

A study of Arabic lexicons, their compilation methods, and development

Mohammad Hassan Edrisi^{1*} and Gholamhossein Sharifi²

1- PhD Candidate of Literature, Isfahan University, Isfahan, Iran

2- Faculty Member, Isfahan University, Isfahan, Iran

Corresponding author: Mohammad Hasan Edrisi

ABSTRACT: A Lexicon is a limited collection of words or lexical unites of a language or a number of languages (bilingual or multilingual dictionaries) which have been chosen with a certain arrangement. The order and arrangement with which the lexicographer choses to present the dictionary leads to the formation of the structure of the dictionary. A challenge of the lexicographer in all languages is that there is no one universal method for this occupation. The method the lexicographer choses depends entirely on the vocabulary presented in their lexicon, and their goal and readers. For example, the method of lexicographers who organize vocabulary based on meaning differs completely from that of lexicographers who organize vocabulary based on their last letter, which is for the use of poets and prosaists. The present essay studies Arabic lexicon categorization and classifies them into 8 categories: irregular order, classified order, topical order, semantic order, syntactic order, etymological order, phonetic order, and alphabetical order.

Keywords: Arabic dictionaries, Lexicography, Arabic language and literature, Bilingual dictionary, Dictionary, Vocabulary, Persian vocabulary.

INTRODUCTION

Arabic lexicographers do not, as yet, have a standardized technique of lexicography. Thus, different methods of Arabic lexicography have been studied in the present paper. In order to study this topic and to achieve an optimal method for Arabic lexicography for non-Arabic speakers, especially Farsi speakers, it is necessary to study Arabic lexicography methods used during the course of Arabic literary history.

By studying the different methods of Arabic lexicography we can recognize 8 lexicography schools. These 8 lexicography schools consist of: irregular order, classified order, topical order, semantic order, syntactic order, etymological order, phonetic order, and alphabetical order.

Different Methods of Word Order and Classification in Arabic Dictionaries

Arabic lexicography schools, as mentioned above, are divided into 8 categories as follows:

1. Irregular order

In this method, words are entered into the dictionary without a certain system and are place one after the other by random. We do not imagine that any modern lexicographers will use this method. As an example of this method we can name the Kitāb al-jīm² by Ishāq ibn Mirār Abū 'Amr al-Shaybānī, deceased in 822. Although Shaybānī divided his Kitāb al-jīm into different chapters, each of which is dedicated to one letter of the alphabet, he has not adhered to any order within the chapters. Therefore, to look for a word beginning with Ā, one has to look through all entries in the chapter on Ā. However, John Haywood believes Shaybānī's Kitāb al-jīm to be based on a particular system which we have been unable to discover as yet.

2. Classified order

This method is specific to lexicons related to a certain book or text. The order of entries in these lexicons is based on the order of the words within the text the lexicon is related to.

This method has mostly been used in books on interpretation of the Quran. To our knowledge, the use of this method dates back to the 8th century. The *Gharīb al-Qur'ān* by Ibn Qutaybah (828–885 CE) may be the oldest example of this type to still exist. This book is divided into three sections; the first and second sections are of irregular order, but in the third section, words are ordered based on the order of the Sura of the Qur'ān they are found in.⁴ This method has also been used by some modern "Gharīb al-Qur'ān" writers.

3. Topical order

This method is divided into two categories of specific and general.

3.1. Specific topical lexicons

These lexicons are assigned to a specific scientific subject. In Arabic literature, these lexicons are recognized as individual essays. The best example of this type of lexicon is the book by Abd al-Mālik al-Aṣma'ī (740-828). He has compiled essays on different topics, such as *Kitāb al-khayl* (The Book of the Horse), *Kitāb al-Ibil* (The Book of the Camel), *Kitāb al-farq* (The Book of Rare Animals), and *Kitāb al-Wuḥūsh* (The Book of Wild Animals). Only 7 of these essays remain today.

In this method, the author divides the terminology based on their subtopics, and then, orders the scientific concepts. For instance, in *Kitāb al-Ibil* he writes:

(المرشح) فإذا القت (الناقه) و لدها فهو ساعه يقع (سليل) فإذا وقع عليه اسم التذكير و التأنيث، فان كان ذكراً فهو (سقب) و ان كان انثى فهو (حائل) فإذا قوي و مشي فهو (راشح) و هي [(وهي (المطفل) مادام ولدها صغيراً فإذا ارتفع عن الرشح فهو (الجادل

As can be observed, these terminology or words are not given in alphabetic or etymological order, but based on subject (subtopics related to main topics).

3.2. General topical lexicons

The different sections of these books are compiled based on general topics and this requires a classification that includes all the various concepts of the universe. Thus, no limit can be set for it, although attempts have been made. Abū 'Ubayd al-Qāsim ibn Sallām was one of the first to use this method in his book 'al-Gharīb al-muṣannaf'. He divided his book into 25 main topics, each of which are divided into subtopics comprising a total of 900 topics.⁷ Many other lexicographers have been influenced by Abū 'Ubayd, the most renowned being Ibn Sidah's (1007–1066) *al-mukhassas fi al-lughah*. In order to use topical lexicons, the reader must know the main topic of the word under study, and then, to study the section related to that topic entirely to find the intended word.

4. Semantic order

In this method, words are categorized based on their semantic background. Each entry consists of all words which semantically related to the main entry. Moreover, the words in each section are not alphabetically ordered, but semantically ordered. Therefore, they are ordered based on their closeness in meaning to the main entry. The main aim of authors of semantically ordered lexicons is not to explain words, but to give synonyms, words of close meaning to the main entry, or words that come to mind when thinking of the main entry.

Differences of semantically ordered and topically ordered lexicons

Topical and semantic lexicons divide their material based on their subject. In topical lexicons, a word is explained completely through its derivatives and Arabic syntax; however, in semantic lexicons, synonyms, antonyms, and common terms of the entry and some of its idiomatic applications are given.

The *Kitāb al-alfāz al-kitābīyah* by 'Abd al-Raḥmān ibn 'Īsā Hamadhānī is one of the first semantic dictionaries of the Arabic language. On the word «السنة» he writes:

يقال السنة و الحول و العالم و الحجه وفي القرآن الكريم

و عام الاول⁸ و يقال السنه و تجرمت و انقضت و يقال كان ذلك عاماً اولاً، ثماني حجج، وفيه: يحلونه عاماً، و فيه حولين كاملين.

5. Syntactic order

In this method, the author organized the entries in the lexicon based on their morphological and syntactic structures. When compiling the entries, the author considers whether a word is a noun or a verb, if it is a verb whether it is an intransitive or transitive verb, and if transitive whether it consists of 4 letters or 3, and thus, in this way the

entries are organized in a syntactic order. An example of this kind of lexicon is the *Dīwān al-adab* by Ishāq ibn Ibrāhīm al-Fārābī (962). This book is divided into 6 chapters:

1- Ketāb Salem 2- Ketāb Mozāaf 3- Ketāb Mesāl 4- Ketāb Ajvaf 5- Ketāb Nāghes 6- Ketāb Hamza

Each book is divided into 2 parts; the first part specific to nouns and the second to verbs. Each section of these 2 parts is divided into other sections based on structure. For Example: (فَعْلٌ و فَعْلَانٌ و فَعْلَانٌ) have been separated into different sections.⁹

In using a syntactic lexicon, one must determine the identity of a word in respect to its gender and syntactic category and then to look it up in the lexicon.

6. Etymological order

In this method, the lexicographer sees terms as a family in which the words are born of the same origin. In this method, entries are organized based on their etymology and derivatives are given under their main root.

Here, lexicographers are faced with the issue of the ordering of derivatives of an entry. This is an issue of etymological lexicons, because no one specific rule for ordering of derivatives exists in Arabic lexicons. Most lexicons are ordered based on morphology.

By looking at the order of derivatives in *al-Mu'jam al-Wasit*, the difficulty of the issue is better understood. The method of ordering present below is the method used by the authors of *al-Mu'jam al-Wasit*; however, other methods have been used in some other etymological dictionaries.

1. Precedence of verbs over nouns
2. Precedence of singular over augmented in verbs
3. Precedence of intransitive over transitive verbs
4. The order of Verbs is as follows:

4.1. Singular verbs based on trilateral roots

فعل يفعل (ضرب يضرب).....فعل يفعل (نصر ينصر)

فعل يفعل (علم يعلم).....فعل يفعل (فتح يفتح)

فعل يفعل (حسب يحسب).....فعل يفعل (شرف يشرف)

4.2. Augmented verbs based on trilateral roots are ordered alphabetically

4.2.1. Augmented by one letter:

1-افعل 2- فاعل 3-فعل 1

4.2.2. Augmented by two letters:

1-افعل 2- انفعال 3- تفاعل 4- تفاعل 5- افعل 1

4.2.3. Augmented by three letters:

1-استفعل استغفر 2- افوعول (اعشوشب) 3- افعأل (احمأز) 4- افوعأل (احلوذ)

In this method, nouns are alphabetically ordered in the lexicon.¹⁰

7. Taghlibi

Order of roots based on the taghlib of roots is a method that al-Khalīl ibn Aḥmad al-Farāhīdī invented and used in his dictionary, *Kitāb al-'Ayn*. This is a system with mathematical precision the goal of which is the collection of all the terms that constitute the Arabic language. This method is effective in the recognition of the great wealth of words of the Arabic language. In this method, from a trilateral root, root with three consonants, for instance ع.ر.ب, 6 other items can be gained [(ع.ر.ب) (ع.ب.ر) (ع.ر.ب) (ب.ع.ر) (ب.ر.ع) (ب.ع.ر)], from a quadrilateral root, roots with four consonants, 14 items can be gained, and from a quinqueliteral root, five root-consonant forms, 120 items. Some of these items are used and others are nonsensical; al-Khalīl ibn Aḥmad informs the reader of the nonsensical items. However, he based this lexicon on explaining used word.

Most researchers have only referred to the phonetic order, which differs from alphabetic order in Arabic, of words in *Kitāb al-'Ayn*. However, the order of words in *Kitāb al-'Ayn* is of secondary importance. The *al-'Ayn* lexicon is based on the taghlib and variation system in the main root of the word. For instance the entry (ع) consists of all words which contain this letter in their root, and (ح) all words which contain this letter in their root except for those given under (ع). This method is based on the theory that the meanings of derivatives of a root are somewhat similar, and thus, are given in the same category. This is what ibn Jinnī named the greatest derivation (الاشتقاق الأكبر).¹¹

It was this theory that caused al-Khalīl ibn Aḥmad to categorize variations on word roots as one item. Some other great lexicographers have also used this method; Abū 'Alī Ismā'īl ibn al-Qāsim Qālī (901-967) in *Kitāb al-Bārī*, Abū

al-Qāsim Ismā'īl ibn 'Abbād Ṣāhib al-Ṭālqānī (died 995) in al-Muḥīṭ fī al-lughah, and Ibn Sīdah (1007–1066) in the al-Muḥkam wa-al-Muḥīṭ al-A'ẓam.

8-Phonetic order

Today, researchers in the field of lexicography use terms such as abjad, alphabet, phonemes, and mu'jam in the order of their books; however, all of these words are of the same meaning.¹²

Phonemes are that which organizes the writing system of each language. Therefore, when the lexicographer categorizes his lexicon based on the alphabet, it can be said that mu'jam is the same as the phoneme, and the alphabet is organized in 3 main methods in the Arabic language.

8.1. Abjad Order

This order is derived from the writing system of the Phoenician alphabet, which was used by all Semitic languages in the Arabian Peninsula, and Syria and Iraq. Its order is as follows:

ا، ب، ج، د، هـ، و، ز، ح، ط، ي، ك، ل، م، ن، س، ع، ف، ص، ق، ر، ش، ت، ث، خ، ذ، ض، ظ، غ.

The Arabs added the letters which were not in the Phoenician alphabet (ث، خ، ذ، ض، ظ، غ).

Although the abjad order is used for teaching in Quranic schools in many Arab and Islamic countries, this method has not been used in any Arabic lexicon and the alphabetic or sound order of letters have been used.¹³

8.2. Alphabetical order

The Arabic alphabet did not implement any diacritical marks (e.g., dots) until the time of Mu āwiya ibn Abu Sufyān. Abu al-Aswad al-Du'alī was the first person to use the harakat (short vowel marks) of fathah, kasrah, and dammah in the form of dots above and below the letters. During the rule of Abd al-Malik ibn Marwān, Nasr bin 'Asim al-Laithī (died 1494) used i'jam (phonetic distinctions of consonants) to differentiate between consonants that looked similar in their written forms (e.g., ب ت ث، يا ج ح، ح خ). This was called the "alephba" based on the first two letters "الف" and "ب". The order of the Arabic alphabet is as follows:

ا . ب . ت . ث . ج . ح . خ . د . ذ . ر . ز . س . ش . ص . ض . ط . ظ . ع . غ . ف . ق . ك . ل . م . ن . هـ . و . ي .

This is the order used until today by lexicographers.

8.3. Phonetic order

When al-Khalīl ibn Aḥmad al-Farāhīdī compiled his lexicon the Kitāb al-'Ayn, the abjad and alphabetic order was known and popular. However, he invented the method he used in his lexicon, which is based on phonetic pattern [pronunciation from the deepest letter of the throat (ع) to that pronounced by the lips (م)]. This pattern is as follows:

ع . ح . هـ . خ . غ . ق . ك . ج . ح . هـ . ض . ص . ط . ظ . ع . غ . ف . ق . ك . ل . م . ن . هـ . و . ي .

Some lexicographers have been influenced by this method. An example is 'Alī Ismā'īl ibn al-Qāsim al-Baghdadī, known as Qālī, who modified this method in his lexicon the al-Bare. He changed it to the following order:

هـ . ح . غ . ق . ك . ح . ض . ج . ش . ل . ر . ن . ط . هـ . ص . ز . ظ . ذ . هـ . ف . م . و . ي .

Of other lexicographers who used this method, Abū Manṣūr Muḥammad ibn Aḥmad al-Azharī (died 980) in Tahdhīb al-lughah, al-Sahib ibn Abbad ibn Abbas (died 995) in al-Muheet fī al-lughah, and Ibn Sīdah in al-Muḥkam wa-al-Muḥīṭ al-A'ẓam can be named.

Thus, it can be concluded that phonetically ordered Arabic lexicons are organized based on phonemes, these phonemes may have occurred historically (Abjad), or based on sound (phonetics) or shape (alphabet).

Phonetic order is a common noun consisting of three types (Abjad, alphabetic, phonetic), only one of which is used in practice.

Phonetic ordered lexicons may be organized based on the first letter, last letter, or first and last letters of a word. Therefore, there are three types of phonetically ordered lexicons.

8.3.1. First order

In this method, the entries are organized based on the first letter of words. Thus, all words beginning with the letter ع are categorized in one chapter, and in this chapter entries are ordered based on the second letter, third letter, fourth letter, and etcetera.

8.3.2. Last order

In this method, also named the rhythmic order, words are organized first based on their last letter, and then, on the first and second letters. For example, in the chapter *الباء* the order of entries is as follows:

ايب ، آيب ، ادب ، ارب ، ارب ، ارب ، اسب ، اشب ، ...

The origin of this method can be found in the *Dīwān al-adab* by Ishāq ibn Ibrāhīm al-Fārābī.

8.3.3. First and last order

In this method, the words are organized based on their first and last letters without considering the letters in between them.

The only example of the use of this method is the *Tuhfat al-arīb bī-mā fī al-Qur'ān min al-gharīb* by Abū Ḥayyān Muḥammad ibn Yūsuf ["al-Nahwi" (the grammarian)]. This book is a small lexicon, in which the lexicographer has entered strange words of the Quran without consideration of their order of entry in the suras, but based on their first and last letters. For instance, in the word *خلاق*, the following are given:

...حسأ ، خبا ، خطب ، خبث ، خرج ، خلد ، خلد ، خمد ، خمد ، خضد

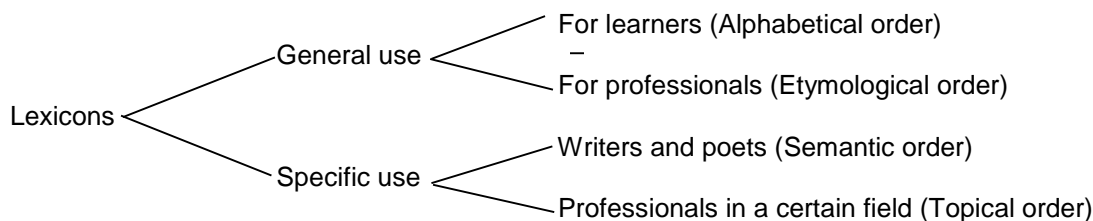
The Combining of different methods

The categorization of lexicons into 8 different categories is mainly a theoretical categorization and it cannot be stated that these methods of ordering are of the same amount or importance in lexicons.

Research shows that some of these methods cannot be used alone to organize the entries of a lexicon, but must be used in combination with other methods. For example, the syntactic order cannot be the basis of organizing a lexicon alone, because syntactic order is within a closed circle in which every word cannot be placed. Categorization based on syntax is limited and hundreds of words are placed in one section; therefore, another method is needed to order the words within each section. Thus, we have no other solution but to use the phonetic order method to organize words in each section.

This is also true for topical ordered lexicons; the words categorized under each topic must also be organized using a second method. The alphabetic order is the only method that can be used alone in Arabic lexicography.

Each of the abovementioned methods can be suitable in different circumstances. However, if we are to identify the best method, we can categorise lexicons from a more general viewpoint and based upon the aim due to which they have been compiled. In respect to purpose and aim, lexicons are divided into two groups of general and specific use. Therefore, lexicons can be categorized in a summarized manner as follows:



The best method

The best method of compiling lexicons in the Arabic language is the alphabetical order. However, this method, like others, has its advantages and disadvantages.

1. Advantages

Although the alphabetic order method separates a word from its derivatives, it helps the reader find their desired word without difficulty. This is the best method of lexicography for learners of Arabic as a foreign language. The lexicographer should provide the root of each word in parentheses in front of the word. It may also be best to provide the main derivatives of words beneath them without their meaning in order to help learners better distinguish them. For example, the word *مكتبة* can be thus written:

[... كاتب ، كتب ، مكتوب ، كتابه [مكتبه : (ك ، ت ، ب)

2. Disadvantages

The alphabetical order causes issue on which lexicographers are not in agreement. Therefore, this has been the cause of much controversy. Some of these issues are discussed below.

1- Lack of consideration of the order of harakat (dammah, fathah, and kasrah) and sukun in homophones and homographs (such as كَتَبَ، كَتَب).

2-Lack of consideration of the precedence of written hamza to letters (such as: (أ، ؤ، ث) and whether these letters (الف (ـ) should be considered as letters or hamza.

3-The issue of the letter (هـ); should it be considered as (هـ) or as tahghigh (تحقيق) in words such as كتابه و شجاعه و رفعه.

CONCLUSION

It seems that, today, the best method for Arabic lexicography is the alphabetical order, because it can be used by both beginners and professionals of Arabic. In this method, the alphabetical letters are ordered based on similarities in their shape; letters of the same shape such as (ب، ت، ث), (ر، ز), (س، ش), and (ع، غ). Moreover, the order of short harakat depend on the order of long harakat (ي، و، ؤ); meaning sukun, fathah, dammah, kasrah, and hamza are ordered based on the order of their letters (ي، و، ؤ), for instance (شأم، شام), (شؤم، شوم), and (شئم، شيم), and alif maqṣūrah (ألف مقصورة, 'limited/restricted alif') precedes the letter ي. Furthermore, letters with tashdid (consonant gemination mark) succeed letters without tashdid.

Notes

1-In the categorizing of lexicons based on the method used by each lexicographer, they have been divided into 4 schools and each has been attributed to their founders 1-the school of Khalil based on the articulators of letters 2-the school of Abū 'Ubaidah based on lexical topics, 3- the Jawharī school based on the last letter, and 4- the Barmakī school based on the alphabet.

(For further information refer to: Imīl Ya'qūb, al-Ma'ājim al-lughawīyah al-'Arabīyah: bidā'atuhā wa-taṭawwuruhā, Dār al-'Ilm lil-Malāyīn, Bayrūt, Lubnān, P. 15)

2-The name of Kitāb al-jīm, unlike Kitāb al-'Ayn by al-Khalīl ibn Aḥmad al-Farāhīdī, does not take its name from the letter it begins with, but has been likened to ج, an abbreviation of ج meaning silk garment (for further information refer to Sanaa al-Ma'ājim al-'Arabīyah li-ghayr al-natiqīn al-'Arabīyah by Ali al-Ghasemī, P. 25)).

REFERENCES

- Al-Zabīdī MIMM. 2001. Tāğ al-'Arūs min Ġawāhir al-qāmūs, Maṭba'at ḥukūmat, Kuwait.
- Al-Ṭālqānī AGIIAS and Āl Yāsīn MH. 1994. al-Muḥīṭ fī al-lughah, 'Ālam al-Kutub, Beirut, Lebanon.
- Al-Farāhīdī KHIA. 1993. Kitāb al-'Ayn, ed. Fīrūzābādī1975., Qom, Iran.
- Al-Bayhaqī AJ. 1988. Taj al-Masader, Ed. Hadi Alamzadeh, Volume 1, Institute for Cultural Studies and Research, Tehran, Iran.
- Al-Azdi MIHID. 1987. Jamhara fi 'l-lughah, Dār al-Elm, Beirut, Egypt.
- Al-Shaybānī IMAA and Ibyārī I. 1975. Kitāb al-jīm, Al-Hay'ah al-'Āmmah li-shu'ūn al-Maṭābī' al-Amīrīyah, Cairo, Egypt.
- Al-Tawwāb RA. 1973. Fuṣūl fī fiqh al-'Arabīyah, Maktabat al-turāth, Cairo, Egypt.
- Al-Daqqāq U. 1972. Maṣadir al-turāth al-'Arabī fī al-lughah wa-al-ma'ājim wa-al-adab wa-al-tarājim, Maktabat Dār al-Sharq, Beirut, Lebanon.
- Al-Samī MAA. 1969. al-Ma'ājim al-'Arabīyah: dirāsah taḥlīlīyah, Dār al-Fikr al-'Arabī, Cairo, Egypt.
- AL-Khasā'is AFOIJ. 1956. 2 vols., Dār al-Kutub al-Misriyya, Cairo.
- Al-Azhari Ebiary I.1924. Tahzib al-Lughah, 1924, Dar al-ghomia al-Arabiya, Cairo, Egypt.
- Fīrūzābādī MIY and Wafā'ī M. 1994. al-Qāmūs al-muḥīṭ, Dār al-Fikr, Beirut, Lebanon.
- John A. 1965. Haywood, Arabic Lexicography: Its History, and Its Place in the General History of Lexicography, Brill Publishers, Leiden, Netherlands, P. 96.
- Hamadhānī ARII. 1885. al-Alfāz al-kitābīyah, Maṭba'at al-Ābā' al-Yasū'īyīn, Beirut, Lebanon, 1885
- Ibn Manzūr. 1993. Lisān al-'Arab, 18 vol., Dār Iḥyā' al-Turāth al-'Arabī: Mu'assasat al-Tārīkh al-'Arab, Beirut, Lebanon.
- Ibn Sallām AUQ. 1989. al-Gharīb al-muṣannaf, Ramaḍān 'Abd al-Tawwāb, 1980, Maktabat al-Thaqāfah al-Dīnīyah, Cairo, Egypt.
- Ibn Qutaybah. 1978. Tafsi'r Gharīb al-Qur'an, Dār al-kutub al-'ilmīya, Beirut, 1978
- Ibn Sīdah. 1958. al-Muḥkam wa-al-Muḥīṭ al-'Aẓam, Muṣṭafa Saqqā, Ma'had al-Makhtuṭāt bi-Jāmi'at al-Duwal al-'Arabīyah. Cairo, Egypt
- Naṣṣār H. 1968. al-Mu'jam al-'Arabī: nash'atuhu wa-taṭawwuruh, Dār Miṣr lil-Ṭibā'ah, Cairo, Egypt.
- Qāsimī A. 1981. Ṣinā'at al-mu'jam al-'Arabī li-ghayr al-nāṭiqīn bi-al-'Arabīyah, Jāmi'at al-Duwal al-'Arabīyah, Rabat, Morocco.
- Qāsimī A. 1975. 'Ilm al-lughah wa-ṣinā'at al-mu'jam, Jāmi'at al-Riyāḍ, Riyadh, Saudi Arabia, 1975
- Ya'qūb I. 1985. al-Ma'ājim al-lughawīyah al-'Arabīyah: bidā'atuhā wa-taṭawwuruhā, Dār al-'Ilm lil-Malāyīn, Beirut, Lebanon.
- Zamakhsharī MIU. 2007. Asās al-balāghah, ed. 'Abd al-Raḥīm Maḥmūd, Dār al-Kutub al-Misriyya, Cairo, Egypt.