Science, Education and Innovations in the Context of Modern Problems
Issue 2, Vol. 9, 2026

RESEARCH ARTICLE

The Role of Computational Linguistics in Teaching the Arabic Language -Automated Parsing as a Model

Hammou Naima
Prof.
University M'Hamed Bougara of Boumerdès
Algeria
Email: hammou.naima1981@gmail.com

Issue web link https://imcra-az.org/archive/392-science-education-and-innovations-in-the-context-of-modern-problems-issue-2-vol-9-2026.html

Keywords Computational linguistics, automated parsing, teaching Arabic, Arabic grammar, AI applications.

Abstract

The current era is characterised by significant advancements in information and knowledge brought about by computers, which have become a standard measure of national progress. Computing has become integrated into human life and society as a result of its applications and programs that enable human activities to become more precise, faster, and more proficient. One of the fields that has greatly benefited from computers is language, as it is a tool for human communication and represents the identity of every nation. The Arabic language has therefore achieved a prominent position in this domain. English, the original language of computer development, has become a prestigious global language. Consequently, it is crucial for the Arabic language to attain a similar status. This necessity has given rise to a new field of study: computational linguistics. This discipline offers avenues for advancing the Arabic language. We have therefore chosen to explore the topic of 'Computational Linguistics and the Arabic Language' in our research, with the aim of highlighting the benefits of automated parsing in Arabic teaching and the associated challenges. Thus, we seek to address the following questions: How can artificial intelligence techniques be used to improve the Arabic language? What role does automated parsing play in achieving this goal? And to what extent can automated parsing applications analyse and parse Arabic sentences accurately compared to human parsing?

Citation

Hammou Naima. (2026). The Role of Computational Linguistics in Teaching the Arabic Language -Automated Parsing as a Model. *Science, Education and Innovations in the Context of Modern Problems*, 9(2), 1-8. https://doi.org/10.56334/sei/9.2.42

Received: 26.07.2025 | Accepted: 05.12.2025 | Published: 20.01.2026 (available online)

Introduction:

Computational linguistics is the study of how computers interact with and process natural language. The aim is to develop and improve programs and systems that handle human language effectively and intelligently. Artificial intelligence applications offer tremendous potential for understanding, analysing and learning the grammatical rules of Arabic. One such application is the parsing tool 'Barki', which analyses Arabic sentences. This study will explore the capabilities of Barki in the field of parsing.

1. Computational linguistics:

Computational linguistics is an important branch of applied linguistics that has transformed research into language teaching practices and procedures, driven by technological advancements. Integrating the disciplines of linguistics and computer science, it acts as an interdisciplinary field connecting the mechanical aspects represented by computers. One of the most prominent definitions of computational linguistics is that it is: 'an interdisciplinary field in which one part pertains to linguistics

and language, and the other part is computational and involves translating language into mathematical symbols that computers can understand.' (Al-Anati, 2005) This allows computers to perform many of the linguistic activities carried out by the human mind.

2. The importance of computational linguistics:

The introduction of the Arabic language to computers has become essential in order to keep pace with modernisation and the technological advancements achieved by the West. Arab scholars have developed specific software for the Arabic language and successfully localised a significant proportion of computer functionalities. While they cannot be said to have fully Arabised the entire system, they have made considerable progress and are providing valuable services to the Arabic language and its sciences. This will yield numerous benefits, which can be summarised as follows: (Bachar, 2020)

- Staying updated: It provides comprehensive awareness of new developments in all fields by integrating Arabic into areas dealing with contemporary issues, particularly in the scientific and technological domains.
- Comprehensive collection: It facilitates the comprehensive collection of the Arabic language, as human efforts alone cannot fully achieve this. Collecting its heritage and history will resolve many developmental, terminological and lexicographical issues, and prepare fertile ground for Arab linguists to complete projects.
- Encouraging production: It encourages Arab scholars and researchers to produce contemporary sciences, including computer science, in Arabic, opening up opportunities for Arabic creativity in computing and information technology. Consequently, Arab scholars will be empowered to innovate in Arabic independently of other languages.
- Facilitating language teaching: It eases the process of language teaching with the help of computers, benefiting both native and non-native speakers.
- Enhancing lexicography: Computational linguistics facilitates the work of lexicographers through statistics, classification, organisation, comparison and automated summarisation.

3. The Impact of Computational Linguistics on the Development and Advancement of the Arabic Language.

Artificial intelligence has achieved significant milestones in language processing by integrating with both linguistics and computer science, leading to the emergence of a new branch of knowledge: computational linguistics. This field has influenced the development of the Arabic language by enabling the learning and differentiation of concepts, the comprehension and acquisition of vast amounts of information, the connection of concepts with facts and conclusions, the quick response with appropriate actions, and the identification and correction of errors. (Sawsan Maafa., 2021)

Using technology to nurture and enhance Arabic language skills among learners in the digital age is fundamental to improving the efficacy and accuracy of the learning process. It has assisted learners in the following ways: (Al-Fitriana)

1. Enhancing educational skills: it has developed the educational skills of Arabic language learners and improved the proficiency of teachers who lack experience.
2. Reducing errors: It has minimised errors in teaching Arabic that may be caused by ignorance or inexperience, as artificial intelligence can serve as an expert in this role.
3. Ease of use: The techniques of artificial intelligence can be easily utilised by both teachers and learners, once programmed by specialists, without the need for an expert. Users do not require prior experience to engage with complex technologies as, by nature, most artificial intelligence techniques do not necessitate expert handling. This enables all teachers and learners to benefit from it.

4. The Concept of Automated Parsing:

Automated parsing is a technical process that uses artificial intelligence and natural language processing techniques to analyse linguistic texts and determine the structure of sentences and linguistic expressions within them. Automated parsing, or automated grammar analysis, breaks down a sentence into its constituent parts and identifies the syntactic and functional relationships between them. It assigns a grammatical position to each word, automatically parsing the words in the sentence both syntactically and morphologically. (Lais, 2012, p. 20.)

In other words, automated parsing involves using computational software to analyse sentences and determine the function of each word. This includes identifying nouns, verbs, adjectives and other parts of speech, as well as determining grammatical cases in languages that use inflection, such as Arabic. In other words, the system specifies how words are arranged in a sentence and how they are related to each other grammatically. It also demonstrates how words are combined to form meaningful sentences and the role that each word plays, e.g. subject, object or predicate.

Furthermore, it identifies the grammatical case of each word (e.g. nominative, accusative, genitive or jussive) based on its position in the sentence and its relationship with other words. The system does not solely focus on identifying grammatical cases; it also aids in determining the correct pronunciation and understanding of words in context.

Contemporary dictionaries define automated parsing as "a software program that automatically shapes Arabic texts according to their rules" (Arabic., n.d.). This means that the process is carried out by software specifically designed to perform certain tasks, such as shaping Arabic texts. In Arabic, shaping refers to the addition of diacritical marks above and below letters to determine the correct movement of words (fatha (ً), damma(ٌ), kasra(ٍ), sukun (ْ) etc.). This process is carried out automatically, without the need for direct human intervention, according to the syntactic and morphological rules of the Arabic language that govern the use of these marks.

The correct shaping of Arabic words relies heavily on the context in which they appear. For instance, consider the word "" (the scale) in the verse from Surah Ar-Rahman (The Beneficent): 'And the heaven He raised high, and He set the balance, lest you transgress the balance.' (Surah Ar-Rahman, 55:7–8), is shaped differently based on its position in the sentence.

5. The Role of Parsing in Understanding Texts and Its Impact on Meaning:

We all know that the fabric of letters and the arrangement of words consist of 29 letters, excluding the hamzah, which are interconnected through morphological, syntactic, lexicon, and semantic rules that enable us to understand texts correctly and align with their meanings. This understanding cannot occur without accurately marking the endings of words.

Parsing has received wide attention from the very beginning of the establishment of "the science of grammar" when errors became prevalent among the Arabs, and out of fear that this might lead to the distortion of the Holy Quran, "grammar" emerged during the time of Abu Al-Aswad Al-Du'ali. Numerous examples have been narrated in this regard that reflect the effectiveness and role of grammatical markers in clarifying and making meaning coherent. Among these examples is what Jalaluddin Al-Suyuti (d. 911 AH) mentioned in his book :The Narrated Accounts on the Reasons for Establishing the Science of Arabic: An Arab approached Umar and asked, "Can anyone recite to me what Allah revealed to Muhammad?" A man recited Surah At-Tawbah (The Repentance) to him, saying: Indeed, Allah is Disassociated from the polytheists and His Messenger" (Surah At-Tawbah, 9:3), with 'His Messenger' in the genitive case. The Arab exclaimed, 'Has Allah disassociated Himself from His Messenger? If Allah has disassociated from His Messenger, then I disassociate from him!' Umar (may Allah be pleased with him) heard the Arab's statement, summoned him and asked, 'Do you disassociate from the Messenger of Allah?'

He replied, 'O Commander of the Faithful, when I came to Medina, I knew nothing of the Quran. I asked someone to recite it to me and he read me this Surah of At-Tawbah, saying, 'Indeed, Allah is disassociated from the polytheists and His Messenger.' So I said, 'Has Allah disassociated from His Messenger?' If Allah has disassociated from His Messenger, then I disassociate from him!' Umar said, "It is not like this, O Arab." The Arab asked, 'Then how is it?'

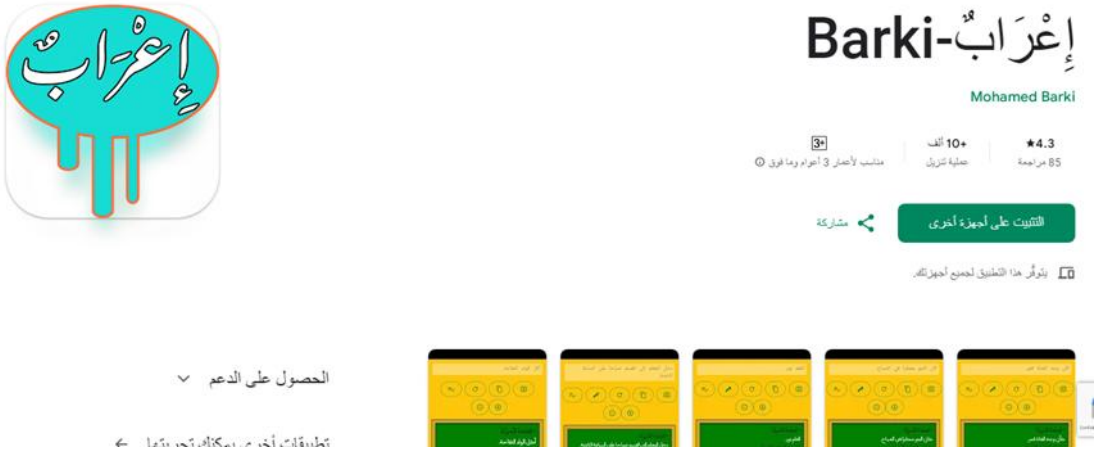
Umar replied, 'Indeed, Allah is disassociated from the polytheists and His Messenger,' with 'His Messenger' in the nominative case. The Arab then said, "By Allah, I disassociate myself from whoever Allah and His Messenger disassociate from." (Al-Suyuti, 1988, p. 30.) Umar commanded that the Qur'an be recited only by those who are knowledgeable in the language.

This clearly shows the important role that parsing plays in clarification and elucidation: meaning cannot be understood without it. Changing the diacritical marks inevitably changes the semantic meaning. This has been affirmed by Ibn Faris, who stated: 'Parsing distinguishes meanings and clarifies speakers' intent. For example, if someone were to say "How beautiful Zayd is!" or "Umar struck Zayd!" without parsing, their intent would not be clear.' (Faris, 1997, p. 143.) In the first example, it is unclear whether the speaker is expressing admiration, denying kindness or asking a question. In the second example, the distinction between subject and object, and who is acting and who is being acted upon, is unclear.

The merit here lies with the grammatical markers; they unlock meanings and their implications. In this context, Abdul Qahir Al-Jurjani states: 'The words are locked to their meanings until parsing opens them and extracts the underlying intent.' (al-Jurjani, 2003, p. 28.) Thus, parsing preserves the Arabic language from error, making it essential for elucidating meaning and understanding texts without ambiguity or confusion.

6. Sentence parsing using the Barki parsing program:

Barki is an educational Android app aimed at simplifying the rules of parsing in the Arabic language. It features colourful, simplified explanations and interactive examples, covering grammar and morphology concepts for various educational levels to facilitate comprehensive learning.



To test the parsing functionality of the application, I provided it with selected literary examples, which it successfully parsed.

- The verbal sentence: The Past Tense

(Model)	Sentence	Words	Parsing
Past Tense	Zayd returned from the outing. (رجع زيد من النزهة)	returned (رجع)	Past tense verb built on the apparent fatha at its end
		zayd (زيد)	Subject in the nominative case, with the marker of elevation being the apparent dammah at its end; it is also possessive
		From(من)	Preposition; a fixed noun in the genitive case as a possessive modifier
		the outing (النزهة)	Adverb of place; an accusative object with the marker of its case being the apparent fatha at its end

Correct:

- (من) :Preposition
- (النزهة) A noun governed by 'with the c ,'من.se marker being the apparent kasrah on the last letter

Commentary:

We note from the parsing provided by this application that it can correctly parse the past tense verb and the subject (doer), but not the preposition or the noun it governs.

- The present tense verb connected with the pronoun of the speaker:

Model	Sentence	Words	Parsing
Present Tense Verb Connected with the Pronoun of the Speaker	It makes me happy that you visit me (يسعدوني أن تزوروني)	يسعدوني	-Present tense verb in the nominative case, with the marker of elevation being the estimated dammah on the letter 'و', prevented from appearing due to difficulty -The connected pronoun (ني) is in the accusative case functioning as a direct object
		أن	The subject is an implied pronoun, interpreted as "you"
		تزور	Particle of negation and source, fixed on sukoon, has no grammatical position The connected pronoun (ني) is in the accusative case functioning as a direct object. The subject is an implied pronoun, interpreted as "you"

From the parsing provided by this programme, we observe that it is incomplete and contains errors in the marker for the present tense verb. Additionally, it lacks the ability to parse grammatically significant compound sentences, as it is limited to parsing individual words.

Correct:

- (يسعد): Present tense verb in the nominative case, with the marker of its elevation being the apparent dammah on its final letter.
- (النون) : For protection (noun of prevention)
- (الياء) A pronoun in the genitive case, functioning as a direct object.

The verbal sentence (أن تزورني)' (that you visit me) serves as the subject in the nominative case.

- The nominal sentence:

Model	Sentences	Words	Parsing
إنّ وأخواتها (Inna and Its Sisters)	- إنّ العلم مفيد Indeed, knowledge is beneficial.	إنّ	A particle of negation, affirmation, and future, built on fatha
		العلم	The noun of "إن" in the accusative case, with the marker of its case being the apparent fatha at its end
		مفيد	The predicate of "إن," in the nominative case, with the marker of elevation being

			the apparent dammah on its end
	- ليت الشباب يعود If only youth would return	ليت	A particle that indicates desire, built on fatha
		الشباب	The noun of "ليت," in the accusative case with the marker being the apparent fatha at its end
		يعود	The predicate of "ليت," in the nominative case with the marker of elevation being the apparent dammah at its end
	- كان الطالب مجتهدا The student was diligent.	كان	An incomplete past tense verb, built on the apparent fatha at its end
		الطالب	The noun of "كان," in the nominative case, with the marker of elevation being the apparent dammah at its end
		مجتهدا	The predicate of "كان ," in the accusative case, with the marker being the estimated fatha on the letter "ا," prevented from appearing due to difficulty
	- ليس العلم صعبا Knowledge is not difficult.	ليس	/

Correct:

1. (إن): A particle resembling a verb that indicates affirmation.

2. (يعود) A present tense verb in the nominative case, with the marker of its elevation being the apparent dammah on the last letter. The subject is an implied pronoun, interpreted as 'he'. The verbal sentence '(يعود) serves as the subject for '(إن).

2. (مجتهدا): The predicate of the verb 'كان' in the accusative case, with the case marker being the apparent fatha on the last letter.

3. '(ليس): A past incomplete verb built on the fatha.

(العلم): The noun form of 'ليس', in the nominative case, with the apparent dammah on its final letter indicating elevation.

(صعبا): Predicate of (ليس), in the accusative case, with the case marker being the apparent fatha on the last letter.

We note that the program sometimes lacks the ability to accurately parse the apparent markers, misinterpreting them as estimated instead. Furthermore, it is sometimes unable to provide detailed parsing, instead offering a basic full parsing. As for the defective verb 'ليس', which belongs to a group of verbs similar to '(كان), the programme does not recognise it at all and thus lacks the capacity to parse it.

- Parsing of the Defective, Incomplete, and Extended Noun According to the Barki Application:

Model	Sentence	Word	Parsing
Defective Noun	(هذي عصاي) This is my staff		The particle 'الهاء' is used for indication.. This (ذه): a demonstrative pronoun in the nominative case, serving as the subject. 'My rod'عصاي: a predicate in the nominative case, indicated by the visible tanween.
Incomplete Noun	(دخل القاضي) The judge entered.		'Entered' (دخل): a past tense verb built on the apparent fatha at the end. 'The judge' القاضي: a subject in the nominative case; the sign of its nominative case is the estimated dammah on the ya, which does not appear due to heaviness.
Extended Noun	(رفعت السماء) The sky was raised.		'Raised' (رفعت): a past tense verb in the passive voice, built on the apparent fatha at the end. 'The sky السماء: a noun acting as the agent in the nominative case, with

			the visible damma at the end indicating this case.

Correct:

(عصاي): Predicate in the nominative case, with the marker of elevation being the estimated dammah on the letter 'a', which is prevented from appearing due to difficulty. (الياء) A connected pronoun in the genitive case acting as a possessive modifier. From the previous examples, we can see that, despite some deficiencies in parsing, the program is capable of parsing the defective noun, incomplete noun and extended noun. However, we also note that the program cannot recognise the estimated marker in the defective noun, incorrectly identifying it as apparent, whereas it does recognise it in the incomplete noun.

Conclusion:

This study concluded that using modern technology and artificial intelligence programmes to teach Arabic, such as the automated parsing programme 'Barki', is an important step towards improving the quality of education and enhancing users' skills.

The Barki application can assist Arab learners to some extent in the area of parsing.

However, the program has deficiencies in parsing: it lacks the ability to parse grammatically significant sentences, and contains some parsing errors. It cannot fully recognise '(ليس), which belongs to a group of verbs similar to '(كان), indicating the need to work on overcoming these deficiencies and improving and developing the programme to meet learners' needs in the parsing domain.

The program benefits users by facilitating the parsing process for Arabic readers, saving them time and effort.

Despite the services computers offer to the Arabic language, the human mind still surpasses them in terms of parsing accuracy.

Ethical Considerations

This research was conducted in accordance with recognized ethical standards in linguistics and educational technology research. The study is based on theoretical analysis, methodological discussion, and the examination of existing computational tools and applications. It does not involve human participants, personal data, or experimental interventions. Therefore, no ethical approval was required. The author adhered to principles of academic integrity, transparency, and proper attribution of sources throughout the study.

Acknowledgement

The author would like to express sincere appreciation to colleagues at the University M'Hamed Bougara of Boumerdès for their academic support and constructive discussions, which contributed to the development and refinement of this research.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of Interest

The author declares that there is no conflict of interest regarding the publication of this article.

References:

- Abdul Qahir al-Jurjani. The Proofs of Miraculous Nature in the Science of Meaning, Al-Maktabah Al-Asriyyah, no edition, Beirut, 2003.
- Abu Al-Hussain Ahmad Ibn Faris. The Companion in the Jurisprudence of the Arabic Language and Its Issues, and the Customs of the Arabs in Their Speech, edited by Omar Farouk Al-Tabba, Al-Maaref Library, 1^(st) edition, Beirut, 1997.
- Ibrahim Bachar. 'Computational Linguistics: Western Foundations and Arab Reception'. Journal of Arabic Language Sciences and Literature, Mohammed Khider University, Biskra, Algeria, Vol. 12, Issue 2, 2020.
- Jalal Al-Din Al-Suyuti. Narrated Accounts on the Reasons for Establishing the Science of Arabic, edited by Marwan Al-Attiyeh, Dar Al-Hijra, 1st edition, Beirut, 1988.
- Joseph Tatius Lais. Informatics, Language, Literature and Civilisation, 1st edition, Modern Publishing Foundation, Beirut, Lebanon, 2012.
- Riyadh Dictionary of Contemporary Arabic. [https://dictionary.ksaa.gov.sa].
- Saif Al-Fitriana. 'The Impact of Artificial Intelligence on Teaching the Arabic Language', Department of Arabic Language and Literature, College of Humanities and Social Sciences, King Saud University, Kingdom of Saudi Arabia.
- Sawsan Maafa. Techniques for Automated Processing of the Arabic Language: A Comparative Study of Models from Arabic Syntactic Analyzers', Thesis submitted for the Doctorate, University of May 8, 1945, Qalmah, Amiar Al-Ayashi, 2021.
- Walid Ahmed Al-Anati. Computational Arabic Linguistics: Concept, Applications, and Feasibility'. Zarqa Journal for Research and Studies, Vol. 7, No. 2, 2005.