

The CJKI Arabic Learner's Dictionary

القاموس العربي الإنجليزي للمتعلمين

Jack Halpern (春遍雀來) يعقوب هالبرن

Introduction

The primary goal of this new dictionary is to enable beginner and intermediate learners to fully understand the core vocabulary of Modern Standard Arabic (MSA). A host of unique features, including the CARS romanization system, clear meanings, illustrative examples, and detailed grammatical information, enable learners to quickly access a wealth of useful information to help them gain a deep understanding of the meanings and usage of the most frequently used Arabic words and phrases.

Why a new Arabic dictionary?

Arabic is the sixth most widely spoken language in the world. Because of the increasingly important role it plays in world affairs, the number of Arabic learners worldwide is rapidly increasing. However, the lack of effective learner's dictionaries puts these learners at a severe disadvantage compared with those of other major world languages such as French and Spanish. Existing dictionaries often suffer from the serious shortcomings described below, making them inadequate for learners.

- Dictionaries are designed for native speakers and are too advanced and complicated for learners to use.
- Poor design, small fonts, and unclear type decrease ease of use.
- The lack of romanization makes it difficult to determine the correct pronunciation.
- Equivalents are overly formal or stilted, are inaccurate, or do not reflect contemporary usage.
- Phrasal verbs and fixed expressions are treated inadequately.
- Locating entries requires an intermediate knowledge of Arabic grammar.
- Entries are often out-of-date, mixing historical meanings with modern ones with no indication of their status.
- Multiword expressions such as *Middle East* are not included systematically.
- Recent words and meanings, especially those related to information technology and the Internet, are not included.



- Most dictionaries do not give example sentences, while comprehensive dictionaries give examples that are too literary for learners.

In a nutshell, most Arabic dictionaries are not designed with the learner in mind. Traditional lookup methods hinder learners from finding entries quickly, the lack of romanization is a source of frustration, and poor design makes many dictionaries user-unfriendly. In addition, since in many dictionaries the meanings are not up-to-date, there is an urgent need for a new learner's dictionary that overcomes these various shortcomings.

Distinctive Features

The CJKI Arabic Learner's Dictionary provides a number of features that distinguish it as an entirely new type of Arabic learner's dictionary. It answers the urgent need for an easy-to-use dictionary, compact enough to be easily carried around, yet detailed enough to serve as an effective reference work that satisfies the practical needs of the serious learner. Compiled with the aid of modern computational lexicography, this unique work enables learners to gain a full understanding of the core vocabulary of MSA.

The key features that make this possible are:

- This is the first dictionary with a variety of features **targeted specifically for beginner and intermediate learners**.
- **Clear and concise** word meanings are listed in order of importance.
- **Headwords and meanings have been carefully selected** and fine tuned to the beginner and intermediate levels.
- Word meanings are **fully illustrated by example sentences** designed expressly to be useful to the learner.
- **Up-to-date entries** include multiword expressions and recent words such as *blog* and *ATM*, while rare or archaic words and senses have been omitted.
- **Alphabetical ordering** ensures that even beginners can quickly locate entries without a knowledge of Arabic grammar (roots and patterns).
- **Cross-references** for common spelling variants increase ease of lookup.
- **All words and examples are fully romanized** using *CARS*, a new romanization system developed specifically for learners.
- For the first time ever, **words stress** and **neutralized vowels** are indicated (e.g. ٱنا́ *'ána*, ٱ where *á* shows the stressed syllable and *á* indicates a *shortened* long *ā*).
- **Part of speech labels** for every sense of every headword also indicate verb transitivity.

- The **Grammar Box** gives details on both regular and irregular inflected forms such as the imperfect, participles, plurals, the model verb number, verbal nouns, and more.
- **Model verb tables** enable the user to correctly and fully conjugate any verb.
- **Subentries** include word combinations such as fixed expressions (إِنْ شَاءَ اللَّهُ 'God willing'), and noun-derived adverbs (بِسْرَعَةٍ 'quickly') that are highly useful to the learner.
- A **user-friendly design** includes an attractive layout with easy-to-read fonts and colors.
- **Compact and lightweight**, the printed dictionary is portable and a pleasure to use.
- A **mobile version** of the dictionary for iOS and Android allows easy access via smartphone or tablet.

How this dictionary is different

The **CJKI Arabic Learner's Dictionary** is a state-of-the-art Arabic dictionary that has been tailor-made to meet the specific needs of beginning and intermediate Arabic learners. Below are descriptions of the key features that set this dictionary apart.

Dictionary Goals

The primary goal of this dictionary is to overcome the various shortcomings found in existing dictionaries by enabling learners to gain a good understanding of the meanings and usage of the core vocabulary of MSA. To this end, the dictionary was designed from the ground up: everything from the selection of headwords and their meanings to the presentation of the English equivalents, the writing of examples, and the design and layout were fine tuned to be maximally useful to the learner.

Target Audience

This dictionary is specifically designed for beginner and intermediate learners of MSA. Such learners have not been adequately served by existing dictionaries, which are too often usable only by advanced learners or native speakers.

Alphabetical Order

This dictionary utilizes a strict alphabetical ordering system rather than the root-based order of traditional dictionaries. In addition, cross-references are provided for common spelling variants. This means that the user needs neither detailed knowledge of Arabic grammar nor knowledge of a word's roots and conjugation patterns to look up a word.



Though it departs from tradition, this rational approach is convenient for quick lookup. In the traditional approach, in order to look up a verb like *انتظر* 'intázara 'wait for', one needs to know that it is Form VIII of *نظر* 'názara 'look at', or at least to know that its root is *نظر* *nẓr*, where *انتظر* is listed along with other verbs like *أنظر* 'ánzara 'grant' (Form IV) and *تأظر* *tanázara* 'face each other' (Form VI). Looking for *انتظر* at the location for *نظر* is not at all self-evident, meaning that beginners, even after repeated time-consuming searches, may fail to find the desired entry.

Up-to-date Coverage

Headwords have been carefully selected to ensure that they are truly relevant to learners of MSA. Traditionally, dictionaries have selected headwords based on classical texts and past dictionaries, including large numbers of archaic words and excluding many newer words, such as those pertaining to modern technology. In order to ensure that all entries are useful to learners, we have selected words instead through statistical analysis of corpora, a technique that has now become the gold standard in modern lexicography. In addition, the results of this analysis have been supplemented with entries drawn from textbooks, other modern dictionaries, and the Web. These include important *multiple word expressions* (MWEs), which are seldom included systematically even in modern dictionaries.

It is important to note, however, that headword selection based on raw frequency alone can be counterproductive. In order to avoid the pitfalls of a pure frequency-based approach, potential entries were evaluated by a team of experienced Arabic language educators and lexicographers, an important step to ensure that there were no omissions or unnecessary inclusions (see monograph at Appendix 2 for details on issues of headword selection). The result of this process is a work composed of practically useful words and MWEs such as those shown below.

الصَّرَافُ الآلِيّ	<i>aṣṣarráfu l'áliyū</i>	ATM
الشَّرْقُ الأَوْسَطُ	<i>ʾash·shárqu l-ʾáw·saṭ·u</i>	Middle East
مُدَوَّنَةٌ	<i>mudawwánatun</i>	blog

Up-to-date Meanings

The word meanings given in this dictionary, listed in order of their importance in contemporary Arabic, are well organized, concise, easy to understand, and accurate. When a new concept first appears in a language it often happens that the meaning of an existing word is extended to cover that concept. In other cases a new loanword is adopted, like *تَلْفُون* *tilifún* 'telephone', which is used alongside the native *هَاتِف* *hátif*,



which originally meant 'one who shouts'. Another example is the word سَاحَة *sāḥa*, which means 'courtyard, field'. Recently this was extended to mean 'forum' in the sense of 'a medium for public discussion'. In this dictionary we have made an effort to include such recent meanings, which are not listed in other dictionaries.

Illustrative Examples

Word meanings are illustrated by example sentences written specifically for this dictionary and designed to be maximally useful to the learner. Merely quoting examples from a corpus, literary works, or newspapers, especially if the examples are long and complicated, can be counterproductive from a pedagogical point of view. Example sentences are given for every sense of every primary entry, enabling the learner to gain a proper understanding of how the entry word is used in context.

CARS Romanization

The innovative **CJKI Arabic Romanization System** (*CARS* for short) was developed specifically to make it easy for learners to pronounce Arabic with ease and precision. *CARS* has several unique features not found elsewhere, including a user-friendly set of symbols that represent Arabic pronunciation unambiguously. For example, in الْحُكُومَةُ الْيَابَانِيَّةُ *'alḥukūmatu lyabānīyyatu* 'the Japanese government', long vowels (as in *kū*) are shown by a macron, word stress by the accent mark (as in *kū́*), and, for the first time, long vowels that are shortened (*neutralized*) in actual pronunciation are indicated by an underline (as in *ḡ*). These and other *CARS* features are described in detail in a separate document (see Appendix 2).

Grammar Box

This gives detailed grammatical information such as the regular and irregular inflected forms for the imperfect (*PR*), the imperative (*IM*), active (*AP*) and passive (*PP*) participles, the verbal nouns(s) (*VN*), the feminine form, sound and broken plurals, and the like. Moreover, the model verb number points to the appendix that contains detailed model verb conjugations.

Form I - Z1 - 123	<i>AP</i> أَكُّ ل <i>'ákil - un</i>
<i>PR</i> يَأْكُلُ ل <i>yá'kuluu</i>	<i>PP</i> مَأْكُولُ ل <i>ma'kúl-un</i>
<i>IM</i> كُ ل <i>kul</i>	<i>VN</i> أَكْلُ ل <i>'ákl-un</i>

Multiple plurals corresponding to different word senses are also shown, e.g., the plural of بَيْتُ *báytun* 'house' is بِيُوتُ *buyūṭun* for the sense 'house' but أَبْيَاتُ *abyátun* for the



sense ‘verse’. In short, learners have at their fingertips detailed information on inflected and derived forms, eliminating the need to laboriously consult grammar books and conjugation tables.

Part-of-speech

Part of speech labels are given for every sense of every headword in an explicit manner. For example, many *nisba* adjectives such as those indicating pertinence or nationality have noun counterparts, such as *يَابَانِيّ* *yabāniyyun* ‘Japanese’ meaning ‘a native of Japan’ as a noun and ‘relating to Japan’ as an adjective. These are given under separate senses of the main headword to show precisely how they are used. Moreover, the transitivity (whether a verb is transitive, intransitive, or both) is indicated explicitly by the part-of-speech labels.

Subentries

Subentries include word combinations that are highly useful to the learner, showing how the word is used in context. These include prepositional phrases such as *مِنَ الْمُمَمَّكِن* *mina lmuḥkīni* ‘possible’, and fixed expressions such as *إِبْنُ أَصْل* *‘ibnu ‘aṣlin* ‘faithful man’. Also included are noun-derived forms, such as adverbs of the form *بِسْرَعَةٍ* *bisur’atin*. Finally, prepositions used with intransitive verbs are made explicit for each sense.

Attractive Design

Another distinctive feature is a layout that is both visually attractive and easy to use. Typographical attention to detail – including an Arabic font customized for this dictionary – were used to achieve an aesthetically pleasing design that stimulates a desire to browse and to truly enjoy the study of Arabic (see Appendix 1 for a sample page). For example, clutter was avoided by placing grammatical information such as important verb inflections and verbal nouns in the Grammar Box rather than within the entry itself. Furthermore, every effort was made to ensure that the dictionary is compact and lightweight without sacrificing any information important to the learner.

Summary

This dictionary is based on a systematic approach and a firm theoretical foundation that emerged from several decades of experience in compiling bilingual learner’s dictionaries that have become widely used standard reference works, such as *The Kodansha Kanji Learner’s Dictionary* (Kodansha, 2013). Modern linguistic theory has been effectively integrated with computational lexicography to produce a powerful tool designed to meet in full the specific needs of Arabic learners. The advanced compilation techniques and the special features of this dictionary distinguish it as the most useful and easy-to-use Arabic learner’s dictionary ever compiled.

This dictionary is scheduled for completion in 2020.

Appendix 1: Page Sample

اليابان

3

بَيْتٌ *báy·t-un*

NOUN m. (ب-ي-ت)

① **house, home, residence** MP1

أَسْكُنُ فِي بَيْتٍ صَغِيرٍ. *'áskunu fī báytin ṣaghírin.* I live in a small house.

② **home** MP1

رَجَعَ إِلَى بَيْتِهِ. *rájæa 'ilá báytihi.* He returned home.

③ **verse** MP2

قَرَأْنَا بَيْتَ الشَّعْرِ بِصَوْتٍ عَالٍ. *qará'ná báyta_šshéeri bišáwtin éálin.* We read the verse of poetry out loud.

MP1 **بُيُوتٌ** * MP2 **أَبْيَاتٌ** *

عَادَ *éá·da*

VI (ع-و-د)

① **to return, go back** ▶ إِلَى to

عَادَ إِلَى بِلَادِهِ. *éáda 'ilá biládihi.* He returned to his country.

② **to repeat, do again**

عَادَ الرَّجُلُ يُدَخِّنُ ثَانِيَةً. *éáda_rrájulu yudákhkhinu thānīyatan.* The man took up smoking again.

③ **to bring, take back** ▶ بِ sth

عَدْتُ بِكُلِّ كُتُبِي إِلَى بَيْتِي بَعْدَ إِكْمَالِ دِرَاسَتِي فِي الْخَارِجِ. *éúdtu bikúlli kútubij 'ilá báyti báeda 'ikmáli dirásatij fī_lkhárij.* I took all my books back home after completing my studies abroad.

SUBENTRIES

لَمْ يَعُدْ *lam yá·éud*

PHRASE [followed by an adjective or imperfect indicative verb]

no longer, not anymore

لَمْ أَعُدُّ أَصَدِّقَكَ. *lam 'áedu 'uṣaddíquka.* I don't believe you anymore.

لَمْ يَعُدْ يَزُورُ هَذَا الْمَكَانَ. *lam yáéud yazúru hádhá_Imakána.* He no longer visits this place.

Form I-H1 ^[w/u]	PR يَعُودُ	IM عُدَّ
AP عَائِدٌ	PP مَعُودٌ	VN عَوْدَةٌ

أَبْدَلَ *'áb·da-la*

VT (ب-د-ل)

to exchange, replace ▶ sth بِ sth with

أَبْدَلَ ثِيَابَهُ. *'ábdala thiyábahu.* He changed his clothes.

Form IV-S1 ^[l]	PR يُبَدِّلُ	IM أَبْدَلُ
AP مُبَدِّلٌ	PP مُبَدَّلٌ	VN إِبْدَالٌ

مُمَثِّلٌ *mu·máth·thi-l-un*

NOUN m. (م-ث-ل)

① **representative, delegate**

سَأَكُونُ مُمَثِّلًا لِبِلَادِي. *sa'akúnu mumáth-thilan libáladj.* I will be a representative of my country.

② **actor**

ذَآكَ الْمُمَثِّلُ وَسِيمٌ جِدًّا. *dháka_Imumáth-thilu wasímun jíddan.* That actor is very handsome.

مُمَثِّلٌ مَسْرَحِيٌّ *mumáththilun masrahíyyun* stage actor

③ **active participle of** مَثَّلَ → مَثَّلٌ

MP مُمَثِّلُونَ	FS مُمَثِّلَةٌ	FP مُمَثِّلَاتٌ
-----------------	----------------	-----------------



Appendix 2: Further reading

1. A detailed description of CARS is available at:
http://www.cjk.org/cjk/arabic/cars/cars_sum.pdf
2. A monograph on dictionary headword selection is available at:
http://www.cjk.org/kanji/dictionaries/cald/Ar_dict_entries.pdf
3. A list of features that make this dictionary unique is available at:
http://www.cjk.org/kanji/dictionaries/cald/how_different.pdf



Jack Halpern

Jack Halpern (春遍雀來), CEO of The CJK Dictionary Institute, is a lexicographer by profession. He was engaged for sixteen years in the compilation of the New Japanese-English Character Dictionary, and, as a research fellow at Showa Women's University (Tokyo), he was editor-in-chief of several kanji dictionaries for learners, which have since become standard reference works.



Born in Germany, Jack Halpern has lived in six countries including France, Brazil, the United States, and Japan, the last of which has been his home for over 40 years. He is an avid polyglot who specializes in Japanese and Chinese lexicography, having studied 15 languages (speaking nine fluently). He has devoted several decades to the study of linguistics and lexicography.

Jack Halpern has published over twenty books and dozens of articles and academic papers, mostly on the Japanese writing system and CJK information processing. He has given over 600 public lectures on Japanese language and culture, and has presented several dozen papers at international conferences.

On a lighter note, Jack Halpern loves the sport of unicycling. Founder and long-time president of the International Unicycling Federation, he has promoted the sport worldwide and is a director of the Japan Unicycling Association. Currently, his passion is playing the quena and improving his Chinese, Esperanto, and Arabic.

The CJK Dictionary Institute



The CJK Dictionary Institute, Inc. specializes in Chinese, Japanese, and Korean lexicography. CJKI is headed by Jack Halpern, editor-in-chief of the New Japanese-English Character Dictionary and various other dictionaries that have become standard reference works for studying Japanese. CJKI is one of the world's prime sources for CJK and Arabic dictionary data, and is contributing to CJK information processing technology with its high-quality lexical resources.

The CJK Dictionary Institute, Inc.

34-14, 2-chome, Tohoku, Niiza-shi, Saitama 352-0001 JAPAN

Phone: 048-473-3508 Fax: 048-486-5032

Email: jack@cjk.org Web: www.cjk.org