

## Contribution of Artificial Intelligence on Arabic Language and Literature: A study

**Dr. Sofiur Rahman**

*Asisstant Professor, Rajiv Gandhi Memorial College,  
Lengtisinga, Distric: Bongaigaon, Assam*

### ABSTRACT

*Arabic literature, one of the richest traditions in world culture, has historically thrived on oral poetry, classical texts, and later modern prose and novels. With the rise of Artificial Intelligence (AI), new dimensions are being added to how Arabic literary works are created, preserved, analyzed, and disseminated. AI is not only aiding in the study of historical manuscripts but also inspiring creative writing and digital humanism in the Arab world. The language is a phonologically rich language, with sounds and emphatic letters that non-native speakers often find difficult to pronounce. AI-based speech recognition systems analyze a learner's speech, compare it with native pronunciation, and provide real-time feedback. Tools such as Google Speech to Text and AI-powered voice assistants help learners practice correct articulation of Arabic letters and words.*

*AI-powered NLP (Natural Language Processing) systems break down Arabic sentences, identify grammatical structures, and explain rules interactively. Intelligent-grammar checkers and writing assistants guide learners in constructing accurate sentences and mastering syntactic rules. Furthermore Arabic has many dialects. AI is being trained to translate works not only in Modern Standard Arabic (MSA) but also in Egyptian, Levantine, Maghrebi, and Gulf dialects, making literature more inclusive.*

*This study will explore the trans-formative impact of Artificial Intelligence (AI) on Arabic language learning, translation, and speech recognition, addressing both the potential and challenges of these technologies. The paper will examine the effectiveness of AI tools in enhancing Arabic language proficiency, the challenges posed by dialectal variations and cultural nuances, and the need for greater cultural sensitivity in AI development. The study highlights the importance of developing more advanced, inclusive, and culturally sensitive AI models to better meet the needs of Arabic speakers.*

### ARTICLE INFO

*Article history:*

**Received** 15 Sep 2025

**Received in revised form**  
15 Oct 2025

**Accepted** 15 Nov 2025

**Keywords:** Artificial Intelligence, Language, Literature, NLP, Gulf and dialects.

---

*Hosting by Innovatus Publishing Co. All rights reserved. © 2025*

---

### Introduction

The technology that allows machines and computer systems to perform tasks that typically require human intelligence, such as learning, problem-solving, and understanding language. It involves creating systems that can process data, recognize patterns and adapt to new information enabling them to assist with or perform complex tasks autonomously is called AI (Artificial Intelligence).

Artificial Intelligence (AI) contributes to Arabic language and literature by enhancing language learning through personalized platforms, providing tools for translation and cross-cultural communication, and offering new avenues for literary analysis, such as NLP-based Quranic study. AI also presents

opportunities to develop cultural awareness and offers support for both native and non-native speakers, although challenges like data accuracy and the preservation of cultural identity require careful consideration.

Language is one of the most important means of transferring knowledge, ideas, and communications between peoples, and it is part of the self-identity of nations. In light of the development of artificial intelligence, scientific paths began to strengthen the tools necessary for natural linguistics in terms of making machines distinguish and mimic human languages. Scientific attempts go back to the mid-1950s. The 20th century was fueled by the need to translate Russian language texts into English automatically and frequently by exploiting artificial intelligence algorithms in language learning.

The idea of building dictionaries and representing electronic knowledge was strengthened by applications that were integrated into the world of commercial competitions, as the concept of machine translation from multiple languages began to become broader and more widespread, as the Fourth Industrial Revolution, including the concepts of artificial intelligence, began to investigate human perception and the capabilities of machine self-sensing. Especially since technological innovations in the second half of the 21st century opened many horizons for growth and progress, which was reflected in the development of electronic capabilities in interpreting and communicating natural languages and developing the capabilities of educational programs.

Language is the defining identity of the indigenous people. The inhabitants of Mesopotamia and the Nile Valley countries were known by the names of their languages. For example, the inhabitants of Iraq and the Levant used to call themselves Arameans because they spoke Aramaic. The Arabic language, “the language of Dhad”, is one of the important languages on the global level because it carries the pillars of the civilization and cultural heritage of the world, which include writing, pronunciation, grammar, and morphology, as the Arabic language consists of 28 letters and 16,000 linguistic roots, compared to the English language, which has 26 and 700 linguistic roots, and the Hebrew language has only 22 letters and 2,500 linguistic roots.

The research problem emerges from the idea: “The extent to which artificial intelligence contributes to learning the Arabic language,” especially since artificial intelligence provides to assist in learning languages in general and Arabic in particular, along with the search for modern automated improvements necessary to ensure the success of their application.

Objective: The research aims to investigate tools related to artificial intelligence for benefit in learning the Arabic language, then highlights the advantages and shortcomings, as well as future scenarios for their use and the required improvements. The research is trying to answer the following questions:

A. What is artificial intelligence and computational linguistics?

B. What are the applications of the Arabic language in artificial intelligence, and what are its obstacles?

### **Hypothesis:**

The research assumes that information development in the 21st century, especially the tremendous progress in computer tools, was reflected in the development of artificial intelligence related to linguistic algorithms, which took a wider space in human uses while expanding the circle of human linguistic interests.

**Data Collection:** Datas are collected to prepare the paper from the secondary sources. It includes from wide range of sources, including web scraping, sensors, social media, public datasets, internal databases, books and journals.

### **Artificial Intelligence and Computational Linguistics:**

The term artificial intelligence refers to something “made up” that is unnatural and has the ability to understand appropriately through algorithms that can perform certain tasks that are close to human performance, and the complex overlap between information sciences and linguistic structures has led to the emergence of what is called “language computing” or “computational linguistics.” These sciences emerging from interoperability have begun to work on building and developing software for machine translation, and with the progress made by generative artificial intelligence in shaping audio, images, text,

and videos, the importance of developing content related to raising the efficiency of the level of artificial intelligence when communicating becomes increasingly important. Accordingly, we divide this requirement into three paragraphs:

- (a) The concept of artificial intelligence,
- (b) The concept of computational linguistics,
- (c) Natural language programming algorithms, and (4) areas of computational linguistics.

### **The Concept of Artificial Intelligence:**

Intelligence is the ability to learn and understand to achieve goals. It is divided into human intelligence and artificial intelligence. Human intelligence is a natural gift from the Creator to man so that he can understand and control the laws of life. Artificial intelligence is one of the computer sciences that search for advanced methods for carrying out analysis and deduction processes within limits that are somewhat similar to the processes of human intelligence. It is defined as the part of computer science that aims to simulate the cognitive ability to replace humans in performing appropriate functions in a specific context that requires intelligence. It is also defined as the efforts made in developing computer information systems so that the computer can behave in a manner similar to human thinking, such as learning languages and using images and cognitive forms to rationalize physical behavior, as well as using accumulated knowledge and experiences in decision-making processes, that is, the ability to understand, deduce, solve problems, and other capabilities that It requires intelligence whenever a person uses it.

Artificial intelligence has developed in philosophical stages since the 17th century when the mathematician and physicist René Descartes appeared, who likened the animal to a machine in that it neither thinks nor is conscious. In the 18th century, the French mechanic Jacques Vaucanson, who invented an automatic musical instrument, and the French physician and philosopher, emerged. Julien de la Mettrie wrote a book entitled: "Human Machine," in which he considered that human actions are automated actions.

In the 1940s, the American mathematician Norbert Vins appeared, who developed the science of cybernetics, which includes how the human brain works by focusing on nerve cells and producing the laws of the human brain's operation by means of a machine. Thus, the nerve cell was invented, and the cognitive approach described thought in an abstract manner that treats symbols in isolation. Regarding financial support, this approach was linked to language as a symbol. Then, in 1950, the British mathematician Alan Matteson Turing conducted a test on a computer machine, which was a game based on simulation.

Research on artificial intelligence has moved in three directions: "Limited artificial intelligence" seeks to simulate human cognitive functions according to a system of rules imposed by humans, such as programming self-driving cars. "General artificial intelligence" is based on simulation to be able to equal human intelligence, such as thinking, planning, programming, problem-solving, and communication. "Artificial superintelligence" seeks to go beyond human intelligence to become more capable than humans in subjects such as the arts, decision-making, and emotional issues.

Artificial intelligence includes a group of applied fields, most notably natural languages, writing and reading, vision and image robots, neural networks, and expert systems. Artificial intelligence also includes a group of characteristics, the most important of which are:

- (i) The possibility of solving the problems presented in the absence of complete information, relying on expertise and experience in understanding and learning.
- (ii) The ability to respond quickly to circumstances and situations and to deal with difficult and complex case.
- (iii) The ability to think and perceive and use right and wrong hypotheses to explore various phenomena.
- (iv) Acquiring and applying knowledge by using experiences and knowledge and employing them in appropriate situations.
- (v) The ability to provide information that supports appropriate administrative decisions.

(vi) Distinguishing the importance and priority of the elements of the presented cases.

Human needs and cognitive overlap have prompted increased interest in artificial intelligence to achieve a set of goals. The most notably are as follows:

- i) Create an organized knowledge database so that information is stored effectively so that employees in the organization, especially knowledge departments, are able to obtain knowledge and learn experimental rules that are not available in books.
- ii) Store information and knowledge related to artificial intelligence so that the organization can protect its knowledge from leakage and loss due to the leakage of employees through resignation, transfer from the organization, or death.
- iii) Create a mechanism that is not subject to human feelings such as anxiety, fatigue, or exhaustion, especially when it comes to work that represents physical and mental danger.
- iv) Generating and finding solutions to complex problems by analyzing the problems and addressing them in an appropriate and short time.

### **The Concept of Computational Linguistics:**

The term “computational linguistics” refers to the science that uses computers to transform information and linguistic texts into digital languages. The computer can analyze and translate by establishing strict formal descriptions that address various linguistic phenomena to provide the computer with the various operations and knowledge present in the language.

The fields of computational linguistics have multiple applications and can be developed according to modern computer technologies. Perhaps the most important applications of computational linguistics in artificial intelligence are applying linguistic statistics through linguistic roots, nouns, verbs, and derivatives and applying automated morphological analysis in extracting word roots and their derivatives. The application of spelling and grammar checking, the application of speech-to-text conversion, the application of teaching the Arabic language to non-native speakers by clarifying the shapes of letters and words, the application of comparison and contrast through the storage capacity and processing speed of the computer, and the application of automatic translation from Arabic to other languages, or vice versa.

### **Applications of Arabic Language in AI and Its Obstacle:**

Artificial intelligence has contributed to the development of the cognitive and learning fields of natural languages, as it is considered the digital gateway to the world of the future. Therefore, launching the machine to typically recognize the features of languages is one of the most important means used in analytical processors and transforming symbols into understandable expressions, despite the obstacles facing the development of computer systems. This simulates human uses regarding written or spoken communication, such as morphological analysis and syntactic disambiguation, but there are many researchers who have been able to develop digital systems that address various linguistic levels.

### **Automated Processing of the Written Arabic Language:**

The writing of Arabic letters is characterized by characteristics, the most prominent of which are the differences in Arabic letters in derivation, grammar, and morphology, as well as the path of writing the letters from right to left. The shape of the Arabic letter also depends on its connection with what is around it. Some letters are similar, and there is no difference between them except in some places, so Optical recognition of written Arabic letters is one of the most popular areas of artificial intelligence in designing Arabic language algorithms. Since the Arabic letters consist of vowel and consonant phonemes, the Arabic language units of speech are divided into five types of letter phonemes, which areas follows:

- i) A phonetic syllable of a vowel, such as (m)
- ii) An audio syllable of an extended letter, such as (ma)
- iii) An audio clip of a vowel and a consonant, such as (ml)
- iv) A phonetic syllable of a vowel and a consonant, such as (mal)

The natural language processing mechanism requires a database of standard language identifiers, which includes a database of letters, pronouns, and their graphemes, as well as the creation of identifiers for the parts of speech of nouns, verbs, and their derivatives, in addition to the creation of special identifiers for the nominal and verbal sentences and their derivatives, and tools for adverbs, adjectives, and links with the aim of analyzing Grammar and grammatical analysis, and there are four steps for language processing:

- i) Rooting: Searching for the root of the word, especially three-letter verbs.
- ii) Rooting: returning the word to its original state after deleting the prefixes and suffixes.
- iii) Branching: returning the verb to the past form or the singular noun in its original form.
- iv) Excesses: Identifying the affix at the beginning or end of a word and using it in analysis.

Despite the technical difficulties facing artificial intelligence in the characteristics of Arabic letters, the researcher, Shawqi Boufnar, used deep convolution neural networks to recognize images of written Arabic letters, and the results of the research showed an accuracy of (32.97) percent, and the research presented data sets containing Various types of written texts through the application of short-term, long-term, and multi-directional memory networks.

### **Automated Processors of Readable Arabic:**

Computer systems experts in the field of automatic speech recognition applications sought to save time and effort by bringing humans to the stage of speaking to the computer and giving it commands without the need to write linguistic texts. The application of automatic spoken speech recognition is defined as the process of converting audio signals into spoken speech or units. Linguistics is done through computer-based text coding with high accuracy.

With the development of technological techniques in the use of artificial intelligence, applications for automatic recognition of spoken Arabic speech appeared, and in the field of deep learning technology, researchers used a recurrent neural network to recognize the spoken Arabic numeral language by emphasizing the properties of the sounds of spoken speech, which are.

- (i) The sound energy feature: It distinguishes between static and vowel syllables. The moving syllables have a high value, and the static syllables have a low value. Sound energy is extracted from the spoken audio syllable by converting the continuous audio syllable into an interrupted one with a fixed time interval.
- (ii) The pitch property: It is a cognitive property that arranges audio clips in the form of a ladder that is linked to oscillations equivalent to Hz per second for the pitches and is used to determine the rise or fall in the pitch of the sound.
- (iii) The frequency characteristic of the sound of spoken speech: These are the frequencies related to the resonance and vibration that occur in the vocal cords during pronunciation and are distinguished by voiced letters more than voiced letters.
- (iv) The Mill Frequency Coefficients feature: It is the approximation of sounds to the way humans hear, as the human ear distinguishes low frequencies, and the Mill method enlarges the range of low frequencies using logarithms ranging from 660 to 1320 samples per frame and is estimated at 15 to 30 milliseconds.
- (v) Discrete wave transformation property: This is the property of representing the signal in the specified time and frequency domain by dividing the signal range of the frequency of the audio clip and arranging this signal from least to most.

### **Challenges of Text Processing in the Arabic Language:**

The artificial intelligence revolution has become part of the development of people. Researchers have been able to develop applications according to a futuristic outlook that may change the human way of life, to give precedence to human ability and replace it with automation. The robotics industry has begun to provide solutions to the problems raised in various areas of life, especially education, which benefits it in

the speed of learning, and with regard to the Arabic language benefits from artificial intelligence; it still faces many challenges, the most notable of which are:

- (i) The lack of adequate spread of the Arabic language on the Internet, which limits the spread of the Arabic language compared to the entry of other languages, especially in the field of education and scientific research, as these fields still depend on foreign languages, so the Arab user relies on these languages to meet his needs. Research, educational, and scientific research calls for more computer programs that help search engines browse the Internet in the Arabic language.
- (ii) Lack of coordination between companies and institutions interested in computerizing the Arabic language.
- (iii) The separation of Arab educational centers between training programs and educational programs specialized in computing and linguistics, in addition to the shortcomings in teaching general and applied linguistics.
- (iv) The inability of some programs to respond to total computing and cognitive and scientific accuracy, such as the automatic formation program, the automatic corrector program, the automatic reader, and automatic translation.
- (v) The inability to compile the Arabic blog and computerize it completely.
- (vi) Absence of a clear linguistic policy in which all Arab countries participate in developing the language and computerizing it.

It is necessary to spread the culture of artificial intelligence in the Arab world by establishing universities that teach artificial intelligence using the Arabic language and disseminating the experience so that the Arab person can engage in it since the educational and scientific reality of the Arabic language is still far from the scientific revolution brought about by artificial intelligence.

### **Conclusion:**

The advantages of AI are efficiency and automation, which handle repetitive tasks quickly, improved accuracy, minimizes human error, enhanced decision-making, as it analyzes vast data for insights 24/7 availability for continuous operation and cost reduction through automation and optimization.

The reality of the Arabic language and its future bets are linked to artificial intelligence and the software and applications it provides that work to computerize the Arabic language and process its levels in digital ways. The Arabic language has benefited from the computerization of some of its linguistic levels, such as the written level, the morphological, syntactic, semantic, and lexical levels, but there are still obstacles, such as the pragmatic level, textual analysis, controlling and computerizing the Arabic blog, and improving the quality of machine translation. Also, studies related to the Arabic language in the field of artificial intelligence, compared to other languages, lack specialists and institutions concerned with computational linguistics and computer science that contribute to reaching results that have cognitive value and scientific additions.

Below is the summary of the conclusions reached from the present study:

There is an absence of projects to process the Arabic language in computers in universities and institutions. Although Arabic is a broad and grammatical language, it can be processed using a computer.

The efforts to process the Arabic language are linked to technical companies in Western countries, and sometimes, there is no cooperation between these companies, computer programmers, and Arabic language specialists in the field of developing language applications.

There is a need for more applications of the Arabic language related to artificial intelligence, as the Arabic language is characterized by geometric and mathematical characteristics, such as roots and morphological weights, which require studies, research, and analysis to bridge the gap between applications of idiomatic structure

There is a need to promote the use of artificial intelligence in processing Arabic. It is preferable to divide projects for building Arabic language processing software into grammatical, morphological, and semantic analyses.

We should focus on reading and recognition programs for Arabic texts, sounds, images, and videos.

## References

1. Al-Qaraawi, S. M., & Mahmood, S. S. (2014). Wavelet transform based features vector extraction in isolated words speech recognition system. 2014 9th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP 2014), IEEE. 847–850. <https://www.proceedings.com/content/023/023907webtoc.pdf>.
2. Al-Shams, K. (2020). Computational Linguistics: Theory and Application. Academic Book Center.
3. Alotaibi, Y. (2004). Spoken Arabic digits recognizer using recurrent neural networks. Proceedings of the Fourth IEEE International Symposium on Signal Processing and Information Technology, IEEE. 195–199. <https://ieeexplore.ieee.org/document/1433720/>.
4. Amina, M., Ikram, T., & Ikram, B. (2021). Application of artificial intelligence and emotional intelligence in decision making. Knowledge Collections Magazine, 7(1), 1–6.
5. Bhatti, M. W., Wang, Y., & Guan, L. (2004). A neural network approach for human emotion recognition in speech. Proceedings of the 2004 IEEE International Symposium on Circuits and Systems (ISCAS). II–181. <https://ieeexplore.ieee.org/document/1329238>.
6. Bonnet, A. (1985). Artificial Intelligence: Promise and Performance. Prentice-Hall International.
7. Boufenar, C., & Batouche, M. (2017). Investigation on deep learning for off-line handwritten Arabic character recognition using Theano research platform. Proceedings of the 2017 Intelligent Systems and Computer Vision (ISCV). 1-6. <https://ieeexplore.ieee.org/document/8054902>.
8. Buchanan, B. G. (2005). A (very) brief history of artificial intelligence. AI Magazine, 26(4), 53–60.
9. Crockett, L. J. (2019). AI ethics-the thin line between computer simulation and deception. In P. Griffiths & M. N. Kabir (Eds.), ECIAIR, 2019 European conference on the impact of artificial intelligence and robotics. (pp. 80–87). Oxford.
10. Diduh, O. (2009). The effectiveness of Arabic computational linguistics. Al-Athar Magazine, 8, 90–91.
11. He, X., & Deng, L. (2008). Discriminative Learning for Speech Recognition: Theory and Practice. Morgan & Claypool Publishers.
12. Kavar, M. M., & Mazen, K. (1991). Processing arabic text in ARC/INFO. Proceedings of the Eleventh Annual Environmental Systems Research Institute User Conference, Environmental Systems Research Institute, University of California, 483–492.
13. Lane, H., Howard, C., & Hapke, H. (2019). Natural Language Processing in Action: Understanding, Analyzing, and Generating Text with Python. Manning Publications.
14. Al-Arif, A. A. H. (2007). [Making use of computational linguistics in the service of Arabic studies]. Journal of the Jordanian Arabic Language Academy for Arabic, 7(3), 1–34.
15. Al-Moataz, R. B. (2019). A Basic Program for Arabic Texts. Publications of the King Abdullah Bin Abdulaziz International Center for Arabic Language Service.