

**The Role of Computational Linguistics in Making the Contemporary
School Dictionary**

Dr. Fatma Saadi^{1*}.

¹ Scientific and Technical Research Center for the Development of Arabic Language,
The Research Unit of Tlemcen, (Algeria) saadifatima14@gmail.com

Date of submission: 01/10/2025

Date of acceptance: 10/12/2025

Date of publication: 27/12/2025

Abstract:

This research seeks to present mechanisms for employing computational linguistics in the creation of the electronic dictionary, using the school dictionary as a model, in response to the educational field's need for modern digital techniques. By adopting a descriptive-analytical approach, introducing the concepts and explanations of the study's terms, and presenting two models of Arabic electronic dictionaries, we conclude that it is possible to develop this lexicographic field through the use of the latest computational methods in selecting and standardising educational material.

Key words: dictionary, school dictionary, electronic dictionary, computational linguistics.

Introduction:

In the Arab world, education has undergone a series of rapid technological transformations that cannot be ignored, especially with the use of computers in the teaching and learning process. Computational linguistics, alongside the digital revolution, has contributed to a qualitative shift in dictionary-making, whether through discs, applications in computers and smartphones, or by being available on the Internet. Thus, this renders contemporary education closely linked to modern digital technology.

In the modern era, computational linguistics, technological advancement, and digitisation have contributed to the development of dictionaries. It is an important branch in applied linguistics, and it is the outcome of what researchers in linguistics have achieved by using computer techniques in human language. This has opened scientific horizons for automatic processing to adapt and facilitate its use by both native and non-native speakers alike. The use of computers in linguistic research helps keep pace with contemporary linguistic challenges, particularly in the field of electronic dictionary-making, and the transition from printed dictionaries, which no longer address the needs of specialists and researchers, to computerised electronic dictionary.

* Corresponding author.

Likewise, the Arabic language has been influenced by this technological transformation, seeking to keep pace with civilised progress and rise to the level of the digital age. Hence, the computer has become an accompanying tool in its various manifestations. It contributes to the development of its instruments and teaching methods. In this framework, computational linguistics aims to employ computer potentials and artificial intelligence to digitise language through specialised software and applications that process language automatically. Consequently, an urgent need has emerged to prepare specialised electronic school dictionaries that meet the needs of learners across different ages and levels. School dictionaries are important pedagogical tools that enrich the child's linguistic and cognitive repertoire, help them understand ambiguous terms, and overcome difficulties in reading and comprehension. They also shape the learner's personality and develop their linguistic and cultural behaviour.

This study sheds light on the role of computational linguistics in developing school dictionaries. It examines the fundamental theoretical concepts and explores the possibilities of employing modern information and communication technologies in practical applications that serve the making of the school dictionary. It addresses a key question: to what extent can computational linguistics develop the school dictionary and make it an effective digital model for learners?

I- Terminological and Conceptual Approaches

This part presents the definitions of common terms related to computational linguistics, computing, and automatic language processing.

1- Computational Linguistics

1-1- Definition of Computational Linguistics

Computational linguistics has been designated by different terms as a result of terminological overlap, particularly related to translation and Arabization. Abdelwahed Dkiki states: *"Neither individual nor institutional Arabic linguistic research has agreed upon a unified definition of computational linguistics. It has been referred to by many names, including automatic, computational, computer, and algorithmic linguistics, and other Arabic terms that encompassed this new linguistic concept introduced into Arab culture"* (Dkiki, 2014, p. 76).

This terminological multiplicity led to a variety of definitions for computational linguistics from the Arab linguists' perspective. Some of them have considered it an interdisciplinary system based on data, with half related to language and the other half to the computer (Al-Anati, Djabr, 2007, p. 23). Others viewed it as an auxiliary tool for handling language via the computer. Still others argued that it is concerned with the computer using natural language, that is, the automatic programming of natural language in the computer.

At its first international conference in 1965, computational linguistics was defined as a *new science which linguistics intersects with a formal device produced by mathematical logic. It is subject to the constraints of machines designed for the*

automatic processing of information, and in which research leads to the creation of an algorithmic model.

David Crystal defined computational linguistics as *“a branch of linguistic studies in which computational techniques are employed to clarify linguistic and phonetic problems.”* (Bekal, 2020, p. 73) Abd Al-Rahman Ben Hassan Al-Aref defined this science briefly, *“it is the science that studies human language as a pliable tool to be processed by the machine. Its principles are general linguistics with all its analytical levels: phonetics, morphology, syntax, and semantics, together with computer science, artificial intelligence, logic, and mathematics.”* (Abd Al-Rahman, 2014, p. 4)

This convergence of knowledge among scientific disciplines, in addition to the researchers' efforts on the project of Arabic language computationalization and the optimal use of linguistic data in applied fields of automatic processing (Abd Al-Rahman, 2014, p. 43), allows computational linguistics to emerge as the scientific study of the linguistic system at all its levels through a computational lens. Its goal is the application of computational models to linguistic competence. (Aied, 2018, p. 25)

Based on the previous definitions, it is clear that computational linguistics is the scientific study of natural languages. These definitions accurately explain the tasks and objectives of computational linguistics, which researchers strive to achieve by all possible means, both linguistic and non-linguistic, since its fundamental mission is the computational modelling of language in its various domains.

1-2- The Foundations of Computational Linguistics

Computational linguistics rests on two complementary pillars: the theoretical and the applied components. None of them can function independently of the other. The theoretical component is concerned with issues in theoretical linguistics and addresses formal theories of linguistic knowledge that humans need to generate and comprehend the language. It also investigates how the electronic brain operates to solve linguistic problems such as machine translation, automatic spelling correction, and automated text diacritization. (Nihad, 2000, p. 54) The applied component is concerned with *“the practical outcome of modelling human language use, and aims to produce programs with knowledge of human language. Such programs are needed to improve human-machine interaction, since the main obstacle in this interaction lies in communication”* (Nihad, 2000, p. 53).

Representing human knowledge in computers is similar to *“representing linguistic knowledge in the human brain. Therefore, if we succeed in identifying how the human brain operates, we can then program computers in a way analogous to the brain's processes,”* (Lazreg, Rabah , 2014, p. 5) thereby enabling a precise description of linguistic knowledge. Thus, computational linguistics is based on a theoretical conception of the computer as if it were a human mind. It seeks to uncover the mental and cognitive processes of the human brain when it produces, understands, and perceives language. However, it acknowledges that the computer is a deaf machine with no creative capacity. For this reason, the linguistic system must be precisely

described to the computer, covering the linguistic problems that humans intuitively grasp, to reach the desired outcome.

Language computationalization is based on three essential technologies: text, image, and speech processing, such as machine translation, automatic handwriting recognition, and converting written text into spoken language. In computational linguistics, these constitute the central concept in language production (Bekal, 2020, p. 3). Within this framework, it can be noted that computational linguistics relies on planning, organisation, and programming.

1-3- Artificial Intelligence and Automatic Language Processing

Automatic processing is the automated application of a corpus of texts by transforming them to generate something new on their basis. These applications may vary across linguistic levels, morphology, phonology, syntax, semantics, and lexis, as well as in the development of machine translation applications, whether written or spoken. It is also the automatic processes performed by machines, as opposed to those carried out by humans. The machine in question is the computer, which was originally invented to perform numerical computations. (Chacha, 2008, p. 13) Automatic processing studies the problems of the automatic generation and comprehension of natural human languages. Natural language generation systems aim to convert the data stored in computer databases into human language that appears natural.

Automatic processing falls within the computational applications of artificial intelligence. This term officially appeared for the first time in 1969 during the first conference on Artificial Intelligence in Washington, although *John McCarthy* had already used it in 1956 at the Massachusetts Institute of Technology. Studies in this field are concerned with two aspects: reproducing human intelligence through the computer, and expanding the computer's capacity by granting it a degree of inference and autonomous action in various fields of knowledge, such as physics, mathematics, engineering, and others (Alain, 2017, p. 11).

Based on what has already been mentioned, it is clear that artificial intelligence is a simulation of natural human intelligence. However, the former is performed by machines through engineering processes, automatic translation, natural language processing, and recognition of statistical phenomena. It assists researchers in overcoming obstacles and solving problems in their computational studies.

II -The Computational Making of the School Dictionary

1- Electronic Dictionaries

The use of computers in the making and classification of dictionaries is considered an ideal phenomenon of linguistic processing. Linguistic studies have achieved tangible progress; their results appear in automated dictionaries used by researchers and learners. The computationalisation of Arabic dictionaries facilitates access to the Arabic lexical repertoire, storing it in ready-made digital databases to serve intended dictionary purposes.

1-2- Definition of the Electronic Dictionary

The term *electronic dictionary* emerged recently when specialists realised the necessity of paying attention to learners and providing them with special dictionaries. Researchers have defined it variously. Some stated that it is: “*a computerised version adapted from the printed edition. It consists of many entries, each containing the information that can be collected about it. This information varies from one dictionary to another, depending on its objectives and users.*” (Lazreg, Rabah , 2014, p. 54) According to Aze-Eddine Bouchikhi, it is “*the product of applying electronics and computer sciences to the field of dictionary making.*” (Aze-Eddine , 2004, p. 22) Accordingly, the electronic dictionary is an automated database of linguistic or non-linguistic entries belonging to a specific field of knowledge.

1-3- Definition and Importance of the Electronic School Dictionary

It is a list of words arranged alphabetically, drawn from the pupil’s actual use of language, accompanied by definitions suited to their level and needs, and made accessible via the Internet. (Bekal, 2020, p. 6) The digital school dictionary is a computerized linguistic corpus and a repository of vocabulary that contains a large stock of words, accompanied by explanations and definitions, and often supported with images, examples, and illustrations. It is arranged in a specific order and made accessible online, in the form of compact discs, or integrated into computers, smartphones, and other devices.

The electronic school dictionary constitutes an important resource from which learners acquire terminology to develop their linguistic repertoire and knowledge base. It enables them to keep pace with modern technological developments and benefits them by making research easier and faster. Moreover, it has many advantages that highlight the importance of dictionary authorship as an educational and pedagogical tool for both teachers and learners, contributing to the improvement of the teaching–learning process and addressing learners’ linguistic and non-linguistic needs. (Mohamed, Djamila , 2021, p. 2)

According to specialists, this aspect of language education should not be neglected, as it enriches the learner’s linguistic and cognitive repertoire, shapes their personality in its behavioural, psychological, and motor dimensions, and fosters communicative linguistic competence between teachers and their students.

2-Foundations and Standards for Making the Electronic School Dictionary

2-1- Linguistic and Technical Standards

The electronic school dictionary constitutes an indispensable source for studying the productivity of language and the interrelations of lexical items. (Ben Yechou, 2015, p. 62) Today, informatics is a more precise science than many others, which is why the advancement of the electronic school dictionary has become a fundamental priority within the educational system. Technical means have led to the expansion of knowledge and have transformed how information is received; the emergence of the computer has triggered a digital revolution in the field of education. The making of electronic school dictionaries requires a set of standards and foundations upon which it is built. To this end, some researchers have proposed the following key steps:

-Developing a morphological analyser based on a lexical database drawn from the dictionary. This analyser serves as the foundation for a knowledge base of formal grammatical rules. Practically, this database is established upon a lexicon of simple words, which in turn are derived from an Arabic root database constructed from traditional Arabic dictionaries.

-Developing a syntactic analyser that is founded on a database of correct linguistic structures, since linguistic forms composed of sequences of lexical items constitute the meaningful units in a text. (Lazreg, Rabah , 2014, p. 10)

-Cataloguing all the original materials that make up the words, along with all derived patterns, while specifying the frequency of each one. They are arranged and classified starting from the word itself in a specific order, and from the concept and the semantic fields, and according to the science or art to which they belong. (Ben Yechou, 2015, p. 65)

-Avoiding linguistic redundancy, such as the overuse of synonyms, and excluding what Al-Hadj Saleh calls “dead elements,” which no longer correspond to the learner’s evolving linguistic reality or the demands imposed by contemporary changes.

-Relying on the educational linguistic repertoire in use, which necessarily responds to the learner’s needs, particularly the new ones in their present context, while also emphasising the need for integration between the selected corpus in the educational linguistic field and that of the dictionary.

-Ensuring the electronic school dictionary matches the learner’s cognitive abilities. It must be sufficient to cover their needs without being so limited as to fall short of the required learning outcomes.

-Adopting a clear methodology in presenting terminologies and definitions, ensuring that the lexical repertoire has a direct link with the learner, meets their needs and interests, and supports them in overcoming their linguistic challenges.

-Appending the dictionary with grammatical and morphological supplements at the end, so that learners can refer to them for conjugation and grammatical rules.

2-2- Prospects for the Computerisation of the School Dictionary

The computerisation of the school dictionary is a significant area in computational linguistics. Scholars strive to adapt language study to the computer and to enable the latter to understand and analyse it. Technological advances have encompassed every aspect of social life, including schooling. Globalisation and the communication and information revolution have redirected the technical tools used in education toward what is known as *informational methodology* in addressing educational issues.

To enable educational curricula to keep up with scientific and intellectual developments, a forward-looking approach to the electronic school dictionary is proposed. It consists of the following points:

-Rapid modernisation of the school dictionary’s structure, to build an interactive digital dictionary once the computational and technical requirements are met. Such a dictionary would serve both the Arab and the foreign learner. (Ben Yechou, 2015, p. 58)

-Developing its use for teaching the Arabic language in particular, and foreign languages in general, as it constitutes a reliable source that meets learners' needs, and enables them to master linguistic skills in any language.

-Recently, the increasing interest in the use of computers in language teaching and learning has motivated specialists to prepare automated school dictionaries. These motivations have converged to invest in computer technologies in the educational field. However, there is a need for cooperation between linguistic academies, regulatory bodies, scientific institutions, and learning centres, since we are living in an era that requires the computerised dictionary.

-Training researchers specialised in terminology and generalising its teaching theoretically and practically at the secondary and university levels. (Lazreg, Rabah , 2014, p. 12)

-Developing electronic school terminological work technically, and shifting it from being an individual activity to an institutional endeavour supervised by governmental centres. These centres are responsible for training lexicographers and educators and establishing specialised research teams in school lexicography and terminology (Ben Yechou, 2015, p. 66).

Therefore, the computerisation of the school dictionary is considered the most in-demand domain of computational linguistics, in response to the requirements of the modern era. It offers several services to the learner researcher, especially since computational linguistics is a recent discipline that employs computers to study diverse subjects, perform automated processes, develop a variety of programs, and establish electronic dictionaries required in the information and technology age.

III- Models of Electronic School Dictionaries

1- Almaany Dictionary

Almaany dictionary is one of the most prominent contemporary electronic dictionaries. It is a multilingual dictionary that provides educational, lexicographic and translation services simultaneously. It has a comprehensive linguistic base, which includes bilingual dictionaries between Arabic and other world languages such as English, Spanish, Portuguese, French, Turkish, Persian, Indonesian, and German. It stands out as both an educational and interactive tool. It not only presents the meanings of words but also offers:

- ✓ Precise linguistic explanations of Arabic vocabulary at different levels (classical and modern).
- ✓ Lexical equivalents between Arabic and other languages, which have made it an important reference for learners, researchers, and translators.
- ✓ Contextual usages of words through illustrative examples that help in understanding subtle meanings.
- ✓ Advanced search options to facilitate access to the desired word for explanation or translation.
- ✓ Updated content that keeps pace with semantic and linguistic developments and meets learners' needs in academic and applied contexts.

Thanks to these features, the Almaany dictionary represents a practical model of interactive electronic dictionaries that combine educational, lexicographic, and translational dimensions. This makes it a good tool for supporting the learning and dissemination of the Arabic language, as well as for enhancing cultural and linguistic communication between Arabic and other languages.

2- Mu‘jam Al-Lughah Al-‘Arabiyyah Al-Tafa‘ulī (The Interactive Arabic Language Dictionary)

It is the result of scientific and research collaboration between a group of researchers, engineers, and experts in the fields of informatics and the Arabic language at the Higher Institute for Applied Sciences and Technology in Damascus. This was carried out within the framework of a partnership agreement between the Institute and King Abdelaziz City for Science and Technology (KACST) in Riyadh, which jointly funded the project. This agreement is a result of the ongoing cooperation between KACST and the Arab League Educational, Cultural and Scientific Organisation (ALECSO), signed on 25/2/1429 AH (corresponding to 3/3/2008 AD), regarding the funding and implementation of the “Open-Source Interactive Computational Dictionary” project.

This dictionary is oriented to native speakers of Arabic and their learners. It is an important reference for students, intellectuals, and writers, due to its distinctive features:

- Being a monolingual dictionary (Arabic–Arabic).
- It includes newly coined vocabulary, structures, and idiomatic expressions.
- Providing precise lexicographic, morphological, and statistical information.
- Enabling the user to search for words and review their meanings, properties, and examples, in addition to offering further information that includes collocations, semantic fields, idiomatic expressions, grammatical and rhetorical notes, and correction of common errors.
- Enriching it with authentic examples that illustrate the correct usage of words.
- Incorporating root, derivative, and standard vocabulary, along with inflecting nouns and verbs in their various morphological and syntactic forms, through its integration with the open-source derivation and inflection system (SARF) developed by ALECSO.
- Allowing specialists to enrich its content with new words, meanings, and semantic links.
- Providing Application Programming Interfaces (APIs) that facilitate integration with Arabic language processing applications.
- Offering spelling check services and suggesting alternatives in case the entered word is absent.
- Relying on multimedia (audio, images, video, and others) in presenting linguistic knowledge, thereby enhancing interactivity (The Interactive Arabic Language Dictionary, s.d.).

Conclusion

In conclusion, we can summarise the results of this study in the following points:

- Computational linguistics is the cornerstone of the educational process. It provides tools and techniques that contribute to the development of the Arabic language, enabling it to keep pace with the digital age, and making it suitable for automatic processing. The electronic school dictionary has benefited from these technologies in organising, cataloguing, analysing, and updating lexical resources, thereby ensuring quick and efficient access to its content.
- Transforming school dictionaries from their traditional printed format into digital versions not only meets learners' needs but also helps in overcoming numerous organisational and technical challenges. It makes the school dictionary an accessible tool for all: learners, teachers, and researchers in an interactive educational environment that leverages digitisation and automatic language processing.
- Computational linguistics provides accurate and simplified linguistic content tailored to the target group. This enhances learners' comprehension and helps them overcome reading and lexical difficulties.
- Dictionary-making is witnessing significant expansion due to digital transformation, which has made it possible to produce more specialised and interactive electronic school dictionaries.

Based on the above results, we propose the following recommendations:

- One of the most promising prospects is the design of digital dictionaries tailored to different educational stages, and their integration into interactive curricula and educational applications.
- This field relies on a range of technologies, including natural language processing, artificial intelligence, lexical databases, morphological and syntactic analysis, and intelligent search systems.
- There is an urgent need to provide digital school dictionaries that keep pace with technological advancements and meet learners' needs for comprehension and communication.
- These dictionaries must be built upon rigorous pedagogical and linguistic foundations, while taking into account learners' characteristics and learning environments.
- The development of electronic school dictionaries requires the combined efforts of linguists, lexicographers, technology experts, and educators.
- It is essential to foster collaboration among specialists and experts to produce effective digital school dictionaries and to encourage scientific research in dictionary making, which has now emerged as an independent discipline that has achieved significant results both within and beyond the Arab world.

Bibliography:

Books:

- Abdelwahed Dkiki, *The Perspective of Dictionary Fundamentals, Laboratory of Discourse Grammar and Pragmatic Rhetoric, Faculty of Letters, Arts and Humanities, (2014), Manouba, Tunisia, 1st ed.*
- Ben Yechou, Djilali. *Studies in Applied Linguistics. 1st ed., Cairo, 2015, Dar Al-Kitab Al-Hadith.*

- Lazreg, Meflah, & Rabah Besseghir. *Arabic Dictionary Making in Light of the Challenges of Computational Language Processing. Symposium on the Investment of Computational Linguistics in the Making of Electronic Dictionaries*, (2018), The Higher Council of the Arabic Language.
- Nihad, Moussa. *Arabic: Towards a New Description in Light of Computational Linguistics. Arab Institute for Research and Publishing*, (2000), Jordan, 1st ed.
- Walid Al-Anati, Khaled Al-Djabr, *A Researcher's Guide to Arabic Computational Linguistics*, (2007), 1st ed, Amman, Jordan, 2007, Publishing and Distribution House.

Review:

- Abd Al-Rahman Ben Hassan Al-Aref, *The Use of Computational Linguistics in Serving Arabic Linguistic Studies*, (2014), *Journal of the Jordan Academy of the Arabic Language*, Issue 73.
- Aied Mohamed Al-Asmarai, *Machine Translation from the Perspective of Computational Linguistics*, (2018), *Journal of Abhath*, Vol. 4, Issue 3.
- Alain, Bonnet. *Artificial Intelligence: Its Reality and Future. Trans. Ali Farghali Sabri, Kuwait*, (2017), the World of Knowledge Series, vol. 17, Issue. 2.
- Aze-Eddine Bouchikhi. *Electronic Arabic Dictionaries and Prospects for Their Development. Conference on Dictionary Industry: Reality and Aspirations, Atlas Global Centre for Studies and Research*, (2004), University of Sharjah, 2004.
- Bekal, Sonia. *The School Dictionary Material between Reality and Aspiration. (2018) Journal of Al-Lisaniyyat, Algeria*, Issue. 16.
- Sonia Bekal, *Computational Linguistics – Theoretical Debates- (2018), Journal of Academic Studies*, Vol. 2, Issue 2.
- Mohamed, Hadj Henni., & Djamila Rougab. *School Dictionaries in the Light of the Lexicography Foundation. Case-study: The School Dictionary 2021 “Motiquin, Djoussour El-maarefa Journal*, vol. 4, Issue. 1.

Thesis:

Fares, Chacha. *Automatic Processing of the Arabic Language: Creating a Morphological-Syntactic Linguistic Model of the Arabic Verb. Magistral Thesis*, 2008 University of Algiers, Algeria.

websites:

- The Interactive Arabic Language Dictionary, <http://ar-di Mu'jam Al-Lughah Al-'Arabiyyah Al-Tafa'uli>.