ISSN 2007-9737

Methods and Trends of Machine Reading Comprehension in the Arabic Language

Mubarak Alkhatnai[1], Hamza Imam Amjad[2], Maaz Amjad[3], Alexander Gelbukh[3]

[1] King Saud University,
Department of English Language and Translation,
Saudi Arabia

[2] Moscow Institute of Physics and Technology,
Russia

[3] Instituto Politécnico Nacional,
Centro de Investigación en Computación,
Mexico

malkhatnai@ksu.edu.sa, {hamzaimamamjad, maazamjad}@phystech.edu

Abstract. Machine Reading Comprehension (MRC) is an essential task in natural language understanding in which the goal is to teach machines to understand the text. Machine understanding can be evaluated through question answering techniques. Limited research has been done in Arabic reading comprehension. This study presents a survey on trends and methods of Machine Reading Comprehension (MRC) methods in Arabic. The survey summarizes recent advances in Arabic reading comprehension, primarily emphasizing on two aspects (i.e., available datasets and methods). The study provides a detailed analysis of MRC techniques and compares various datasets used to study Arabic reading compression.

Keywords. Machine reading comprehension, low-resource language, arabic language.

1 Introduction

Machine Reading Comprehension (MRC) has been an essential task in Natural Language Processing (NLP). An enormous impact on humanity would emerge if computers could understand what they command to read. Moreover, machines help to do multiple tasks. For example, summarizing information, providing answers to our queries, and emerging new ideas. Finally, there are growing appeals for the MRC techniques in low resource languages, in particular, the Arabic language which is now being spun out into commercial applications.

Machine Reading Comprehension is an AI-complete practice that needs a Q&A system. To understand the text, the machine first processes the given text to understands it and then extracts the answer to the user question. Computer reading comprehension is a simple task of addressing a textual query, in which each query is given the meaning from which to infer the answer.

The goal of MRC is to derive the right solution from the context in question or even to produce a more nuanced context-based response [1]. MRC is promising to close the void in understanding natural language between humans and computers.

There is a deficiency of MRC based applications in low resource languages. However, the machine mimics human abilities tremendously by processes the span of text and transects the answer.

This machine's ability gained enormous attention by incorporating various deep learning techniques over the past few years, and a significant amount of data is required to train machines to do sophisticated tasks. Nevertheless, the emerging issue with MRC is the unavailability of a large scale corpus in various languages, mainly in Arabic.

Substantial work on popular languages like English [2, 3, 4], Chinese [5], and in a multilingual aspect [6] have administered great success. However, low-resource language, in particular, the Arabic language, is in dreadful needs of tools and resources. Besides, Arabic lacks large scale corpora, and resources are insufficient to develop statistical models for multiple tasks such as text summarization, text plagiarism, and fake news detection [39].

Arabic is one of the worldwide spoken languages in Gulf areas having more than 200 million speakers[1] and is used as a first language in more than 30 countries. It is also a low resource language in terms of the unavailability of NLP tools and limited labelled datasets. This language has a considerable number of speakers, and a significant number of native Arabic speakers are unable to speak English. Therefore, they prefer to talk in their indigenous language. This demands to overview the current advancements in the Arabic language, especially for MRC to find the gap in the literature review.

It is also important to mention that, according to the BBC report[2], half of the world's languages will be extinct by the end of this century. Thus, it is essential to establish and advance the current NLP techniques to low resource languages to preserve them. Similarly, UNESCO[3] started an endangered languages program to preserve and highlight the importance of low resource languages. The UNESCO report mentioned that it is vital to preserving unwritten and undocumented languages to avoid cultural heritage loss and valuable ancestral knowledge. I addition to this, WHO4 started a Universal Health Coverage (UHC) initiative for less-resourced languages. Therefore, survey studies in low resources language like Arabic, especially on NLP topic such as MRC is significant.

Current models trained on Arabic data pose various challenges and deliver low-quality performance since most recent pre-trained models are available only in English. This problem is compounded in the sense of social media, where contact using languages without an adequate model must be made possible to interpret passages written in Arabic [7]. Furthermore, in Arabic, little work has been proposed in MRC using deep learning techniques such as utilizing the neural networks and transformers based models such as BERT.

No previous study is conducted to survey the machine reading comprehension task in the Arabic language. Thus, this absence of exploration in machine reading comprehension spurs this study.

The rest of the paper is organized as follows. In section 2, we discuss the related works in Arabic reading comprehension task. Section 3 shows the different MRC datasets available in the Arabic language. Section 4 discusses the recent development in MRC techniques for the Arabic language. MRC in Arabic. In section 5, we shed light on some challenges in Arabic reading comprehension task. Section 6 talks about open challenges. Section 7 is about future development in MRC. The paper is concluded in Section 8.

2 Related Work

2.1 MRC in Machine Learning

As deduced by the research, there are two main approaches to NLP in low-resource languages, in particular, Arabic language settings. Firstly, a traditional technique that mainly focuses on collecting data for a single or a variety of languages. In this approach, the data collection and annotation need to be done by an expert. However, due to each language features and complexity, a data collection strategy from scratch is required.

Secondly, deep learning neural networks, especially transfer learning-based approaches which are used to tackle various NLP tasks. Transfer learning focuses on storing data and knowledge gained when searching for one problem and applying the acquired knowledge in a different but related question. For example, suppose that a machine learning model is trained to find plagiarism in the text.

[1] https://vistawide.com/languages/top_30_languages.htm

[2] https://www.bbc.com/future/article/20140606-why-we-must-save-dying-languages

[3] http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CLT/pdf/FlyerEndangeredLanguages-WebVersion.pdf

[4] https://www.who.int/medical_devices/innovation/compendium/en/

Similarly, that trained machine learning algorithm can also be used to find the authors of the text (author attribution). Therefore, transfer learning is used to learn about a specific language pattern, i.e., English and use the gained knowledge to solve NLP tasks in another language like Arabic.

Deep learning comes in very handy when dealing with word embedding, or word vectors as they are the cornerstone of many NLP approaches. Word embedding is a learned representation for a particular text where words with the same meaning can be represented the same way. It is a technique where each word is assigned a real-valued vector in a predefined vector space. An individual word is mapped against a single vector, and the values are learned in a way that it resembles a neural network. The phenomenon is considered a breakthrough in deep learning as it directly addresses various NLP's problems.

2.2 MRC in Non-Arabic Languages

Machine Reading Comprehension has been used in various languages. This task has been addressed in English to create multiple applications ranging from educational testing services to enterprises. For example, a study [8] reported that MRC techniques were used to display text on a screen by TOEFL Listening examination. Moreover, Microsoft uses MRC techniques to answer enterprise domain-specific questions[5]. However, low-resource languages such as Arabic is in dire need of such NLP tools and resources to deliver benefits on a large scale.

Different deep learning models, especially attention-based models, are introduced for machine-reading comprehension task in the English language. For example, a self-attention mechanism was introduced [9] to capture the information in the text. Similarly, the previous model [9] was improved [10] by establishing a relationship between the text and the question on SQuAD dataset. However, BERT [11] and QANet [12] provide state-of-the-art results for machine-reading comprehension task in the English language.

Deep learning is a phenomenon that enables machines and tools to mimic the human brain and process the data given to it. After processing the input data, machines can detect objects, recognize speech, translate languages, and make decisions on their own without human supervision.

Deep learning incorporates neural networks within its architecture. Neural networks have become a hot development in machine learning, as it consists of several algorithms that can learn more complex functions. Just like the human brain uncovers patterns and connections in a dataset, neural networks enable machines to learn hidden patterns to make correct decisions.

MRC has been studied in many different languages, primarily in the English language [13]. MRC techniques are applied in various applications and industries.

For example, it is used to study human behaviour [14] using an adaptive evaluation system [15]. Moreover, it has been used in mobile applications [16], web browser [17], in educational objectives such as in spoken content [18], and novice code comprehension skills [19].

Furthermore, it is also used in multiple learning programs form disable (special) students [20] and teaching mathematics [21]. Likewise, MRC techniques also have been used in designing applications like chrome applications [22]. However, all these applications are in non-Arabic languages. Therefore, it is essential to initiate advancement in MRC methods for the Arabic language.

3 MRC Datasets in Arabic

In this section, we summarized different datasets used for Arabic reading comprehension task. Most of the previous studies on reading comprehension task in the Arabic language used automatic machine translation.

For example, several studies used google translator to translate existing datasets in English such as SQuAD and used google translator to translate into Arabic for the development of end-to-end MRC algorithms. Datasets from English to Arabic. The reason is that creating resources in the form of a dataset is an expansive and time taking task.

[5] https://www.microsoft.com/en-us/ai/ai-lab-machine-reading

Table 1. The question answering datasets available in the Arabic language, where p: paragraph, q: question and a: answer

Dataset	Source	Formulation	Size
ArabiQA [34]	Wikipedia	q,a	200
DefArabicQA [35]	Wikipedia and Google search engine	q, a with documents	50
Translated TREC and CLEF [36]	Translated TREC and CLEF	q,a	2,264
QAM4MRE [37]	Selected topics	document,q and multiple answers	160
DAWQUAS [39]	Auto-generated from a web scrape	q,a	3205
QArabPro [40]	Wikipedia	q,a	335
Quranic Corpus Ontology [7]	Holy Quran	q,a	7500
Arabic-SQuAD [21]	Translated SQuAD	p,q,a	48,344
ARCD [21]	Arabic Wikipedia	p,q,a	1,395

However, a recent study [21] used crowdsourcing to create a high-quality dataset for Arabic reading comprehension task.

The datasets were extracted from several sources such as Wikipedia, Factoid articles, stories, cQA, Holy Quran. The machine learning-based approach for MRC in Arabic is different from the rule-based MRC approach. The reason is that previous studies restricted to a few words.

For example, during machine translation, the length and position of words in paragraphs from English to Arabic translation were different. Therefore, to tackle multi-sentence question answers with a courtesy fragment, and lexical and semantic dependencies consideration and different Arabic QA scheme model was proposed by Romeo et al. [23]. This research was based on a support vector machine (SVM) through tree kernels with advanced text representations achieved better results.

Table 1 shows the statistics and summary of different datasets used for Arabic reading comprehension task.

4 MRC in Arabic Language

Arabic is the first official language in more than thirty counties, but limited research used deep learning for Arabic reading comprehension task.

Most of the previous works in Arabic reading comprehension considered rule-based techniques.

For example, a recent study by Emad Al-Shawakfa [24] presented a rule-based QA approach for Arabic questioning and answering.

In common languages like English, regardless of the type of question being asked, every question falls in the category of who, what, why, when, where, and how (5Ws and 1H). However, that is not the case with Arabic, as there are more than ninety rules defined for a detailed question-answering in Arabic.

In addition, most studies translated questions and context into the English language for text processing. The text was segmented into tiny parts, and a semantic tree was built. After a semantic tree was constructed on Stanford parser's base, the text was translated from English to Arabic. Finally, the candidate answers were extracted and ranked based on the allocated score.

Several studies used different techniques to tackle Arabic reading comprehension task. Some studies used question answering techniques. For example, a study [31] presented a question answering system called Al-Bayan to tackle basic multi-sentence questions. Although a highly accurate QA system was proposed by Abdelnasser et al. [25] to address general to domain-specific questions.

Nonetheless, the QA system, called Al-Bayan, was explicitly designed for the Holy Quran, and it was based on three primary modules. The first module focused on question analysis, including pre-processing and classification of information.

In the second module, the information retrieval process was applied to the verse containing the required answer. Entity recognition and feature extraction were performed in the third module. All in all, 6236 verses of the Holy Quran were classified into 1217 concepts by taking Arabic questions as input and returning the relevant verse as an output. While it may sound simple for a human to choose the right choice that fits the blank, it is not that simple for machines to do so. To address the problem related to multiple-choice questions and answer tasks from a piece of text, Trigui et al. [26] proposed an Arabic MRC question answering approach based on Information Retrieval (IR). The approach presented a shallow process that was developed for Arabic test understanding. Whenever the system needed to infer text, the system used IR methods to conclude. In such an NLP system, anaphora is a common challenge that needs to be addressed.

Information retrieval (IR) and other MRC techniques were also combined to comprehend Arabic text. For example, Hammo et al. [26] proposed an Arabic MRC QA system called QARAB, zeroed in on data extraction, alongside a distinctive language handling approach. The catchphrase coordinating strategy was joined with the coordinating structure derived from the query, and the candidate text recovered by the IR technique. The IR framework was utilized to look through all the documents explicitly according to the inquiry while the NLP framework was used to make Arabic dictionaries for tokenization, checking, and extraction of highlights and acknowledgement of appropriate names. The reason for the proposed system was the Arabic content separated from the Arabic documents.

Another study by Mozannar et al. [27] proposed an Arabic MRC system called (SOQAL). The system extracted articles from Wikipedia for Arabic questions and removed answers by translating using SQuAD dataset. The study proposed a system called SOQAL, which was a two-step system used for open-domain Arabic QA tasks using Wikipedia as the knowledge base.

The first step involved document retrieval through the Hierarchical TF-IDF mechanism. In the second step, the MRC model using a pre-trained bidirectional BERT transformer was used. This enabled the system not to carry out a full-on search, and therefore, as a result, saved a considerable processing time.

Likewise, another Arabic dataset named ARCD was presented [21] which contained 1,395 questions taken from Wikipedia by the automatic machine translation of Arabic-SQuAD. The ARCD dataset contained 155 articles, and each article had 250 characters. In addition to this, additionally, 235 articles were translated using Google NMT translator, which included 48344 questions from the 10354 paragraphs obtained from them. In total, roughly 269k sections had 233 characters in each of them. In the end, the answers from the ARCD dataset were manually categorized into numerical and non-numerical form, and the questions were labelled on the word match, syntactic variation, ambiguous and multiple sentences.

Most of the studies used three metrics to evaluate the performance of the systems. The first is an exact match (EM) that quantify the percentage of correctly predicted ground-truth answers. Macro_F1 score (macro-averaged) is used as a second measured to evaluate the performance of the systems, which is the average overlap between the prediction tokens and the ground truth answer tokens. Finally, a sentence match (SM) metric is used as a third metric which calculates the percentage of the predicted sentence that fall in the paragraph containing the ground truth answer. Table 2 shows the performance of various models of Arabic reading comprehension.

5 Challenges in Reading Comprehension in Arabic

We identify the following challenges in reading comprehension in Arabic:

- **Lack of Data**: Accessibility of enormous datasets is a prerequisite for the training of language models. For example, deep learning models such as RNN, LSTM and GRU require a substantial amount of annotated data for

Table 2. Comparison of the different document reader modules in the Arabic language

Method	Arabic-SQuAD			ARCD		
	EM	F1	SM	EM	F1	SM
Sliding Win. + Dist.[32]	0.00	5.80	29.2	0.07	14.2	58.4
Embedding Fast Text [33]	0.04	6.96	43.1	0.36	15.3	73.1
TF-IDF Reader [21]	0.27	2.41	49.2	0.22	5.6	75.3
QANet fastText [33]	29.4	44.4	61.7	11.0	38.6	83.2
BERT [11]	34.1	48.6	66.8	19.6	51.3	91.4

model training so that the significant results can be achieved. However, a limited of annotated data is available in the Arabic language, which makes Arabic reading comprehension even more challenging.

- **The semantic problem:** Semantic issues in multilingual displaying where just punctuation is considered during the interpretation cycle. At the point when MRC is performed in some low-resource languages by multilingual showing the standard datasets are considered in English, and the transformation of text from a low-resource language to a rich resource language causes semantic issues due to the variety in the linguistic type of the two dialects and affect the understanding of the content.

- **Insufficient knowledge base:** Arabic has insufficient information base, such as Wikipedia. In contrast, Wikipedia for English is available in enormous amounts for various areas. The absence of data is a significant impediment to the creation of large datasets for low resource dialects such as Arabic.

- **No standard benchmark:** A typical benchmark is needed to test and advance the examination network. PC Vision advanced after the improvement of the ImageNet dataset, where different models proposed could be analyzed for execution evaluation. Arabic does not have a benchmark dataset for the MRC task.

- **Scalability Issues**: There are adaptability issues for different data domains because of the dependence on rich language models, such as English. For example, a system is required to identify "Islamic hadith" that is in Arabic. Yet the machine learning model prepared by multilingual demonstrating cannot perform hadith as there is no information on hadith in English.

- **Lack of capitalization**: Contrasted with English, Arabic does not have an upper casing that makes it difficult to perceive proper nouns, abbreviations and acronyms.

6 Open Issues

A literature study in Arabic reveals many issues with machine reading comprehension (MRC) that remain unsolved. These following open issues create a gap between MRC applications and their users:

- **Robustness of Arabic MRC Systems:** Based on the literature analysis, most of the MRC models in Arabic proved ineffective on overlap context [28]. The MRC model called Stanford Question Answering Dataset (SQuAD) drops the performance of Arabic comprehension tasks dramatically in adversarial questions.

- **Lack of Inference Ability**: Machine reading comprehension (MRC) systems retrieve the answers based on semantic matching between the question and given data. Based

on the literature review in this study, Arabic comprehension systems contain ineffective inference ability [29].

- **Difficulty in Interpretation**: MRC systems work in a black box, i.e., insufficient information is available about how does a system retrieve the answer. Lack of interpretability is its primary drawback that drops its efficiency. However, some MRC models such as HotpotQA [30] are introduced in English to enhance the interpretability of MRC systems, but such models are not available in Arabic.

7 Future Development

As day-to-day computing plays an increasingly prominent role in our everyday lives, the reader's interpretation is of critical importance for human development [23]. Guld countries such as Saudi Arabia, UAE, Qatar, and Egypt have a large population.

However, many people who live in these countries are not well versed in English and like to communicate and learn in their language. For such purposes, they need to find the right information that is not possible without having an MRC system in Arabic. Thus, this study aims to provide an analysis of what has been done in Arabic and raises research questions to develop robust MRC models for Arabic comprehension.

8 Conclusion

This article provides a survey of machine reading comprehension models in the Arabic language. Based on a detailed review of recent works, we include clear descriptions of the tasks of the MRC in Arabic and compare them in detail. We also discuss the recent advancement of MRC models in Arabic.

This study also highlights limitations in Arabic comprehension models and research questions. However, the course of the MRC is changing very rapidly, and it is challenging to incorporate all works in Arabic MRC in this report. We hope that this analysis will ease the connection to recent MRC developments and inspire more researchers to work in Arabic reading comprehension.

Acknowledgment

The authors express their gratitude to the Deanship of Scientific Research at King Saud University, Saudi Arabia, and the Research Center at the College of Languages & Translation for their support with the publication of this article.

References

1. **Gibbons, J. (2016).** *Comprehending ringads.* A List of Successes That Can Change the World. pp. 132–151, Springer, Cham.
2. **Sun, K., Yu, D., Yu, D., & Cardie, C. (2018).** *Improving machine reading comprehension with general reading strategies*. arXiv:1810.13441.
3. **Horbach, A., Aldabe, I., Bexte, M., de Lacalle, O. L., & Maritxalar, M. (2020).** Linguistic Appropriateness and Pedagogic Usefulness of Reading Comprehension Questions. *Proceedings of The 12th Language Resources and Evaluation Conference*, pp. 1753–1762.
4. **Zhang Xin, Yang An, Li Sujian & Wang Yizhong (2019).** *Machine Reading Comprehension: a Literature Review*. arXiv:1907.01686.
5. **Cui, Y., Che, W., Liu, T., Qin, B., Wang, S., & Hu, G. (2019).** *Cross-lingual machine reading comprehension*. arXiv:1909.00361.
6. **Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., & Stoyanov, V. (2019).** *Unsupervised cross-lingual representation learning at scale*. arXiv:1911.02116.
7. **Liu, S., Zhang, X., Zhang, S., & Wang, H. (2019).** Neural Machine Reading Comprehension: Methods and Trends. *J. Applied Sciences*, Vol. 9, No. 18, pp. 3698. DOI: 10.1016/j.apsusc.2018.11.214.
8. **Tseng Bo-Hsiang, Shen Sheng-syun, Lee Hung-yi & Lee, Lin-Shan (2016).** *Towards Machine Comprehension of Spoken Content: Initial TOEFL Listening Comprehension Test by Machine*, pp. 2731–2735. DOI:10.21437/Interspeech.2016-876.
9. **Wang, W., Yan, M., & Wu, C. (2018).** Multi-granularity hierarchical attention fusion networks for reading comprehension and question answering. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Vol. 1, pp. 1705–1714. DOI: 10.18653/v1/P18-1158.

10. **Wang, W., Yang, N., Wei, F., Chang, B., & Zhou, M. (2017).** Gated self-matching networks for reading comprehension and question answering. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics,* Vol. 1, , Association for Computational Linguistics, pp. 189–198. DOI: 10.18653/v1/P17-1018.
11. **Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2018).** *Bert: Pre-training of deep bidirectional transformers for language understanding*. arXiv:1810.04805.
12. **Yi Yang, Wen-tau Yih, & Meek, C. (2015).** Wikiqa: A challenge dataset for open-domain question answering. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 2013–2018. DOI:10.18653/v1/D15-1237.
13. **Guo, J. (2019).** Research on the Application of Machine Reading Comprehension in Adaptive Evaluation System. *Conference: ICDTE 2019: 2019 The 3rd International Conference on Digital Technology in Education*. DOI:10.1145/3369199.3369241.
14. **Jia, R. & Liang, P. (2017).** *Adversarial Examples for Evaluating Reading Comprehension Systems*. arXiv:1707.07328
15. **Lee, C.H., Lee, H., Wu, S.L., Liu, C.L., Fang, W., Hsu, J.Y., & Tseng, B.H. (2018).** Machine Comprehension of Spoken Content: TOEFL Listening Test and Spoken SQuAD. *IEEE/ACM Transactions on Audi,* Vol. 27, No. 9. DOI:10.1109/TASLP.2019.2913499.
16. **Lewis, P., Oğuz, B., Rinott, R., Riedel, S., & Schwenk, H. (2019).** *Evaluating Cross-lingual Extractive Question Answering*. arXiv:1910.07475.
17. **Li, Q., Morris, M., Fourney, A., Larson, K., & Reinecke, K. (2019).** The Impact of Web Browser Reader Views on Reading. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems,* No. 524, pp. 1–12. DOI:10.1145/3290605.3300754.
18. **Liu, A., Qu, L., & Xu, Z. (2019).** Machine Reading Comprehension: Matching and Orders. *Conference: the 28th ACM International Conference*. DOI: 10.1145/3357384.3358139.
19. **Liu, S., Zhang, S., Zhang, X., & Wang, H. (2019).** R-Trans: RNN Transformer Network for Chinese Machine Reading Comprehension. IEEE Access, Vol. 7, pp. 27736 – 27745 DOI:10.1109/ACCESS.2019.2901547.
20. **MI, C., XIE, L., & Zhang, Y. (2020).** Loanword Identification in Low-Resource Languages with. *ACM Transactions on Asian and Low-Resource Language Information Processing*, No. 43. DOI:10.1145/3374212.
21. **Mozannar, H., Hajal, K., Maamary, E., & Hajj, H. (2019).** *Neural Arabic Question Answering*. arXiv:1906.05394.
22. **Pannim, P., Suwannatthachote, P., & Numprasertchai, S. (2018).** Investigation of Instructional Design on Reading Comprehension Affect the Demand for Mobile Application for Students with Learning Disabilities. *Conference: the 2nd International Conference*. DOI:10.1145/3291078.3291100.
23. **Riloff, E., Chiang, D., Hockenmaier, J., & Tsujii, J. I. (eds.) (2018).** *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing.* Association for Computational Linguistics.
24. **Nishida, K., Saito, I., Otsuka, A., & Asano, H. (2018).** Retrieve-and-Read: Multi-task Learning of Information Retrieval and Reading Comprehension. *CIKM´18,* pp. 22–26. DOI:10.1145/3269206.3271702.
25. **Romeo, S., Martino, G., Belinkov, Y., Barrón-Cedeño, A., Eldesouki, M., Darwish, K., & Moschitti, A. (2017).** Language processing and learning models for community. *Information Processing & Management,* Vol. 56, No. 2, pp. 274–290. DOI:10.1016/j.ipm.2017.07.003.
26. **Sucena, A., Machado, A., Freitas, H., Marques, C., Santos, A., Santos, I., & Rangel, I. (2019).** I Read: reading and writing skills promotion software. *ARTECH´19: Proceedings of the 9th International Conference on Digital and Interactive ArtsOctober´19*, No. 50, pp. 1–4. DOI:10.1145/3359852.3359884.
27. **Thapa-Chhetry, B. & Keck, T. (2019).** A Chrome app for improving reading comprehension of health information online for individuals with low health literacy. *IEEE/ACM 1st International Workshop on Software Engineering for Healthcare (SEH).* DOI: 10.1109/SEH.2019.00018.
28. **Trigui, O., Belguith, L., Rosso, P., Amor, H., & Gafsaoui, B. (2012).** Arabic QA4MRE at CLEF´12: Arabic Question Answering. LEF Online Working Notes/Labs/Workshop, Vol. 1178, *CEUR Workshop Proceedings, CEUR-WS.org*.
29. **Xiao, H. (2018).** *Machine Reading Comprehension: Learning to Ask & Answer*.
30. **Zhang, W., Wang, H., Zhang, S., Zhang, X., & Liu, S. (2019).** Neural Machine Reading Comprehension: Methods and Trends. *J. Applied Sciences*, Vol. 9, No. 18, pp. 3698. DOI:10.1016/j.apsusc.2018.11.214.

31. **Abdelnasser, H., Ragab, M., Mohamed, R., Mohamed, A., Farouk, B., El-Makky, N., & Torki, M. (2014).** Al-Bayan: An Arabic Question Answering System for the Holy Quran. *Proceedings of the EMNLP 2014 Workshop on Arabic Natural Language Processing (ANLP).* DOI: 10.3115/v1/W14-3607.
32. **Richardson, M., Burges, C.J., & Renshaw, E. (2013).** Mctest: A challenge dataset for the open-domain machine comprehension of text. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 193– 203.
33. **Yu, A.W., Dohan, D., Luong, M.T., Zhao, R., Chen, K., Norouzi, M., & Le, Q.V. (2018).** *Qanet: Combining local convolution with global self attention for reading comprehension*. arXiv:1804.09541.
34. **Benajiba, Y., Rosso, P., & Lyhyaoui, A. (2007).** Implementation of the ArabiQA question answering system's components. *Proc. Workshop on Arabic Natural Language Processing, Information Communication Technologies Int. Symposium, (ICTIS´07)*, pp. 3–5.
35. **Trigui, O., Belguith, L.H., & Rosso, P. (2010).** DefArabicQA: Arabic definition question answering system. *Workshop on Language Resources and Human Language Technologies for Semitic Languages, 7th LREC*, pp. 40–45.
36. **Abouenour, L., Bouzouba, K., & Rosso, P. (2010).** An evaluated semantic query expansion and structure-based approach for enhancing Arabic question/answering. *International Journal on Information and Communication Technologies*, Vol. 3, No. 3, pp. 37–51.
37. **Peñas, A., Hovy, E.H., Forner, P., Rodrigo, Á., Sutcliffe, R.F., Forascu, C., & Sporleder, C. (2011).** Overview of QA4MRE at CLEF´11: Question Answering for Machine Reading Evaluation*. CLEF Notebook Papers/Labs/Workshop*, pp. 1–20.
38. **Sidorov. G. (2019).** *Syntactic n-grams in computational linguistics*. Springer.
39. **Ismail, W.S. & Homsi, M.N. (2018).** Dawqas: A dataset for Arabic why question answering system. *Procedia Computer Science*, Vol. 142, pp. 123–131. DOI: 10.1016/j.procs.2018.10.467.
40. **Akour, M., Abufardeh, S.O., Magel, K., & Al-Radaideh, Q. (2011).** QArabPro: A rule based question answering system for reading comprehension tests in Arabic. *American Journal of Applied Sciences*, Vol. 8, No. 6, pp. 652–661.

Article received on 18/07/2020; accepted on 25/09/2020.
Corresponding author is Alexander Gelbukh.