Collocation and the Arabic-English Dictionary:
Ideas for Better Dictionaries

Abdul-Fattah Abu-Ssaydeh
University of Sharjah

Abstract: Collocation, thus far, has been infrequently used to clarify senses in Arabic-English dictionaries. However, this lexical phenomenon can be used more effectively to improve these dictionaries in many ways: it can help determine if a given word belongs to Modern Standard Arabic (MSA) or Classical Arabic; it can identify the various lexical contexts in which the headword may appear, thus enabling the lexicographer to use frequent ones in the dictionary and separate collocations from multi-word units; it can indicate cases of undertranslation and overtranslation; senses given to a certain headword can be verified against the collocational range of the headword in order to decide if such English equivalents are indeed valid senses or not and, finally, the systematic use of collocates in the bilingual dictionary as well as the development of collocational bilingual dictionaries can be demonstrably of great help to the language learner.

1. What is collocation?

Basically, collocation is a type of lexical relation that typically obtains between two lexical items; it is a “… recurrent combination of words that co-occur more often than expected by chance and that correspond to arbitrary word usages" (Carter 1987; Smadja 1993:143; Abu-Ssaydeh forthcoming a) The words arms akimbo, foot the bill, powerful explosion and heavy rain are examples of collocations. Furthermore, collocation encompasses a fairly vast domain in lexis covering the lexical expanse between (the somewhat fixed) idioms and the free lexical associations, that is, cases where lexical items operate as part of an open system. Within collocation, however, the strength of the bond between a word and its collocate is a matter of degree; at the upper end we find what lexicographers call strong collocations (e.g. strong tea, heavy smoker, stiff breeze) which frequently overlap with the members of the idiom category. At the lower end, collocations gradually fade away into combinations that are characterized by total unpredictability where even oxymorons (a combination of two collocationally incompatible words such as thunderous silence, conspicuous by his absence, a definite maybe and original copy) may be encountered.
According to Smadja (1993:147), collocations have four properties: they are arbitrary, domain-dependent, recurrent and form cohesive lexical clusters. They fall into two major categories: lexical collocations and grammatical collocations. A typical example of a lexical collocation is a pair of frequently co-occurring verb + noun combination (foot the bill), an adjective + noun combination (bloodless coup, cushy job, investigative journalism), noun + noun (a piece of advice, chocolate bar, the ghost of a smile), an adverb + verb combination (behave abominably, clearly demonstrate, admit willingly) and adverb + adjective (strikingly beautiful, horribly scarred, politically incorrect). A grammatical collocation according to Hoogland (2004) is a phrase that consists of a dominant word (noun, adjective, verb) and a preposition or grammatical structure such as an infinitive or a clause. Examples would include verb/adjective + preposition (knock at/on the door, aims to, account for, proud of, dependent upon).

Stubbs (1996) lists several applications for collocations: a) natural language generation, b) computational lexicography c) parsing and d) corpus linguistic research (qtd. in Manning & Schutze 1999:142). But the significance of collocation to language pedagogy and translator training far outweighs those applications; hence the abundance of research in this particular area (see for example Farghal & Shakir 1992; Abu-Ssaydeh 1995; Lewis 2000; Zughoul & Abdul-Fattah 2003). This interest arose from an awareness that collocation is central to the process of foreign language learning and translation. Research focussed on the importance of the analysis of this complex lexical phenomenon (Kjellmer 1987; McCarthy 1990), the nature of the problem encountered by learners and how they circumvent it by resorting to other strategies such as transfer, avoidance and paraphrase (Zughoul & Abdul-Fattah 2003) as well as ways of incorporating collocations in pedagogical materials (Bahns 1993; Kasuya 2000; Lewis, 2000; Shei & Pain 2000). These studies were paralleled by the production of a series of monolingual collocational dictionaries which began with the BBI Combinatory Dictionary (1986) and culminated in Oxford Collocations Dictionary for Students of English (2002). However, only two bilingual dictionaries have thus far been produced; Horst Bogatz’ English-German Collocational Dictionary (ARCS) (1997) and A Dictionary of English Collocations (English/Chinese) published in 1990. No other bilingual dictionaries of any significance are available on the market. This lack of interest in the place of collocation in bilingual dictionaries is also mirrored in the scarcity of research carried out in the area of collocation in bilingual lexicography. To redress this imbalance, this paper will, therefore, seek to explore...
ideas that maximally use collocation in bilingual lexicography and to show how a bilingual dictionary that systematically utilises collocation can be far superior to existing dictionaries which tend to use collocation in a random fashion.

2. Collocation and Lexicography

The starting point in the lexicographical study of a word is usually the identification of the lexical environments in which that word may appear in the language. This survey of a lexical item's distribution will ensure that meanings, cases of polysemy (a word with multiple meanings), homography (the same spelling but different meanings), frequent grammatical and lexical collocations and multi-word units (binomials, idioms, speech routines, etc) will all receive their fair share of representation in the dictionary. Collocations, on their part, have been a permanent fixture of both general monolingual and bilingual dictionaries throughout the ages. But the arrival of *Corpus Linguistics* has changed the scene dramatically. With the help of online lexical corpora and concordancers, the lexicographer can effectively rely on the formulae of mutual information and Z-score to determine which lexical substrings he should extract to identify different meanings of a given word, to discriminate the differences in usage, particularly between synonyms and, finally, to identify collocations that are statistically significant enough to merit citation in the dictionary. This has resulted in a more prominent role for collocation in general dictionaries; almost invariably, the more recent the dictionary, the more collocations it will contain. Compare, for example, the number of collocations in Hornby, Gatenby and Wakefield’s *The Advanced Learner’s Dictionary of Current English* (1958) and those cited in *Longman Dictionary of Contemporary English* (1995), which, incidentally, is based on the British National Corpus. Take the word **abandon**, which is the first commonly used word in the dictionary, and you will spot the difference immediately. Hornby et al. cites no collocations with this entry. *Longman* (1995), by comparison, lists **abandon**: - hope (of something); - yourself to; - ship. From the second page of both dictionaries, the word **ability** appears with no collocations in Hornby et al., but with several in *Longman* (1995): **ability**: - to do something; have the - to do something; of great, exceptional etc -; high, low, average -; mixed – class. The verb **accept** appears with no collocations in Hornby et al. but appears with **gift, offer, invitation; plan, suggestion, advice; an idea, statement, explanation; a situation, problem; blame, responsibility; somebody’s apology** in *Longman* (1995). Thousands of other examples can be cited to this effect, but I believe these three are quite indicative of the status collocation enjoys in both dictionaries.
3. Collocation in Arabic-English Dictionaries

While recent English monolingual dictionaries continue to upgrade their methodologies and content, benefiting from the great strides in the field of Corpus Linguistics, Arabic-English dictionaries have lagged behind on almost every front. Generally speaking, the latter category comprises only two dictionaries that are worth consideration: Wehr’s A Dictionary of Modern Written Arabic (1961) and R. Baalbaki’s Al-Mawrid (1988). Wehr’s dates back to an Arabic-German manuscript which was completed in the middle of the last century, and no attempt has been made to update its contents. Baalbaki’s Al-Mawrid (1988), which is less than two decades old, has unfortunately failed to benefit from the recent developments that have impacted English monolingual dictionaries such as the use of frequent collocations, citation of multi-word units, stating differences between spoken and written items and the total reliance on lexical corpora.

As far as collocation is concerned, it can be safely stated that neither dictionary assigns this phenomenon the status it has been given in recent English monolingual dictionaries or that it deserves in this type of dictionaries. While it is true that some words are amply provided with collocations, the majority of lemmas are cited on their own. To illustrate, take the word مَدْ in both dictionaries and compare the collocations given to its senses:

**Wehr (1961):**
- مَدْ: extend, distend, expand, dilate; stretch, stretch out, crane (the neck); draw out, protract; spread out; lay out, lay (tracks, pipeline); spread (a net); lengthen, elongate, prolong; grant a respite or delay; rise (flood, river); help, aid, assist; support (by or with), supply, provide with; reinforce (an army); fertilize, manure (the soil); مَدْ عُمْرَهُ: prolong his life (of God); مَدْ البصر: turn his eyes to, direct his glance to; مَدْ جَذْراً في الأرض: strike roots (tree); مَدْ سَفَعَهُ: prick up one’s ears; مَدْ فِي المَانِثَة: set the table; مَدْ الفوَاسِر: lay pipe; مَدْ إِلَيْهِ يَدَهُ: extend one’s hand to s.o. (896-897).

**Baalbaki (1988):**
- مَدْ: extend, stretch (out), outstretch, spread (out); expand, distend, dilate; extend, lengthen, elongate, prolong, protract, draw out, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to; respite, grant a respite or delay; help, aid, assist, support; supply with, provide with, furnish with, equip with; مَدْ الأُنْسَنُ: protract, stretch out, add to;
Impressive as this list might look, and though both lexicographers apparently recognize the significance of collocation as the examples demonstrate, there seems to be no agreement as to which collocations should be included. Nor is there agreement on whether such collocations should be cited along with the Arabic headword or with its English equivalent (Hoogland 2004). Thirdly, synonymous sets are not differentiated by the citation of an adequate number of collocations. To complicate things, the two lexicographers disagree on which words should be cited as members of the same synset.

Nonetheless, the abundance of collocations this example displays should not blind us to the fact that the majority of headwords in the two dictionaries are given with a minimal number of collocations or with none at all as the following examples show:

ii) Wehr (1961):

قوتي : strong; vigorous; potent; mighty, powerful, forceful; intense, violent, vehement; firm, solid, robust, hardy, sturdy (803).

Baalbaki (1988)

قوتي : strong, powerful, forceful, mighty, vigorous, potent; sturdy, husky, burly, stout, robust, hardy, tough, able-bodied; heavy-duty; intense, violent, vehement, intensive, keen, acute (878).

iii) Wehr (1961):

استقلال : to become dreadful, terrible, momentous, serious, difficult (for s.o.; affair); to get out of control, become excessive, become irreparable (damage) (698).

Baalbaki (1988)

استقلال : to become serious, grave, critical, drastic, severe, intense, extreme, excessive, aggravated, exacerbated; to worsen, become worse; to spread dangerously, reach alarming proportions; to be or become massive, rampant, progressive; to intensify, build up, grow, increase (96).

Even if we agree that English collocations are important for the Arabic-English dictionary user, the number of English collocations cited in the Arabic-English dictionary remains extremely abysmal. In an earlier study (Abu-Ssaydeh 2005), strong collocations in the first ten pages of Longman (1995) were listed and their occurrence in Al-Mawrid was tested; the result showed that out of sixty-three collocations given in Longman, only three are found in Al-Mawrid, that is, less than ten
percent. This statistic demonstrates how alarmingly low the number of collocations is in *Al-Mawrid* despite their significance.

This lack of collocations in the Arabic-English dictionary will lead to many types of errors. Firstly, the translator, assuming that the lexicon of the English language is an open system where combinability is a purely lexical, rather than collocational, matter, may pick any member of a synonymous group to create a collocation. It is as likely as not that the translator will select the wrong word, in which case, for example, he would generate the following English strings:

*deliver* a reproach, do graduate studies, *perpetrate* suicide, commit treachery and make estimations (Benson 1985:64).

These collocations, we are told by Benson, are all erroneous in English. The following, though, are completely natural:

*deliver/administer* a rebuke, commit treason, commit suicide, commit/perpetrate fraud, bake/make a cake, make pancakes, do graduate work.

Such collocations, continues Benson, are “… arbitrary and non-predictable. Non-native speakers cannot cope with them” (*ibid*).

A second type of likely error is that some translators would take the English equivalent to have the same collocational distribution the Arabic word has. For the Arab learner, a man may possibly *hunt* both *a lion* and *a fish*, because both verbs are translatable by بطاطس واصطاد اصطاد. Teachers of English and translation will testify to the abundance of errors of this type in their students’ compositions and translations.

4. Improving Quality of Arabic-English Dictionaries Through Collocation

The occasional use of collocation to discriminate between the different synonymous sets or to enable the language learner or translator to find the equivalent of an Arabic word represents only a fragment of the potential collocation can have in bilingual lexicography. In fact, there are numerous ways in which collocation can be used to improve the quality of the Arabic-English dictionary that we, as lexicographers, can place at the disposal of both the translator and the language learner.

a. Differentiating between Classical Arabic and Modern Standard Arabic (MSA)

One of the major problematic (and often debated) areas of bilingual dictionaries in the Arab World is the fact that, quite frequently, no distinction is made between the various historical periods of the Arabic language (Abu-Ssaydeh 1994). *Al-Mawrid* (1988), for example,
abounds in words that belong to Classical Arabic as do Wehr’s and Hasan Al-Karmi’s Al-Mughni Al-Akbar (ibid.). Examples from the first of these include مهراق، مهفهف، مهفوت، مهراق، مهك all of which are taken from page 1133. This practice should be avoided as it will serve neither variety, and one method of determining if a certain word belongs to MSA or not is to analyze its collocational range in an online lexical corpus. Due to the lack of such a corpus, we shall use the lexical data available to us through the search engine Google (henceforth lexical data). Let us look at the following examples which are taken from Baalbaki (1988):

1. The Arabic verb ﺑَﺪِ in the sense of run wild, be/become wild/untamed as in the collocation ﻣُﻬِّﺎً appears also as a lexical entry in classical dictionaries, but not as a real language combination in the lexical data.

2. ﺛَﻼِثٍ which is a MSA collocation appears twenty-one times compared to ﺑَﺬِرُأٍ ﻣُﻬِّﺎً which appears zero times in the data.

3. The word ﻟَوْدِي golden appears only twice not as an adjective but, interestingly enough, as the name of a racehorse in the UAE. By comparison, the more modern adjective ﺩَهْﻣِي golden appears scores of times including cases where it is used metaphorically as in ﺗَوْردُأٍ ﺩَهْﻣِي.

4. ﺗَأْرِي ﺑَﺬِرُأٍ ﻣُﻬِّﺎً ﻣُﺴَﺖُ which appear many times in the lexical data along with their collocants: ﺑَﺎَّسَمُ ﻣُﺴَﺖُ ﻣُﺘَرِي ﻣُﺴَﺖُ ﻣُﺘَرِي ﻗَﺒْﺮ. ﺑَأَذِرُأٍ ﻣُﻬِّﺎً ﻣُﺴَﺖُ ﻣُﺘَرِي ﻣُﺴَﺖُ ﻣُﺘَرِي 

5. ﺗَأْرِي which in the sense of policeman never occurs in MSA. The same word meaning ﺩَهْﻣِي hazelnuts appears in the data three times. This contrasts sharply with the word ﺑَذِرُأٍ which appears over 14,000 times in the corpus along with collocations such as ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً ﻣُﻬِّﺎً 

6. ﺗَأْرِي ﺑَذِرُأٍ ﻣُﻬِّﺎً ﻣُﺴَﺖُ which appears over 14,000 times in the corpus along with collocations such as ﺑَذِرُأٍ ﻣُﻬِّﺎً ﻣُﺴَﺖُ ﻣُﺘَرِي ﻣُﺴَﺖُ ﻣُﺘَرِي ﻣُﺴَﺖُ ﻣُﺘَرِي 

7. ﺗَأْرِي ﻣُﻬِّﺎً ﻣُﺴَﺖُ which appears over 14,000 times in the corpus along with collocations such as ﺑَذِرُأٍ ﻣُﻬِّﺎً ﻣُﺴَﺖُ ﻣُﺘَرِي ﻣُﺴَﺖُ ﻣُﺘَرِي 

8. The phrasal verb ﺗَأْرِي in the sense of turn away from, dissuade from, discourage from appears twenty-nine times, all of which are either part of religious discourse or are cited in religious texts based on ancient religious or historical texts such as Ibn Ul-Athir and Al-Tabary.

9. ﺗَأْرِي meaning (the skin) to shrivel or wrinkle, be shriveled or wrinkled: no examples whatsoever of the collocation ﺗَأْرِي ﻣُﺴَﺖُ ﻣُﺘَرِي ﻣُﺴَﺖُ ﻣُﺘَرِي 

75
be found in the data, though it is cited in the majority of online Arabic dictionaries.

The fact that none of these lemmas appears in MSA in collocational settings suggests that they belong to another era and, therefore, to another type of dictionary.

b. Determining the collocational range

The use of collocation can and should determine as exhaustively as possible the majority of contexts a certain word may occur in. To illustrate, take the Arabic word استشرى which is found in the following lexical contexts:

The verb استفعل also has the following collocates:

The identification of an exhaustive range of collocations for a lemma is essential for two reasons:

i) The list will determine the Arabic collocates of the word, thus enabling the lexicographer to pinpoint the different contexts (e.g. strong collocations, idioms, spoken expressions, phrasal verbs) in which a given lexical entry is used. As stated above, this type of information is crucial in the process of dictionary compilation.

ii) It is only through the establishment of such a list that a full representation of the various senses of the word can be provided. This is particularly important since the collocational ranges of Arabic and English lemmas are bound to differ. If such a list is not established, it is likely to mislead both the lexicographer and the language student into mapping the collocates of the Arabic word onto those of its English equivalent or vice versa. Moreover, the identification of such a range will help the lexicographer to suggest phrasal equivalents in cases where single-word equivalents may be untenable. For example, Al-Mawrid (1988) suggests reach alarming proportions, spread dangerously and become aggravated/ critical/ drastic/ exacerbated/ excessive/ extreme/ grave/ intense/ serious/
severe/ intense as equivalents for the Arabic word استَفْحَلَ on the basis of collocation.

c. Determining statistically significant collocations
A systematic analysis of the collocational distribution of a given word will show which collocates are particularly frequent and are, therefore, worth citation in the dictionary either as examples of collocates or as semi-unique collocations that merit listing on their own in the dictionary. Take, again, the word استَفْحَلَ and see how frequently collocates occur with it in the corpus (out of about 150 hits):

- أمره: 48
- المرض: 12
- الفساد: 8
- الخلاف: 7
- الخطر, الداء,��ل, وضُعَ: 5
- اليأس: 3
- الاستبداد, الصراع, السحّاق, العذّاب, العلم: 2

The results demonstrate that the collocation استَفْحَلَ أمره does indeed merit a separate listing. The collocates استَفْحَلَ المرض, استَفْحَلَ الفساد, استَفْحَلَ الخلاف, استَفْحَلَ الخطر, الداء,��ل, وضُعَ would prove handy if the dictionary compiler wants to identify the statistically significant collocates found in the vicinity of the verb استَفْحَلَ, a practice that is competently illustrated by the methodology used in Longman (1995). Imagine how improved the lemmas in 3.iii in both Wehr’s and Baalbai’s above would be if such collocations were cited.

d. Verifying equivalence in current dictionaries
With the collocational contexts exhaustively identified, collocations can then be used to verify the accuracy of the English equivalents in the Arabic-English dictionary. Take the Arabic lexical lemma أدَأَرَ: شغل, أعمل (الآلة, محرّك, الخ) in the sense of start up (an engine, a car, a motorbike). Longman Dictionary of Contemporary English (1995) defines start up as follows:

Start

10.► CAR ◄ also start up: if you start a car or engine or if it starts, it begins to work: The car wouldn’t start this morning.

But Al-Mawrid (1988) also lists the verbs operate, move, set in operation or motion, run and work as synonyms to the verb start (up) in the list of senses given to the Arabic entry. However, if we take one of these verbs, namely operate, we will find that the Longman Dictionary
of Contemporary English (1995) gives it the following senses and collocations:

1 ► MACHINE a) to use and control a machine or equipment b) if a machine operates in a particular way, it works in that way. [The closest Arabic equivalent here would be (شغل)].

Unfortunately, neither the sense nor the collocational distribution of the English verb operate justifies its use as an equivalent to the Arabic verb أدار; the British National Corpus does not cite even a single example where the noun car and the verb operate occur as a collocation. The same principle, in fact applies to the other verbs listed with operate above. Therefore, it is perhaps in lexical contexts other than car that the verb operate may be a valid English equivalent.

Take another example. Al-Mawrid (1988) cites the headwords شفѧ and استكصبѧ which share the following two sets of synonyms: seek to discover, try to find out and explore, scout, reconnoiter. But a careful examination of the meaning of the verb discover, its collocations and the collocations of the verb استكصب would show that discover is a very good candidate for equivalence:

explores: 1. To travel through or examine an area in order to find out what is there or what it is like. 2. To examine or discuss something carefully in order to find more about it.

explore: an area, a castle, a cave, experiences, the facts, a gorge, new horizons, issues, limitations, a model, the mysteries, the past, a problem, a property, relations, roots, other solutions, a town


Reconnoiter, on the other hand is best translated in Arabic neither by استكصب nor استكصب as its English meaning indicates:

Reconnoiter: to try to find out the position and size of your enemy's army, for example by flying planes over land where their soldiers are. (ibid.:1183)

On this basis, the closest equivalent in Arabic would be: استطاع ( مواقع, دفاعات, تخصصات العدو). Hence, it should be removed from the list of equivalents and placed in its proper place in the dictionary. Let us turn to the verb استجلاء/now. In Arabic, this verb may be found in the following collocational frames:

استجلاء: - الأبعاد, الجوانب, الأهداف, الدروس, الرواي, الخبير, الفضائي, مواطن العظمة, الغواشم, الحقائق, الموقف, الوضع, الأمر.
Judging by these collocations, the verb استكشَفَ may be translatable by explore because the two verbs do share some collocational frames such as explore facts, issues, the mysteries, a problem, etc.

The verb scout, however, is a different kettle of fish; according to Longman (1995:1276), this verb means to look around in order to find something, check out a place, mostly for military purposes or to find about potentially good players a university or club can recruit. Collocationally, it occurs in the following contexts:

- scout around the shed/the room
- scout any suggestion
- scout around auction rooms
- a patrol scouts south of the border
- scout opinions
- we scouted till daybreak but found nothing
- scout ahead to look for a possible camp site
- his eyes scouted the room
- scout the cottage
- he scouted the surrounding faces
- he scouted for werewolf (ibid.).

This is on the one hand. On the other, Baalbaki’s Al-Mawrid (1999:821) lists the following Arabic equivalents for the verb scout:

- scout:
  1) يمهمدٍ (١) يَبَحثٍ (٢) يَكْتَشِفٍ، (٣) يَبَحثٍ (٤) يَكْتَشِفٍ (٥) يَقْمَ بِنِشَاطٍ (٦) يَبَحثٍ بِنِشَاطٍ (٧). بِنِشَاطٍ 

Accordingly, it would be acceptable to cite scout as an equivalent for استكشَفَ as both share some collocations, but using it as one of the senses of the Arabic verb استجِلَ would be ruled out on such grounds.

In many cases, the English equivalents given to an Arabic lexical entry are phrasal due to the lack of single-word equivalents. Using the verb استخفَ for illustration, we will find that the following senses have been given as English equivalents in Al-Mawrid (1988): become grave; become serious; become severe; become intense.

Let us reconsider the first of these. Longman Dictionary of Contemporary English (1995:622) gives the following two senses for the adjective grave:

1. very serious and worrying: this decision may have very grave consequences. | I have grave doubts about his ability. 2. Looking or sounding quiet and serious, especially because something important or worrying has happened: Turnbull’s face was grave as he told them about the accident.

It may be said that the first of the two senses cited here is close enough to the meaning of the Arabic verb استخفَ. This claim, however, will be acceptable only if it is validated by lexical data. In other words, we have to demonstrate that the combination be/become grave does have a
collocational set that corresponds fully or significantly to that found with the Arabic verb استطلغ. The British National Corpus cites 2071 examples of the word grave, including its use as a noun (in the sense of tomb). The fifty examples we have examined yielded the following collocations quoted as sentence fragments to save space:

i) grave and widespread violations
ii) a grave lack of frantic drivenness
iii) expressed grave and prophetic doubts about
iv) there existed grave doubts about
v) watched with grave concern
vi) so grave was the extent of his injury
vii) another grave discrepancy between substance and show
viii) and a grave hazard to
ix) brought grave problems
x) matter may turn out to be very grave
xi) solve this grave difficulty
xii) their lives were in" grave danger"
xiii) This meeting views with grave concern

Now, the Arabic equivalents one may suggest to the different contexts in which the adjective grave is being used may include 

اﺑﺻﺎﺑة, انﺗﻬاﮐ, نﻘﺺ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْتِء, ﺗَﻨَﺎﻗِﺾ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻞ, ﺗَﻨَﺎﻗِﺾ, ﻣُﺳْﺘَفَﺣِﻠ
e. Overtranslation and Undertranslation
Close to the notion we have thus far been discussing is the case where the English equivalent cited against the Arabic lexical entry is either too general or too specific as a result of the lack of full correspondence between the underlying semantic representations of the two terms in question. The use of grow and increase as meanings of استفحلاً, for instance, is an example of undertranslation; they both are general equivalents that would collocate with a wide range of nouns, both positive and negative:

grow: bitterness, a campaign, exports, foreign debt, output, revenues, resentment-
s
increase: amount, consumption, net income, manufactured goods, a number, poverty, rates, revenues, risk, shipments, traffic -

The verb استفحلاً, by comparison, collocates exclusively with undesirable things (check the collocates above) or things that are “perceived” to be undesirable by the speaker. Take the Arabic words علمانية (secularism) and استيطان (settlement). In the first case, it is the use of the verb استفحلاً that indeed expresses distinctly the political convictions of the speaker; evidently, secularism is viewed as an evil thing that his community can do without. The same political orientation underpins the use of the word استيطان (settlement) which according to some is a religious duty and a right regained but to others a usurpation of land and an act of grave injustice. The English word “colonization” is defined in Longman (1995) in terms so neutral that they would anger any self-respecting citizen of the Third World countries that had to experience the ferocity, inhumanity and degradation that accompanied colonization in the recent history. This is no more evident than it is in the collocational contexts in which this word appears in Arabic: نير الاستعمار (the yoke of colonialism), التحرر من الاستعمار (freedom from colonialism), مقاومة الاستعمار (resistance against colonialism), الثورة ضد الاستعمار (the revolution against colonialism), الكفاح ضد الاستعمار (struggle against colonialism), ارث التماثل (inheritance of colonialism), and أناناس الاستعمار (servants of colonialism). It is obvious that neither increase nor grow would sufficiently convey the depth of resentment embodied in the Arabic verb استفحلاً when it collocates with such negative words. The English equivalents given for the word باند (extinct, dead; past, bygone) are so diluted that they fail to capture the resentment and hatred implicit in the word when it is used in the frames of a past regime, cultural heritage, thought, colonialism, reign: النظام الطاغوت/العنصرية الباند, موروث متألف باند, فكر باند, الاستعمار الباند, العهد الباند. It is the collocation here that can save the day and demand that such equivalents be ignored or modified in order to cater for the various semantic components and shades of meaning the Arabic word may have.
f. Metaphors

Despite the immense wealth of human language, it is perhaps man’s desire to be creative and to express himself in a picturesque way or by approximation that underlie the predominance of metaphors in the human language. Perhaps it is in the nature of language as a living thing that the process of life and death continues as a cycle. Be that as it may, metaphors and collocations are germane to bilingual lexicography in many ways. To start with, metaphors are largely language-specific. This means that if an English word is used metaphorically, there is no guarantee whatsoever that its equivalent in Arabic will necessarily be used metaphorically. The word pairs flood: ﻓﻴﻀﺎن, mountains: ﺟﺒﺎل, seas: ﻭﺮ, oceans: ﺑﺤﺎطات illustrate this point clearly. While these English words are frequently used in a metaphorical way, creating in the process collocational networks for themselves that are quite distinct from those associated with their literal senses, the author can rarely find any cases where the Arabic equivalents are used metaphorically. Secondly, even if the English word and its Arabic equivalent are used metaphorically, again there is no lexical evidence to sustain the claim that the two words will have the same metaphorical meaning or the same collocational network. [For a full discussion of this point, see Abu-Ssaydeh 2003 and Abu-Ssaydeh 2005.] Collocationally speaking, the Arab translator and/or language learner is unlikely to identify the metaphorical potential of the English vocabulary in a systematic manner or to identify their collocational ranges on his own unless he is deliberately made aware of them. More often than not, he is apt to assume that if an Arabic word is used metaphorically, its English counterpart will certainly have the same network of collocations.

This being the case, it is crucial that bilingual lexicographers recognize this fact and express it in their dictionaries explicitly for the benefit of the user. The first aspect to be taken into account is recognition of similar metaphors; for example, the English word intoxicated and its Arabic equivalents ﺎﺳﮏ أﺳ، ﻋﻠ، ﺎﺛﻤ and its Arabic equivalents are used metaphorically and in the same sense, hence the similarity of the collocational networks they demonstrate:

intoxicated: by the enthusiasm, by success, with the knowledge, with pleasure, with power

أﺳﮏ أﺳ: ﺳو، ﺓ، ﺑ، ﺳ، ﻣ، ﺳ، ﺟ، ﺑ - (Abu-Ssaydeh 2005).

The same principle applies to the English word rivers and the Arabic equivalent ﺎﻧﻬﺎر:

rivers: - of blood, refugees, lava, champagne, rock

أﻧﻬﺎر: - ﻣ الندم، ﺳ، ﺓ، ﺓ، ﺓ، ﺑ - (Abu-Ssaydeh 2005).
The words cradle (of civilization), glimmer (of hope), ghost (of a smile) and dawn (of civilization) also belong to this category.

Added to that, Arabic has, over the past few decades, borrowed from English quite a few metaphors along with their collocational ranges. An example of such borrowing is the word wave. According to the Oxford English Dictionary (2001), the literal sense of the word wave was current around the sixteenth century. Only in the second half of the nineteenth century do we begin to have any records of the collocations wave of vehicles, soldiers and aircraft. The collocation crime wave also dates back to the same period. It is the metaphorical potential that generates the following English collocations with the word wave:

wave: ~ of crime, applications, attacks, condemnation, dissent, immigrants, inflation, kidnappings, killings, nationalism, optimism, planes, popularity, strikes, troops, unrest (Abu-Ssaydeh 2005).

Lisanu L-Arab (711 H. corresponding to 1333 A.D) also defines موجة in a way identical to that given to its literal sense in the Oxford English Dictionary and in Longman (1995). On the other hand, Al-Ghani (an Arabic-Arabic dictionary compiled in the second half of the last century) goes beyond this to include the collocations موجة الانفعال (in transmission), موجة قصيرة (a short wave), موجة الغلاء (a wave of high prices), موجة البرد (a wave of cold) as illustrations of how the word is actually used in Arabic. Such an inclusion in Al-Ghani reflects the awareness of the lexicographer of the interaction taking place between the two languages. Words of a similar nature include puppet (government), piracy (of software), school (of thought), enthusiasm evaporate (الحماس تبخر), حمى (الانفجارات) and most probably horizon (فهد) and picture صورة (مثلا ووضعه في الصورة). [For a fuller treatment of this subject, see Abu-Ssaydeh 2005.]

The majority of metaphors are, for the most part, language-specific and exhibit, cross-linguistically, complex patterns of collocational distribution. The English words dynamite, pig, leg (of a journey, a tour, a piano), mouth (of a bottle, a river, valley, a cave) and foot (of the mountain) are indeed very different from their Arabic counterparts both in meaning and in the collocational frames they occur in, though the collocates of both leg and ساق may actually overlap in certain cases (leg of a piano, table ساق البيانو/الطويلة) as may those of mouth and قم الفتقية (mouth of a river, قم النهر, mouth of a bottle). The English word torrent and its Arabic equivalent سبيل also represent an interesting, yet complicated,
case for the lexicographer. The following represents a sample of the lexical environments in which they occur (Abu-Ssaydeh 2005):

torrent: - of abuse, advertisement, condemnation, complaints, destructive attention, defiance, documentation, emotions, interviews, new laws, new products, passion, publicity, research, tears, weeping, words

These patterns suggest that when dealing with the English word, the Arabic equivalent سيل will definitely not be adequate; we have to use other equivalents including قيض, مقدار كبير, أعداد هائلة, كم هائل, هائل, لإتهام، etc. This fact has to be reflected in the Arabic-English dictionary; when looking for English equivalents for such Arabic words and expressions, the English word torrent will stand out as the best candidates. Conversely, when the lexicographer seeks equivalence in English for the Arabic word سيل, he will have to look beyond the word torrent and consider accurate translations of the combinations appearing above. Thus, in addition to the list of equivalents given in, say, Al-Mawrid (1988) which include torrent, flood, deluge, shower, hail, stream, fusillade and volley, an adequate Arabic-English dictionary should also cite rivers (of refugees), avalanche (of applications, claims, complaints, documents, orders, words, etc), bombardment (of images, icons), mass (of allegations), onslaught (of attacks), plethora (of books, cases, issues, letters, schemes) and rain (of missiles).

Another example; the word إنهمر appears in Al-Mawrid (1988) along with the equivalents be poured out, be shed; to pour down, fall heavily, flow in torrents; to rain heavily. If we examine this list, we will find that it is unable to accommodate the fact that this Arabic verb may appear with the collocates قاذف فائض and قاذف فائض. Moreover, when it collocates with دموع دموع, it may also have other possible equivalents in English. These facts are expressed in the following entry:

انهمر: (bombs, shells) rain down (on); (tears) run/ stream down (his cheeks; burst into tears.

These, and many other, examples, demonstrate that the metaphorical potential of Arabic and English can, from a lexicographer’s point of view, be best served if collocational considerations are taken into account, for unless this is done in a methodical and systematic manner,
the dictionary will always fail to capture this rudimentary fact about lexis.

5. A Sample Dictionary

The acquisition of metaphors and their collocational networks can be effected either by formal instruction or by means of specialized or carefully designed dictionaries. So far, the only lexicographic work to cater for this need is the author’s *Dictionary of Semantic Extensions, Similes and Conceptual Metaphors (English-Arabic)*. But metaphors remain a fairly small fragment of the larger picture. Unfortunately, very little research has been done on the implications of collocation to lexicography in the bilingual context. Arab lexicographers have also shied away from producing bilingual dictionaries, perhaps because of the complexity and fluidity of the phenomenon. The only references in the literature to work in progress include scattered and mostly undocumented references to Hassan Ghazala, Mohammad Hilmy Heleyil and this author whose bilingual collocational dictionary has been in the making for the past twenty-two years.

Like metaphors, collocations are language-specific. In English, people *smoke a cigarette*, while in (colloquial) Arabic, a smoker *drinks his cigarette* as well as a *tablet of medicine*. You *eat your soup* if you are English but *drink it* if you are an Arab. In (colloquial) Arabic, you *open the radio, the TV, the switch, investigation, the door, a subject, a tap, an account, the curtains, a road, fire, fortune, a country, the appetite, a department*, etc. The fact that this verb is translatable by several verbs in English (such as *switch on, open, conquer, start, draw, turn on*, etc) is the bane of language and translation teachers who encounter, almost daily, errors resulting from the failure of the Arab student to distinguish between the different senses of this verb. To make things worse, bilingual dictionaries rarely provide adequate information that distinguishes between the different synonyms or even the different senses they may list against a given word. To illustrate, let us take, yet again, the word 

\[
\text{استفحل} (\text{Al-Mawrid} 1988:96) \text{ lists the following senses against it:}
\]

\[
\text{استفحل: be /become serious/ grave/ critical/ drastic/ severe/ intense/ extreme/ excessive/ aggravated/ exacerbated; to worsen, become worse; to spread dangerously, reach alarming proportions; to be/ become massive/ rampant, progressive; to intensify, build up, grow, increase}
\]

Let us view these equivalents from the language learner’s and translator’s point of view: are they differentiated in a manner that would permit the target user to use them accurately? Most language students and translators would probably answer in the negative. Secondly, how do we match these senses with the Arabic nouns that are collocates of
the verb المُستَثْمَر. Which of these synonyms, if any, would overlap in their distribution with the different nouns? The answers to these questions cannot be found in traditional general bilingual dictionaries. Only three sources can help here; online lexical corpora, a general bilingual dictionary that provides adequate collocations and bilingual collocational dictionaries. The use of online corpora is an awkward and time-consuming task that needs constant access to the Internet. Practically, it would serve the needs of lexicographers and language researchers, but not translators and language students. The latter may occasionally access the corpora perhaps for verification purposes or for finding a specific collocation, but the corpora will not be a viable alternative to the dictionary. On the other hand, a general Arabic-English dictionary that places more emphasis on collocation can go a long way in meeting the needs of the learner and translator. Compare the quality of the information available below to the list of single items given above by Al-Mawrid (1988):

(ب) المرض، النزاع، الفقر، الأزمّة مثال: (ئل) (a conflict) deepen, grow مثال: (ام) (a conflict) grow، (د) (corruption) be انتشرى مثال: (ص) (a problem) get worse. The answers to the questions can help here; online lexical corpora, a general bilingual dictionary that provides adequate collocations and bilingual collocational dictionaries. The use of online corpora is an awkward and time-consuming task that needs constant access to the Internet. Practically, it would serve the needs of lexicographers and language researchers, but not translators and language students. The latter may occasionally access the corpora perhaps for verification purposes or for finding a specific collocation, but the corpora will not be a viable alternative to the dictionary. On the other hand, a general Arabic-English dictionary that places more emphasis on collocation can go a long way in meeting the needs of the learner and translator. Compare the quality of the information available below to the list of single items given above by Al-Mawrid (1988):

Take another example, the verb المُستَثْمَر: مثال: (ع) (a conflict, a dispute, struggle) escalate مثال: (ج) (corruption, a disease, piracy, prostitution) become rampant مثال: (د) (corruption, terrorism be) rife مثال: (ط) (a person) become enormously/ highly influential (Abu-Ssaydeh forcoming b).

86
(corruption) (ز) (race, terrorism be/ become) rampant

(Abu-Ssaydeh forcoming b) (a problem) be aggravated

Another complementary approach to the question of collocation is to list the verbs along with their nouns under the noun category. For example, the fact that the word crisis collocates with several of these verbs can be indicated by the following entry:

- a - be/ becomes aggravated/

The noun corruption, on the other hand, may appear with the following collocations:

Yet a more thorough solution may be proposed, one that can actually help the language student and translator produce accurate, idiomatic English combinations; a fully collocational bilingual dictionary that lists words along with their appropriate collocations. Take the Arabic word مَوْجَة; instead of listing just the verbs cited above, the word can appear as a lexical entry in the following manner:

- مَوْجَةَ (اسم): (لا) - زّرَكَبُ الْمَوْجَةَ - صَارِعُ الْمَوْجَةَ

(Abu-Ssaydeh forcoming b).
6. Conclusion

The compilation of dictionaries must be informed by the needs of the target user. In the case of the Arabic-English dictionary, such needs will be deemed to have been met if the information available to the user is adequate enough to enable him to use the English language accurately through the medium of Arabic. In this paper, we have sought to show that if the lexical phenomenon of collocation is used optimally, it can certainly help serve this purpose. Collocation, we have shown, can be used to separate words that belong to Classical Arabic from those whose place is Modern Standard Arabic, because the confusion of the two varieties will not serve either of them. It will identify the common lexical contexts in which a certain word may appear, thus enabling the lexicographer to separate weak collocations from strong ones, select the strong ones, cite them and identify contexts in which the lexical item is part of a multi-word unit. English equivalents can be selected on the basis of collocations, a step that will serve to eliminate the confusion arising from the general practice of lumping synonyms and near-synonyms together. Once a sense is determined to be collocationally unsubstantiated, it can be eliminated from the dictionary and replaced by another equivalent whose status is validated by data-backed collocational distribution. An Arabic-English dictionary where collocations are cited as part of the methodology will be far superior to any dictionary currently in use in the Arab World. But perhaps the most significant contribution lexicographers can make to Arabic-English dictionaries is the compilation of a collocational dictionary of the type explained above.

References


