Towards the Making of a Contextual English-Arabic Dictionary

Hameed Ahmed Khallaf

Department of English, Faculty of Education, Damanhour University

hameed_ak2002@yahoo.com

Abstract

English-Arabic lexicography has received little attention in the literature. It is a fact that most of general-purpose English-Arabic dictionaries are products of single authors; consequently, they suffer from some weaknesses. They are based on a sense-enumeration technique, which assumes that a multi-sense word has a fixed set of meanings, which makes it impossible for such types of dictionaries to account for meanings of a word in novel contexts. Moreover, most of these dictionaries provide decontextualized equivalents to the different senses of their lexical entries; therefore, translators, learners of English as a second/foreign language, and teachers consulting English-Arabic sense enumerative dictionaries find difficulties in choosing suitable equivalents due to the absence of contextual clues. Thus, this paper aims at discussing weaknesses of English-Arabic dictionaries and proposing a way of making a contextual English-Arabic dictionary. The paper focuses on two main issues: lexical coverage and meanings in contexts. The proposed dictionary suggests making use of information technology, as it is unimaginable, especially with this worldwide desire for digitization, to compile a dictionary without resorting to available sources of IT. The suggested dictionary could be made by eclectic compilation of lexical entries from the specialized dictionaries of the *Academy of the Arabic Language* in Cairo, in addition to some English-English corpus-based dictionaries, and some other English-Arabic sources. Moreover, two major corpora can be used to provide necessary contextual information: The British National Corpus (BNC) and the Coronavirus Corpus. To avoid limitations of folio editions, the suggested dictionary is an electronic one.

Keywords: Lexicography - Contextual Information - Lexical Coverage - Dictionary-making - the British National Corpus

1. Introduction

The Arabs have pioneered dictionary-making. This is manifested in the affluence of monolingual Arabic dictionaries that have been professionally compiled with

exceptional care given to both lexical entries and word sense distinctions. Nevertheless, some English-Arabic dictionaries abound with weaknesses and shortcomings that need to be spotlighted so that suitable remediation for such weaknesses and shortcomings can be recommended. Current English-Arabic dictionaries are not bad, but making an authoritative, contextual dictionary that tells dictionary users the exact meaning of a lexical entry in a given context is so laborious for a single author. Thus, it is time to regard dictionaries as academic products that are to be made under the supervision of eminent scholars and under the aegis of government academic institutions. Therefore, there is a pressing need for conducting studies that attempt to reduce the gap between Arabic lexicography as a craft and lexicology as a science.

The study of lexicology and lexicography is of paramount importance to all other functional parts a living language consists of. "Any living language consists of four independent but exceedingly interrelated parts" (Sanni, 1992, p.1). These four parts include vocabulary, phonology, grammar, and semantics. "Vocabulary or lexicon is perhaps the most important of these functional parts, as it supplies the materials with which the other three levels operate" (Sanni, 1992, p.1). Etymologically speaking, "the term lexicology is composed of two Greek morphemes: *lexis* meaning 'word, phrase'… and *logos* which denotes 'learning', a department of knowledge" (Ginzburg et al., 1979, p.7). Thus, it could be said that *lexicology* is the theory of writing words or 'the science of words'. Unlike lexicology, lexicography consists of *lexico-* (prefix meaning 'speech or words') + *graphy* (suffix meaning 'something written about a specific subject') (Wiktionary available at https://wiktionary.org). Thus, lexicography can be described as applied lexicology. It is concerned with the writing of words in some concrete form i.e. compiling of dictionaries. As noted by Ginzburg et al. (1979):

> Lexicography is a practical application of lexicology so that the dictionary-maker is inevitably guided in his work by the principles laid down by the lexicologist as a result of his investigations. It is common knowledge that in his investigation, the lexicologist makes use of various methods (p.11).

Thus, lexicography as a craft is closely related to lexicology as a science. Both of them deal with problems of meaning, form, and usage of vocabulary.

Moreover, lexicography has now become a well-established field of research. Many associations and societies interested in lexical matters have been established: *Dictionary Society of North America* (DSNA) (2017), *European*

Association for Lexicography (Euralex) (1983), *African Association for Lexicography* (Afrilex) (1995), *Australasian Association for Lexicography* (Australex) (1990), and *Asian Association for Lexicography* (Asialex) (1997). As for the Arab world, the *Academy of the Arabic Language* in Cairo (1932), took the lead in the field for its immense contribution in building up both Arabic monolingual dictionaries and specialized English-Arabic dictionaries. Besides the *Academy of Arabic Language* in Cairo, so many other academies have been established in the Arab world: *Arabic Language Academy of Damascus* (1918), *Academy of the Arabic Language* in Khartum (1993), *Jordanian Academy of Arabic* (1924), and *the Iraqi Academy of Sciences* (1948). By examining the dates of establishment of these academies for the Arabic language, it is clear that these academies have been founded decades earlier than other associations in non-Arab countries. Moreover, this abundance of global associations and societies for lexicography and the flux of regional academies for the Arabic language verifies the widespread interest in lexicographical issues and emphasizes the importance of making dictionaries under the umbrella of recognized institutions.

2. Objectives of the Study

The overall objective of this study, as its title indicates, is to discuss the possibility of making a contextual English-Arabic dictionary (CEAD), both the macrostructure and microstructure[1] of which are different from those of general-purpose English-Arabic dictionaries available so far. In this sense, the dictionary aims to show how a word is used in a specific context rather than what it generally means. For this purpose, weaknesses of English-Arabic dictionaries will be identified and solutions for these weaknesses will be set forth. Furthermore, sources for compiling the proposed dictionary will be pinpointed and the mechanism of collecting lexical data will be elucidated. To avoid limitations of printed dictionaries, the dictionary discussed in this paper is an electronic one. The choice of an electronic version of a dictionary is always preferred to a folio edition for many reasons. Firstly, it can store huge amounts of data. Secondly, unlike hard-copy dictionaries that provide one way of searching for information, usually alphabetically, an electronic dictionary offers various routes a user can follow to find the information for which he/she is consulting a dictionary. Thirdly, an electronic dictionary allows searching "by 'chain', or 'hyperlink', a search mechanism by which a double click on a word on the screen will call up a dictionary entry for that word (Nesi, 1999, p.61, cited in Sterkenburg, 2003, p.5). Fourthly, an electronic dictionary is highly extendable as it "can also be edited on

a daily basis, allowing changes to be made, neologisms to be added and obvious errors to be corrected" (Sterkenburg, 2003, p.6).

3. Arabic Lexicography

Tracing back Arabic lexicography is an interesting field of study per se. Lexicographical practice in the Arab world is deep-rooted: "Arabic lexicography has a centuries-old tradition" (Arts, 2014, p.109). Based on the plethora of Arabic monolingual dictionaries, most of which go back to the eighth century, it could be said that the Arabs were pioneers in the field. Haywood (1960) argues that "a dozen or more major dictionaries, besides many vocabularies, both general and specialized, bear witness to their pre-eminence in this field, at a time when such works were almost unknown in Western Europe" (p.1). Arabic monolingual dictionaries surpass their counterparts in many respects. The Arabs "were obsessed by the copiousness of the language and were very mathematically-minded in this matter. In this, they differed from the earlier lexicographers of other nations, whose chief aim was to explain rare and difficult words" (Haywood, 1960, p.2).

Al-Khalil ibn Aḥmad al-Farahi:di, the founding father of Arabic lexicography (died 170/786) has sown the first seeds of Arabic monolingual dictionaries, as he was the first to compile a dictionary in the Arabic language known as *Kitab al-ʕayn. Al-Khalil* arranged his dictionary sometimes permutatively and other times anagrammtically. *Kitab al-ʕayn* "was a comprehensive descriptive record of the lexicon of the contemporary Arabic language at the time" (Attia et al., 2011, p.101). This groundbreaking lexicographical work marks the beginning of Arabic lexicography as "it did not just record the high-level formal language of the Koran, Prophet's sayings, poetry, memorable pieces of literature and proverbs, but it also included a truthful account of common words and phrases as used by Bedouins and common people" (Attia et al., 2011, p.101). According to Haywood (1960) *Kitab al-ʕayn "*is a landmark not only of Arabic lexicography but in the history of world lexicography" (p.27).

Kitab al-ʕayn unleashed a revolution of Arabic monolingual dictionaries. It kick-started Arabic lexicography and the huge number of Arabic monolingual dictionaries following *al-ʕayn* is quite remarkable. The following are some salient Arabic monolingual dictionaries that followed *al-ʕayn*: *al-Jim* (al-Huru:f), Abu ʕamr al-ʃaybani (died 206/821), *al-ɣari:b al-Muṣannaf,* Abu ʕubayd Qasim Sallam (died 223/837)*, al-Alfaẓ,* ibn al-Sikkit (died 244/858), *Kitab al-Jamharah,* Ibn Durayd (died 321/933), *Diwan al-Adab* (Mizan al-Luɣah wa Miʕyar al-Kalam), al-Farabi (died 350/961), *Tahzi:b al-Luɣah,* al-Azhari (died 370/981),

Muxtasar al-ʕayn, al-Zubaydi (died 379/989), *al-Muhi:t,* al-Sahib bin ʕabbad (died 385/995)*, al-Sihah,* al-Jawhari (died 394/1003), *al-Muhkam wa al-Muxassas,* Ibn Sidah (died 459/1066), *Asas al-Balaɣah,* al-Zamaxʃari (died 538/1143), *Muxtar al-Sihah,* al-Razi (died 666/1268), *Lisan al-ʕarab,* ibn Manzu:r (died 711/1311), *al-Misbah al-Muni:r,* al-Fayyumi (died 770/1368), *al-Qamu:s al-Muhi:t,* al-Firu:z ʕabbadi (died 817/1414), *Taj al-ʕaru:s,* al-Zabidi (died 1205/1790), *Muhi:t al-Muhi:t,* Butrus al-Bustani (died 1300/1883), *Aqrab al-Mawarid fi al-Fasi:h wa al-ʃawarid*, Saʕid al-ʃartuni (died 1330/1912), *al-Mundjid,* al-Ab Luis Maʕlu:f (died 1365/1946), *Fakihat al-Bustan,* ʕabd Allah al-Bustani (died 1348/1930), *Matn al-Luɣah,* Ahmad Rida (died 1372/1953), *al-Muʕjam al-Wasi:t, Academy of Arabic Language* in Cairo (see Kocak, 2002). In addition to *al-Muʕjam al-Wasi:t*, the *Academy of Arabic Language* in Cairo has compiled two other renowned lexicographical works: *al-Muʕjam al-Waji:z* and *al-Muʕjam al-Kabi:r*. The latter, as Meisami and Starkey (1998) believe "is the most ambitious lexicographical understanding of this kind so far" (p.469).

Based on this enumeration of Arabic monolingual dictionaries, from the old era until present, it is crystal clear that the lexicographical practice is deep-rooted in the literature. Arabic monolingual lexicography has been so influential even in the compilation of Arabic bilingual dictionaries. As admitted by Edward William Lane in the preface of his dictionary that his dictionary, *Arabic-English Lexicon*, is indebted to previous Arabic monolingual dictionaries especially *Taj al- ʕaru:s*, by al-Zabidi.

However, as noted by Jamiu (2014) "Arabic bilingual dictionaries were few in number in the Arab world" (p.60). Among the most famous bilingual English-Arabic dictionaries are *The Dictionary: English-Arabic*, by Research and Study Centre (2004), *al-Mawrid al-Hadi:θ* (2014), by Munir Baʕalabki and Ramzi Muni:r Baʕalabki, *Elias' Modern Dictionary*, by Elias A. Elias (1974). *The Unified Dictionary of Anatomy*, by the World Health Organization (2005), *Al-Muɣni al-Akbar* (Librairie du Liban Publishers), and *al-Muhi:t Oxford Study Dictionary* (2003). The English edition of *al-Muhi:t Oxford Study Dictionary* was compiled by Joyce M. Hawkins, John Weston, and Julia C. Swannel, whereas the English-Arabic edition was prepared by Mohamed Badwi, and revised by Mohamed Debs. Finally comes *the Oxford Arabic Dictionary*. This dictionary is the crown jewel of English-Arabic dictionaries known so far. It is based on real modern evidence of Modern Standard Arabic and English. In addition to these general-purpose English-Arabic dictionaries, the *Academy of Arabic Language* in Cairo took English-Arabic lexicography to a higher level by making a large

reservoir of specialized English-Arabic dictionaries: *Dictionary of Medical Terms* (1999), *Dictionary of Chemistry and Pharmacy* (1983), *Mathematics Dictionary* (1995), *Dictionary of Geology* (1982), and *Dictionary of Computer Terms* (2012), to name but a few.

Based on the above discussion of Arabic monolingual dictionaries and English-Arabic dictionaries, it is clear that lexical practice in monolingual Arabic dictionaries surpasses its counterpart in English-Arabic dictionaries. It is also noted that contextual English-Arabic dictionaries are rare to find. Thus, to achieve the goal of preparing a contextual English-Arabic dictionary, shortcomings of general-purpose English-Arabic dictionaries in current use have to be initially expounded. Instead of focusing on the weaknesses of all English-Arabic dictionaries, this paper will focus only on the weaknesses of general-purpose dictionaries.

4. Weaknesses of General-Purpose English-Arabic Dictionaries

Weaknesses and shortcomings of general-purpose English-Arabic dictionaries in current use are in abundance. Dolezal and McCreary (1999) show that in a study conducted to evaluate students' use of the different types of dictionaries (al-Badry 1990), 493 Arab students of ESL were surveyed on their use of both English learners' dictionaries and bilingual English-Arabic dictionaries. "El-Badry's study showed that Arabic students found that most monolingual and bilingual dictionaries were difficult to use because the definitions were ambiguous, the examples were insufficient, and the usage was unclear" (Dolezal & McCreary, 1999, p.32).

It could be said that the first weakness of English-Arabic dictionaries in use is that most of them are products of a single-author. According to Asfour (2003) "since AL-Khalil's pioneering work was completed in the eighth century, dictionaries in the Arab world have largely been the fruit of the labours of single souls working against enormous odds" (p.41). Putting in mind that dictionary-making is a huge and demanding task, it is expected for most English-Arabic dictionaries, especially those compiled by single authors, to suffer from some sorts of weaknesses. Asfour (2003) sees that "single-author dictionaries are not necessarily bad, but modern lexicography is increasingly becoming a too arduous task for a single individual to be able to do an adequate job" (p.41).

The second weakness of some English-Arabic dictionaries is reduced coverage. Coverage is a quintessential issue for dictionary users. In some English-

Arabic dictionaries, lack of coverage is most notable at a lexical-entry level and sense-definition level (i.e. macrostructure and microstructure). Al-Ajmi (2004) who has developed an English-Arabic parallel corpus notes that "the corpus was compared with *al-Mawrid English-Arabic Dictionary* (2003), indicating about 10000 Arabic equivalents and senses are missing in the dictionary" (p.329).

Moreover, so many general terms are not in English-Arabic dictionaries. For example, upon looking up *clitic, endocentric*, *lockdown*, *trailblazing*, and *exocentric,* which are monosemous, it was found out that all of them are missing in *al-Muɣni al-Akbar online dictionary*. By the same token, *browse* which is a polysemous word does exist but its meaning as "to look for information on the internet" is missing. Furthermore, not all potential forms of a lexical entry are available, for example, *Spanish Influenza* exists, but *Spanish Flu* does not. Besides, by looking up the meaning of *grammaticality* and *oddness* in LDLP[2] online dictionaries, the search gives no Arabic equivalents at all.

Unlike some English-English dictionaries, such as Oxford Dictionaries, which update their content regularly and at short intervals, English-Arabic dictionaries, due to restrictions imposed by printed formats, are unable to regularly add new terms associated with social changes. There is some truth in this because words such as *Covid 19*, *WFH* (work from home), *self-isolate*, *self-isolation*, *infodemic*, *social distancing*, *elbow bump*, *foot shake*, *PPE* (Personal Protection Equipment), *self-quarantine*, and *covidiot*, which came to dominate the global discourse during coronavirus pandemic (Covid 19), are missing in English-Arabic dictionaries, though they were added to Oxford Dictionaries a few weeks after the outbreak of Covid 19. Moreover, some English-Arabic dictionaries extract their English from old texts, as Arts (2014) argues "the most popular work is *al-Mawrid* (Baʕalbaki, 2013), which is updated regularly but has certain major flaws. The English used seems mostly derived from very old texts, and its policy for including new words in updated editions is not clear" (p.110).

The third weakness of English-Arabic dictionaries is the inability to account for the creative use of words in novel contexts. Although context-based meaning is an inveterate practice in Arabic monolingual lexicography, it is not always the case in modern English-Arabic dictionaries. Arabic monolingual dictionaries are unprecedentedly contextual, as the inclusion of encyclopedic information is a basic component of such dictionaries. Citations from the Holy Qur'an, Prophet's sayings, and poetry in addition to idioms, proverbs, and collocational combinations are included for lexical entries to show the proper use of a word, e.g., *Tag al-ʕaru:s* by al-Zabidi is exceedingly contextual. On the contrary,

English-Arabic dictionaries are based on a sense enumeration technique in which a headword is given a list of decontextualized meanings in the target language. Such dictionaries ignore elaborate sense distinctions and lack contextualization clues that give evidence about the appropriate use of a lexical entry. As a result of this absence of contextual information, translators and learners of English as a second/foreign language are hampered to get the appropriate meaning of a word. Therefore, they, sometimes, are compelled to choose any appropriate translation for the headword, which is so detrimental to the acceptability of the chosen equivalent. This absence of contextual information sometimes forces a dictionary user to form odd collocations. Drawing on the potentially odd collocations of word senses, it becomes so palpable that without an appreciation of the syntactic structure of a language, the study of lexical semantics is bound to fail. "There is no way in which meaning can be completely divorced from the structure that carries it" (Pustejovsky, 1991, p.410). Fillmore (1982) sees that "it is possible to think of a linguistic text, not as a record of 'small meanings […,] but rather as a record of tools that somebody used in carrying out a particular activity" (p.112).

Due to this lack of information about the occurrences of a lexical entry in a given context, senses of lexical entries in English-Arabic dictionaries become permeable, i.e. there are no clear-cut distinctions between the different senses of a lexical entry. If we were to assume that sense enumerative dictionaries were adequate as a descriptive mechanism, it is not always obvious how to select the correct word sense in any given context. For example, upon consulting *al-Muɣni al-Akbar* online dictionary for the possible meanings of two English words: *abnormal* and *abnormality*, the dictionary gave their meanings as شاذ، غير معهود، /ʃað/, /ɣayru maʕhu:d/, حالة شاذة، شذوذ، خروج عن المعهود، خارج عن العادة (أو) المعهود /ḥala ʃaðah/, /ʃuðu:ð/, /xuru:g ʕan al-maʕhu:d/, /xarig ʕan al-ʕadah (or) al-maʕhu:d/ respectively. Upon considering these two lexical entries, it becomes so conspicuous that the user finds himself/herself on the horns of a dilemma. The Arabic equivalents of *abnormal* and *abnormality* as being شاذ and شذوذ respectively may lead the user, especially if he/she lacks experience, to think of them as synonyms of *homosexual* and *homosexuality*. The same holds for the lexical entry of *press* (v): ضغط = [كبس]. لز، ازدحم، ملس (الثِّياب) = كواها، عصر، ضم، شد /daɣaṭa/،= /kabasa/, /lazza/, /izdaḥama/, /mallasa/ (clothes) = كواها (iron clothes) /kawaha/, /ʕṣara/, /ḍamma/, /ʃadda (ʕala)/, /akraha (ʕala)/, أكره (على)، حث /haθθa/, Upon examining the lexical entry of *press* and its Arabic equivalents as stated in *al-Muɣni al-Akbar*, it is obvious that sense permeability is attributed to a lack of full definitions and full example sentences that determine the different

meanings of a headword according to its contexts. This lack of sense definitions and example sentences opens the door for so many semantically odd collocations.

The fourth weakness of English-Arabic dictionaries is the yawning gap between linguistic theory and lexicographical practice. "Lexicographers need linguistic theory to help them analyse the linguistic data effectively and draw useful conclusions from it; and we have to understand the needs of our target audience if we are going to produce a language description that is accessible and relevant to the people who will use it" (Atkins & Rundell, 2008, p.2). In his preface to *A Practical Guide to Lexicography*, Sterkenburg (2003) argues that "recent developments have facilitated new theories combining language as a system with the way in which language manifests itself. Lexicographers have taken cognizance of the most recent models developed in semantics and pragmatics".

Though Arabic lexicography has substantially affected Arabic grammatical theory, modern English-Arabic lexicographical practices manifest a big gap between dictionary-making and modern syntactic theories. "Lexicography in the Arab world has had important effects on the development of the Arabic language. The origin and subsequent development and refinement of traditional Arabic theory […] had intimate links with the practice of writing dictionaries" (Benzehra, 2012, p.83). *Kitab al-ʕayn*, the first fully-fledged, Arabic monolingual dictionary, which is viewed as a foundation for Arabic grammatical thought, paved the way for the works of famous Arabic grammarians such as Sibawayh. "It is now assumed by most linguistic frameworks that much of the structural information of a sentence is encoded from a lexicalized perspective", as noted by Pustejovsky (1995, p.5).

Thus, ignoring subcategorization and contextual information makes most English-Arabic dictionaries as registers of lexical entries associated with decontextualized meanings which are proved to be mostly unserviceable without recourse to an English-English contextual dictionary to get an idea about how a lexical entry is used in a given context and the meanings associated with such a context.

5. Methodology of Compiling CEAD

Zgusta, the twentieth-century godfather of lexicography, believes that one of the best definition of the term dictionary was given by C.C. Berg:

a dictionary is a systematically arranged list of socialized linguistic forms compiled from the speech habits of a given speech community and commented on by the author in such a way that the qualified reader understands the meaning" (in Zgusta, 1971, p.197).

This definition describes how a dictionary is a reflection of the speech habits of a speech community. Yet, this definition, is rather 'elitist', as it describes a dictionary as a book that can be understood only by a 'qualified reader'. Restricting a dictionary to the province of a highly educated reader makes such a definition a little bit incomplete and fragile. "Bo Svensen (1993, pp.3-4) provides a less fragile and much more explicit definition. To him, a dictionary is a book that in the first place contains information on the meaning of words and their usage in specific communicative situations" (Sterkenburg, 2003, p.4). Bo Svensen's definition best describes how a dictionary should be, i.e. a dictionary is a book that tells the user what a given word means in a specific context, how it is used, and what words it collocates with. "The most neutral is to say that a dictionary should provide information on the meaning of the lexical units included and information on their usage in specific language situations" (Sterkenburg, 2003, p.8).

Inspired by Bo Svensen's definition, the researcher suggests that CEAD be carried out in line with the new trend in dictionary-making in which corpus data represent the main source for providing all possible senses of a headword, determining the different meanings of these senses, and expanding the headwords of the dictionary. Thus, lexical entries in CEAD are divided into one or more senses in which the senses are given full definitions, and equipped with full example sentences to illustrate typical uses of the headword in different contexts.

6. Data Collection for CEAD

Comprehensive, authoritative dictionaries are not born every day. It is hugely expensive to produce them from scratch. It is notable that most new dictionaries still owe much to some earlier incarnation. Therefore, the main sources of data for the proposed dictionary are the specialized dictionaries compiled by the *Academy of the Arabic Language* in Cairo, and some English-Arabic dictionaries in current use. Some salient corpus-based English-English dictionaries can also be made use of as sources for English lemmas. Based on the fact that corpora have become the norm in modern lexicography, CEAD benefits from the *British National Corpus* (BNC) and the *Coronavirus Corpus*. The former was created by Oxford University Press in the 1980s, and it holds 100 million texts from a wide range of

genres, while the latter was created to be a definitive record of the social, cultural, and economic impact of the coronavirus (COVID-19) and beyond. The corpus is currently about 392 million words in size, and it continues to grow by 3-4 million words each day (see english-corpora.org). In this sense, CEAD will be substantially hybrid, because, as a rule of thumb, different dictionaries provide inconsistent sense definitions for the same lemma.

7. Lexical Coverage and Clues for Contextual Meanings

One of the main goal of the proposed dictionary is to achieve lexical coverage. Corpora, either annotated or unannotated, gave vent to a mainstream in dictionary-making in which corpus-data play a pivotal role. Thus, since computerized corpora enhance lexical coverage, why does reduced coverage still mark most English-Arabic dictionaries? This question gives rise to another question that is inextricably linked with it: is that lack of coverage attributed to the unavailability of lexicographical resources or merely to a voluntary process of inclusion and exclusion on the part of the authors? With the advent of the computer, many corpora are now available online for free. "Since the release of COBUILD[3], corpora proved to be excellent resources for developing new dictionaries. In addition to lexical information, corpora can provide more useful information that would enrich dictionaries such as idioms, phrases, collocates, word sketches, and thesauri of words" (Sawalha, 2019, p.36). Thus, based on the availability of lexical resources, especially English-Arabic parallel corpora, it could be said that the problem of the lack of coverage, though not stated by the authors of English-Arabic dictionaries, is not attributed to lack of lexicographical materials, but rather to a desire on the part of the authors to include specific lexical entries and exclude others. Thus, the first goal of the proposed dictionary, namely lexical coverage, can be achieved by the utilization of available corpora that provide invaluable sources of lexical data.

7.1 Example Sentences as Clues for Contextual Meanings

The second main goal of the suggested dictionary is to provide contextual meanings for the different senses of a headword. Pre-corpus lexicographers, with the absence of corpus evidence, used to depend on their intuitions and on other dictionaries to describe lexical items, but with corpus data being freely accessible, "it is important for lexicographers to research and describe the conventions associated with each lexical item individually" (Hanks, 2009, p.217). The corpus

helps lexicographers to study lexical items compositionally by providing example sentences of the different occurrences of the lexical entry in text corpora.

The importance of using example sentences to identify the exact meaning of a lexical entry gives rise to a very crucial question: what is more beneficial in showing how a word is typically used – examples invented by lexicographers or ones extracted from a corpus of data or just selecting examples from the Web? It seems to me that authentic examples extracted from a corpus of data top lexicographer-made examples because the latter are so exhaustive, time-consuming, and less natural than the former. Furthermore, corpora are now available for hundreds of languages, growing larger and more diverse. Moreover, the multifarious anarchy of the Web makes it less reliable than corpora as a source for full example sentences. For this purpose, the researcher suggests that CEAD be compiled in line with the main stream of making modern dictionaries in which corpora are basically used to provide example sentences for lexical entries. "There is hardly any alternative to corpora as the primary and main source for lexicographers now, and the number of corpus-based dictionaries is steadily growing" (Cermak, 2003, p.21). Equipping the dictionary with authentic example sentences will help dictionary users to identify the valency of each lexical item as well as its collocations. In this way, English-Arabic dictionary users can get evidence of how the members of a speech community communicate with one another linguistically correct and socially appropriate. As Tony et al. (2006) maintain, there are so many advantages of corpora: "The greatest advantage of using corpora in lexicography lies in their machine-readable nature […]. The second advantage of the corpus-based approach […] is the frequency information and quantification of collocation which a corpus can readily provide" (p.60).

The two corpora used in the proposed dictionary, namely the *British National Corpus* and the *Coronavirus Corpus*, in addition to some corpus-based English-English dictionaries will be so beneficial in providing example sentences which pinpoint the exact meaning of a word based on its co-text. Consider the following co-occurrences of *abnormality* as it occurs in the BNC:

Table 1. Some co-occurrences of *abnormality* in the BNC

Term	Frequency	Example Sentences
abnormality	281	- Miscarriage can be caused by *a foetal abnormality*. - The court must, lastly, be satisfied that the *mental abnormality* substantially impaired his mental responsibility for the killing. - A minor *abnormality* in the *chromosomes* - Children with *developmental abnormality* - It is believed to be probably caused by an *organic abnormality* of the brain. - to be the only specialists in the world with the expertise to tackle the *retina abnormality*

Contrary to the ambiguous meanings given to the word *abnormality* in most English-Arabic dictionaries, in which the term is translated as شذوذ/حالة شاذة /ʃuðu:ð/, /ḥala ʃaðah/, the *British National Corpus* gives example sentences which identify the exact meanings of the word in question in addition to so many nontraditional uses of this word. We have, for example, *foetal abnormality* تشوه الجنين, /taʃauuh al-jani:n/, *mental abnormality* الاختلال العقلي /al-ʔixtilal al-ʕaqli/, *abnormality in the chromosomes* خلل في الكروموسومات /xalal fi al-kromosomat/, *developmental abnormality* خلل في النمو /xalal fi al-numu/, *organic abnormality* خلل عضوي /xalal ʕuḏwi, *retina* abnormality عطب في شبكية العين /ʕutb fi ʃabakiyat al-ʕayn/. In addition to these example sentences found in the BNC, the researcher has combed *Cambridge Dictionary* (online version) and an additional meaning for *abnormal* was found: *Because he has to work abnormal hours, Jack doesn't often eat with the rest of the family.* The meaning of *abnormal* in this example sentence can be translated as غير منتظم (ساعات عمل غير منتظمة) /ɣayru muntaẓim (saʕa:tu ʕamalin ɣayru muntaẓimah/

Furthermore, after looking up the word *encyclopedic* in the BNC, it was found out that its meanings are not restricted to the common meanings specified in most English-Arabic dictionaries. For example, *encyclopedic* is described in *Thorndike Dictionary* (En-Ar) as "1. covering a wide range of subjects; possessing wide and varied information, 2. having to do with an encyclopedia." موسوعي: شامل المعرفة، موسوعي، /mawsu:ʕi:/ /ʃamilu al-maʕrifati/, /mawsu:ʕi:/ /xa:ṣ بموسوعة خاص bi: mawsu:ʕatin/. With the absence of example sentences that identify the collocational ranges of this lexical entry, the information provided by the sense

definition is not definitive enough to help the user to determine whether *encyclopedia* collocates with all information resources other than people or people can safely collocate with *encyclopedia*. Resorting to the BNC, the examples encountered make it so palpable that human beings can also be described as encyclopedic. The following table illustrates this note:

Table 2. Some co-occurrences of *encyclopedic* in the BNC

Term	Frequency	Example Sentences
encyclopedic	28	-...is an *encyclopedic book*… - The danger of using *encyclopedic informants* - *Encyclopedic knowledge* of … - The *encyclopedic John Grant* mainly offers …

In addition to specifying the exact meaning of an adjective based on the context of occurrence, the suggested dictionary provides clues about the syntactic behaviour of an adjective by a multitude of examples that show whether an adjective is used predicatively and/or attributively or postpositively. Features such as ATT (attributive), PRED (predicative), or POST (postpositive) can be employed. Consider the following example sentences extracted from the BNC which show precisely how *utter*, *loath* and *general* are used:

Table 3. Some co-occurrences of *utter*, *loath,* and *general* in the BNC

Term	Frequency	Example Sentences
utter	646	- This is so *utter* nonsense. (ATT) - This has been an *utter* fiasco. (ATT)
loath	101	- He was very *loath* to part with them. (PRED) - I said I was *'loath'*, not 'loathe'. (PRED) - He paused to draw on the cigar, still *loath* to reveal what he had come to say. (PRED) - He held up his hands as if *loath* that they…(PRED)
general	37610	- It should be available for *general* use. (ATT) - … in accordance with the *general* principles and *general* terms. (ATT) - She has noted that the attorney *general* is an independent constitutional officer. (POST)

Based on these co-occurrences of *utter*, *loath* and *general* in the BNC, the first will be tagged in CEAD as ATT only, the second as PRED only while the third as ATT and POST. Providing English-Arabic dictionary users with example sentences which determine the exact meaning of an adjective along with some syntactic features which show how an adjective behaves syntactically will be of a great help for translators and learners of English as a second/foreign language, especially if they are beginners.

Moreover, the proposed dictionary will give a detailed description of adverbs, a part of speech which, sometimes, is treated unfairly in some dictionaries. It is noted that adverbs in some English-Arabic dictionaries pass unnoticed; therefore, there was no systematic way of dealing with them, thus the classical phrase "in an x way" (e.g., in a bad way, in a sad way, etc. as paraphrases of "badly", and "sadly") is found in most dictionary as a description of adverbs of manner. This lack of systematic description of adverbs may lead an English-Arabic user to overgeneralize, i.e. apply the rule of "in an x way" to all adverbs of manner. To illustrate how English-Arabic dictionaries handle adverbs of manner, consider the following lexical entries (sadly and hopefully) adopted from LDLP online dictionaries:

Sadly[1]

1. in a sad way: Jimmy nodded *sadly*.

1. على نحو حزين /ʕala naḥwin hazi:nin/

2. unfortunately: *sadly*, the concert was cancelled.

2. لسوء الحظ، للأسف /lisu:ʔi al-ḥaẓi/, /lilʔsafi/

3. in a way that seems wrong or bad: politeness is *sadly* lacking these days. You are *sadly* mistaken if you think Anne will help.

3. على نحو خاطئ أو سيء /ʕala naḥwin xaṭiʔin/, /sayʔin/

<div align="right">(Longman Dictionary of Modern Arabic, En-Ar)</div>

Sadly[2]

على نحو حزين، على نحو محزن، على نحو مؤسف /ʕala naḥwin ḥazi:nin/, /ʕala naḥwin muḥzinin/, /ʕala naḥwin muʔsifin/

<div align="right">(Thorndike Dictionary, En-Ar)</div>

Sadly[3]

بكآبة، بحزن، بكدر /bikaʔabatin/, /biḫuznin/, /bikadarin/

(al-Muɣni al-Akbar, En-Ar)

Hopefully[1]

1. a word used at the beginning of a sentence when you are saying what you hope will happen: *Hopefully,* the letter will be here by Monday.

1. على أمل أن، راجياً /ʕala ʔamalin ʔan/, /ra:giyan/

2. in a hopeful way: *"Can we go to the zoo tomorrow?"*, he asked *hopefully.*

2. بطريقة آملة /bi ṭariqatin a:milatin/

(Longman Dictionary of Modern Arabic, En-Ar)

Hopefully[2]

بأمل، برجاء، باستبشار، بما يبعث على الأمل /biʔamalin/, /bistibʃarin/, /bima yabʕθu ʕala alʔamali/

(al-Muɣni al-Akbar, En-Ar)

Upon considering the above definitions and examples, only *Longman Dictionary of Modern Arabic* refers, obliquely, to the correct usage of *hopefully*. The absence of such notes of the appropriate usage may get dictionary users to handle *hopefully* and *sadly* as filling the same syntactic slots. Yet, the opposite is true: while it is correct to paraphrase *Sadly, she committed suicide* as *It is sad that she committed suicide*, it is not the case that *Hopefully, she will finish the task before next week*, can be paraphrased as *It is hopeful that she will finish the task before next week*. Therefore, CEAD will focus on the lexical compositionality of each lexical item rather than focusing on the idiosyncratic assumptions associated with it. The proposed dictionary, being based mainly on text corpora, will handle lexical item compositionally, providing a detailed description of each entry, lexically and syntactically.

Providing novel usages of words is not the only job of a corpus of data. The use of corpora helps in identifying word collocations and giving clues about the morphological productivity of words and their subcategorization frames which

helps dictionary users to fill so many gaps in their lexical knowledge. For example, text corpora provide rich information on verb collocates which enrich the dictionary with invaluable sources for multiword constructions such as phrasal verbs, idioms, light verb constructions, stretched verb constructions, serial verb constructions, etc. e.g. occurrences of *dash*, *dawn*, and *zap* in the BNC provide a variety of multiword constructions that these words can be part of (e.g. dash out, dash down, dash off, dash back, dash along, zap through, from dawn to dusk, etc.). Consider the following example sentences extracted from the BNC which provide multiword constructions that can extensively enrich the proposed dictionary:

Table 4. Multiword constructions extracted from the BNC

Term	Frequency	Example Sentences
dash	723	- *Dash out* and build a house. - Cup fans *dashed down* Baker St. - You can *dash off* to Spain for a week's bed and breakfast - And then it will *dash back* to her side for security and reassurance. -…he *dashed a long*…
dawn	2316	- *From dawn to dusk*
zap	51	- Lucker fills up and checks the car whilst I buy some food, *zap it through* the microwave…

8. Updating the Proposed Dictionary

Updating a dictionary depends heavily on text corpora which are updated regularly with new terms that find its way daily to the language in addition to adding nontraditional usages to old terms. The importance of using corpora is attributed to its richness with up-to-date texts which include coinages or new meanings of old terms. Coronavirus-related terms which are missing in most English-Arabic dictionaries and most English-English dictionaries consulted to the date of writing this paper are frequently occurring in the *Coronavirus Corpus*. The following table shows the frequency of occurrences of some of these terms, some of them are monosemous and some are polysemous.

Table 5. Co-occurrences of some coronavirus-related terms in the Coronavirus Corpus

Term	Frequency	Example Sentences
infodemic	455	- *infodemic*, the term given by WHO for the flood of fake news on COVID-19. - *infodemic* … an overabundance of information
social-distancing	3671	-…*social-distancing* practices -… *social-distancing* measures
work from home	17670	- The company told employees who travelled to China to *work from home*.
stay-at-home	18122S	- I wouldn't want to be a *stay-at-home* mother. - …is likely to cement the *stay-at-home* mentality.
self-isolate	10743	- Anyone who has returned to Britain from Wuhan should *self-isolate* and stay indoors.
self-isolation	13929	- The man's family is also in *self-isolation*. -….has been in regular contact with the woman during her *self-isolation* period. -…*go into self-isolation* for 10 days. -…were told to *undergo self-isolation*.
shelter-at-place	1	- More than fifty percent of America's sleep patterns have been negatively affected since *shelter-at-place* orders came into effect.

The *Coronavirus Corpus*, besides including new terms, it also provides new meanings of a word. For example, *lockdown*, which traditionally meant *to put prisoners in a cell where no one is allowed out of his cell*, has achieved a new and much wider usage to describe the situation of sheltering-in-place, i.e., to avoid the risk of being infected with a coronavirus. Moreover, by example sentences, the corpus provides information on the syntactic behaviour of these new terms. For example, *self-isolation* occurs in different syntactic slots: *is in self-isolation*, preceded by a copula, *during her self-isolation*; preceded by a preposition, *go into self-isolation*; preceded by a prepositional verb, *undergo self-isolation;* preceded by a lexical verb. Based on this plethora of example sentences, the syntactic behaviour of novel terms to be added to the dictionary can be meticulously identified.

9. Results

Reliable, contextual English-Arabic dictionaries are no longer possible without the use of text corpora. A modern dictionary in the first place is regarded as a reflection of social changes and is generally used to find systematized information quickly. The different occurrences of new terms in a corpus of data, besides identifying the meanings of these lexical entries in novel contexts provide full subcategorization properties of them. The frequency of occurrences and the creative use of words show the importance of working with authentic language data in lexicography. Furthermore, corpora provide material for selecting potential headwords in addition to identifying the different senses and uses of that headword. By the same token, corpora provide reliable data for the lexical compositionality and syntactic behaviour of words, especially the collocational behaviour and the lexical entry's ability to form multiword units including idioms, phrasal verbs, prepositional verbs, prepositional phrasal verbs, light verb constructions, etc. In this way, lexicographers can identify all combinative potentials of the headword. Nevertheless, it should be stressed that authenticity alone is not substantially a guarantee that an example is suitable to be included in a dictionary entry, i.e. it is recommended that authentic example sentences extracted from a corpus of data undergo a process of revision and scrutiny to be decided as suitable for inclusion.

10. Conclusion

Making a contextual English-Arabic dictionary is anything but easy. It is a toilsome task that needs expertise, effort, and time. Though making such a dictionary is a demanding task, yet it is attainable. The dictionary could be made by eclectic compilation of lexical entries from the specialized dictionaries of the *Academy of the Arabic Language* in Cairo, in addition to some English-English corpus-based dictionaries, and some other English-Arabic sources. Two major corpora are to be used to provide the dictionary with example sentences and help in extracting nontraditional usages of lexical entries: the *British National Corpus* and the *Coronavirus Corpus*. Finally comes the process of adding Arabic equivalents to the lexical entries based on their sense definitions and example sentences that identify the meaning of the lexical entry based on its co-text. The dictionary, in this sense, is not just a reference book but rather a storeroom for English words and their Arabic equivalents. It would be an illusion to think that this could be captured in full if the process of compiling the proposed dictionary did not involve a team of expert advisors or a team of scholars working under the

umbrella of a well-known academic institution. The processes of data collection, compilation, identifying syntactic behaviour, adding real example sentences, and providing contextualized equivalents to lexical entries can be regarded as a strenuous task. However, being based on gigantic amounts of example-based contextualized terms, the copiousness of the proposed dictionary will be quite satisfactory and the final product will mostly provide a good degree of comprehensiveness, authoritativeness, and real usage guidelines.

Endnotes

[1] Macrostructure is the list of all words that are described in a dictionary, while microstructure is all the information given about each word in the macrostructure of a dictionary.

[2] LDLP (Librairie du Liban Publishers) has a collection of specialized and general dictionaries available online. This collection includes some corpus-based dictionaries, the most important of which is *Longman Dictionary of Modern Arabic*. *Al-Muɣni al-Akbar online dictionary* that is frequently cited in this paper is one of these online dictionaries.

[3] COBUILD is an acronym for *Collins Birmingham University International Language Database*. When COBUILD was published in 1987, it revolutionized learner's dictionaries. It was the first of a new generation of dictionaries that were based on real examples of English rather than on compilers' intuition.

References

Al-Ajmi, H. (2004). A New English–Arabic Parallel Text Corpus for Lexicographic Applications. *Lexikos* 14 (AFRILEX-reeks/series: *14:* 326-330.

Al-Muɣhni al-Akbar. Accessed November 15, 2020 at http://www.ldlp-dictionary.com/home/dictionary/.

Arts, T. (2014). The Making of a Large English-Arabic/Arabic-English Dictionary: The Oxford Arabic Dictionary. *Paper presented at the xvi EURALEX Congress*, Italy: EURAC Research.

Asfour, M. (2003). Problems in Modern English-Arabic Lexicography. *Journal of Arabic Linguistics*. 42:41-52. Retrieved January 11, 2021, from https://jstor.org.

Atkins, B. T., & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. New York: Oxford University Press.

Attia, M., Pecina, P., Toral, A., Tounsi, L. & Genabith, J. V. (2011). A Lexical Database for Modern Standard Arabic Interoperable with a Finite State Morphological Transducer. In C. Mahlow and M. Piotrowski. (eds) CCIS 100, 98-118. Berlin: Springer. Accessed January 2021, at http://www.attiaspace.com/Publications/2011-SFCM_MSA_Database.pdf.

Benzehra, R. (2012). Issues and Challenges for a Modern English-Arabic Dictionary. *Journal of the Dictionary Society of North America*. 33, 83-102.

British National Corpus. Accessed February 20, 2021, at http:www.//english-corpora.org/bnc/.

Cambridge Dictionary. Accessed March 24, 2021, at https:www.//dictionary.cambridge.org/dictionary/english/lexis.

Cermak, F. (2003). Source Materials for Dictionaries. In *A Practical Guide to Lexicography*. In Piet Van Sterkenburg (ed.). Amsterdam: John Benjamin's Publishing Company.

Coronavirus Corpus. Accessed January 20, 2021, at http://www.*english-corpora.org/corona/*.

Dolezal, F. T. & McCreary, R. (1999). *Pedagogical Lexicography Today: A Critical Bibliography on Learners' Dictionaries with Special Emphasis on Language Learners and Dictionary Users.* De Gruyter.

English-corpora. Accessed December 23, 2020, at https://www.english-corpora.org/bnc/.

Fillmore, C. J. (1982). Frame Semantics. Linguistics in the Morning Calm, Selected Papers from SICOL. *The Linguistics Society of Korea*, 111-137. Seoul: Hanshin Publishing Company.

Ginzburg, R. S., Khidekel, S. S., Knyazeva, G.Y. & Sankin, A. A. (1979). *A Course in Modern English Lexicology*. Moscow: Vyssaja Skola.

Hanks, P. (2009). The Impact of Corpora on Dictionaries. In Paul Baker (ed.) *Contemporary Corpus Linguistics*. 213-236. London: Continuum.

Haywood, A. J. (1960). *Arabic Lexicography: Its History and Its Place in the General History of Lexicography*. Leiden: E.J. Brill.

Jamiu, F. O. (2014). Science of Arabic Lexicography: A Survey of its Emergence and Evolution. *Journal of Education and Practice*. Vo l5, No. 14.

Kocak, A. Y. (2002). *Handbook of Arabic Dictionaries*. Berlin: Verlag Hans Schiler.

Lane, E. W. (1968). *An Arabic-English Lexicon*. Beirut: Librairie Du Liban.

Lexico. Accessed January 20, 2021, at https://www.lexicon.com.

Longman Dictionary of Modern Arabic. Accessed March 23, 2021, at https://www.ldlp-dictionary.com/home/dictionary/.

McEnery, T., Kiao, R. & Tono, Y. (2006). *Corpus-Based Language Studies: An Advanced Resource Book*. London and New York: Routledge.

Meisami, J. S. & Starkey, P. (1998). *Encyclopedia of Arabic Literature*. London: Routledge.

Meriam-Webster. Accessed July 24, 2020, at https://www.merriam-webster.com/dictionary/.

Onelook dictionary. Accessed December 12, 2020, at https://www.onelook.com.

Pustejovsky, J. (1991). The Generative Lexicon. *Computational Linguistics*. MIT Press. Vo 17. No. 4.

… (1995). *The Generative Lexicon*. Massachusetts: MIT Press.

Rundell, M. (2006). More than one Way to Skin a Cat: Why Full-Sentence Definitions Have not Been Universally Adopted. *A paper presented at the XII, EURALEX International Congress*. Torino: Edizioni dell'Orso.

Sanni, A. (1992). The Arabic Science of Lexicography: State of the Art. *Islamic Studies*, Vol 31, No 2. 141-168. Islamic Research Institute, International University of Islamabad.

Sawalha, M. (2019). The Design and the Construction of the Traditional Arabic Lexicons Corpus (The TAL-Corpus). *Modern Applied Science*. Vol. 13, No. 2. Canadian Center of Science and Education.

Sterkenburg, P. V. (2003). A *Practical Guide to Lexicography*. Amsterdam: John Benjamin's Publishing Company.

Thorndike Dictionary. Accessed March 15, 2021, at https://www.ldlp-dictionary.com/home/dictionary/.

Wiktionary. Accessed May 21, 2021, at wiktionary.org.

Zgusta, L. (1971). *Manual of Lexicography*. Prague: Publishing House of the Czechoslovak Academy of Science.

نحو بناء معجم سياقي ثنائي اللغة من اللغة الإنجليزية إلى اللغة العربية

د. حميد أحمد خلاف

قسم اللغة الإنجليزية، كلية التربية، جامعة دمنهور

hameed_ak2002@yahoo.com

المستخلص

كان للعرب السبق في الصناعة المعجمية، حيث استهلها الخليل بن أحمد الفراهيدي بمعجمه "كتاب العين"، وتلى هذا العمل المعجمي الفريد في عصره سلسلة من المعاجم العربية أحادية اللغة. ورغم أن المعاجم العربية أحادية اللغة تعد فريدة من حيث غزارة المدخلات المعجمية ودقة معاني تلك المدخلات وذلك لاستنادها إلى الاستشهاد بعبارات أصلية توضح معنى المدخل المعجمي وتجليه وتبعد عنه أي لبس ربما يقع فيه القارئ، إلا أن المعاجم ثنائية اللغة والتي تكون فيها اللغة الإنجليزية هي اللغة المصدر بينما اللغة العربية هي اللغة الهدف تفتقر إلى هذا الثراء في المدخلات المعجمية وذلك يعزو إلى انتهاجها القالب التقليدي وهو الشكل الورقي للمعجم وإغفالها الثورة التكنولوجية الهائلة التي وفرها الحاسب الآلي سواء في المعالجة الحاسوبية للمحتوى المعجمي أو القالب الذي يقدم من خلاله المعجم. كما أن تلك المعاجم تفتقر إلى التفريق الدقيق بين المعاني المختلفة للمدخل المعجمي، وذلك لإغفالها المعلومات السياقية للفظة. لذا يناقش هذا البحث إمكانية إعداد معجم ثنائي اللغة (إنجليزي-عربي) يكون شاملا في محتواه، سياقياً في معاني مدخلاته. ومن أجل تحقيق ذلك الهدف، يناقش البحث نقاط ضعف المعاجم ثنائية اللغة المتاحة حاليا، وكيفية معالجة تلك النقاط، كما يحدد مصادر البيانات المعجمية التي يمكن الاستناد إليها في بناء المعجم المقترح، وكذلك الآلية التي يمكن من خلالها صياغة تلك البيانات داخل المعجم ، كما يناقش البحث آلية تحديث المعجم بشكل دوري لإضافة أي مفردات جديدة تدخل اللغة سواء من خلال النحت أو استحداث معان جديدة لمفردات لغوية قديمة، ولذلك يقترح الباحث أن يتم إعداد المعجم المقترح من قبل هيئة أكاديمية مرموقة، كما يقترح الباحث الاستفادة القصوى من الإمكانات التي وفرتها الثورة التكنولوجية الحديثة، لذا يقترح الباحث أن يكون المعجم إلكترونيا؛ حتى يتم تلاشي القيود التي يفرضها القالب الورقي في الصناعة المعجمية.

الكلمات الدالة: معلومات سياقية، الصناعة المعجمية، المدخلات المعجمية، المكنز اللغوي، المعجم

Appendix

List of Arabic Phonetic Symbols

Arabic Consonants	Symbol
أ	/ʔ/
ب	/b/
ت	/t/
ث	/θ/
ج	/g/
ح	/h̲/
خ	/x/
د	/d/
ذ	/ð/
ر	/r/
ز	/z/
س	/s/
ش	/ʃ/
ص	/s̲/
ض	/d̲/
ط	/t̲/
ظ	/z̲/
ع	/ʕ/
غ	/ɣ/
ف	/f/

ق	/q/
ك	/k/
ل	/l/
م	/m/
ن	/n/
هـ	/h/
و	/w/
ى	/y/

Arabic Vowels	**Symbol**
diacritic (damma)ُ	/u/
diacritic (fatha)َ	/a/
diacritic (kasra)ِ	/i/
آ , ى	/a:/
و	/u:/
ي	/i:/