



**EURALEX XIX**  
Congress of the  
European Association  
for Lexicography

**Lexicography for inclusion**

**7-11 September 2021**  
Ramada Plaza Thraki  
Alexandroupolis, Greece

[www.euralex2020.gr](http://www.euralex2020.gr)

**Proceedings Book  
Volume 1**

Edited by Zoe Gavriilidou, Maria Mitsiaki, Asimakis Fliatouras

**EURALEX Proceedings**

ISSN 2521-7100

ISBN 978-618-85138-1-5

Edited by: Zoe Gavriilidou, Maria Mitsiaki, Asimakis Fliatouras

English Language Proofreading: Lydia Mitits and Spyridon Kiosses

Technical Editor: Kyriakos Zagliveris



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

**2020 Edition**

# Arabic Loanwords in English: a Lexicographical Approach

Fournier P., Latrache R.

Sorbonne Paris Nord University, France

## Abstract

This article deals with Arabic loanwords in English from a lexicographical perspective. To create a representative corpus of Arabic loanwords in English, items are extracted from the *Oxford English Dictionary* database (henceforth *OED*) with an etymological advanced search. Among the criteria affecting the etymological tagging, the concept of two languages of origin is probably the most difficult one for lexicographers to deal with. This study presents some of the issues lexicographers are faced with in the dictionary-making process. Following that, Arabic loanwords are classified according to semantics, along with the date of their first attestation in the *OED* database. This quotation dating work that the *OED* systematically performs is not only an immense task, but also an essential one, as it enables researchers to determine the semantic spheres these corresponding loanwords are integrated into, as well as the cultural relationship between Arabic-speaking countries and English-speaking countries.

**Keywords:** Arabic loanwords; contemporary English; dictionary-based study

## 1 The Arabic Language

Arabic is considered as one of the major languages with a tremendous cultural impact in the world (Sapir 1921). According to Salloum & Peters (1996) 6500 Arabic loanwords are attested in the English language, though many of them have been introduced through the Spanish language (in Thawabteh 2011: 104). Indeed, Serjeantson (1935: 213-220) determines that Arabic loanwords can be either direct (i.e. with no intermediate language between Arabic and English) or indirect (with notably French or Spanish as transitory languages). She underlines the impact of Arabic on the English language (1935: 213): “It is from Arabic that English has borrowed the greatest number of Eastern loan-words, though it is true that a considerable proportion of them have not come to us direct”. The best explanation is linked to science (mathematics, astronomy...), business factors (especially during the 14<sup>th</sup> century in North Africa) and exploration as well. If older loanwords were rooted in science, the latter ones were representative of everyday life, in zoology and religion as Thawabteh’s description of semantic fields clearly demonstrates (2011). Darwish (2015: 106) offers a historical perspective of the Arabic loanwords in English: “Wilson (2001) notices that by the eighth century in North Africa, Arabic had ousted Latin as the dominant language; by the eleventh and twelfth centuries, Arabic civilisation had fully spread through Spain.” Thawabteh (2011:114) indicates that “[t]he development of Arab architecture, particularly in Granada, Seville, and Cordova was a catalyst for numerous borrowings.” The most prolific periods of borrowing were during the Middle Ages and the Renaissance when “English speakers came into contact with the prestigious intellectual centres of the Arab World” (Darwish 2015: 107).

The result was:

a flow of borrowings from Arabic into English, primarily in the fields of chemistry, medicine, philosophy, mathematics, astronomy, optics, physics, botany, literature, religion (chiefly Islam), music, warfare, shipping, trade, architecture, geography, government and sovereignty (Daher 2003 in Darwish 2015:107).

## 2 Corpus Building: Methods, Problems and Limits

Several factors can affect the creation of a representative database of Arabic loanwords in English. The corpus is elaborated thanks to the *OED* database using the advanced research tool. When mentioning “Arabic” as the language of origin and the main criterion of the advanced search, this first extraction attests 511 items.<sup>1</sup> Yet it appears that 102 words should be discarded because they do not truly correspond to the notion of “Arabic loanword.” First, some extracted words are morphologically composed of stems/roots of Arabic origin and English suffixes (examples: *Bohairic*, *Fatimite*, *Hanbalite*, *Saadian*). Such hybrid words are not representative of Arabic loanwords, but they show that these items can be truly integrated into the English lexicon, so that they can be affected by the word-formation rules of English. Approximate translations or rewriting of Arabic words (examples: *Hobson-Jobson*, *mockery*, *nugger*, *sheregrig*) are not kept in our analysis. Finally, items which are ultimately of Arabic origin, but whose transfer into English was made thanks to intermediate languages, are not kept either. The direct borrowing parameter is an ambiguous one and there may be controversy about it. Indeed, Serjeantson (1935) admits that Arabic loanwords can be either direct or indirect. Mossé (1943) assumes, concerning French loanwords in English, that whatever the ultimate origin of loanwords may be, the language which provides the new item should be considered as the source language. Well-known examples exist such as

<sup>1</sup> 511 items at the time of creating the database. The *OED* frequently adds more items that need further research.

*giraffe*, which is an Arabic word. However, it is considered to be a French loanword in English because it was transmitted to English through the French language. Therefore, researchers must be careful not to be mistaken by etymological information provided by dictionaries, which can include precise details concerning the ultimate origin of words and the source language/target language question. This analysis of Arabic loanwords is dictionary-based and the etymological tagging of the *OED* is adopted. However, some examples of non-selected words are listed below along with the etymological information provided by the *OED*. Some of the emblematic Arabic words are, interestingly enough, not tagged as Arabic loanwords because Arabic is not the transmission language.

The following examples are particularly relevant because they belong to the religious semantic domain of Islam, and since Arabic is the language of Islam and of its holy book, the Qur'an, the ultimate origin of these examples is necessarily Arabic. Hence as an example, the word *Islam* is not tagged as a direct borrowing from Arabic, but as being of multiple origins.

(1) '*Islam*'. (ID 99980)

Origin: Of multiple origins. Partly a borrowing from Turkish. Partly a borrowing from Arabic.

Etymons: Turkish islām; Arabic islām.

Frequency (in current use): 6

The word *Islam* had probably been transferred from Arabic into Turkish long before the first contact between the two languages, English and Arabic, was made. Thus, it is highly probable that English speakers may have had contacts with both Turkish and Arabic, which is the reason why the *OED* database could not determine exactly where the word had originated. The *OED* suggests some linguistic contact with the "strongly Arabized Ottoman Turkish language":

Contact between English speakers and Islam in the early modern period was chiefly through the institutions and peoples of the Ottoman Empire, which controlled south-eastern Europe, the Middle East, and North Africa and used the strongly Arabized Ottoman Turkish language.

Another example is related to the two major Muslim feasts; '*Eid-al-Adha*' and '*Eid-al-Fitr*', which are Arabic words, but which are tagged as follows by the *OED*:

(2) *Eid*. (ID 242685)

Origin: Of multiple origins. Partly a borrowing from Persian. Partly a borrowing from Arabic. Partly formed within English, by compounding.

Etymons: Persian 'īd ; Arabic 'īd ; Eid-al-Adha n., Eid-al-Fitr n.

Frequency (in current use): 4

(3) *Eid-al-Adha*. (ID 243484)

Origin: Of multiple origins. Partly a borrowing from Persian. Partly a borrowing from Arabic",

Etymons: Persian 'īd-i aẓḥā; Arabic 'īd al-aḏḥā

Frequency (in current use): 2

(4) *Eid-al-Fitr*. (ID 91137)

Origin: Of multiple origins. A borrowing from Arabic; modelled on a Persian lexical item.

Etymons: Arabic 'īd al-fīṭr.

Frequency (in current use): 2

The above *OED* tagging is interesting. These words are semantically related, yet in the last item it appears that the origin is accounted for differently. The same argument evoked concerning Turkish can be used for the first two loanwords (*Eid* and *Eid-al-Adha*). First, '*Eid*' in Arabic means "a celebration of". Its listing as a borrowing from Persian suggests that it is a word from the pre-Islamic period. However, all the quotations used show that the word is related to Islam. Moreover, Persian uses other words to translate "celebration" and '*Eid*' is not one of them. It proves that Persian borrowed those two words from Arabic and that these loanwords may have been integrated into English through the exposure to the Persian language. However, the reference to the Persian origin is not consistent, as the last one is listed as "modelled on a Persian lexical item". *Eid-al-Fitr* is said to be an Arabic loanword, modelled on a Persian lexical item. This argument is problematic as it refers to two different processes. If this is an Arabic loanword, it means that Arabic is considered to be the source language. Yet as it is ultimately an Arabic word, how could it possibly be modelled on a Persian lexical item? The word *Ramadan* is another example of etymological tagging, showing that when lexicographers are confronted with deciding between several possible languages of origin, those languages are simply listed.

(5) *Ramadan*. (ID 157727)

Origin: Of multiple origins. "Of multiple origins. Partly a borrowing from Persian. Partly a borrowing from Turkish. Partly a borrowing from Arabic."

Etymons: Persian ramaẓān, Turkish ramaẓān; Arabic ramaḏān.

Frequency (in current use): 4

*Ramadan*, being the holy month for Muslims, can only be of Arabic origin. Its etymons in Persian and Turkish are probably mere pronunciation variations of the Arabic word.

The corpus of Arabic loanwords this study is based upon is composed of 409 items (see the whole corpus in Annex 1 along with the list of non-selected words in Annex 2). The Arabic loanwords will be analysed through different angles in the next sections: dating work, semantic domains and frequency. The last part deals with the *OED* quotations and the

impact data collectors may have on the influx of Arabic loanwords in English.

### 3 Influx of Arabic Loanwords Through Centuries

Initially when the dates of the first attestation of the Arabic loanwords in the *OED* were examined, it was apparent that the influx of Arabic loanwords per century was inconsistent. This method, which consists of determining the influx of loanwords thanks to the *OED* quotations, obviously has some limitations. These restrictions are linked to the *OED* data collectors (Gilliver 2015) (this point will be dealt with in Section 6). Table 1 classifies the number of Arabic loanwords integrated into English century by century.

14 <sup>th</sup>	15 <sup>th</sup>	16 <sup>th</sup>	17 <sup>th</sup>	18 <sup>th</sup>	19 <sup>th</sup>	20 <sup>th</sup>
1	6	29	73	76	162	62

Table 1: Number of Arabic loanwords integrated into English through centuries.

It turns out that the beginning of the phenomenon can be dated back to the 17<sup>th</sup> century with a peak in the 19<sup>th</sup> century with 162 loanwords out of 409 (i.e. 39.6%), which had been integrated into the English language. The correlation between these figures and the study of semantic domains in the next section may explain why the number of Arabic loanwords reached such a culmination in the 19<sup>th</sup> century.

### 4 Semantic Domains

A detailed survey of the semantic domains of the 409 Arabic loanwords in English along with the number of loanwords borrowed per century is presented in Table 2. Only semantic areas with at least 8 loanwords are integrated into the table. The other semantic domains are considered as minor ones. The corpus covers seven centuries, from 1393 to 1996, a period during which the *OED* database attested 113 items belonging to the religious semantic domain and 53 items belonging to the science semantic domain. Those two domains are unquestionably the most prolific ones. Science is the oldest semantic field Arabic loanwords are related to in the *OED*. The first direct Arabic loanword referenced by the *OED* is *Aldebaran*, which dates back to 1393, belonging to the semantic field of astronomy. Indeed, science brings together such domains as botany (example: *sebesten*), astronomy (example: *Al Nath*) or chemistry (example: *azoth*). If at first, only scientific Arabic words were borrowed, this phenomenon changed from the 16<sup>th</sup> century. Indeed, from then on, Arabic loanwords covered a large range of semantic areas such as religion, commerce, food and drinks and so on... Furthermore, it is worth mentioning that in the 17<sup>th</sup> century, the religious semantic domain began to prevail over the other semantic areas and that this field generated more loanwords than the totality of the loanwords from all the other categories.

This table raises essential issues. First, one can wonder why the borrowing process between Arabic and English, which was very limited before the 16<sup>th</sup> century, suddenly accelerated and covered various semantic fields. Then, Table 1 above shows that there was a massive influx of Arabic loanwords in the 19<sup>th</sup> century. How could these two phenomena be accounted for? The survey of *OED* quotations might be of interest in that respect.

	14 <sup>th</sup> c.	15 <sup>th</sup> c.	16 <sup>th</sup> c.	17 <sup>th</sup> c.	18 <sup>th</sup> c.	19 <sup>th</sup> c.	20 <sup>th</sup> c.	total
religion	/	/	5	30	20	48	10	113
science	1	4	8	6	7	25	2	53
clothing	/	/	2	1	7	14	1	25
food / drinks	/	/	1	3	2	9	9	24
titles	/	/	2	6	5	6	2	21
commerce	/	/	1	3	4	2	3	13
arts	/	/	/	2	4	6	/	12
weight	/	/	5	/	3	/	/	8
desert	/	/	/	1	1	3	3	8
sailing	/	/	/	3	/	3	2	8
geopolitics	/	/	/	/	/	/	8	8

Table 2: Number of Arabic loanwords classified by semantic domains along with centuries.

But before dealing with *OED* quotations, the next section investigates the frequency of Arabic loanwords in English.

### 5 Frequency

The *OED* frequency tool is used to investigate frequency in Arabic loanwords (see Annex 3 for the whole list of

loanwords sorted out with frequency numbers). It classifies words from 0 to 8, 0 corresponding to obsolete words and 8 to words very frequently used in English. For example, the most used Arabic loanwords in English are *coffee* and *Muslim*, which are classified 6 out of 8 on the frequency scale. The majority of Arabic loanwords are ranked 2 or 3 out of 8 and it corresponds to 303 items (i.e. 74.1% of the whole corpus). Some words in these categories are used in everyday life (examples: *hammam*, *falafel*, *harissa*), but there are also words, which are used in such restricted semantic areas that native English speakers are probably not familiar with them. The 4 out of 8 classification is subjectively considered to be the clear-cut separation between rare words and more attested ones. Indeed, such words as *giraffe* or *Sahara*, which are well-known Arabic loanwords, belong to this category. Table 3 gives examples of Arabic loanwords, from the less frequent items to the most attested ones.

-	- / +	+
<i>cabeer, caroteel / carotel, fana, gaiassa, ghazeeyeh, halawi, iggri / iggry, kuphar, nil, quaiss kitir, rottol, shereefa, tahalli, Takbir, Zarnich</i>	<i>adhan, alim, ardeb, arrack, bejel, burgoo, daman, doum, fennec, fitna, hadj, hakim, halal, halfa, hamza, hawala, hijab, ihram, jebel, jihadi, jinnee, jol, kantar, kazi, khamsin, khan, khatib, khor, khula, khutbah, kohl, lablab, lebbek, loofah, madhhab, madrasah, maghrib</i>	<i>diss, Iraqi, Koran, Qur'an, Saudi, Sufi, Sunni, sunt, Swahili, coffee, Muslim</i>

Table 3: Examples of Arabic loanwords based on frequency parameters.

## 6 Working with the OED Quotations

The last section deals with the *OED* quotations and the impact this tremendous work of data collecting may have on the perception of Arabic loanwords. A detailed analysis shows that a consequential number of quotations came from the same sources. For instance, 50 quotations (first and second citations) in the 19<sup>th</sup> century came solely from the works of two famous Orientalists, Edward W. Lane and Richard F. Burton. The *OED* does not give details about the suppliers of those quotations, but it is clear that they relied on the same sources; which is common according to Gilliver (2015: 51): “This pattern of a small group of readers producing the lion’s share of the quotations [...] has recurred throughout the Dictionary’s history.” If the loanwords massively integrate English through the translations of a few books, the meaning of words can be biased. One example would be the different translations of the *Arabian Nights* used as sources by the *OED*.

The most famous direct translations of the *Arabian Nights* from Arabic into English were done by the Orientalists Edward W. Lane (1801-1876), Richard F. Burton (1821-1890) and John Payne (1842-1916). Indeed, our research reveals that eleven Arabic loanwords in English were introduced through their different translations and editions of the *Arabian Nights* or *Thousand & One Nights*; eight items were introduced through Edward W. Lane’s works (as a first and / or second citation), while two citations (second and third) were from Richard F. Burton’s works. Only three items were used in third citations from John Payne’s works.

It is worth mentioning that it was Jean-Antoine Galland (1646-1715) who first introduced *Les Mille et une nuits* (known as *The Arabian Nights* in English) to Western countries by translating them directly from Arabic into French. According to Irwin Robert (1994:16), “Galland used a three-or four-volume manuscript, dating from the fourteenth or fifteenth century, as the basis for his translation”. The first volume of *Les Mille et une nuits* was published in 1704, the twelfth and final volume was published in 1717 (BNF Essentiels Littérature). The book was an immediate success, first in Europe, and then in America and in Australia. As a result, some parts of Galland’s work were translated into English soon after their publication. However, little is known about the translator or the exact date of the first English translation (1706? 1708?) (Knipp 1974: 52; Irving 1994:19).

In spite of the well-known great success of the *Arabian Nights*, our detailed study of the citations reveals that the largest number of Arabic loanwords in English has been included into the *OED* database prior to the publication of the *Arabian Nights*. Indeed, 40 items were used in quotations from the works of the same two Orientalists (who also translated the *Arabian Nights*). These two major books are *An Account of the Manners and Customs of the Modern Egyptians* (1836) by Edward W. Lane and *Personal Narrative of a Pilgrimage to El-Medinah and Meccah* (1855) by Richard F. Burton. Lane’s book introduced 30 items, including 19 first citations and 9 second citations, while Burton’s book introduced 10 items, including 4 first citations and 5 second citations. This phenomenon might explain why the influx of Arabic loanwords culminated in the 19<sup>th</sup> century. Furthermore, Burton’s work, which deals with religion, accounts for the massive number of religious Arabic words that entered the English language. The data collectors selected those translations from Arabic into English probably because they knew those words were direct translations from Arabic and the first evidence ever of these items in English.

The consequences of direct translation from Arabic into English are multiple. The most obvious one is the reduction of meaning of some words. The *OED* does not always offer a wide range of meanings for words such as *madrassa* for instance.

### (6) *Madrassa*. (ID 112073)

Origin: Probably of multiple origins. Probably partly a borrowing from Turkish. Probably partly a borrowing from Persian. Probably partly a borrowing from Urdu. Partly a borrowing from Arabic.

Etymons: Turkish *medrese*; Persian *madrasa*; Urdu *madrasa*, *madarsa*; Arabic *madrasa*.

Frequency: (in current use): 4

Definition:

1. In Muslim countries: a school of Islamic theology and law; (also more generally) a school (esp. a secondary school) or institution of higher Islamic education.
2. In other Muslim communities (esp. South Africa, in form *madressa* South African /mə'dresə/): a Muslim school, operating after normal school hours and teaching children subjects such as Islamic history, Islamic belief, and the reading, memorizing, and reciting of the Qur'an.

The two definitions provided by the *OED* give the word a religious dimension. However, *madrasa* does not necessarily mean a “religious school”, it could also mean “a school” that provides any kind of non-religious education.

Thawatbeh (2011:111-112) also gives the example of *Jihad* which is defined as “a religious war of Muslims against unbelievers, inculcated as a duty by the Qur'an and traditions” whereas it also has other meanings such as “a struggle against one's self” or “stating the truth forcibly” or “refraining from [doing/saying] bad things.” Thawatbeh (2011: 112) argues that this “limited view of *Jihad* seems to be ideologically motivated. The foreign text is imprinted with values specific to the target culture.”

It is true that the selection of quotations can be subjective and therefore reflects the opinion of the suppliers rather than the different meanings of the words that can change and evolve according to the context, the period and the society in which they are used. The two examples of *Madrasa* and *Jihad* show that the current geopolitical context emphasizes one specific sense of those words. Gilliver (2015: 71) highlights the importance of quotations and their context when he writes:

I constantly find myself needing quotation evidence (...) recent evidence, for example, which demonstrates that a particular sense of a word (...) is still current, or an earlier example of a particular use of a word that can only be confirmed to be an example of that use by careful reading of the extended context.

Some words show that the *OED* sometimes provides several definitions reflecting the context in which a word is used. For instance, today, *fatwa* has a negative connotation as it is generally used in the sense of “a death sentence delivered by a Muslim authority”. Not only does the *OED* give another meaning of the word *fatwa*, but it also explains the reasons behind its negative connotation:

(7) *Fatwa*. (ID 69642)

1. a. Islam. A formal, authoritative ruling on a point of Islamic law; a scholarly opinion given (typically in writing) by a mufti or other Muslim juridical authority in response to a question posed by an individual or a court of law.
  - b. irregular. A declaration or decree by a Muslim authority calling for a person to be put to death, typically as a punishment for blasphemy or apostasy; a death sentence.
- Modern use in this sense appears to be influenced by the frequent (though inaccurate) glossing of fatwa as 'death sentence' in media reports of a fatwa issued by Ayatollah Khomeini in 1989, which called for the killing of Salman Rushdie, whose 1988 novel The Satanic Verses was considered by some to be blasphemous and insulting to Islam.*
2. An edict or statement issued by a religious authority belonging to a faith other than Islam. Also occasionally in colloquial or trivial use: a forcefully expressed opinion, judgement, or condemnation; a decree.

Many Arabic loanwords in English are exclusively defined according to religious criteria when broader definitions are attested in Arabic. However, the semantic areas those loanwords cover are restricted and conditioned by translators when integrating the English language. Indeed, especially in the 19<sup>th</sup> century, direct translations from famous Arabic books, whose topics mainly deal with religious matters, necessarily restrict the scope of the meanings of the original Arabic words in English. The topics of translated books are therefore determining parameters in the survey of semantic properties of loanwords. It appears that this translation phenomenon from literary sources can account for the great majority of Arabic loanwords in English.

## 7 Conclusions and Future Research

The complexity of a dictionary-making process is reflected in this case study of Arabic loanwords in English. Determining the origins of words can be challenging for lexicographers because it appears that the etymological tagging of the *OED* can raise questions. The *OED* database provides interesting information related to semantics and frequency. It has also been possible to measure and date the influx of Arabic loanwords in English through the quotation system. The diversity of quotations in terms of periods and sources is also another challenging, yet determining, aspect that must be taken into consideration. Writing dictionaries is a multidimensional process that should be considered through an interdisciplinary lens. This preliminary research on Arabic loanwords in English needs further investigation on several issues. It would be interesting, from a geopolitical perspective, to work on the views English-speaking countries may have on Arabic-speaking ones through the integration of loanwords in institutionalized references as well as on the literary sources the *OED* quotations are extracted from. The last section on the *OED* quotations has only tackled a few points related to translation and semantics, but it would be interesting, especially when dealing with 20<sup>th</sup> century loanwords, to try to connect the integration of Arabic loanwords with geopolitical events that have taken place in the world. In that the massive influx of Arabic loanwords in the 19<sup>th</sup> century can legitimately be assumed to be the result of literary translation; it still has to be determined whether more recent loanwords have been integrated into the English language through the same prism.

## 8 References

- BNF Essentiels Littérature. *Les Mille et Une Nuits*. Galland (1704). Accessed at: <https://gallica.bnf.fr/essentiels/galland/mille-nuits> [25/05/2020].
- Darwish, H. (2015). Arabic Loan Words in English Language. In *IOSR Journal of Humanities and Social Science*, 20(7), pp. 105-109.
- Gilliver, P. (2015). The Quotation Collectors: A Conspectus of Readers for the *Oxford English Dictionary*. In *Dictionaries*, 36, pp. 47-71.
- Irwin, R. (1994). *The Arabian Nights: A Companion*. London: Allen Lane.
- Knipp, C. (1974). The "Arabian Nights" in England: Galland's Translation and Its Successors. In *Journal of Arabic Literature*, 5, pp. 44-54.
- Mossé, F. (1943). On the Chronology of French Loan-Words in English. In *English Studies*, 25, pp. 33-40.
- Salloum, H. & Peters, J. (1996). Arabic Contributions to the English Vocabulary: English Words of Arabic Origin; Etymology and History. Librairie du Liban Publ.
- Sapir, E. (1921). *Language. An Introduction to the Study of Speech*. New York: Harcourt, Brace.
- Serjeantson, M.S. (1935). *A History of Foreign Words in English*. Routledge & Kegan Paul.
- Simpson J. & Weiner, E. (1989). *Oxford English Dictionary. Online Edition*. Oxford: Clarendon Press. [20/04/2020]
- Thawabteh, M. (2011). The Other Side of the Coin of Lexical Borrowing from Arabic into English. In *TransculturAl*, 1, pp. 103-122.

### Appendix 1: Corpus of Arabic Loanwords in English (409 items)

aba, abaya, abjad, Adeni, adhan, agal, Ahmadiyya, Alawi, Al Borak, Aldebaran, Al-Hajj, Alhamdulillah, alim, Allah, Allahu Akbar, al Qaeda, Altair, Amal, ameer, Ansar, Ansayri, ardeb, argan, argel, argileh, ariel, arrack, askar, athanor, atlas, azoth, Baath, baba ganoush, bahar / barr(e)<sup>2</sup>, Bedu, bejel, ben, bint, bismillah, burgoo, Caaba, cabeer, cabob, cadi, cafila, camise / camiss, caratch, caroteel / carotel, coffee, coffle, cossid, cowle, dabuh, dahabeeyah / dahabiah, daman, darbuka, deloul, Deneb, dibs, dieb, dirhem / dirham, diss, douar / dowar, doum, Druse / Druze, durra / dhurra, Eid-al-Fitr, El Nath, emir, ezan, faki, fakir, fana, Fatah, Fatiha(h), faufe(l), fedayeen, feddan, falafel, fellagha, fellah, fennec, fils, fingan / finjan, fiqh, fitna, fluce / floose, Fomalhaut, fonduk, frasilah, freekeh, futah, gaiassa, galabiya, gandoura, Garshuni, Gelalaeen, ghaffir, ghawazee, ghazeeyeh, Ghazi, gholam, ghoul, Ghuzz, gibli, gimbri, giraffe, girba, gourbi, grab, gufa, gundi, habara, haboob, Hadith, hadj, hadji / hajji, haik / haick, hakeem / hakim, hakim, halal, halala, halawi, halfa, Hamas, hammada, hammal / hummaul, hammam / hummaum, hamza, Hanif / Haneef, haram, haram<sup>3</sup>, harissa, harka, hashish / hasheesh, hawala, henna, hijab, Howeitat, hummum, huzoor, hygeen / hajeen, iggri / iggry, ihram, ijtihad, ilb, imam, imshi, inshallah, intifada, Iraqi, ithel, izar, jann, jebel, jelab, jerid / jereed, jerm, jihad, jihadi, jinn, jinnee, jol, jubbah, Kababish, Kabyle, kadish, Kaffir, kali, kantar, kat, kazi, kef / keif / kief, keffiyeh, khamsin, khan, kharaj / kharatch, khatib, Khilafat, khimar, khor, khula, khutbah, kibbeh, kiblah, kibrit, kissar, Kiswa, Kitab, kitar, kohl, Koran, Koreish, kourbash / koorbash, kuphar, Kuwaiti, lablab, leban, lebbek, litham, loof, loofah, madhhab, madrasah, mafeesh, maghrib, Maghribi, mahaila, mahalla, Mahdi, mahmal, mahr, Makhzan, malem, malesh, Maliki, Mameluke, Mamur, mancala, mandarah, mandil, maqam, marid, mashallah, mastaba, maulana, maund, mauz, mawla, mawlid, medina, mellah, Mendoub, merissa, Metawileh, Metran, mhorr, mihrab, millime, minbar, Miraj, mithqal, mohair, Muallaqat, mudhif, mudirieh, muezzin, muhajir, Muhammad, Muharram, Mukhabarat, mukhtar, mulai, murid, Murji'ah, murshid, Muslim, Muslimah, Mussulmin, Mussulmin, mut'a, mutawwa, nabk, naib, Naqib, Nasara, naskhi, Nasrani, nil, niqab, Nusayri, omda, Othman, oud, pastilla, qasida, Qatari, quaiss kitir, quies, Qur'an, qursh, raad, rafiq, rayah, rebab, reg, resalgar, riad, riba, riqq, robin, roc, rottol, Roumia, Rwala, sabkha, Sahara, Sahrawi, sakia, salaam, Salafi, Salafiyya, salat, saluki, samn, santir / santour, sarsar, Saudi, sebesten, seif, semsem, Senussi, Senussia, serab, serir, seyal, shadda, shadoof, shahada, shahid, Shaitan, shamal, Shammar, Sharia, shawarma, sheikh, Sheikha, Sherarat, Sherari, shereef, shereefa, sherifi, sherwal, Shiah, shisha, shott, shufti / shufti, shura, Siddi, sief, simoom, sim-sim, Subbhanallah, sudd, Sufi, sulham, sultany, Sunna, Sunni, sunt, sura, Swahili, syce / sais, tabbouleh, tabl, taboot, tahalli, tahina, taj, tajine, takaful, Takbir, talak, talha, tarada, tarboosh, tarfa, tariqa, tawaf, tazia, tecbir, tell, terjiman, tezkere / teskere, tibbin, tobe, torba, tuba, tumbak, tumbaki, ulema, umma, umrah, urs, wadi / wady, Wafd, Wahabi / Wahabee / Wahhabi / Wahabee, wakf / waqf, wali, wazir, weli / wely, worral, Yahudi, yashmak, Yemeni, yimkin; Yunani, zaatar, Zaidi, Zar, zarf / zurf, zariba / zareba, zarnich, zawiya, Zendik, zibib, ziczac, zikr, ziraleet

### Appendix 2: Non-Selected Words (102 items)

afion, afrit, Alawite,alconde, Alcoran, alfaqui, Algebar, Alhaji, alif, alkanet, almeh, askari, Bohairic, bulbul, burkundaz / burkundauze, cabaan / caban, cazimi, chermoula, dewan, dghaisa, dinar, dubba / dubber, Eid, Eid-al-Adha, Fatimite, fatwa, fedai, felucca, fistic, gerfaunt, halva, Hanafite, Hanbalite, hardun, harem, Hezbollah, Hobson-Jobson, hookah, Hubshee, iddat, Islam, Kadarite, kalian / kalioun, karrozzin, keiri, kharif, Kurd, lampuki, Lyhianic, Liwianic, liwa, Madan, maid, majlis, mangal, manzil, Marinid, mashrabiyya, masjid, mattamore, mazar, mimation, minaret, mocker, mocked up, Moslemah, mudir, mufti, mujahid, mujahidin, mullah, muqaddam, Mussulman, Mutazilite, nabi, nugger, nunation, oojah, otto, ouguiya, Ramadan, razzia, redif, reis, rial, Rifai, Roubi, rubai, rufiyaa, Saadian, Sabian, sahib, Sahidic, Salafist, sandal, sayyid, Shafiite, sheregrig, spitchered, tabasheer, tamasha, tamboura, vizierate

<sup>2</sup> The items which are separated by slashes correspond to the different spellings.

<sup>3</sup> *Haram* has two distinct entries in the dictionary.



**Appendix 3: Frequency of Arabic Loanwords in English**

No data = ameer, Ansar, cadī, dabuh, faufe(1), freekeh, mandil, mauz, Mussulmin, Othman, resalgar, robin, sief, sultany, terjiman, worral, zikr (17 items)

1/8 = cabeer, caroteel / carotel, fana, gaiassa, ghazeeyeh, halawi, iggri / iggry, kuphar, nil, quaiss kitir, rottol, shereefa, tahalli, Takbir, Zarnich (15 items)

2/8 = abjad, agal, Al Borak, Aldebaran, Alhamdulillah, Allahu Akbar, Altair, Amal, Ansayri, argan, argel, argileh, askar, athanor, azoth, baba ganoush, bismillah, Caaba, cabob, cafila, camise / camiss, caratch, cossid, cowle, dahabeeyah / dahabiah, darbuka, deloul, Deneb, dibs, dieb, Eid-al-Fitr, El Nath, ezan, faki, Fatah, fellagha, fingan / finjan, fluce / floose, Fomalhaut, fonduk, frasilah, futah, galabiya, gandoura, Garshuni, Gelalaeen, ghaffir, ghawazee, gholam, gibli, gimbri, girba, gourbi, gufa, gundi, habara, haboob, haik / haick, hakeem / hakim, halala, Hamas, hammada, hammal / hummaul, harka, Howeitat, hummum, huzoor, hygeen / hajeen, ilb, imshi, inshallah, ithel, izar, jann, jelab, jerid / jereed, jerm, jubbah, kadish, keffiyeh, khimar, kibbeh, kiblah, kibrit, kissar, Kiswa, kitar, kourbash / koorbash, leban, litham, loof, mafeesh, mahaila, mahmal, malem, malesh, mancala, mandarah, marid, mashallah, Mendoub, Merissa, Metawileh, Metran, mhorr, Muallaqat, mudhif, mudirieh, Murji'ah, Muslimah, mutawa, nabk, Nasara, naskhi, Nasrani, niqab, omda, qursh, raad, rafiq, riad, riqq, Roumia, sakia, samn, santir / santour, sarsar, sebesten, semsem, Senussia, serab, serir, shadda, shadoof, shamal, shawarma, Sherarat, Sherari, sherifi, sherwal, shisha, shufti / shuftu, sim-sim, Subbhanallah, sulham, taboot, takaful, talha, tarada, tarfa, tecbir, tezkere / teskere, tibbin, torba, tumbak, tumbaki, umrah, Yahudi, yimkin, zaatar, zarf / zurf, Zendik, zibib, ziczac, ziraleet (165 items)

3/8 = abaya, Adeni, adhan, Ahmadiyya, alim, ardeb, ariel, arrack, bahar / barr(e), Bedu, bejel, burgoo, coffle, daman, douar / dowar, doum, durra / dhurra, Fatiha(h), falafel, fennec, fitna, Ghuzz, grab, hadj, hadji / hajji, hakim, halal, halfa, hammam / hummaum, hamza, Hanif / Haneef, haram, haram, harissa, hawala, hijab, ihram, jebel, jihadi, jinnee, jol, Kababish, kantar, kazi, kef / keif / kief, khamsin, khan, kharaj / kharatch, khatib, khor, khula, khutbah, kohl, Koreish, lablab, lebbek, loofah, madhhab, madrasah, maghrib, Maghribi, mahalla, mahr, Makhzan, Mamur, maqam, mastaba, maund, mawla, mawlid, medina, mellah, millime, minbar, Miraj, mithqal, muhajir, Muhammad, Mukhabarat, mukhtar, mulai, murid, murshid, Muslimin, mut'a, naib, Naqib, Nusayri, oud, pastilla, qasida, Qatari, quies, rayah, rebab, riba, Rwala, sabkha, Sahrawi, salaam, Salafi, Salafiyya, salat, saluki, seif, Senussi, seyah, shahada, shahid, Shaitan, Shammar, Sheikha, shereef, Shiah, shott, shura, Siddi, simoom, sudd, sura, tabbouleh, tabl, tahina, taj, tajine, talak, tarboosh, tariqa, tawaf, tazia, tell, tuba, urs, weli / wely, yashmak, Yunani, zariba / zareba, zawiya (138 items)

4/8 = aba, Alawi, Al-Hajj, Allah, al Qaeda, atlas, Baath, ben, bint, dirhem / dirham, Druse / Druze, emir, fakir, fedayeen, feddan, fellah, fils, fiqh, Ghazi, ghoul, giraffe, Hadith, hashish / hasheesh, henna, ijti had, imam, intifada, jihad, jinn, Kabyle, Kaffir, kali, kat, Khilafat, Kitab, Kuwaiti, Mahdi, Maliki, Mameluke, maulana, mihrab, mohair, muezzin, Muharram, reg, roc, Sahara, Sharia, sheikh, Sunna, syce / sais, tobe, ulema, umma, wadi / wady, Wafd, Wahabi / Wahabee / Wahhabi / Wahabee, wakf / waqf, wali, wazir, Yemeni, Zaidi, Zar (63 items)

5/8 = diss, Iraqi, Koran, Qur'an, Saudi, Sufi, Sunni, sunt, Swahili (9 items)

6/8 = coffee, Muslim (2 items)