THE DEVELOPMENT OF ARABIC LEXICOGRAPHY: FROM EARLY PHONETIC CLASSIFICATION TO MODERN LEXICOGRAPHIC STANDARDIZATION

Madina Nizomova, Master's student,
Department of linguistics: Arabic language,
Tashkent State University of Oriental Studies,
Tashkent, Uzbekistan.
madyikrm@gmail.com
+998908059282

Abstract: Arabic lexicography spans nearly thirteen centuries of continuous, self-aware methodological development. This paper traces that development through three successive macrostructural systems for organizing dictionary entries: an early phonetic-articulatory order, introduced in the eighth century by al-Khalil ibn Ahmad al-Farahidi's Kitab al-'Ayn, which arranged entries by the point of articulation of their initial sound; a rhyme-based (qafiya) order, in which entries were grouped by the final root letter to assist poets in locating rhymes, exemplified by al-Jawhari's al-Sihah, Ibn Manzur's thirteenth-century Lisan al-'Arab, al-Fayruzabadi's al-Qamus al-Muhit, and al-Zabidi's eighteenth-century Taj al-'Arus; and a strict alphabetical order organized by the first root letter, which became dominant from the nineteenth-century Nahda reform movement onward and underlies virtually all modern dictionaries, including the Cairo Arabic Language Academy's Al-Mu'jam al-Wasit (1960) and subsequent academy and commercial lexicons. The paper argues that each transition was driven not by arbitrary preference but by a change in the primary intended user and use-case of the dictionary: from grammatical-theological exegesis, to poetic composition, to mass literacy and rapid reference in an age of print journalism and, eventually, digital search. Understanding this trajectory clarifies why certain structural features persist in even

the most modern Arabic dictionaries and provides historical context for the ongoing transition toward digital and computational Arabic lexicography.

Keywords: Arabic lexicography, Kitab al-'Ayn, Lisan al-'Arab, dictionary macrostructure, root-based ordering, language academies

1. Introduction

Arabic lexicography is among the oldest continuously documented lexicographic traditions in the world, predating the great European monolingual dictionaries by many centuries. Unlike traditions that developed a single stable macrostructure early and retained it, Arabic dictionary-making passed through at least three distinct systems for organizing headwords, each dominant for several centuries before being displaced by the next. These systems are not merely technical curiosities: the choice of macrostructure reflects, in each period, a different assumption about who the dictionary's primary user was and what task the dictionary was meant to support.

This paper offers a historical-descriptive account of that evolution, organized around three stages identified in the specialist literature on Arabic lexicography: a phonetic-articulatory stage inaugurated by al-Khalil ibn Ahmad al-Farahidi in the eighth century; a rhyme-based (qafiya) stage that dominated the classical and medieval periods and produced the tradition's most celebrated works, including Ibn Manzur's Lisan al-'Arab; and a strict alphabetical stage that emerged alongside the nineteenth-century Nahda reform movement and became the near-universal standard for modern dictionaries, including those produced by twentieth-century Arabic language academies. A companion line of scholarship has examined the most recent phase of this trajectory, the shift from print to digital and online terminology resources; the present paper instead focuses on the earlier, print-era methodological transitions that set the stage for that later digital shift. The paper proceeds as follows: Section 2 describes the historical-comparative method and sources used; Section 3 presents the three stages in turn, with a comparative summary table; Section 4 discusses the forces driving each

transition and situates them within the broader history of Arabic language planning; Section 5 concludes.

2. Materials and methods

This paper uses a historical-comparative method common to studies in the history of lexicography, comparing the macrostructural organization, scale, and stated purpose of representative dictionaries across three historical stages, supported by both classical primary works and modern secondary scholarship. Primary sources include al-Khalil ibn Ahmad al-Farahidi's Kitab al-'Ayn (8th century), Ibn Manzur's Lisan al-'Arab (completed 1290), al-Fayruzabadi's al-Qamus al-Muhit (14th century), al-Zabidi's Taj al-'Arus (completed 1774), and the Cairo Arabic Language Academy's Al-Mu'jam al-Wasit (1960). Secondary scholarship includes Haywood's (1960) history of Arabic lexicography, Baalbaki's (2014) comprehensive study of the Arabic lexicographical tradition from the eighth to the eighteenth century, and a recent structural-historical survey of Arabic dictionary features covering phonetic, rhyme-based, and root-based ordering systems.

For each dictionary, four descriptive parameters were recorded where available: the ordering principle used for headwords (phonetic-articulatory, rhyme/final-radical, or strict alphabetical/first-radical); the approximate scale of the work in number of entries or volumes; the stated or inferable primary audience (grammarians and exegetes, poets and literary scholars, or general readers and students); and the institutional context of production (individual scholar, court-sponsored project, or language academy). These parameters were then compared across the three historical stages to identify the functional logic underlying each macrostructural transition.

3. Results

The reviewed material confirms a three-stage trajectory in the macrostructural organization of Arabic dictionaries, each stage associated with a distinct primary user and a distinct scale of undertaking.

3.1. Stage one: phonetic-articulatory ordering (8th century)

The first Arabic dictionary, Kitab al-'Ayn ("The Book of the Letter 'Ayn"), was compiled by the Basran grammarian al-Khalil ibn Ahmad al-Farahidi (c. 718-791), who is also credited with founding the study of Arabic prosody and introducing the standard system of Arabic vowel diacritics (harakat). Rather than arranging entries alphabetically in the order familiar today, al-Farahidi organized Kitab al-'Ayn according to the physiological place of articulation of each consonant, beginning with the deepest pharyngeal sound, 'ayn (ع) — from which the work takes its name — and proceeding outward through the vocal tract to the labial consonants pronounced at the lips. Within this phonetic framework, al-Farahidi further organized entries by root pattern (reduplicated bilateral roots, sound trilateral roots, weak trilateral roots, quadrilaterals, and so on), reflecting his broader theoretical interest, evident also in his grammatical and prosodic work, in the systematic combinatorial structure of the Arabic language. This ordering system, sometimes described in the specialist literature as "anagrammatical," was explicitly designed less as a quick-reference tool than as an exposition of the language's underlying phonological and morphological system, consistent with its primary audience of grammarians and religious scholars engaged in Qur'anic exegesis. Virtually all subsequent Arabic lexicographic works, even those that abandoned phonetic ordering, drew on Kitab al-'Ayn's foundational insight that entries should be organized by consonantal root rather than by the surface form of individual inflected words.

3.2. Stage two: rhyme-based (qafiya) ordering and the classical tradition (10th-18th centuries)

Beginning around the tenth century, Arabic lexicography shifted toward an ordering system built around the poetic needs of its users. Al-Jawhari's al-Sihah introduced a structure in which headwords were grouped first by their final root radical and, within that group, by their initial radical — the reverse of ordinary alphabetical intuition, but

highly practical for a poet searching for words that would rhyme, since Arabic poetry is organized around a shared end-rhyme (qafiya) across a poem's lines. This rhyme-based macrostructure was adopted by the two most influential dictionaries of the pre-modern period. Ibn Manzur's Lisan al-'Arab, completed in Cairo in 1290 and comprising roughly eighty thousand entries compiled from earlier sources such as al-Azhari's Tahdhib al-Lugha and Ibn Sidah's al-Muhkam, organized each final-letter "section" into sub-classes by initial letter, explicitly continuing al-Jawhari's rhyme-oriented model. Al-Fayruzabadi's fourteenth-century al-Qamus al-Muhit ("The Encompassing Ocean") likewise followed rhyme-based ordering while compressing definitions considerably relative to Lisan al-'Arab, making it a more portable working reference. The tradition culminated in Murtada al-Zabidi's Taj al-'Arus min Jawahir al-Qamus, an eighteenth-century expansion of al-Qamus al-Muhit begun in 1760 and completed in 1774 after fourteen years of work, drawing on some 115 consulted sources and comprising approximately 120,000 entries — the largest dictionary in the history of the Arabic language.

Across this six-century stage, the intended user shifted from the grammarian-exegete of the phonetic era to the poet, literary scholar, and later the encyclopedic compiler, and dictionary scale grew substantially, from Kitab al-'Ayn's more modest scope to Taj al-'Arus's 120,000 entries. This growth reflects both the accumulation of citations from an expanding corpus of classical poetry and prose and the increasingly encyclopedic ambitions of later compilers, who sought not merely to record vocabulary but to preserve, cross-reference, and correct the entire preceding lexicographic tradition.

3.3. The critique of rhyme ordering and the shift toward alphabetical access

By the nineteenth century, rhyme-based ordering had become a significant practical obstacle for a readership whose primary need was no longer poetic composition but rapid reference in an expanding world of print journalism, translation, and modern education. The specialist historical literature on Arabic lexicography identifies this

period as the site of open "rebellion" against qafiya ordering, associated with the Nahda-era reformer and journalist Ahmad Faris al-Shidyaq — the same figure credited, alongside Butrus al-Bustani, with helping standardize the modern Arabic press register. The core objection was straightforward: locating a word by its final rather than initial letter requires the reader to already know the word's full root, defeating the purpose of a reference tool for the very words a reader is least likely to know. This critique set the stage for the third and current macrostructural regime.

3.4. Stage three: strict alphabetical ordering and the language-academy dictionary (20th century-present)

The twentieth century saw strict alphabetical ordering by initial root letter become the near-universal standard, adopted both by Arabic language academies and by Orientalist lexicographers producing bilingual dictionaries such as Hans Wehr's Dictionary of Modern Written Arabic (1961). The Academy of the Arabic Language in Cairo, founded in 1932, undertook a coordinated, multi-tier dictionary program reflecting this shift toward accessibility: Al-Mu'jam al-Wasit, published in two volumes in 1960 by a committee including Ibrahim Mustafa, Ahmad al-Ziyat, Hamid Abd al-Qadir, and Muhammad al-Najjar, was explicitly designed to give general readers and students clear, accessible entries organized primarily by root but with problematic derived forms cross-referenced alphabetically; the more concise Al-Mu'jam al-Wajiz followed in 1980; and the encyclopedic, historically oriented Al-Mu'jam al-Kabir, planned since the Academy's founding and begun in earnest in 1946 with its first published volume (covering the letter hamza) appearing in 1956, remains in progress today, illustrating that the ambition for a comprehensive historical dictionary on the scale of Taj al-'Arus persists even within the modern academy system. Later single-author works such as Ahmad Mukhtar Omar's Mu'jam al-Lugha al-'Arabiyya al-Mu'asira (2008), a 3,367-page dictionary of contemporary Arabic, continued this alphabetical, accessibility-oriented model while updating vocabulary

coverage to include modern scientific, technical, and media terminology of the kind discussed in the broader literature on Arabic terminology planning.

3.5. Comparative overview

Table 1. Representative Arabic dictionaries across three macrostructural stages

Dictionary	Author / institution	Date	Ordering principle	Approx. scale
Kitab al-'Ayn	Al-Khalil ibn Ahmad al-Farahid	8th century	Phonetic-articulatory	First comprehensive Arabic lexicon
Al-Sihah	Al-Jawhari	10th century	Rhyme (final radical)	—
Lisan al-'Arab	Ibn Manzur	Comple ed 129(	Rhyme (final radical)	~80,000 entries, 20 vols.
Al-Qamus al-Muhi	Al-Fayruzabadi	14th century	Rhyme (final radical)	Compact, portable
Taj al-'Arus	Murtada al-Zabidi	1760-1774	Rhyme (final radical)	~120,000 entrie
Al-Mu'jam al-Wasi	Academy of the Arabic Language, Cairo	1960	Strict alphabetical (first radical)	2 volumes
Mu'jam al-Lugha a 'Arabiyya al-Mu'asira	Ahmad Mukhtar Omar	2008	Strict alphabetical	3,367 pages

4. Discussion

The three-stage trajectory traced above is best understood not as a sequence of arbitrary editorial fashions but as a series of adaptations to changing readerships and

use-cases. Al-Farahidi's phonetic ordering in Kitab al-'Ayn served an eighth-century audience of grammarians and religious scholars for whom the dictionary was as much a theoretical exposition of Arabic's phonological system as a reference tool; its organizing logic mirrors the same systematic, combinatorial thinking that produced al-Farahidi's contributions to prosody and vowel notation. The centuries-long dominance of rhyme-based ordering reflects the centrality of poetry to classical and medieval Arabic intellectual life: a macrostructure that appears deeply counterintuitive to a modern reader was, for a poet searching for a rhyming word, the single most useful possible arrangement. The eventual rebellion against that system, associated with Nahda-era figures such as al-Shidyaq, coincided precisely with the same nineteenth-century expansion of print journalism, mass education, and translation activity that, as discussed in the broader literature on Arabic media and language planning, also drove the standardization of Modern Standard Arabic as a shared written register. A rhyme-ordered dictionary is simply incompatible with the reference speed a journalist on deadline or a student consulting an unfamiliar word requires; strict alphabetical ordering by initial letter is not a self-evidently "better" system in the abstract, but it is the system best matched to rapid, deadline-driven, or self-directed lookup.

This functional reading of the historical transitions has a direct bearing on the current, still-unfolding shift toward digital and online Arabic lexicographic resources. Just as the nineteenth-century shift to alphabetical ordering was driven by a change in primary use-case — from poetic composition to rapid mass reference — the ongoing shift toward searchable digital databases and terminology platforms is driven by a further change in use-case, from human page-by-page lookup to machine-queryable, programmatically searchable data intended for use in translation software, natural-language processing, and, increasingly, artificial-intelligence systems. In this sense, the digital transition documented elsewhere in the literature on Arabic terminology databases can be read as a fourth macrostructural stage, continuous with the three

traced here rather than a wholly novel departure: each stage answers the same underlying question — how should the language's vocabulary be organized to serve its readers — differently, because each stage's readers are different.

This study has clear limitations. As a historical-comparative synthesis rather than a systematic corpus study, it relies on a small number of canonical, well-documented dictionaries rather than an exhaustive survey of the many regional and lesser-known works produced across thirteen centuries; a fuller account would incorporate specialized and bilingual dictionaries, dialectal lexicons, and non-Cairene academy dictionaries (Damascus, Baghdad, Rabat) in more detail. In addition, the three-stage framework, while well supported in the specialist literature, necessarily simplifies a more gradual and regionally uneven process; strict alphabetical ordering, for instance, did not displace rhyme-based ordering everywhere at the same pace, and specialized or regional dictionaries sometimes retained older conventions well into the twentieth century.

5. Conclusion

Arabic dictionary-making has passed through three major macrostructural regimes over nearly thirteen centuries: an eighth-century phonetic-articulatory system designed for grammarians and exegetes; a rhyme-based system, dominant for roughly six centuries and responsible for the tradition's most monumental works, designed for poets and literary scholars; and a strict alphabetical system, emerging from nineteenth-century Nahda-era reform pressure and institutionalized by twentieth-century language academies, designed for the mass reading public of a print and, later, digital age. Each transition was driven by a change in who the dictionary was for and what task it needed to support, rather than by any inherent superiority of one ordering principle over another. Viewed in this light, the current shift toward searchable digital and machine-readable Arabic terminology resources is not a break from this history but its logical continuation — a fourth stage answering the same enduring question of Arabic

lexicography in terms appropriate to a new class of readers, both human and computational.

References

Al-Fayruzabadi, M. ibn Y. (14th c./2005). Al-Qamus al-Muhit [The encompassing ocean] (Rev. ed.). Mu'assasat al-Risala.

Al-Khalil ibn Ahmad al-Farahidi. (8th c.). Kitab al-'Ayn [The book of the letter 'Ayn]. Dar al-Kutub al-'Ilmiyya. (Modern ed. 2003)

Academy of the Arabic Language in Cairo. (1960). Al-Mu'jam al-Wasit [The intermediary dictionary] (Vols. 1-2).

Baalbaki, R. (2014). The Arabic lexicographical tradition from the 2nd/8th to the 12th/18th century. Brill.

Haywood, J. A. (1960). Arabic lexicography: Its history and its place in the general history of lexicography. Brill.

Ibn Manzur, M. ibn M. (1290/1997). Lisan al-'Arab [The tongue of the Arabs] (Reprint, 15 vols.). Dar Sadir.

Omar, A. M. (2008). Mu'jam al-lugha al-'Arabiyya al-mu'asira [Dictionary of contemporary Arabic language]. Alam Al-Kutub.

Wehr, H. (1979). A dictionary of modern written Arabic (J. M. Cowan, Ed., 4th ed.). Harrassowitz.

Al-Zabidi, M. (1774). Taj al-'Arus min jawahir al-Qamus [The bride's crown from the jewels of al-Qamus].