

The Arabic Dictionary Between Linguistic Identity and Computational Identity: A Terminological Approach

Dr. Azzouz. Saida

Mohammed Seddik Benyahia University- Jijel- Faculty of Literature and Languages

Department of Arabic Language and Literature

Email: saida.azzouz@univ-jijel.dz

Submission: 20/12/2025 Acceptance: 21/04/2026 Publication: 28/08/2026

Abstract:

This article seeks to interrogate the concept of the Arabic ‘dictionary’ (mu‘jam) in its historical movement from its traditional linguistic identity to its contemporary computational identity, through a terminological-foundational approach that surveys the lexical and terminological definitions of the term, traces the emergence of the Arabic dictionary and its religious and civilizational motivations, analyses its four structural elements (linguistic material, entry, ordering, definition), examines the classical Arabic lexicographic ordering schools in their various classifications, and classifies dictionary types according to multiple criteria, before discussing the methodological problems facing contemporary Arabic lexicography. The study concludes that the computational identity of the dictionary is not an epistemic rupture with its linguistic identity, but rather a functional extension of its deep structure, one that has recast its traditional elements in a new digital, procedural form, while the problem of terminology and standardization remains open.

Keywords: dictionary, linguistic identity, computational identity, terminology, lexicographic industry, ordering schools.

Introduction

The dictionary, or lexicon, is a linguistic and cognitive tool that expresses the cultural level of a nation; through it, language is consolidated and its status and continuity are secured, and through it all social forms of communication are made possible. It is the noblest of books, a vessel for a shared treasure, and a vast space of referential knowledge — no longer confined to public libraries or consulted only when necessary, but rather a pressing daily need, a constant companion to literary and scholarly works alike. In this sense it is a text in which other texts intersect, transcending the bounds of language as such to open onto the manifestations of life and the fields of work and knowledge.

The earliest Arabic lexicographic attempts were no more than lists of words, scarcely rising to the level of the comprehensive dictionary in its strict scholarly sense, until computational linguistics — with the techniques and tools it provided, such as spell-checkers, transliteration tools, and systems for analysing and processing Arabic texts of all kinds, alongside Arabic-capable search engines built on massive databases — brought about a radical shift in the trajectory of lexicographic composition. Dictionaries thus diversified from the general to the specialised, and from dictionaries for teaching Arabic to native speakers to those for teaching

it to non-native speakers, and Arabic lexicography moved toward building digital linguistic dictionaries that allow the automatic storage and retrieval of linguistic information.

Between these two poles — the dictionary's traditional linguistic identity on the one hand, and its emergent computational identity on the other — a terminological and conceptual problem arises that merits rigorous scholarly interrogation: is the traditional Arabic term (*mu'jam*), with its original semantic load built on the opposition between obscurity and clarification, still capable of accommodating this shift from the closed paper book to the composite digital structure of databases, processing algorithms, and interactive interfaces? And what has actually changed in the dictionary's elements and mechanisms as it moved from its linguistic to its computational identity, and what has remained constant?

This study proposes that the computational identity of the Arabic dictionary is not an epistemic rupture with its traditional linguistic identity, but rather a functional extension of the deep structure established over centuries of lexicographic composition; the structural elements on which the paper dictionary was founded — linguistic material, entry, ordering, and definition — carry within themselves a latent readiness for procedural digital reformulation, owing to the openness and methodological flexibility that characterised them from their earliest inception.

To address this problem, the study adopts a descriptive-analytical method drawing on Arabic and Western lexicographic and terminological sources, structured across six axes: the first surveys the concept of the dictionary lexically and terminologically, as the foundation against which all its later transformations will be measured; the second traces the emergence of the Arabic dictionary and the historical stages of its development; the third presents its core structural elements; the fourth examines the Arabic lexicographic ordering schools in their various classifications; the fifth classifies dictionary types according to multiple criteria; and the sixth discusses the methodological problems facing contemporary Arabic lexicography, in preparation for interrogating the threshold of computational identity in the article's synthesising conclusion.

Axis One: The Dictionary Lexically and Terminologically . The Linguistic Identity of the Term

First: The Dictionary in Lexical Usage

Arabic dictionaries record several definitions of the term (*mu'jam*) from the root (‘-j-m). Ibn Manẓūr (d. 711 AH) states in *Lisān al-‘Arab* that the root (‘-j-m) denotes al-‘ujm and al-‘ajam as opposed to al-‘urb and al-‘arab, that ‘ajamī is the opposite of ‘arabī, and that a man described as a‘jam is one whose speech is neither eloquent nor clear¹. In the same vein, Ibn Jinnī (d. 392 AH) states in *al-Khaṣā’iṣ* that wherever the derivation (‘-j-m) occurs in the speech of the Arabs, it denotes obscurity, as opposed to clarity and eloquence²; and al-Zabīdī (d. 1205 AH) follows the same line in *Tāj al-‘Arūs*, defining al-‘ujma as an impediment (*ḥubsa*) of the tongue³.

Aḥmad Mukhtār ‘Umar confirms in *Ṣinā‘at al-Mu‘jam al-Ḥadīth* that the root (‘-j-m) lexically denotes obscurity and vagueness; in *Lisān al-‘Arab*, al-a‘jam is one who neither speaks eloquently nor makes himself clear⁴; accordingly, the meaning of a‘jama becomes: to remove obscurity, vagueness, or ambiguity. Ḥusayn Naṣṣār likewise quotes Ibn Jinnī in *al-Mu‘jam al-*

‘Arabī — Nash’atuhu wa-Taṭawwuruḥu: know that (‘-j-m) occurs in the speech of the Arabs only to denote obscurity and concealment, as opposed to clarity and eloquence; al-‘ujma is an impediment of the tongue, and hence a man is described as a‘jam and a woman as ‘ajmā’ when neither can make their speech clear, and the dumb beast (al-‘ajmā’) is so called because it cannot articulate what is within it⁵.

In sum, the root (‘-j-m) denotes obscurity and vagueness⁶; when the glottal stop (hamza) is prefixed to the verb (‘ajama) to form (a‘jama), the verb acquires a new sense from the hamza’s function, namely negation, denial, and removal — as in the pair qasaṭa/aqsaṭa: the first denotes injustice, the second denotes justice, or the removal of injustice, as in the Qur’anic verses: ‘As for the unjust (al-qāsiṭūn), they will be fuel for Hell’ (al-Jinn 72:15), and ‘Indeed God loves those who act justly (al-muqsiṭīn)’ (al-Mā’ida 5:42, part). It is likewise said: a‘jama the letters, meaning he dotted them, since dotting removes obscurity and ambiguity from the letters⁷, since diacritically marking the letters removes ambiguity and vagueness and renders the word or text clearer and easier to read and understand.

It follows from this that all these definitions converge on two opposed meanings: the first, the original sense, denotes vagueness and lack of clarification, as in the Qur’anic verse: ‘The tongue of those they refer to is foreign (a‘jamī), while this is a clear Arabic tongue (‘arabī mubīn)’ (al-Nahl 16:103, part), where a‘jamī is the semantic opposite of ‘arabī, while the qualifier mubīn (clear) restricts and specifies the meaning of ‘arabī, denoting clarity and elucidation. An example of the same sense is the expression a‘raba fulānun ‘an ra’yihi (‘so-and-so declared his opinion’), i.e., he made it explicit and clear — the reverse of the first, original sense. It is from this that the term mu‘jam (dictionary) arose. What, then, does it mean terminologically?

Second: The Dictionary as a Technical Term

The word (mu‘jam) is applied to a book that gathers the vocabulary of a language, or most of it, arranged in a particular order, explaining its meaning and clarifying its derivatives, dialectal forms, and usages⁸. Al-Mu‘jam al-Wasīṭ, for its part, follows a concise approach, defining the dictionary simply as: a register (dīwān) of the vocabulary of a language, arranged according to its alphabet⁹. Both definitions are notably concise, focusing on the vocabulary of the language as the dictionary’s core, on condition that it be arranged in a particular order, without elaborating further.

Among the scholars who offered a concise definition of the term, ‘Alī al-Qāsimī, in ‘Ilm al-Lughā wa-Ṣinā‘at al-Mu‘jam, defines it as: a book containing a selection of words, usually arranged alphabetically, with an explanation of their meanings and other related information, whether that explanation is given in the same language or in another¹⁰. In a similar vein, Mukhtār ‘Umar defines it as: the book that gathers the words of a given language, explains them, clarifies their meaning, and arranges them in a particular manner¹¹. Others define the language dictionary as: a book that gathers the vocabulary of a language in order to remove ambiguity from it, clarify it, fix its pronunciation and usage, and explain its meaning — its task thereby being clarification and elucidation — on condition that this vocabulary is subsumed

under so-called linguistic materials or roots, themselves arranged according to alphabetical order.

‘Abd al-Qādir ‘Abd al-Jalīl, influenced in his study by the structuralist method, holds that the dictionary as a book comprises three components: linguistic units, the classificatory (entry) system, and semantic definition; on these three pillars the dictionary rests in its general form, as a vessel that preserves the corpus of a language rather than one of its systems, since lexical meaning is part of the language’s broader semantic system¹²; the lexical unit remains the dictionary’s pivot, its activity revolving around derivation, usage, and diachronic development. This conception parallels the notion of the dictionary in Western thought, which has likewise gone beyond a mere word-list to become a reference tool for organising and classifying the vocabulary of a language, as Hartmann and James note: a dictionary is ‘a type of reference work in which the words of a language, language variety, speaker or text are listed and explained, either in alphabetical or in thematic order’¹³.

It may be concluded from the foregoing that the terminological sense of (dictionary) approximates its lexical sense, and achieves two things at once: clarification (*bayān*) and ordering (*tartīb*). Clarification concerns elucidating the words the dictionary contains, explaining their meanings, and fixing their pronunciation; ordering concerns arranging the lexical units alphabetically, semantically, or thematically. The dictionary is thus not a mere repository holding the greatest possible number of words with their meanings and pronunciation and a particular arrangement, but rather — as it has been described — an open system with two doors: one through which new, coined words enter, and another through which obsolete words and usages exit, marginalised and abandoned for various reasons; it is thus in constant change and development¹⁴. It is precisely from this characterization, as an open system rather than a closed book, that, as will become clear in the article’s conclusion, the possibility of its absorbing a computational identity will proceed.

Axis Two: The Emergence of the Arabic Dictionary The Historical Roots of Identity

The Arabs were not the first among nations to invent lexicographic composition; other nations, such as the Chinese and the Greeks, preceded them. It is noted that Arab lexicographers were considerably slow to attend to language as actually used — both old and new — and this concern appeared only among a few dictionary compilers at the end of the nineteenth and the turn of the twentieth century, such as the Bustānī family and Iskandar Ma‘lūf, though this remained limited and unsystematic¹⁵. This historical precedence, however, does not mean the Arabs imitated these other nations in lexicographic work; rather, they devised their earliest dictionaries with originality and creativity, driven by purely Arabic motives to serve and protect the Arabic language from foreign intrusion.

It appears that the primary impetus behind many of the earliest attempts at linguistic study across different parts of the world was religious: among the Chinese, the emergence of dictionaries is traced to the study of Buddhist religious texts; among the Hindus, linguistic research began in service of their sacred text, the Vedas; while among the Greeks, one of the

oldest dictionaries, that of Julius Pollux, resembles Ibn Sayyidih's al-Mukhaṣṣaṣ, both relying on the study of religion as an impetus for lexicographic and linguistic composition.

The Arabs, for their part, knew no dictionaries in the pre-Islamic era, until the advent of Islam created the need to investigate the meaning of obscure and difficult words — as in Abū al-Aswad al-Du'alī's attempt to vocalise the Qur'anic text, instructing his scribe to take the maṣḥaf and use ink of a colour different from black to mark dots on the letters, following the movements of his mouth and lips. This leads directly to the connection between the emergence of the Arabic dictionary and the Qur'an, through the attempt to explain its vocabulary: the earliest lexicographic compositions accordingly bore the title *Gharīb al-Qur'ān* ('The Obscure Words of the Qur'an'), among them those of Abū Sa'd ibn Tha'lab ibn Rabāḥ al-Bakrī and Ibn Qutayba. Numerous works were likewise compiled on *Gharīb al-Ḥadīth* (the obscure vocabulary of the Prophetic tradition), most notably those of Abū Ishāq al-Ḥarbī and al-Khaṭṭābī.

The motive underlying the Arab scholars' undertaking of dictionary composition was the sense of the danger posed by the encroachment of linguistic error (*lahn*) upon the Qur'an, after it had already begun to affect the speech of the Arabs¹⁶. Accordingly, the aim of gathering and preserving the language was not, as in modern dictionaries, simply to establish meaning, but rather to preserve the speech of the Arabs and derive from it the rules and laws that govern it; nor was this gathering confined to vocabulary alone, but extended to poetry, legal rulings, proverbs, and prose. Four overarching aims may thus be discerned in this early lexicographic endeavour: preserving the Qur'an and protecting it from error, mispronunciation, and misreading; clarifying rare and obscure meanings and bringing them closer to the understanding; protecting the language from corruption and extinction; and gathering the linguistic wealth of the language through explanation and citation, thereby contributing to teaching it to non-native speakers.

Before proceeding to the stages of collection, it is worth noting the linguistic epistles (*rasā'il lughawiyya*), which constituted the earliest nucleus of the linguistic material later gathered into various dictionaries; these epistles continued alongside dictionaries proper down to the ninth century AH, with al-Fayrūzābādī, and thus represent not so much an independent stage of collection as a form it took. Among the earliest and most notable authors of such epistles was Ibn Yazīd al-A'rābī, in his *Kitāb al-Ḥaṣharat* ('Book of Insects'), and 'Abd al-Malik ibn Qurayb al-Aṣma'ī, who wrote on camels, human anatomy, horses, and plants, among other subjects.

The process of collection unfolded across three successive stages: first, the stage of oral transmission and audition, in which scholars 'gathered the words of the language however they happened upon them, without following any fixed method — the scholar would travel to the desert and hear a word relating to rain, another relating to the sword, another to crops and plants, recording each as heard, with no order but that of audition itself'¹⁷; second, the stage of compilation (*tadwīn*), namely 'confining the words related to a single subject in the form of epistles, such as the vocabulary relating to rain, camels, animals, sheep — these forming the dictionaries of meaning (*mu'jamāt al-ma'ānī*)'¹⁸, a stage in which the mechanism of

lexicographic collection appeared in two forms, general collection (words gathered in scattered, unordered epistles) and specific collection (words gathered and ordered under major semantic fields, as al-Aṣmaʿī did); and third, the stage of the complete dictionary, which ‘comprises every word of the language accompanied by an explanation of its meaning, its derivation, its pronunciation, and citations illustrating its use’¹⁹. It is in this stage that al-Khalīl ibn Aḥmad al-Farāhīdī’s (d. 175 AH) Kitāb al-ʿAyn emerges as the first comprehensive Arabic dictionary, followed by others: Ibn Duraid’s (d. 321 AH) Jamharat al-Lugha, Abū ʿAlī al-Qālī’s (d. 356 AH) al-Bārīʿ, al-Azharī’s Tahdhīb al-Lugha, Ibn Fāris’s (d. 395 AH) Maqāyīs al-Lugha, and Ibn Sayyidih al-Andalusī’s (d. 458 AH) al-Mukhaṣṣaṣ and al-Muḥkam²⁰ — by which the mechanism of lexicographic ordering emerged fully formed, and the features of the Arabic dictionary in its mature configuration were completed.

Axis Three: The Elements of the Dictionary - A Structure Amenable to Computation

The essential elements of the dictionary may be summarised in four components that together underlie its construction: linguistic material, by which is meant ‘the words or lexical units that the lexicographer collects, then orders and explains the meaning of’²¹; lexicographic entries, the unit under which the remaining units are subsumed — for example, the entries muʿjam, iʿjām, and mutaʿajjim all belong to the root (ʿ-j-m); ordering, meaning ‘the arrangement of entries, as well as of derivatives within language dictionaries under a single root or entry’²², which takes several methods sanctioned by lexicographic practice, and is subdivided into internal ordering (distinguishing ordering by shared root and ordering by grouping of related forms) and external ordering (concerning the entry itself, i.e., ordering by the first or last letter, or alphabetically); and finally, definition or lexicographic explanation, namely ‘stating the meaning of the word, the essential function of any dictionary’²³.

The element of definition requires clarifying a word’s meaning according to a set of governing conditions, chief among them: fixing the correct pronunciation of the word, noting its common usage while clarifying its original sense prior to any figurative one. Other complementary elements enrich the lexicographic entry’s structure, including grammatical information clarifying the word’s gender (masculine, feminine) or category (verb, particle, noun); examples and citations drawn from texts such as poetry or Qur’anic verses, to illustrate the word’s use in a particular context; and, in some dictionaries, historical information tracing the evolution of the word’s meaning over time.

On these four pillars the dictionary rests in its general form, as a vessel preserving the corpus of the language rather than one of its systems, the lexical unit remaining its pivot, around which its activity revolves in derivation, usage, and diachronic development²⁴. Through the integration of these four elements a complete dictionary results — and it is precisely these elements, as will become clear in subsequent axes, that will migrate from their paper formulation to the formulation of data fields amenable to automatic processing, without losing their essential function.

Axis Four: The Ordering Schools of the Arabic Dictionary

Given that the dictionary is a valuable source for understanding, analysing, and organising a language and documenting its vocabulary, it is one of the most important tools on which researchers and linguists rely in studying a language and tracing the development of its vocabulary according to a particular system of ordering. The question of ordering is accordingly among the most important issues in the history of Arabic dictionaries — namely, the method or approach the lexicographer follows in organising the linguistic wealth selected from morphemes, words, and contextual and idiomatic expressions, and presenting it in the dictionary in a manner that allows the user to consult it easily, without effort or waste of time; a fixed system is thus indispensable to the construction of any dictionary, whatever its type.

With the diversity and multiplicity of language dictionaries, lexicographic orientations have varied since the earliest beginnings of lexicographic work down to the present, according to the methods of individual compilers, so that scholars have treated each ordering method as a distinct school with its own approach and pioneers; and views have consequently varied in defining these schools according to the diversity of scholars and their positions on lexicographic composition. The following presents the most prominent of these classifications.

1. Aḥmad Mukhtār ‘Umar’s Classification

Aḥmad Mukhtār ‘Umar, in *Ṣinā‘at al-Mu‘jam al-Ḥadīth*, holds that Arabic dictionaries fall into three schools: the school of phonetic (articulatory) ordering, the school of alphabetical ordering by the first letters of words, and the school of alphabetical ordering by the last letters of words. These schools fall, in his classification, under two types of dictionary: dictionaries of words and dictionaries of meanings; if the researcher knows the word and wishes to obtain something unknown about it — its meaning, pronunciation, etymological derivation, or degree of usage — his access to the dictionary is through the word, and he consults a dictionary of words; if he knows the general meaning or subject and wishes to obtain the words, expressions, or terms that fall under it, he consults a dictionary of meanings²⁵. This classification reflects ‘Umar’s adoption of a principle of simplicity and logical clarity, grounded in a foundational reference for ordering lexicographic entries, and reveals a deep understanding of the realities of Arabic lexicographic practice.

2. Ḥusayn Naṣṣār’s Classification

Ḥusayn Naṣṣār, in *al-Mu‘jam al-‘Arabī — Nash’atuhu wa-Taṭawwuruḥu*, sets out a classification of lexicographic schools without assigning them specific names, using instead simple designations (‘the first school’, ‘the second school’...), each representing a group of dictionaries sharing a single ordering system. His division into four schools runs as follows: the first school, phonetic ordering by the places of articulation (al-Khalīl’s *al-‘Ayn*, al-Azharī’s *Tahdhīb al-Lughā*); the second, termed grammatical ordering or ordering by morphological pattern, relying primarily on alphabetical ordering with the system of patterns and permutations as a secondary ordering principle (**Ibn Duraid’s *al-Jamhara*, Ibn Fāris’s *al-Maqāyīs***); the third, alphabetical ordering by the last letters of words, also called the rhyme (*taqfiya*) school,

taking the final root letter as the primary basis for ordering, then the first letter, then the second (al-Jawharī's al-*Ṣiḥāḥ*, Ibn Manẓūr's *Lisān al-ʿArab*, al-Fayrūzābādī's al-Qāmūs al-Muḥīṭ); and the fourth, alphabetical ordering by the first letters of word roots, without the taqfiya and permutation systems, encompassing both older and more recent dictionaries — the latter distinguished by tighter organisation, sound methodology, and logical ordering of entries according to their roots, exemplified by al-Muʿjam al-Wasīṭ issued by the Arabic Language Academy in Cairo.

Comparing Ḥusayn Naṣṣār's classification with earlier views, one notes that his approach lacks comprehensiveness despite the importance of its alphabetical distinction; he adopted the branch of alphabetical ordering built on the pattern system, thereby neglecting pattern-based dictionaries and alphabetical dictionaries without such abstraction — a gap subsequently addressed by Aḥmad Mukhtār ʿUmar's more precise lexicographic classification.

3. ʿAbd al-Qādir ʿAbd al-Jalīl's Classification

ʿAbd al-Qādir ʿAbd al-Jalīl, in al-Madāris al-Muʿjamiyya — *Dirāsa fī al-Bunya al-Tarkībiyya*, identifies these orientations under four headings: the school of permutative-articulatory ordering (al-Khalīl's al-ʿAyn, al-Qālī's al-Bārīʿ, al-Ṣāḥib ibn ʿAbbād's al-Muḥīṭ fī al-Lughā, Ibn Sayyidih's al-Muḥkam wa'l-Muḥīṭ al-Aʿẓam); the school of pattern-based alphabetical rotation (Ibn Duraid's *Jamharat al-Lughā*, Ibn Fāris's *Mujmal al-Lughā* and *Maqāyīs al-Lughā*); the school of root-based alphabetical ordering (Abū ʿAmr al-Shaybānī's al-Jīm, al-Zamakhsharī's *Asās al-Balāgha*, al-Fayyūmī's al-Miṣbāḥ al-Munīr); and the school of rhyme-based ordering (al-Bandanījī's al-Taḥfiya fī al-Lughā, al-Jawharī's *Tāj al-Lughā wa-Ṣiḥāḥ al-ʿArabiyya*, Ibn Manẓūr's *Lisān al-ʿArab*, al-Fayrūzābādī's al-Qāmūs al-Muḥīṭ, and al-Zabīdī's *Tāj al-ʿArūs*).

4. A Reconciled Four-School Classification

Reconciling these views, and in agreement with Aḥmad ʿAbd al-Ghafūr ʿAṭṭār's classification in *Muqaddimat al-Ṣiḥāḥ* — who divided the lexicographic schools into four for those seeking detail, and into two (the school of meanings and the school of words) for those seeking a summary account — a more precise four-school classification may be adopted, combining the substance of the foregoing views, detailed below with the distinctive methodological features of each school.

a) The phonetic or articulatory school (al-Khalīl's school): the first school in the history of Arabic lexicography, named after al-Khalīl ibn Aḥmad al-Farāhīdī, founder of the first rigorously methodical Arabic linguistic dictionary, *Kitāb al-ʿAyn*. The essence of his school lies in ordering material by the letters according to their places of articulation, dividing the dictionary into 'books', subdividing the books into 'chapters' by morphological pattern, massing words within the chapters, and permuting each word into its various possible forms, while discarding unattested forms²⁶. His method thus rests on a phonetic theory: ordering the dictionary's letters by their place of articulation, from the deepest point in the throat to the lips. As for the ordering of patterns, al-Khalīl adopted a statistical-mathematical method, beginning

his dictionary with the letter ‘ayn as the clearest and most distinct of the letters, and treating patterns as fourfold: biliteral, trilateral, quadrilateral, and quinqueliteral; in the biliteral pattern alone he counted 756 possible forms, the product of multiplying the twenty-eight letters of the alphabet by themselves after discounting repetition of a letter with its own kind²⁷, adopting the system of permutations he himself devised: two permutations for biliteral roots (e.g. barr/rabb), six for trilateral roots (e.g. raja‘a, ra‘aja, jara‘a, ja‘ara, ‘araja, ‘ajara), and one hundred and twenty for quinqueliteral roots.

b) The pattern-based ordering school: represented by Ibn Duraid’s (d. 321 AH) al-Jamhara, the second dictionary after al-‘Ayn, so named because Ibn Duraid selected the ‘commonly used’ (jumhūr) vocabulary of the Arabs, i.e. common usage, excluding rare and obscure words. The method of this school rests on the pattern system, comprising an external ordering — Ibn Duraid divided his dictionary into main chapters according to biliteral, trilateral, and quadrilateral patterns — and an internal ordering, subjecting all patterns to alphabetical arrangement while adhering to the system of permutations.

c) The alphabetical school, itself divided into two branches. The first is alphabetical ordering by the first letters, also called the root-alphabetical school, which adopted ordering the dictionary by the letters of the alphabet from hamza to yā’, taking into account the second, third, and fourth letters after stripping the word of any affixes; Abū ‘Amr al-Shaybānī is regarded as the pioneer of this school with his dictionary al-Jīm, followed by al-Zamakhsharī in Asās al-Balāgha, al-Bustānī in Muḥīṭ al-Muḥīṭ, and Ma‘lūf in al-Munjid. The second is alphabetical ordering by the last letters, called the rhyme (taqfiya) school or al-Jawharī’s school, after the author of Tāj al-Lughā wa-Ṣiḥāḥ al-‘Arabiyya, who innovated the principle of taking a word’s final letter as a ‘chapter’ (bāb) — on the grounds that it is more stable than the word’s other letters — and its first letter as a ‘section’ (faṣl), while attending to roots and setting aside affixes. Among the dictionaries that followed this method are al-Zabīdī’s Tāj al-‘Arūs and al-Fayrūzābādī’s al-Qāmūs al-Muḥīṭ.

d) The thematic (semantic) school: also known as the school of dictionaries of meaning, or Abū ‘Ubayd’s school, which orders the roots of vocabulary according to their meanings or subjects, and is also referred to as the school of synonymic (mutawārid) or semantically-clustered dictionaries; it is named after its pioneer, Abū ‘Ubayd al-Qāsim ibn Sallām, author of al-Gharīb al-Muṣannaf, which comprised twenty-five thematic sections, each containing chapters (on human anatomy, food, clothing, and so on). This school is regarded as the earliest nucleus of Arabic lexicographic composition, since it places every word relating to a given subject under a single heading. Among the most prominent dictionaries of this school is Ibn Sayyidih’s al-Mukhaṣṣaṣ, comprising seventeen books arranged with a gradual thematic progression.

This four-school classification may be summarised in the following table:

School	Basis of Ordering	Leading Figures and Dictionaries
Phonetic (articulatory)	Places of articulation, throat to lips	al-Khalīl ibn Aḥmad — al-‘Ayn

Pattern-based ordering	Biliteral/triliteral/quadriliteral patterns, then alphabet	Ibn Duraid — al-Jamhara
Alphabetical by first letters	First letter, then second, then third	al-Shaybānī (al-Jīm), al-Zamakhsharī (Asās al-Balāgha)
Alphabetical by last letters (rhyme)	Last letter as chapter, first as section	al-Jawharī (al-Ṣiḥāḥ), al-Fayrūzābādī (al-Qāmūs al-Muḥīṭ)
Thematic (semantic)	Meaning or subject, not the word form	Abū ‘Ubayd (al-Gharīb al-Muṣannaf), Ibn Sayyidih (al-Mukhaṣṣaṣ)

Axis Five: Types of Arabic Dictionaries

Given that form (lafz) and meaning (ma‘nā) are two inseparable pillars of language — every word comprising both a phonetic aspect and a semantic one — the traditional Arabic dictionaries may be divided into a logical dyad: dictionaries of words, which proceed from form to meaning, and organise their material phonetically or alphabetically, treating the word and explaining its signification and everything connected to it according to a particular ordering method²⁸; and dictionaries of meanings or subjects, which gather the words revolving around a single meaning or subject²⁹, the most notable example being Ibn Sayyidih’s al-Mukhaṣṣaṣ. The contemporary classification of dictionaries has diversified with the multiplication of the fields covered by lexicographic work and the target audiences addressed, according to four principal criteria:

- 1) By the specificity of the lexicographic material: dictionaries divide into general language dictionaries, encompassing the widest range of shared, general vocabulary (e.g. al-‘Ayn); specialised or terminological dictionaries, addressing a limited segment of intellectual activity — literary, scientific, social, and so forth — gathering the vocabulary and terms of a particular field or discipline³⁰ (e.g. al-Jawālīqī’s al-Mu‘arrab wa’l-Dakhīl); and encyclopedic dictionaries, containing a vast body of linguistic, scientific, and other information, extending their explanations to detailed non-linguistic content such as proper names, cities, and historical information³¹ (e.g. Ibn Manẓūr’s Lisān al-‘Arab).
- 2) By the language of entries and glosses: dictionaries divide into monolingual dictionaries, in which entries and glosses share a single language (Arabic–Arabic); bilingual or multilingual translation dictionaries, using two different languages, one for entries and the other for glosses³²; and pocket dictionaries, small in size and few in pages, more useful for spelling and pronunciation than for detailed semantic explanation, sometimes containing helpful encyclopedic information such as small maps and statistics³³.

3) By semantic methodology: dictionaries divide into synchronic (descriptive) dictionaries, which describe a language, gather it, and order it via entries and usages within a defined period and place, without departing from them³⁴; historical or diachronic dictionaries, concerned with the development of a word across ages — in form, meaning, or orthography — recording the point of its entry into the language, its etymological roots, and tracing its development³⁵, themselves subdivided into etymological dictionaries, highlighting the origin of each word, and general historical dictionaries, describing the changing linguistic corpus across successive periods; and normative (prescriptive) dictionaries, the most widespread type in the modern era, adhering to criteria of both time and place, since Arab linguists set conditions for classical (fuṣḥā) Arabic covering both a temporal criterion and a spatial one — isolation deep in the desert, away from foreign contact — and restricted citation to literary texts, namely the Qur'an, the Prophetic ḥadīth, and poetry³⁶.

Other modern dictionaries have also emerged to meet the needs and demands of the age, such as student dictionaries, school dictionaries, medical dictionaries, and technological dictionaries — all of which confirms that Arabic dictionaries have diversified to a degree unmatched by the dictionaries of other languages, their subdivisions multiplying under each type according to differences of method, aim, target audience, subject matter, and specialisation.

Axis Six: Methodological Problems of Contemporary Arabic Lexicography

Given the technological transformations the world has undergone today, including artificial intelligence technologies that have touched every field and discipline, foremost among them the Arabic language sciences, and given the vast expansion of knowledge and the increasing linguistic hybridity and conceptual inflation, the need has grown for comprehensive, precise Arabic dictionaries to safeguard the Arabic language and keep pace with contemporary developments in knowledge and the sciences. The copious influx of terminology the language has struggled to contain, together with the inflation of linguistic information, have become obstacles impeding the modernisation of Arabic dictionary content and the standardisation of its terminology, directly affecting the quality of the Arabic dictionary and its uses. Four principal issues may be identified in this regard.

a) The Difficulty of Arabic Lexicographic Work

Arabic lexicographic work requires knowledge of the language and its sciences, the boundaries of words and their meanings in different contexts, and how they are used; it likewise requires knowledge of the computational dimension and the methods of organising, classifying, and editing information, demanding a complete and clear picture of the dictionary's user, their thinking, and their purpose — something no individual can achieve alone, making it necessary to entrust lexicographic work to a team of researchers combining linguists and technical specialists. Arabic lexicography is beset by numerous difficulties, foremost among them the difficulty of mastering the general system of the Arabic language and identifying all the properties appropriate to its lexical units; the lexicographer 'deals with an open phenomenon that never settles into a fixed state, and thus any attempt to enumerate the words of any living

language is a goal that is difficult, if not impossible, to attain fully, the lexicographer remaining in a state of constant questioning as to the degree of comprehensiveness his dictionary achieves and how close or far it is from a complete collection of the language's material³⁷. It is thus improper to adopt a principle of linguistic fixity; rather, its developments across the ages must be tracked, and the lexicographer must define meaning precisely and explain it clearly. The lexicographic researcher likewise faces further difficulties relating to gathering material, extracting entries, and the manner of fixing, ordering, and refining them from among millions of words and expressions, up to the point of producing the dictionary in its final form.

b) Semantic Variation Across Words

The question of semantic usage varies from one dictionary to another — as in dictionaries of meanings, dictionaries of rhetoric, dictionaries of philology, and others — as well as varying between dictionaries within the same category, owing to differences in the compilers' methodologies in viewing the language and the meanings of its vocabulary, and to the legacy of the scientific and logical theories held by individual linguists, in addition to the impact of social and cultural differences on the meaning of words and their evolution. This does not mean, however, that semantic divergence will impede the process of computerising the dictionary; rather, methodological solutions must be devised to reconcile these divergent senses, through the adoption of modern linguistic methods and precise techniques from computer science and modern linguistics, in order to develop the Arabic dictionary and reduce its semantic gaps.

c) The Coining and Standardisation of Arabic Terminology

A term (*muṣṭalaḥ*) is defined as a word or compound expression possessing a specific signification within a specific usage; Maḥmūd Fahmī Hījāzī defines it as: any single word or compound phrase whose meaning has stabilised, been employed, and been clearly determined — a narrow expression specific to its specialised signification, as clear as possible, having an equivalent in other languages, and always occurring within the context of its particular system³⁸. A term is thus whatever moves from the sphere of the general to the specific with a fixed sense — that is, it becomes clearer and more precise within its own specific context — yet it suffers from the difficulty of securing universal agreement on a single, precise definition, owing to contextual diversity and cultural difference, so that it becomes hard to determine a term's specific meaning without reference to the context in which it was coined. The term thus suffers from numerous problems: multiple equivalents across dictionaries and translations — appearing at times through traditional Arabic equivalents, at others as an Arabised or borrowed form — alongside multiple pronunciations and usages; the absence, moreover, of Arabic bodies responsible for coining and standardising terminology has negatively affected the quality of Arabic lexicographic work, driving some lexicographers to adopt unfamiliar terms, or ones that have not achieved consensus among Arabic linguistic authorities.

To curb the disorder surrounding the coining of Arabic terminology, it is necessary to 'establish cooperative scholarly bodies, and specialists in the various modern sciences who work to

supply it with the scientific terminology that supports the contemporary Arab renaissance and assists in translating foreign works into Arabic, and in composing and teaching the various sciences in Arabic³⁹. It has thus become necessary to devise precise methodologies for coining and standardising Arabic terminology and adapting it to modern scientific development, while working to modernise this terminology in keeping with Arabic identity and its authenticity, without compromising standards of precision and eloquence — since computerising the Arabic dictionary rests on multiple methodological foundations and rules for achieving the precise fixing of Arabic terminology.

d) The Deployment of Modern Technologies and Tools

Modern technologies contribute to improving lexicographic work in terms of classification, storage, and digital processing, leading to the conversion of lexicographic texts into an integrated electronic format, alongside equipping electronic dictionaries with digital images and interactive audio to provide interactive content for the user. This underscores the importance of contemporary technical tools and means, and the necessity of deploying them in the field of automated lexicographic processing in order to produce rigorous and precise dictionaries.

Conclusion: The Threshold of Computational Identity - A Synthesis of Findings

If the traditional dictionary, as established in the first axis, is an open system rather than a closed book⁴⁰, then it is precisely this quality that rendered it capable of migrating to a computational identity without a radical rupture from its linguistic origin; the four elements on which the paper dictionary was founded — linguistic material, entry, ordering, and definition — have not been abolished in the computational dictionary but rather procedurally reformulated: linguistic material has become a digital corpus; the entry has become a data field amenable to indexing; ordering — once accomplished manually according to the schools of al-Khalīl, al-Jawharī, or Abū ‘Ubayd — has become a search-and-retrieval algorithm executed in fractions of a second; and definition has become a text field capable of hyperlinking to other dictionaries and to audio citations and digital images.

This functional extension is not, however, free of terminological difficulty. When the term (mu‘jam) is rendered into its Western counterparts — the ‘machine dictionary’ or the ‘lexical database’ — it becomes clear that the computational concept exceeds the traditional definition that confined the dictionary to being a ‘book’; the computational dictionary is no longer a book in the literal sense, but rather a hierarchical or relational data structure whose contents are dynamically invoked. This raises the question: can the term (mu‘jam), with its traditional semantic load — requiring a ‘particular ordering’ and an ‘explanation’ that presuppose a fixed paper medium — alone denote this changing, interactive structure?

Tracing the trajectory of the concept, the answer appears to lie not in replacing the term but in broadening its signification: just as the word (mu‘jam), in its historical transformation from scattered linguistic epistles to the complete dictionary, absorbed a major functional shift without losing its identity, it is capable, in principle, of absorbing its computational

transformation as well, so long as it continues to fulfil its dual essential purpose: clarification and ordering. Yet this absorption remains contingent on overcoming the practical problems detailed in the sixth axis: the difficulty of mastering the general system of the Arabic language; the variability of semantic usage across dictionaries; the disorder surrounding the coining, generation, and standardisation of Arabic terminology; and the necessity of deploying modern technologies methodically rather than merely cosmetically.

None of this implies placing on the dictionary alone the burden of developing, modernising, and updating the Arabic language; rather, it means that the computational identity of the Arabic dictionary is a joint linguistic-technical project, requiring a team that unites the linguist and the technologist, rather than the individual endeavour that characterised the traditional lexicographic schools. In light of the foregoing, this study concludes with the following findings:

- The term (mu‘jam) rests semantically on a fundamental opposition between obscurity and clarification, and has moved from this lexical sense to a terminological one achieving two things: clarification and ordering — an opposition that continues to govern the term even in its computational form.
- The Arabic dictionary emerged from a religious-functional impulse tied to protecting the Qur’an from linguistic error, then developed through three stages (audition, compilation, the complete dictionary) to form a firmly established traditional identity resting on four elements: linguistic material, entry, ordering, and definition.
- The multiplicity of lexicographic ordering schools (phonetic, pattern-based, alphabetical in its two branches, thematic) reveals a structural flexibility within the traditional Arabic dictionary — the very flexibility that later enabled its adaptation to the logic of computation.
- The diversity of dictionary types according to criteria of material, language, and semantic methodology reflects an early maturity in Arabic classificatory awareness, preceding the age of computation by centuries and preparing the ground for it.
- The computational identity of the dictionary is not a rupture with its linguistic identity but a functional extension of it: the entry has become a data field, ordering has become an algorithm, and definition has become hyperlinked text; yet this extension continues to raise a terminological problem requiring the broadening, not the replacement, of the term (mu‘jam).
- Achieving a sound computational identity for the Arabic dictionary is contingent on addressing the problems facing its production: the difficulty of lexicographic work, the variability of semantic usage, the disorder of terminology, and the necessity of deploying technology methodically rather than superficially.

This approach remains an initial gateway to a deeper interrogation to be undertaken in subsequent research, addressing in detail the mechanisms of constructing the digital Arabic dictionary, its linguistic and technical methodology, its contemporary applied models, and the role of artificial intelligence in developing and enriching Arabic dictionaries.

References:

¹Ibn Manẓūr, *Lisān al-‘Arab*, entry (‘-j-m).

- ²-Ibn Jinnī Abū al-Faṭḥ ‘Uthmān, al-Khaṣā’iṣ, Maṭba‘at Dār al-Kutub al-Miṣriyya, Egypt, 4th ed., vol. 3, 1957, p. 52.
- ³-Muḥammad Murtaḍā al-Ḥusaynī al-Zabīdī, Tāj al-‘Arūs, ed. a group of specialists, Ministry of Guidance and Information, Kuwait, vol. 33, entry (‘-j-m), p. 67.
- ⁴-Aḥmad Mukhtār ‘Umar, Ṣinā‘at al-Mu‘jam al-Ḥadīth, ‘Ālam al-Kutub, Cairo, 1st ed., 1988, p. 19.
- ⁵-Ḥusayn Naṣṣār, al-Mu‘jam al-‘Arabī — Nash‘atuhu wa-Taṭawwuruḥu, Dār Miṣr li’l-Ṭibā‘a, 4th ed., 1988, vol. 1, pp. 9–10.
- ⁶-Ibid., p. 19.
- ⁷-Ibn Ḥuwaylī al-Akhḍar Mīdnī, al-Mu‘jamiyya al-‘Arabiyya fī ḍawī Manāhiḥ al-Baḥṭh al-Lisānī wa’l-Naẓariyyāt al-Tarbawiyya al-Ḥadītha, Dār Hūma, Algeria, 2010, p. 64.
- ⁸-Sharaf al-Dīn al-Rājihī, Fī ‘Ilm al-Lughā ‘inda al-‘Arab wa-Ra’y ‘Ilm al-Lughā al-Ḥadīth, Dār al-Ma‘rifa al-Jāmi‘iyya, Alexandria, 1st ed., 2001, p. 139.
- ⁹-Arabic Language Academy in Cairo, al-Mu‘jam al-Wasīṭ, Maktabat al-Shurūq al-Dawliyya, 4th ed., 2004, entry (‘-j-m).
- ¹⁰-‘Alī al-Qāsimī, ‘Ilm al-Lughā wa-Ṣinā‘at al-Mu‘jam, Dār al-Ma‘ārif, Saudi Arabia, 2nd ed., 1991, p. 3.
- ¹¹-‘Ayd Muḥammad al-Ṭayyib, al-Mu‘jamāt al-Lughawiyya wa-Dalālāt al-Alfāz, Dār al-Zahrā’ li’l-Naṣh wa’l-Tawzī‘, Riyadh, 1st ed., 2007, p. 8.
- ¹²-‘Abd al-Qādir ‘Abd al-Jalīl, al-Madāris al-Mu‘jamiyya — Dirāsa fī al-Bunya al-Tarkībiyya, Dār Ṣafā’ li’l-Naṣh wa’l-Tawzī‘, Amman, Jordan, 1st ed., 2010, p. 38.
- ¹³-Hartmann, R.R.K., & James, G. Dictionary of Lexicography, Routledge, London, 1998, p. 86.
- ¹⁴-Batūl ‘Abd al-Kāzim al-Rubī‘ī, al-Mu‘jamiyya al-‘Arabiyya fī Fikr al-Duktur ‘Alī al-Qāsimī, Markaz al-Kitāb al-Akādīmī, Amman, Jordan, 1st ed., 2018, p. 29.
- ¹⁵-‘Abd al-Raḥmān al-Ḥājj Ṣāliḥ, al-Mu‘jam al-‘Arabī wa’l-Isti‘māl al-Ḥaqīqī li’l-Lughā al-‘Arabiyya, Majallat al-Majma‘ al-Jazā’irī li’l-Lughā al-‘Arabiyya, no. 18, 2013, p. 14.
- ¹⁶-Ḥilmī Khalīl, Muqaddima li-Dirāsāt al-Turāth al-Mu‘jamī al-‘Arabī, Dār al-Nahḍa, Beirut, Lebanon, 1st ed., 1997, p. 103.
- ¹⁷-Emīl Ya‘qūb, al-Ma‘ājim al-Lughawiyya Binā’uhā wa-Taṭawwuruhā, Dār al-Malāyīn, Beirut, Lebanon, 2nd ed., 1985, p. 27.
- ¹⁸-Rajab ‘Abd al-Jawwād Ibrāhīm, Dirāsāt fī al-Dalāla wa’l-Mu‘jam, Dār al-Gharbī, Egypt, 1st ed., 2001, p. 148.
- ¹⁹-‘Abd al-Ḥamīd Muḥammad Abū Sukayn, al-Ma‘ājim al-‘Arabiyya — Madārisuhā wa-Manāhijuhā, Dār al-Fārūq li’l-Ṭibā‘a wa’l-Naṣh, Cairo, 2nd ed., 1981, p. 8.
- ²⁰-See: Fā’iz al-Dāya, ‘Ilm al-Dalāla al-‘Arabiyya bayna al-Naẓariyya wa’l-Taṭbīq, Dār al-Fikr, Damascus, Syria, 2nd ed., 1996, pp. 208–209.
- ²¹-Ḥilmī Khalīl, Muqaddima li-Dirāsāt al-Turāth al-Mu‘jamī al-‘Arabī, p. 21.
- ²²-Ibid., p. 22.
- ²³-Muḥammad ‘Alī ‘Abd al-Karīm Rudīnī, al-Mu‘jamiyyāt al-‘Arabiyya, Dār al-Hudā, Algeria, 2nd ed., 2006, p. 2.
- ²⁴-‘Abd al-Qādir ‘Abd al-Jalīl, al-Madāris al-Mu‘jamiyya, p. 38.

²⁵Aḥmad Mukhtār ‘Umar, *Şinā‘at al-Mu‘jam al-Ḥadīth*, p. 36.

²⁶Aḥmad ‘Abd al-Ghafūr ‘Aṭṭār, *Muqaddimat al-Şiḥāḥ*, Dār al-‘Ilm li’l-Malāyīn, Beirut, Lebanon, 3rd ed., 1984, p. 90.

²⁷*Ibid.*, p. 96.

²⁸Fahd Khalīl Zāyid, Muḥammad Şalāḥ Rummān, *al-Ma‘ājim wa’l-Dalāla*, Dār al-I‘şār, Amman, Jordan, 1st ed., 2015, p. 101.

²⁹‘Alī ‘Abd al-Wāḥid Wāfī, *Fiqh al-Lugha*, Dār Nahḍat Mişr, Cairo, Egypt, 9th ed., 2004, p. 277.

³⁰Ḥāmid Şādiq Qunaybī, Muḥammad ‘Arīf al-Ḥarbāwī, *al-Madkhal li-Maşādir al-Dirāsāt al-Adabiyya wa’l-Lughawiyya wa’l-Mu‘jamiyya*, Dār Ibn al-Jawzī, Amman, Jordan, 1st ed., 2005, p. 57.

³¹Muḥammad Rashād al-Ḥamzāwī, *al-Mu‘jamiyya — Muqāraba Naẓariyya wa-Muṭabbaqa Muşṭalaḥātihā wa-Mafāhīmiḥā*, Markaz al-Nashr al-Jāmi‘ī, Tunis, 1st ed., 2004, p. 269.

³²Ḥilmī Khalīl, *Muqaddima li-Dirāsāt al-Turāth al-Mu‘jamī al-‘Arabī*, p. 15.

³³Aḥmad Mukhtār ‘Umar, *Şinā‘at al-Mu‘jam al-Ḥadīth*, p. 54.

³⁴Muḥammad Rashād al-Ḥamzāwī, *al-Mu‘jamiyya*, p. 56.

³⁵Aḥmad Mukhtār ‘Umar, *Şinā‘at al-Mu‘jam al-Ḥadīth*, p. 58.

³⁶Adnān al-Khaṭīb, *al-Mu‘jam al-‘Arabī bayna al-Māḍī wa’l-Ḥāḍir*, Maktabat Lubnān Nāshirūn, Lebanon, 2nd ed., 1994, p. 59.

³⁷Aḥmad Mukhtār ‘Umar, *al-Baḥth al-Lughawī ‘inda al-‘Arab*, p. 161.

³⁸Maḥmūd Fahmī Ḥijāzī, *al-Uşş al-Lughawiyya li-‘Ilm al-Muşṭalaḥ*, Maktabat Gharīb, 1st ed., 1993, p. 11.

³⁹Adnān al-Khaṭīb, *al-Mu‘jam al-‘Arabī bayna al-Māḍī wa’l-Ḥāḍir*, p. 59.

Batūl ‘Abd al-Kāzim al-Rubī‘ī, *al-Mu‘jamiyya al-‘Arabiyya fī Fikr al-Duktur ‘Alī al-Qāsimī*, p. 29. 40-