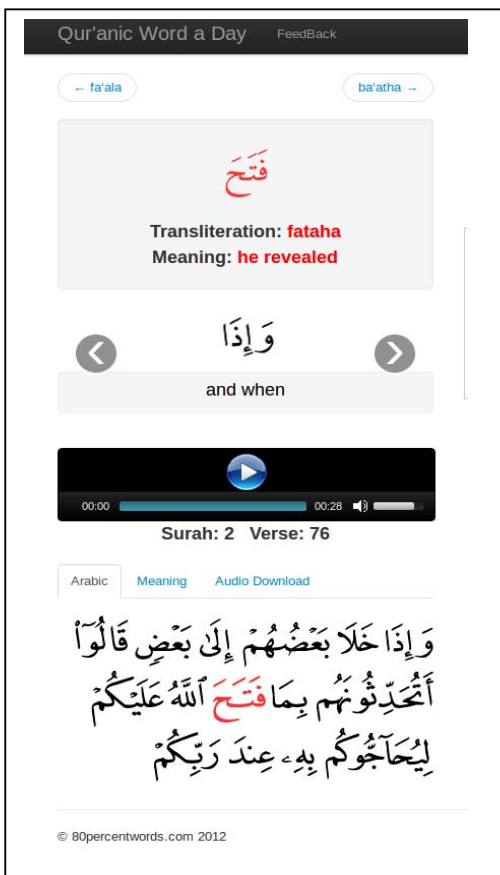


Figure 1: HTML5 implementation



Figure 2: Adobe Flash implementation

In the second phase of the implementation a Qur'anic word-a-day component was introduced to increase user interaction. Each day, a word with its meaning and an example in the context of the Qur'an was sent to the user. The example had both the interlinear and phrase translation so that incidental vocabulary learning could occur. Implementation of this phase was automated by pre-populating a Wordpress blog with the relevant data and then email notification was sent to the users using Feedburner and Facebook page. The Facebook page then posted the information to the Twitter feed.

In the third phase, some of the initial user information was taken as a guideline and the word-a-day component was re-implemented to support mobile devices. Since the initial website was designed with a desktop browser as a requirement and the main component was built based on Adobe flash player the new implementation had to be redesigned from scratch. Desktop user interaction

elements like mouse-over and mouse-click had to be replaced with events like user-tap in a touch-screen environment. The variation in different mobile screen sizes had to be taken into account. This new site was designed based on the HTML5 infrastructure. The same push-infrastructure was used to deliver the content to the user's device as before. Figure 1 and Figure 2 show the different implementation of the word-a-day component.

Data Analysis and Discussion

Methodology

The participants were invited through message boards and email lists to use the software and sign up for the face book fan page. For the initial implementation the user perception was judged by feedback emails, blog comments, Facebook comments, and forum replies. The user and environment data were collected from Google Analytics and Facebook page statistics.

Table 2: User visit by top-ten countries

Country/Territory	Visits
United States	37,748
United Kingdom	22,343
India	8,196
Canada	7,334
Pakistan	6,695
United Arab Emirates	5,801
Saudi Arabia	5,236
Malaysia	4,893
South Africa	2,946
Bangladesh	2,581

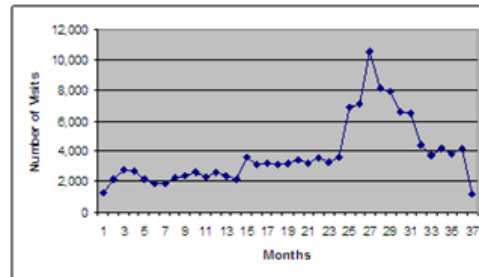


Figure 1 Number of visits over 36 months

Table 3: User subscription by top-ten countries

Country/Territory	Subscribed
Malaysia	1418
Pakistan	1245
Bangladesh	522
India	488
United States	466
United Kingdom	390
Saudi Arabia	367
United Arab Emirates	254
Indonesia	226
Thailand	106

Table 4: Age and gender of users

Age Range	Male	Female
13-17	2.7%	2.6%
18-24	25%	25%
25-34	15%	13%
35-44	4.9%	4.2%
45-54	1.7%	1.7%
55+	1.4%	1.3%

Table 5: User Operating System

OS	Visits
Windows	85.42%
Macintosh	7.51%
Android	1.47%
Linux	1.40%
iPhone	1.35%
(not set)	1.06%
iPad	0.88%
iPod	0.23%
SymbianOS	0.23%
Others	0.45%

Table 6: User Operating System

Country	Visits
United States	33.50%
United Kingdom	22.85%
Malaysia	5.01%
Singapore	4.74%
India	4.12%
Canada	3.61%
United Arab Emirates	3.15%
Australia	2.79%
Saudi Arabia	2.25%
Others	17.98%

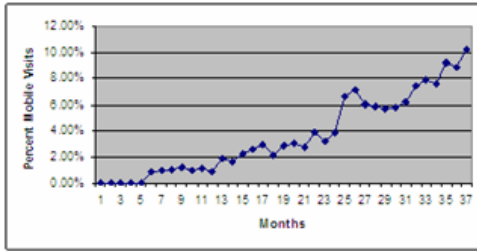


Figure 4: Growth of mobile visits

Table 7: Top-ten mobile device information

Mobile Device Info	Visits
Apple iPhone	20.46%
Apple iPad	16.86%
(not set)	12.33%
Apple iPod Touch	2.99%
Samsung Galaxy Tab	2.60%
SonyEricsson LT15i Xperia Arc	1.79%
Samsung GT-I9100 Galaxy S II	1.58%
HTC Desire	1.44%
HTC G2 HTC Sappire	1.44%
Verizon Droid2	1.08%

Traffic Information

In this paper we want to share 3 years of data collected from June 9th, 2009 to June 9th, 2012. There were 82,655 unique visitors who visited the site and total number of pageviews during this period was 174,586. Figure 3 shows the number of page views over the 36 month period. The spike in the plot of more than 100% starting in May 2011 was due to the second phase of the project where the word-a-day and the social component were integrated in the software. The month in which over 10,000 users visited the site was during Ramadan, 2011. When the word-a-day component was stopped, the monthly users dropped to 3-4K. As of June 9th, 2012, 6,660 users are part of the Facebook fan page and 1,026 users are subscribed via Feedburner, a total of 7,686 users receiving the daily feed.

User Demography

We gathered some basic demographic data from both the users who visited the site and the users who subscribed to the Facebook page. Here we can look at users who visited the main site vs. users who subscribed. Table 2 shows the top ten countries where the visitors came from and Table 3 shows the top ten countries of user subscription. This refers back to our assumptions about the target demography: the site visitors include a lot of incidental visitors but among the people who subscribed to the site, the majority are from Muslim countries where the native language is not Arabic and the tradition of teaching children how to recite the Qur'an is very strong. As Table 4 shows our primary subscribers, almost 80%, are in the age range of 18-34. However, among that group, the age range of 18-24 constitutes about 50% of all the subscribers. This indicates that young people are early adopters of this site.

User environment

The user environments are predominantly desktop browsers, as we can see from Table 5, because the Adobe Flash player is not supported on mobile devices. The major shift we are seeing in the user behavior is towards mobile devices. Although the site did not support mobile browsing the percentage of visitors trying the site with mobile devices was gradually increasing. Figure 4 shows the increase in the percentage of visits using mobile devices over time. Table 6 shows the top ten countries where the mobile visits originate and Table 7 shows the top-ten devices. In all the cases "not set" refers to the data that could not be detected on the server side.

User perception

A total of 162 comments were collected from the users, including emails, blog comments, Facebook comments, and forum replies. Table 8 shows the different categories of comments. Since this was a pilot project most of the feature requests were for more content to learn. The second set of feature requests were for making the site accessible from mobile devices which are in sync with the trend we saw in the traffic statistics. There are also user requests to store user data, social network connectivity and offline access.

HTML5 vs. Flash implementation

Since the HTML5 implementation was done on the first day of Ramadan 2012 we have limited data points to analyze the impact of this change. We see similar increase in data points as Ramadan 2011.

Even in nine days (Figure 5) of data points we can see that the mobile device users now constitute almost 30 percent of the total traffic. This is in accordance with the initial user modeling.

Table 8: User comments

Category	# of comments
Positive feedback	140
Usability issues	27
Feature requests	47
Bug reports	4
Negative feedback	2

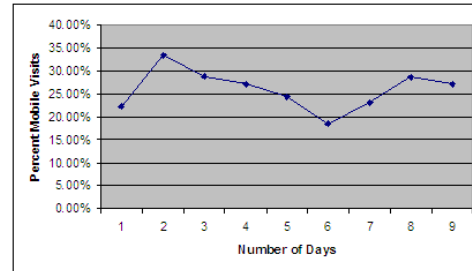


Figure 5: HTML5 Mobile Visits

Conclusion and Future Work

The main focus of this study was to examine the potential for an e-Learning system for learning the Arabic of the Qur'an keeping our target demography in mind. The user perception and the traffic analysis of the initial prototype of a ubiquitous web based language learning software indicate the potential of this approach and highlight the areas where improvement is required. Some suggestions and directions that came out of this study were:

- An overwhelming need and potential for software of this nature.
- The growing need for mobile accessibility for the purpose of e-learning.
- The distributed nature of the user location and the diversity of their system requirements.
- The need for user modeling to customize the user experience and give users an opportunity to track their progress.
- A deeper integration with the social media so that the user has control over sharing the user data with their friends on the social network to create an environment of social learning

To get more representative results, an in-depth user study based on CALL and LSP principles needs to be conducted by utilizing qualitative and quantitative methods. While this study was exploratory in nature and had a limited scope, it is hoped that the findings will provide the basis for designing a more rigorous study to understand the user and task modeling for this purpose.

Acknowledgment

This work was partially supported by NDC Lab, KICT, IIUM.

References

- Arab people. (2012, July 26). In Wikipedia, The Free Encyclopedia. Retrieved 03:32, July 29, 2012, from http://en.wikipedia.org/w/index.php?title=Arab_people&oldid=504333230
- Moore, L. (2006). Learning by Heart in Qur'anic and Public Schools in Northern Cameroon. *Social Analysis, vol 50(3), pp 109-126.*
- Bransford, J. & Johnson, M. (1973). *Visual information processing*, W. Chase (Ed.), New York: Academic Press.
- Byrnes, H. (1984). The Role of Listening Comprehension: A Theoretical Base. *Foreign Language Annals, vol 17, 1984, pp. 317-329.*
- Buck, G. (1995). *A guide for the teaching of second language listening*. San Diego: Dominic Press.
- Dita, G. (2007). *Languages for Specific Purposes: Searching for Common Solutions*. UK: Cambridge Scholars Publishing.
- Levy, M. (1997). *CALL: context and conceptualisation*. Oxford: Oxford
- Computer-assisted language learning. (2012, July 23). In Wikipedia, The Free Encyclopedia. Retrieved 03:43, July 29, 2012, from <http://en.wikipedia.org/w/index.php?title=Computer->

assisted_language_learning&oldid=503828015

Shaikh, S. & Khatri, K. (2007). *The Glorious Quran: Word-for-Word Translation to facilitate learning of Quranic Arabic*. New Delhi: Al Huda Publication.

Baig, Z. (2002). *The Holy Qur'an: WORD FOR WORD ENGLISH TRANSLATION*. Karachi: Baitul Quran.

Mujahid, A. (1999). *Study the NOBLE QUR'AN Word-for-Word*. Ryad: Darrussalam.

Ali, M. (2003). *A Word For Word Meaning of the Qur'an*. Suffolk: Jam'iat 'Ihyaa 'Minhaaj Al-Sunnah.

Nadwi, A. (1996). *Vocabulary of the Holy Qur'an*. Skokie: Iqra' International Educational Foundation.

Penrice, J. (2005). *A Dictionary and Glossary of Koran*. New Delhi: Adam Publishers and Distributors.

Azmi, A. (2003). *A Glossary of the Qur'an*. New Delhi: Goodword Books, 2003.

Parekh, A. (2003). *Complete Easy Dictionary of the Qur'an*. Kuala Lumpur: A. S. Noordeen.

Dukes, K., Atwell E. & Habash, N. (2012) Supervised Collaboration for Syntactic Annotation of Quranic Arabic. *Language Resources and Evaluation Journal (LREJ)*. *Special Issue on Collaboratively Constructed Language Resources*.

Project Root List (2012) Project Root List. Retrieved from: <http://www.studyquran.co.uk/PRLonline.htm>.

Dror, J., Shaharabani, D., Talmon R., & Wintner S. (2004) Morphological Analysis of the Qur'an. *Literary and Linguistic Computing*, 19(4):431-452.

Talmon, R. & Wintner, S. (2003) Morphological Tagging of the Qur'an. *Proceedings of the Workshop on Finite-State Methods in Natural Language Processing, an EACL'03 Workshop, Budapest, Hungary*.

Nadwi, A. (1997). *Learn the Language of the Holy Quran*. Karachi: Darul-Ishaat.

Jones, A. (2005). *Arabic Through the Qur'an*. Cambridge: The Islamic Text Society.

Thackston, W. (2000). *An Introduction to Koranic and Classical Arabic*. Bethesda: IBEX Publishers.

Raheem, A. (2012). Understand Quran - The Easy Way - Free Online Courses. Retrieved from: www.understantquran.com.

Surty, M. (2012). Towards Understanding Qur'anic Arabic. Retrieved from: <http://www.qaf.org.uk>.

Magnicode, Inc (2012) Quranic Words for iPad -- Understand the Arabic Qur'an for iPad on the iTunes App Store. Retrieved from: <http://itunes.apple.com/us/app/quranic-words-for-ipad-understand/id384913391?mt=8>.

Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Apache Flex. (2012, July 26). In Wikipedia, The Free Encyclopedia. Retrieved 03:51, July 29, 2012, from http://en.wikipedia.org/w/index.php?title=Apache_Flex&oldid=501168779