



Universidad Nacional de Catamarca - Facultad de Humanidades
Instituto de Investigación en Teorías del Arte y Estética
Revista Rigel – ISSN 2525-1945.

The Artistic and Aesthetic Values in Algerian Folk Proverbs: An Analytical and Descriptive Study

Hanou Soumia¹, Hanaa Belabase², Abdlkader Makhdoua³

¹University of Djilali Bounaama, Khemis Miliana, Algeria
Somia.hanou@yahoo.com

²University of Djilali Bounaama, Khemis Miliana, Algeria
h.belabbas@univ-dbkm.dz

³University of Djilali Bounaama, Khemis Miliana, Algeria
mekhdouaabd@gmail.com

Received: 27/06/2026

Accepted: 11/09/2026

Published: 07/10/2026

Abstract

The world is experiencing an unprecedented digital revolution. This contemporary era is characterized by rapid change and transformation, which have affected various areas and fields of life. No sector has remained untouched by the impact of this revolution, whether positively or negatively. The Arabic language, too, has not been immune to these developments. Rather, the relationship between Arabic and modern technologies, including digital content, electronic publishing, and artificial intelligence, has become a pressing topic that has imposed itself on the research agenda, prompting scholars to address it from various perspectives with the aim of serving and developing the Arabic language. This study seeks to examine the current state of the Arabic language and the extent to which it keeps pace with the requirements of the contemporary era. It demonstrates that Arabic has been influenced by the technological advances, techniques, and programs developed through modern science, which have significantly contributed to enhancing the readability and accessibility of the Arabic language. In this regard, laboratories, councils, and language academies are working to integrate and employ modern media and technologies in the service of Arabic, with the aim of transforming it into a scientific and global language.

Keywords: Language; Arabic Language; Digitalization; Artificial Intelligence.



1. Introduction

The Arabic language is considered a symbol of pride and identity. It is one of the oldest living Semitic languages in the world. Arabs use it for communication and for expressing their purposes and intentions. For Muslims, Arabic is also an essential tool for performing religious rituals and obligations, as they recite Qur'anic chapters during obligatory and voluntary prayers, which reinforces the use of Arabic. Its use, however, is not limited to Arabs and Muslims, as it is also the liturgical language of many Christian churches, particularly among Eastern Orthodox communities and others.

This research paper proceeds from the following central research problem: How can modern technologies be directed toward serving the Arabic language? This main question gives rise to several sub-questions: What is the current state of Arabic language use? What is the future of the Arabic language in light of the accelerating technological boom? To what extent has Arabic benefited from modern technologies? And how has artificial intelligence been employed in the fields of linguistics and Arabic language studies?

1. Definition of Terms and Concepts

1.1 Definition of Language

1.1.1 Linguistic Definition

The term language is defined in Al-Qamus Al-Muhit as follows: "Sounds through which every people express their purposes. Its plural forms are lughat and lughat. Lagha laghwan means to speak, to fail, and tharidatuhu: he narrated it with fat. Alghahu: he disappointed him. Al-laghwa and al-lagha, like al-fata: that which is insignificant and that which is not taken into consideration in speech and other matters. Thus, al-laghwa, like sukra, refers to something that is not taken into account in transactions. 'Allah will not hold you accountable for idle oaths' (Al-Baqarah, 225), meaning that He will not hold you accountable for the sin involved in an oath if you expiate it. Lagha fi qawlihi, like sa'a, da'a, and radiya, means lagha, laghiyah, and malghah: to make a mistake. A laghiyah word means an obscene expression. Al-laghwa: the confused noise of the sandgrouse. Laghiya bihi, like radiya: he became accustomed to it; and with water: he consumed a great deal of it, yet it did not quench his thirst." (al-Fayruzabadi, 2003, p. 1330)

Al-Jawhari, in his dictionary Al-Sihah, states that the origin of the term language is: "Lagha laghwan, meaning something false or vain. It is said: laghawtu bil-yamin and al-lagha: sound. The original form of al-lughah is lughy, meaning laghw, with the letter ha' serving as a substitute. Its plural is lughan, similar to burrah, and it also has the plural form lughat." (Al-Jawhari, 1990, p. 343)

In Lisan al-Arab, under the root, it is stated: "Al-lughah is one of the defective nouns, and its original form is lughwah, derived from lagha meaning 'to speak'. Al-lughah refers to languages, and its definition is that it consists of sounds through which every people express their purposes. It follows the pattern fi'lah, derived from laghawtu, meaning 'I spoke'. Its original form is lughwah. Qillah and kathrah are all forms whose final radical is waw. It has also been said that its origin is laghy or laghw, with the ha' serving as a substitute for the final radical of the verb. Its plural is laghan, similar to burrah or bara, and its plurals also include lughat and lughuun." (Manzur, n.d, p. 252)

Grammarians also defined language, most notably Ibn Jinni in his book Al-Khasa'is, where he stated: "As for its morphological derivation and the identification of its letters, it is fi'lah, derived from laghawtu, meaning 'I spoke'. Its original form is lughwah, whether indefinite or diminutive, and all of its final radicals are waw."

It can be observed that the term language carries several connotations and meanings identified by lexicographers and grammarians. Language may refer to falsehood, which is the opposite of truth; it may also denote error or obscene speech. However, the meaning agreed upon by the majority of scholars is that language refers to sound. It is a system of sounds adopted by individuals as a means of expression and as a tool through which they communicate their intended meanings. The singular form is lugha, while its plural forms include lughat, lughan, and lughuun. (Jinni, n.d, p. 03)

1-1-2. Terminological Definition:

Arab and Western scholars have theorized the concept of language, and consequently, their definitions of this term have varied due to considerations primarily related to the relationship between language and various fields of knowledge. Among the prominent definitions of this concept is the one found in Al-Khasa'is (The Characteristics). According to Ibn Jinni, language is "sounds by which every people expresses their purposes." Thus, language is the human being; it is what distinguishes humans from animals, and it is a result of thought and a product of the intellect. This definition is based on a set of elements upon which language is founded, foremost among them being sounds, which operate according to a precise system. These sounds are used for communication among members of society, thereby fulfilling a social function. They are also employed to express intentions, opinions, and viewpoints, making language an expressive tool. Moreover, language is not uniform throughout the world; rather, languages differ from one region to another. (Jinni, n.d, p. 03)

Among the early scholars who defined language was Ibn Sinan al-Khafaji, who stated that "language is what a people have mutually agreed upon in terms of speech, or it may be divinely ordained." He further divided the matter into two views: some argued that language is an inspiration from God,

citing His saying: "And He taught Adam the names—all of them" (Al-Baqarah, 2:31); while others maintained that language is conventional, meaning that speakers have agreed or reached a convention to designate everything by a particular name. (al-Khafaji, n.d, p. 44) Based on this statement, language can be traced back to its origins in the conventional and consensual theory. The first view attributes the establishment of this noble language to God Almighty, who taught the father of humanity, Prophet Adam, peace be upon him, all the names. The second view attributes its formation to scholars and speakers collectively agreeing upon the names assigned to things through mutual convention.

Ibn Khaldun regards language as a faculty when he defines it as follows: "Know that language, according to common usage, is the expression of the speaker of what he intends. Such expression is a linguistic act arising from an intention to convey speech; therefore, it must become a repeatedly exercised faculty in the organ responsible for it, namely the tongue, and this differs in every community according to its conventions." (Khaldun, n.d, p. 03) This statement reveals a degree of agreement with Ibn Jinni's definition, particularly in considering language as a means of communication among people for expressing needs and purposes. It is also an intellectual activity realized through individual intention and governed by linguistic competence.

This is also emphasized by Ibn Taymiyyah, who states that language is "a means of communication and expression of what a person conceives and feels. It is a vessel for transmitted content, whether its source is revelation, perception, or reason. It is also a means of examining the validity of knowledge and regulating the principles of sound communication." (al-Shaykh, 2002, p. 08) This statement broadens the range of functions attributed to language. It does not limit language merely to communication and expression; rather, it regards it as a means of exchanging knowledge and discerning its validity, as well as a tool for understanding the foundations of proper discourse. Language, therefore, is closely connected with both perception and reason.

We can conclude that, according to the ancient Arab scholars, language is a linguistic faculty and an intellectual activity based on a precise phonological system. It is used among human beings to fulfill various purposes, foremost among which are communication and the expression of intentions, purposes, needs, feelings, and ideas. Language is not uniform throughout the world; rather, it varies according to different societies.

Western linguists have addressed this concept and devoted considerable attention to it. The linguist Ferdinand de Saussure considers language to be "a social product of the faculty of speech, and a set of necessary conventions adopted by a community to enable its members to exercise this faculty... Language is a system of signs, in which each sign is related to the others." This definition presents a

number of characteristics of language. It is a social phenomenon because it is produced by societies and facilitates communication through speech. Saussure also considers language to consist of signs whose combination conveys a particular meaning, which he conceptualizes through the binary opposition between the signifier and the signified, with each sign possessing its own specific meaning.

Bloomfield considers language to be "the speech or specific sounds uttered by human beings under the control of a particular stimulus, which varies according to human groups. Different groups speak different languages... Every child who grows up within a particular human community acquires these speech habits and responses during the early years of life." (al-Najjar, 1988, pp. 23-27) This definition points to an essential characteristic of language, namely that it is acquired rather than inherited. It is an acquired habit resulting from interaction with the members of a particular community rather than with others. Thus, language is both a social phenomenon and a speech habit based on a process of mutual influence and interaction. It consists of symbols that may be heard or spoken and that vary from one people or community to another.

André Martinet also addressed the subject of language and provided a definition that does not differ substantially from the preceding definitions, stating: "It is an instrument of communication, which analyzes human experience differently in each human community, through units possessing semantic content and a sound expression (monemes). This sound expression is, in turn, articulated into distinctive and successive units, the phonemes, whose number is limited in each language." (Zakaria, 1992, p. 28) Language is therefore regarded as a means of communication based on a set of sound units referred to as phonemes and monemes, which perform an expressive and communicative function and carry specific meanings and connotations that vary from one environment to another.

Noam Chomsky argues that language is "an innate faculty in speakers of a language for forming and understanding sentences." (Ramadan, n.d, p. 24) At birth, human beings are endowed with a set of faculties, language being one of them. It is innate and enables every individual to communicate with others. Language is thus an infinite system of sentences and is a distinctive characteristic of human beings.

Modern scholars have also addressed the concept of language and regarded it as a means employed by individuals to communicate and express their thoughts and feelings. It constitutes a system of various signs and symbols, both spoken and heard, and consists of phonological units and an infinite set of sentences referred to as monemes or morphemes.

Definition of the Arabic Language

1-2. Definition of the Arabic Language

1-2-1. Linguistically:

Arabic dictionaries define the term as follows: "It is a word derived from 'araba, ya'ribu, 'araban, meaning to speak eloquently after having previously spoken with difficulty; 'aruba, 'urūban, 'urūbah, 'arābah, and 'urūbiyyah likewise mean to speak eloquently. It is said: 'aruba lisānuhu (his tongue became eloquent). A'raba fulān means that a person was eloquent in Arabic, even if he was not of Arab origin. With regard to speech, it means to clarify it, express it in accordance with the rules of grammar, and apply grammatical rules to it while conveying its intended meaning. It also means to express something clearly and without ambiguity. Concerning another person, it means to speak on his behalf and provide an argument. It is also said: 'arraba 'anhu lisānuhu, meaning that he expressed and articulated it clearly; with regard to speech, it means to clarify it; and with regard to a person, it means to teach him Arabic. An Arabicized foreign word means adapting it to Arabic; and with regard to speech, it means refining it by removing grammatical errors. Ta'arraba means to imitate the Arabs, reside in the desert, and become a Bedouin. It is also said: ta'arraba fulān ba'da al-hijrah (so-and-so became Arabized after the Hijrah). Ista'raba means to become an outsider among the Arabs and to consider oneself one of them." (Mustafa, n.d, p. 195)

The term Arabic is thus associated with the trilateral root 'r-b, which denotes eloquence and conformity to the rules of grammar. This was a characteristic of the genuine Arab, distinguished by the eloquence of his speech. The Arab represents the original, whereas the musta'rib refers to an outsider or foreigner who sought to become one of the Arabs. Accordingly, Arabic may be defined linguistically as a language characterized by the eloquent speech of the Arabs.

1-1-2. Terminologically:

The origins of the Arabic language are traced back to Shem (Sām), the son of Prophet Noah, peace be upon him. It is "one of the Semitic languages, named after Shem, son of Noah, peace be upon him, who settled with his descendants in Western and Southern Asia, particularly in the Arabian Peninsula. Among these Semitic languages are Canaanite, Nabataean, and Ethiopic. Arabic has managed to survive, whereas only traces of those other languages remain, engraved on rocks here and there." (al-Sayyid, n.d, p. 18)

The Arabic language has endured, and its speakers have preserved it through continuous use and circulation. It is considered "one of the largest languages of the Semitic language family in terms of the number of its speakers, and one of the most widely spoken languages in the world. It is spoken by more than four hundred and twenty-two million people, whose speakers are distributed

throughout the region known as the Arab world, in addition to many other regions.” (Anis, 1965, p. 44)

Arabic is spoken in various parts of the world; however, it remains the language of the Arab world, extending from the Atlantic Ocean to the Arabian Gulf. It was spoken by the inhabitants of the Arabian Peninsula and subsequently by the peoples of the Maghreb following the Islamic conquests, which contributed to the expansion of the geographical scope of the Arab world to encompass both the Mashriq (the Arab East) and the Maghreb (the Arab West).

Arabic became closely associated with human beings, serving as a tool for thought even before communication, as it is “a living human characteristic that does not merely express ideas but rather shapes them. Thinking is nothing but silent language, and language generates thought. It is a precise system that requires extensive knowledge and skills, because the process of communication between the speaker and the listener goes through several highly precise stages.” (al-Qara’i, 2006, pp. 01-02) This definition focuses on the relationship between thought and language, as they are interrelated and interconnected. Thought precedes language in the sense that it constitutes a form of language without symbols or sounds; rather, it is a mental process requiring perception and skill, which is subsequently translated into language and communication between the sender and the recipient.

“...whatever conforms to the language of the Arabs and does not depart from the conventions of the people of eloquence and literature. In order to correct it, grammar was established, and in order to compile it, books on language were written, in which its commonly used forms, as well as the anomalous and neglected forms, were mentioned. It is incumbent upon one who grows up among the Arabs to follow their language, not to depart from the entirety of their vocabulary, and not to allow himself to contradict them in such a way that they would regard his speech as erroneous and grammatically incorrect. Grammatical error is whatever contradicts the Arabic language, departs from the usage of its speakers, or violates the principles upon which its grammatical inflection is based. Such error is generally considered objectionable among men of letters, particularly for those who strive to speak with correct grammatical inflection. Whoever uses rare or unfamiliar expressions from the language of grammatical eloquence is considered even more blameworthy. It is narrated that ‘Umar ibn al-Khaṭṭāb, may Allah be pleased with him, used to punish grammatical errors. As for the Arabs, when one of them committed a grammatical error... his status among the people of language declined, and his language was rejected and disregarded.” (Ja’far, 1982, pp. 142–143)

The Arabic language had its guardians who showed no tolerance toward error, mistakes, or slips of the tongue. They applied grammatical rules of nominative, accusative, and genitive inflection

whenever necessary, because they were fully convinced that language is far more than a means of conveying intended meanings; rather, it is a symbol of existence and enduring continuity. The Arabs did not limit their lexical records to words that were merely in common use; they also mentioned anomalous and neglected forms, recognizing that if such neglected forms became established on people's tongues, they would eventually become part of common usage. The same applies to our Arabic language, which thrives through use and withers through neglect. I would not say that it becomes extinct, because God has undertaken to preserve it, as He says in His revealed Book, after saying, "In the name of Allah, the Most Gracious, the Most Merciful": "Indeed, We have sent down the Reminder, and indeed, We will preserve it." (Qur'an, 15:9). Thus, it has been preserved in hearts and in written texts, and together with it, this noble and sacred language has also been preserved.

"The Arabic language possesses its own phonological, morphological, grammatical, and syntactic systems, just as its words have their own specific meanings. Scholars have considered every departure from this integrated linguistic system to constitute a form of grammatical error ,particularly any violation of one of the fundamental elements of its linguistic structure that distinguishes it from other human languages." (Nayif, 1998, p. 25)

1. The Arabic Language and Modern Technologies

Contemporary linguistic research is undergoing a radical strategic transformation that has shifted it from the realm of abstract theoretical reflection to the domain of "instrumental linguistics," where the computer is no longer merely an auxiliary tool, but has become an irreversible methodological necessity for formalizing linguistic theories and implementing them computationally. As Pierre Nugues argues, computational linguistics lies at the heart of formulating mathematical models of linguistic structures, thereby opening up new prospects for the development of linguistic theories that draw upon immense computational capabilities (Bashir Boujjra, 2024, pp. 505–519)

The computational processing of the Arabic language has emerged as a civilizational necessity imposed by the requirements of the digital age. The tremendous speed of processing and the flexibility of accessing massive amounts of data are essential for preserving, disseminating, and facilitating access to the Arab intellectual and knowledge heritage. Nevertheless, a significant research problem arises from the knowledge gap between "language engineering," which computer scientists master, and the profound "linguistic knowledge" possessed by linguists. Bridging this gap requires establishing computational linguistics as a well-grounded and stable interdisciplinary field that transcends the divide between rigid rule-based approaches and statistical modeling, thereby ensuring effective automatic processing capable of restoring Arabic to its position as a leading language in the technological domain (Bashir Boujjra S. &, 2024, pp. 505–519)

2-1. From Traditional Linguistics to Computational Linguistics

Computational linguistics has evolved since the mid-1960s into a scientific field concerned with designing mathematical models that simulate linguistic structures. This field has undergone a pivotal shift toward quantitative approaches (mathematical and statistical methods) as a result of the dominance of big data, which has revealed linguistic patterns that could not be perceived through unaided observation. As John Sinclair points out, "language looks quite different when we look at a lot of it at once," a fact demonstrated by computational processing through its ability to uncover structural linguistic variations that had remained unnoticed within traditional linguistic approaches (Bashir Boujjra S. &, 2024, p. 521)

1-2. Applications of Automatic Processing of the Arabic Language (Linguistic Levels)

Computer systems rely on an organic interconnection between the different levels of linguistic analysis. Morphological analysis provides syntactic analysis with the necessary data, while both contribute to semantic analysis within an integrated framework.

2-2-1. Morphological and Syntactic Analysis:

Applications of morphological tagging and syntactic trees constitute the fundamental basis for understanding sentence structure. In automatic diacritization, a distinction can be made between:

The traditional approach: This approach relies on strict grammatical rules and dictionaries.

The modern approach: This approach employs the Viterbi algorithm and dynamic programming to select the most probable linguistic sequence based on mathematical probabilities derived from diacritized corpora (Atiyah, 2011)

2-2-2. Semantic Analysis and Semantic Networks:

Semantic networks, such as WordNet and ConceptNet, play a vital role in representing relationships between concepts through knowledge-graph structures. These networks contribute to improving the accuracy of machine translation and sentiment analysis by enabling machines to extract the contextual meaning and underlying content of texts (Bashir Boujjra S. &, 2024, pp. 505–519)

1-2. Artificial Intelligence and Large Language Models:

The revolution in Deep Learning has brought about a qualitative leap through artificial neural networks that simulate aspects of human behavior. The accuracy of these networks increases with the number of hidden layers and the size of the training data, as texts are represented through feature vectors, which transform words into mathematical values within a multidimensional space (Bashir Boujjra S. &, 2024, pp. 505–519)

Large Language Models (LLMs), such as ChatGPT, represent a revolution in context understanding and text generation. However, they do not understand language as a “rule-based system” but rather as “complex probabilistic contexts.” The relationship between these models and the Arabic language remains governed by a “pragmatic relationship,” in which the availability of digital resources determines the model’s success. It is a reciprocal relationship whose centers of gravity shift according to the extent to which each party benefits from the other within the digital space .(Ben Omar, 2025, pp. 123–142)

2-4. Linguistic and Technical Challenges in Arabic Language Computing

Arabic is considered a “morphologically rich and computationally challenging” language, and it faces a number of specific structural challenges:

Absence of diacritics: Ambiguity in words such as complicates statistical and semantic processing . (Rashwan, 2019)

Dialectal variation: The gap between Modern Standard Arabic (MSA) and colloquial dialects, coupled with the absence of standardized writing conventions for the latter, limits the efficiency of unified language-processing systems.

Orthographic and graphemic challenges: Arabic requires more than 190 graphemes, compared with approximately 80 graphemes in English. The problem of “segmentation and recognition” is also particularly significant: accurate recognition requires the identification of boundaries, while boundary identification requires knowledge of the symbols. This creates a circular problem for the system, particularly because of the “horizontal connectivity” and “vertical overlap” characteristic of Arabic script (Attia & Rashwan, 2011).

Data scarcity: The limited availability of annotated linguistic corpora hinders the training of modern neural models compared with European languages.

2-5. Institutional Role and Available Linguistic Resources

Institutional efforts constitute a fundamental driving force behind the development of Arabic computational resources.

King Salman Global Academy for the Arabic Language: It leads initiatives under the “Global Research Track” aimed at supporting Arabic language computing and enriching technical content through rigorous interdisciplinary studies. (Rashwan, 2019)

Available Resources:

The Arabic Thesaurus Project: A comprehensive textual repository designed to facilitate the computational processing of both Arabic heritage and contemporary Arabic texts (Bashir Boujjra S. &, 2024)

The Qur'anic Corpus: It provides three levels of analysis—morphological, syntactic, and semantic—along with highly precise semantic ontologies, owing to the distinctive characteristics of the Qur'anic text.

Arabic ICALL System: Developed at the Faculty of Computers and Information at Cairo University, this system is based on a "rule-based approach" to provide detailed analysis of learners' errors and pedagogically oriented correction (Atiyah, 2011)

Conclusion:

This study concludes that Arabic Natural Language Processing (NLP) has witnessed significant development through theoretical and applied frameworks established by pioneers of computational linguistics, notably Nabil Ali, Nihad Al-Mousa, and Salwa Hamada. Their research has focused on morphological and syntactic analysis, machine translation, indexing, and information extraction. This technical achievement has been further strengthened by the emergence of modern Large Language Models (LLMs), such as ChatGPT and Gemini, as well as specialized Arabic language models such as Jais and ALLaM, which have brought about a qualitative transformation in the understanding, generation, and intelligent processing of Arabic texts.

However, anticipating the digital future of the Arabic language continues to face structural and technological challenges. These include the intrinsic characteristics of Arabic, such as the absence of diacritics in the majority of digital texts, the complexity of its root-and-pattern derivational system, and the context-dependent multiplicity of meanings. These challenges are compounded by the gap between Modern Standard Arabic and the considerable diversity of local dialects, as well as the scarcity of well-documented linguistic corpora. Within this context, institutional initiatives led by the King Salman Global Academy for the Arabic Language and the Arab League Educational, Cultural and Scientific Organization (ALECSO) contribute to the development of digital linguistic resources and the establishment of ethical frameworks for scientific research and artificial intelligence.¹ Accordingly, advancing the Arabic language in the digital sphere requires a transition toward hybrid models that integrate linguistic rules with deep statistical learning, alongside the development of open-source corpora, in order to preserve linguistic identity and consolidate the position of Arabic within the global technological landscape.

The future of the Arabic language depends largely on the extent to which it responds to the digital revolution and transitions from being a language that merely consumes technology to one that actively produces formalized knowledge. To achieve this, we recommend the following:

Interdisciplinary Training: There is a need to train a new generation of “linguistic engineers” who combine language engineering with the foundational sciences of linguistics.

Development of Standards: Efforts in digitization and linguistic annotation should be harmonized in order to build large-scale annotated corpora.

Investment in Content: The data gap should be addressed to enhance the performance of Large Language Models (LLMs) and ensure accurate representation of the Arabic linguistic identity.

Conclusion

This study concludes that the automatic processing of the Arabic language has witnessed significant development through theoretical and applied frameworks established by pioneers in computational linguistics, particularly Nabil Ali, Nihad Al-Mousa, and Salwa Hamada. Their research has focused on morphological and syntactic analysis, machine translation, indexing, and extractive summarization. This technological achievement has been further strengthened by the emergence of modern large language models such as ChatGPT and Gemini, as well as specialized Arabic language models such as Jais and ALLaM. These models have brought about a qualitative transformation in the understanding, generation, and intelligent processing of Arabic texts.

However, anticipating the digital future of the Arabic language continues to face structural and technological challenges. These include intrinsic characteristics of Arabic, such as the absence of diacritics in the majority of digital texts, the complexity of the root-and-pattern derivational system, and context-dependent polysemy. Other challenges include the gap between Modern Standard Arabic and the wide diversity of local dialects, as well as the scarcity of well-documented linguistic corpora. In this context, institutional initiatives led by the King Salman Global Academy for the Arabic Language and the Arab League Educational, Cultural and Scientific Organization (ALECSO) contribute to the development of digital linguistic resources and the establishment of ethical frameworks for scientific research and artificial intelligence.

Accordingly, advancing the Arabic language in the digital sphere requires a transition toward hybrid models that combine linguistic rules with deep statistical learning, as well as the development of open-source corpora. Such efforts would contribute to preserving linguistic identity and strengthening the position of Arabic within the global technological landscape.

The future of the Arabic language depends largely on the extent to which it responds to the digital revolution and transitions from being a language that merely consumes technology to one that actively produces formalized knowledge. To achieve this objective, the following recommendations are proposed:

Interdisciplinary Training: There is a need to train a new generation of “linguistic engineers” who combine language engineering with a solid foundation in linguistic sciences.

Development of Standards: Efforts in digitization and linguistic annotation should be unified in order to build large-scale, annotated linguistic corpora.

Investment in Content: The data gap should be addressed in order to enhance the performance of large language models and ensure an accurate representation of the Arabic linguistic identity.

FOOTNOTES

-The Holy Qur'an

1-Majd al-Din Abu Tahir Muhammad ibn Ya'qub al-Fayruzabadi: *Al-Qamus al-Muhit* [The Comprehensive Dictionary], Egyptian General Book Organization, n.p., Egypt, 2003, p. 1330.

2-Al-Jawhari: *Taj al-Lughah wa Sahih al-'Arabiyyah* [The Crown of Language and the Correct Arabic], ed. Ahmad 'Abd al-Ghafur 'Attar, Dar al-'Ilm lil-Malayin, 1st ed., 1990, p. 343.

3-Abu al-Fadl Jamal al-Din ibn Makram Ibn Manzur: *Lisan al-'Arab* [The Arabic Language], Dar Sadir, 4th ed., Beirut, Lebanon, Vol. 1, p. 252.

4-Abu al-Fath 'Uthman ibn Jinni: *Al-Khasa'is* [The Characteristics], ed. Muhammad 'Ali al-Najjar, 'Alam al-Kutub, n.p., n.d., Vol. 1, p. 3.

5-Ibid., same page.

6-Abu Muhammad ibn Sa'id ibn Sinan al-Khafaji: *Sirr al-Fasahah* [The Secret of Eloquence], Al-Riwaq website, n.p., Cairo, n.d., Vol. 1, p. 44.

7-Abd al-Rahman ibn Khaldun: *Al-Muqaddimah* [The Introduction], ed. Khalil Shahadah, Dar al-Fikr, n.p., Beirut, Lebanon, n.d., Vol. 1, p. 3.

8-Abd al-Salam Ahmad al-Shaykh: *Al-Lughawiyat al-'Ammah* [General Linguistics], International Islamic University, Dar al-Tajdid, 2nd ed., n.p., 2002, p. 8.

9-Nadia Ramadan al-Najjar: *Al-Lughah wa 'Ilm al-Lughah Qadiman wa Hadithan* [Language and Linguistics: Past and Present], cited in: Ferdinand de Saussure, *'Ilm al-Lughah* [Course in General Linguistics], trans. Malik al-Mutallib, Bayt al-Mawsil, n.p., 1988, p. 27.

10-Ibid., p. 23.

11-Michel Zakaria: Buhuth Alsuniyyah 'Arabiyyah [Arabic Linguistic Studies], University Foundation for Studies and Publishing, 1st ed., Lebanon, Beirut, 1992, p. 28.

12-Nadia Ramadan al-Najjar: Al-Lughah wa 'Ilm al-Lughah Qadiman wa Hadithan [Language and Linguistics: Past and Present], cited in: Ferdinand de Saussure, 'Ilm al-Lughah [Course in General Linguistics], op. cit., p. 24.

13-Ibrahim Mustafa et al.: Al-Mu'jam al-Wasit [The Intermediate Dictionary], supervised by 'Abd al-Salam Harun, Arabic Language Academy, n.p., Misr Press, Egypt, 1380 AH/1960, Vol. 2, Letter 'Ayn, p. 195.

14-Anim Karam al-Sayyid: Al-Lughah al-'Arabiyyah wa al-Sahwah al-'Ilmiyyah al-Hadithah [The Arabic Language and the Modern Scientific Awakening], Sinai Library, n.p., Egypt, n.d., p. 1.

15-Ibrahim Anis: Fi al-Lahajat al-'Arabiyyah [On Arabic Dialects], Anglo Library, n.p., Cairo, 1965, p. 44.

16-Ahmad Yusuf al-Qara'i: "Lughatuna al-'Arabiyyah Wi'a' Thaqafatina" [Our Arabic Language as a Vessel of Our Culture], Al-Thawrah Magazine, Yemen, Sana'a, 2006, pp. 1–2.

17-Qudamah ibn Ja'far: Naqd al-Shi'r [The Criticism of Poetry], Dar al-Kutub al-'Ilmiyyah, n.p., Beirut, 1982, pp. 142–143.

18-Ma'ruf Nayif: Khasa'is al-Lughah al-'Arabiyyah wa Tara'iq Tadrisiha [Characteristics of the Arabic Language and Methods of Teaching It], Dar al-Nafa'is, 5th ed., Lebanon, Beirut, 1998, p. 25.

19-Bashir Boujjra, S., & Salmi, A.: "Computational Linguistics: Between Modern Techniques for Automatic Language Processing and Traditional Linguistic Practices," Aleph Journal, 11(3-1), 2024, pp. 505–519.

20-Bashir Boujjra, S., & Salmi, A.: "Computational Linguistics: Between Modern Techniques for Automatic Language Processing and Traditional Linguistic Practices," Aleph Journal, 11(3-1), 2024, pp. 505–519.

21-Al-Fifi, A. B. Y.: Introduction to Computational Linguistics, King Salman Global Academy for Arabic Language, 2017.

22-Bashir Boujjra, S., & Salmi, A.: "Computational Linguistics: Between Modern Techniques for Automatic Language Processing and Traditional Linguistic Practices," Aleph Journal, 11(3-1), 2024, pp. 505–519.

- 23-Atiyah, M., & Rashwan, M.: Basic Applications in Arabic Language Processing, King Salman Global Academy for Arabic Language, 2011.
- 24-Bashir Boujjra, S., & Salmi, A.: "Computational Linguistics: Between Modern Techniques for Automatic Language Processing and Traditional Linguistic Practices," Aleph Journal, 11(3-1), 2024, pp. 505–519.
- 25-Bashir Boujjra, S., & Salmi, A.: "Computational Linguistics: Between Modern Techniques for Automatic Language Processing and Traditional Linguistic Practices," Aleph Journal, 11(3-1), 2024, pp. 505–519.
- 26-Atiyah, M., & Rashwan, M.: Basic Applications in Arabic Language Processing, King Salman Global Academy for Arabic Language, 2011.
- 27-Ben Omar, W., & Zagagh, A.: "Artificial Intelligence Language Models and the Arabic Language," Al-Muqaddimah Journal for Humanities and Social Studies, 10(2), 2025, pp. 123–142.
- 28-Rashwan, M., et al.: Computational Linguistic Resources, King Salman Global Academy for Arabic Language, 2019.
- 29-Atiyah, M., & Rashwan, M.: Basic Applications in Arabic Language Processing, King Salman Global Academy for Arabic Language, 2011.
- 30-Rashwan, M., et al.: Computational Linguistic Resources, King Salman Global Academy for Arabic Language, 2019.
- 31-Bashir Boujjra, S., & Salmi, A.: "Computational Linguistics: Between Modern Techniques for Automatic Language Processing and Traditional Linguistic Practices," Aleph Journal, 11(3-1), 2024, pp. 505–519.
- 32-Atiyah, M., & Rashwan, M.: Basic Applications in Arabic Language Processing, King Salman Global Academy for Arabic Language, 2011.