QURANICA, International Journal of Quranic Research, Vol. 10, Issue. 1, June 2018, Pp. 41-52

DEVELOPMENT OF ARTIFICIAL INTELLIGENCE BASE SYSTEM, FOR TEACHING CLASSICAL ARABIC & QUR'ANIC VOCABULARIES FOR URDU LANGUAGE USERS(*)

Haq Nawaz[1]

ABSTRACT

"Classical Arabic" or "Fusha Arabic" is very important because the Holy Qur'an was revealed in Classical Arabic. The Holy Qur'an describes itself as an "Arabic Qur'an" (Surah Yusuf:2). Secondly, it is the language of the Last and Final Prophet and Messenger of Allah Muhammad (s.a.w) and core Islamic literature is written in Classical Arabic. Arabic is one of the Semitic languages in which different variations of verbs, nouns can be obtained by prefixing or suffixing different characters in root word. Information technology (IT) has a vital role in modern era, learning process has undergone great changes and teaching institutes are using Virtual Learning Systems for teaching purpose. The existing systems for the Arabic language do not meet these modern changes. There is need to improve the state of the art by using advanced techniques to fill this gap. This paper focuses on development of teaching system of Arabic conjugations for non-Arabs; initially it is designed for the Urdu users, with extensions to other languages also planned. This system is developed using techniques from Natural Language Processing (NLP) and Artificial Intelligence (AI). This system till now covers the teaching of "علم الصرف" components. Towards this end, different algorithms have been designed that generate conjugations from the given root words and can additionally generate Urdu translations. The system also generates the runtime examinations as per learning syllabus of the student, and also recommends the weak areas of student where improvement is required.

Keywords: *Qur'an , Artificial Intelligence, Natural Language Processing.*

(*) This article was submitted on: 15/02/2018 and accepted for publication on: 06/05/2018.

[1] University of Central Punjab (MSCS Schollar), Pakistan, email: haqnawaz99@yahoo.com

1. INTRODUCTION

The Qur'an is the holy book of Muslims that was revealed to Muhammad(s.a.w), Qur'an is the complete guideline for all Muslims for all eras. Language of the Qur'an is Arabic, and this form of Arabic used in Qur'an is known as "Classical Arabic" or "Fusha Arabic". This form of Arabic language is also used in literary texts written by early Arabic scholars.Allah mentioned in many verses the importance of the Arabic language.

﴿قُرْآنًا عَرَبِيًّا غَيْرَ ذِي عِوَجٍ لَّعَلَّهُمْ يَتَّقُونَ﴾

"(It is) a Qur'an in Arabic without any crookedness (therein): in order that they may guard against Evil" (Zumar:28).[1]

Qur'an also says:

﴿كِتَابٌ فُصِّلَتْ آيَاتُهُ قُرْآنًا عَرَبِيًّا لِّقَوْمٍ يَعْلَمُونَ﴾

"A Book whereof the verses are explained in detail a Qur'an in Arabic for people who understand"[2]. (Fusilat: 2).

So it is necessary for all Muslims to learn classical Arabic for clear understanding of Qur'an.

When large number of people from non-Arabic regions entered into Islam, then it was realized that there is need to develop Arabic Grammar rules. There are different narrations regarding who started these rules: according to most scholars, these rules were written first by AbulAswad ad Du'ali after being tasked by Ali (r.a), to write down the rules of Arabic grammar. AbulAswad ad Du'ali was also the first person that placed dots on Arabic letters and wrote about Arabic grammar.

Muslim scholars all around the globe are trying their best to reach out for facilitation of Qur'anic learning. Different approaches exist for this purpose. Some regions stick to classical methodologies like "*Madrasah*" system in Pakistan, India and Bangladesh. The system usually has extended curriculum that covers all the major areas of Islamic Shariah. Another upcoming approach since last few decades is to teach Qur'anic Arabic via short courses. Since the curriculum is condensed, this approach is being widely used by the common populace and huge numbers of people are benefiting from it.

IT has a vital role in social and professional landscape of modern era. In the domain of Qur'anic learning; there are many well-known E-Learning

[1] Abdullah, Y., Ali. (2010). The Meaning of the Howly Qur'an: Madison Park.
[2] *Ibid.*

institutes around the globe, offering variety of beneficial opportunities to diverse people irrespective of the barriers of nationalities, race and age, etc. The approaches and methodologies employed by these institutes have various limitations, with regards to IT strategy as well as the target scope.

In this paper we are focusing only on Arabic lexicology "*ilm sarf*", the proposed approach employs automated techniques using different algorithms which generates conjugations of given root for different form used for teaching purpose and similarly the algorithm is used to evaluate the level of student in respective test. The system generates the conjugations on the base of given roots and it also generates the Urdu translations with the help of algorithms. System provides the facilities to choose the lesson as per their respective requirements, and choose the appropriate mode, that is, teaching or examination. The user can begin the lesson or exam as per requirements of available time. The automated system will be made online available for students 24/7. Purposed system will be beneficial for students to learn Arabic and to gain the direct knowledge of Quran.

2. LITERATURE REVIEW

Arabic is a Semitic language. In Semitic languages, words are formed according to defined patterns: similarly, formation of verbs and conjugated nouns also follow these measures or patterns. Arabic conjugation has rules for deriving words by adding affixes and suffixes to the root with specific rules for obtaining different forms of gender, number, case and mood, etc.

This pattern base mode of Arabic morphology attracts the attention of researchers; we find a lot of research in this area that has focused on the development of different approaches to analyze the Arabic language at the morphological level. Techniques have been developed to remove all the extra characters from a word and for analyzing what extra affixes or suffixes were added to form any word. These techniques bring help in the form of an application which is used for information retrieval, text categorization, dictionary automation, text compression, data encryption, vowelization and spelling aids, automatic translation, and computer aided instructions. Mainly there are two approaches used for Arabic conjugations. The first approach, comprising the stem base algorithms (Buckwalter's stemmer), removes the suffixes and prefixes of the word and generates the roots. The other approach is root base stemmers which procure the roots from generated stems (Beesley, Al-Shalabi, and Khojastemmers).

Urdu language is very rich for Islamic literature, there is also large number of Quran translations and commentaries are available. Apart from Quran and Hadith, Urdu Islamic literature covers all the major areas of Islamic knowledge. Urdu Islamic literature is translated in different languages around the globe. Although there are many efforts for Computational Arabic Morphology, but I was not able to find even a single effort of automation based system for generating and mappingUrdu morphology with Arabic morphology. According to the best of our knowledge, this work is the very first effort in this area.

3. TEACHING METHODS

Teaching of Classical Arabic is very important many institutes around the globe focused on teaching Classical Arabic. In Pakistan, Bangladesh and India, the Madrasa system is teaching Classical Arabic language along with other Shariah subjects. Universities have their Islamic departments in which different Islamic subjects including Arabic morphology are being taught. The Madrasahs typically employed traditional pedagogical methods for teaching. Although modern universities have embraced advanced teaching tools, the tools are typically restricted to the design of web pages or the use of slides for teaching. There is need to use the advanced computational techniques for teaching purpose to fill this gap. In particular, there is need to use advanced techniques from the fields of NLP and AI towards this end. The proposed system is an effort in this direction. The system uses different algorithms for teaching lessons and examination. This system will facilitate teachers in preparing their lessons while also helping students by providing enriched opportunities of learning Classical Arabic.

4. METHODOLOGY

In this section, approaches used for teaching and evaluation software are explained, first the teaching system and then the evaluation system.

4.1. Problems and Challenges

There is a huge gap between the computer sciences academia and classical Arabic sciences. They know little about each other's respective domains. Hence, the resulting lack of computational research in the Arabic sciences

community has led to an absence of standardized tools in this domain. Requirement analysis and domain knowledge are very important for development of any system, stronger grip on requirement will results in the form of good system.

Learners belong to all age groups and linguistic backgrounds, thereby creating a unique requirement. Any learning system, therefore, must cater to the challenge of differential learning. Moreover, the beginners need an easier learning curve whereas the advance learners need a good challenge. They require expansive data for research and further analysis.

4.2. Automated Teaching System

The artificial intelligence based system designed for teaching purpose covers all the required aspects. The user of the system can easily choose the lesson plan as per requirement. The system generates the conjugations for past tense "الفعل الماضى" according to the given root and door[1]. The system generates the Arabic conjugations and its translation, Figure 1 shows the data generated by the system against root "س م ع" from trilateral door for past tense.

[1]Same Arabic root letter have different patterns according to given door.

Figure1: The root "s.m.a." (س.م.ع).

Figure 2 shows the data generated by the system against root "س ک ت" from trilateral door for present tense.

Figure2: The root "s.k.t." (س.ک.ت).

Subsequently, the system creates animated lesson according to the selection of the user. Previous two pictures are related to the early stage of Arabic morphology. The occurrence of weak letters[1] in root words changes the conjugation patterns hence the presence of weak letters requires special care and treatment. Arabic conjugation system for these verbs have already been designed by us. Figure 3 shows the conjugation for weak word "قوم" and the conjugation is generated for door "استفعال".

[1]Weak letters in Arabic are "ا","و" and "ی"

Figure3: The conjugaison of the root "q.w.m." (ق.و.م).

Figure4 below shows the results for a student who is learning "اسم الفاعل" and he is learning it for the "باب الاستفعال". Facility is provided to the user to choose the different root words and the system will create lesson on the base of provided algorithms.

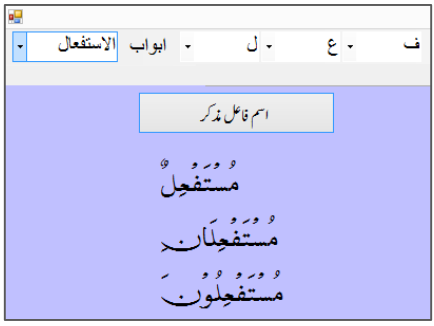

Figure4: The subject form (the doer).

4.5. Automated Examination System

The system also provides facility to the students of to evaluate the current capability of the student, and also recommends the improvements for better understanding. Again it is an algorithm based approach. The flow diagram of the examination of past tense is Figure5.

Figure5: The Past tense diagram.

Just one example is shown here to depict the working of our proposed automated examination system for past tense active voice. System has been provided the list of the root words; the conjugations for past tense active voice are created on the basis of these root words.

The examination system will also create and maintain more records as the number of aspects will be increased. The table1 shows the statistics for number of records to generate and maintain the right answer.

Roots	Doors	Count	Aspects	Answers
200	1	2800	3	8400
200	1	2800	4	11200
200	10	28000	3	84000
200	10	19600	4	112000

Table1: The Statistics of the Records.

5. EVALUATION SYSTEM

The evaluation system will help to find out the strengths and weaknesses of the particular person for a particular test. The system will provide facility to the teachers to find out what are the strong areas of the students and what are the areas where students need more improvement. Figure6 shows the results of a test where 20 questions were asked to the different students including three aspects, that is, the gender, number and person.

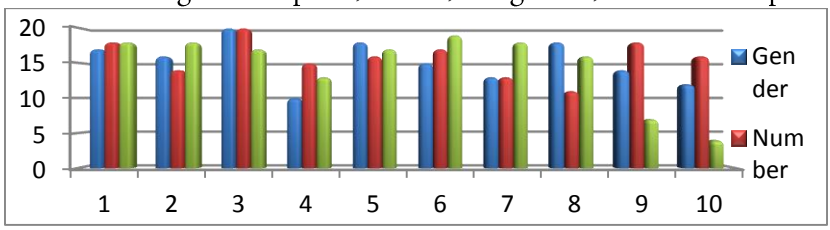

Figure6: Results of the Students.

5.1. Detailed analysis of Errors

The system not only mentioned the number of correct or wrong answers but it also provides the facility to show the comparison of right and wrong answers. For all the wrong answers, the system shows both the wrong answer and the right answer. The detailed analysis of the test is shown in Figure 7.

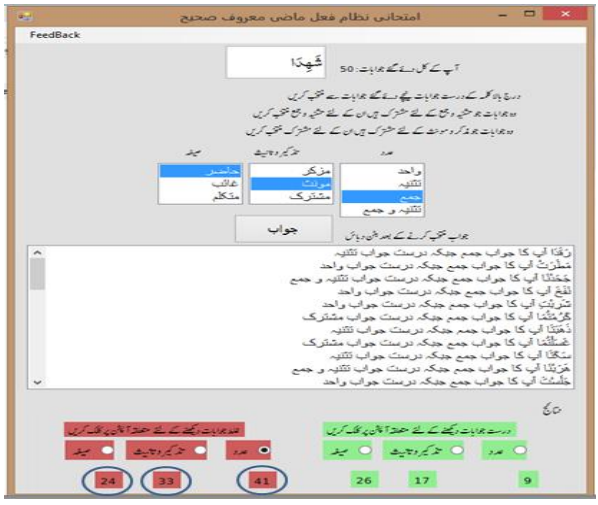

Figure7: Detailed report mentioning right and wrong answers.

6. CONCLUSION

In this paper, we have proposed system for teaching Classical Arabic conjugation for the users of Urdu language. Our system is based on algorithmic techniques borrowed from the field of Artificial Intelligence (AI) and Natural Language Processing (NLP). Our proposed system is both dynamic and intelligent. It is primarily designed and developed to offer teaching and of Qur'anic sciences and student assessments. Native language support is available, starting with Urdu. Differential learning is also possible using the same data sets. The research has shown early adoption by computer literate learners and teachers finding assistance for their distance learning programs. Moreover, the self-evaluation provides rapid learning and increased productivity.

REFERENCES:

A. Alsuwaynea, Information Retrieval in Arabic Language. SA: King Fahad National Library, 1995.

A. Gully (1995). Grammar and semantics in medieval Arabic: A study of Ibn-Hisham's 'Mughni l-Labib. *Curzon Press.* Richmond, Surrey.

A. Jones (2005). Arabic Through the Qur'an. Islamic Texts Society.

Abdul Baquee Muhammad. FIRSTYEARTRANSFERREPORT, Annotation of Conceptual Co-reference and Text Mining the Qur'an 2009.

Abu Shawar, Bayan; Atwell, Eric. 2004. An Arabic chatbot giving answers from the Qur'an in: Bel, B &Marlien, I (editors) Proc TALN04: XI Conference sur le TraitementAutomatique des LanguesNaturelles, Fez, Morocco.

Al-Kharashi, I.A. & Evens, M.W. 1994.Comparing words, Stems, and Roots as IndexTerms in an Arabic Information RetrievalSystem. *Journal of the American Society forInformation Science.***45**(8): 548-560.

Al-Omari, H. 1994. ALMAS: An Arabic Language Morphological Analyzer System. National University of Malaysia.Bangi, Selangor.

Al-Sughaiyer and I. Al-Kharashi, "Arabic morphological analysis techniques: a comprehensive survey," J. Am. Soc. Inf. Sci. Technol., vol. 55, no. 3, pp. 189–213, 2004.

AnasBoubas, Leena Lulu, BoumedieneBelkhouche, SaadHarous GENESTEM: A Novel Approach for an Arabic Stemmer Using Genetic Algorithms Faculty of Information Technology UAE University, Al Ain, UAE

El-Sadany, T.A. & Hashish, M.A. 1988.Semi-Automatic Vowelization of Arabic Verbs. Proceedings of 10th National Computer Conference: 45-56.

J. Dawson, "Suffix removal and word conflation," Association for Literacy and Linguistic Computing Bulletin, vol. 2, no. 3, pp. 33–46, 1974.

J. Lovins, "Development of a stemming algorithm," Mechanical Translation and Computational Linguistic, vol. 11, pp. 22–31, 1968.

K. Dukes and T. Buckwalter (2010). A Dependency Treebank of the Quran using Traditional Arabic Grammar. *Proceedings of the 7th International Conference on Informatics and Systems (INFOS)*. Cairo, Egypt.

M. Giolfo (2010). *yaqumvsqama*in Classical Arabic Conditional Context: A Relativistic Interpretation of the Frontier between the Prefixed and the Suffixed Conjugations. *Foundations of Arabic Lingustics (FAL)*.Cambridge, United Kingdom.

Sawalha, Majdi; Atwell, Eric. 2010a. *Fine-Grain Morphological Analyzer and Part-of-Speech Tagger for Arabic Text* in Proc LREC'2010: Language Resources and Evaluation Conference, Valetta, Malta.

S. Khoja and R. Garside. (1999) Stemming Arabic text.[Online]. Available: www.comp.lancs.uk/computing/users/khoj aJstemmer.ps/

T. Buckwalter, Buckwalter Arabic Morphological Analyzer Version 2.0, Linguistic Data Consortium, 2004.