VARIATION IN MULTI-WORD UNITS: THE ABSENT DIMENSION

ABDUL-FATTAH ABU-SSAYDEH

University of Sharjah, United Arab Emirates

ABSTRACT

Despite the frequent claims most linguists make about the fixedness of *multi-word units* (henceforth MWU's) and the almost uniform representation they receive in dictionaries, the range of variation across categories as well as within the members of the same category is much greater than has hitherto been recognized. The complexity of variation patterns represents a serious challenge to the translator whose mother tongue is not English. This paper will, therefore, attempt to provide a thorough examination of the types of variation in MWU's in general and the potential variations that typically characterize each category. Explanations different linguists have suggested to account for variation are also reviewed and their value to translators and possible contribution to lexicography are assessed. The paper concludes by asserting that explanations are inadequate for purposes of MWU's acquisition and that modifications should be made to current lexicographical methodology in order to help translators identify the changes a given MWU may undergo.

1. Introduction: What are MWU's?

1.1. Definition and significance

MWU's are "... lexical phenomena ... which are conventionalized form/function composites that occur more frequently and have more idiomatically determined meaning than the language that is put together each time" (Nattinger and DeCarrico 1992: 1). According to Moon (1998), the most salient features of these units are:

i)   institutionalization/conventionalization,

ii) lexicogrammatical fixedness, and

iii) semantic non-compositionality. MWU's have been studied under a plethora of designations: "lexical phrases, multi-word units, fixed phrases, formulaic

phrases, chunks, preassembled chunks, prefabricated units, holophrases, and so on" (Willis 1997). They straddle both the lexical level and the syntactic level, ranging from a single phrase (*pipe dream, green thumb*) to compound sentences (*look after the pennies and the pounds will look after themselves*); from binomial fixed phrases (*beck and call; knife and fork, pepper and salt*) to "slot-and-filler frames" (as *... -er, ...-er,* e.g. *the more, the merrier*), even proverbs (*there is no smoke without fire*) (Lavelle 2003). Furthermore, they interact with textuality and serve a multitude of pragmatic and social functions (Abu-Ssaydeh, forthcoming). MWU's are crucial for foreign language learning and communicative competence and they represent probably close to half the lexis of the English language (Fillmore 1979; Jackendoff 1997; Widdowson 1989; Fellbaum 1998; Sag *et al.* 2003; Lewis 1993).

1.2. Major categories

For the purposes of this study, we shall classify MWU's into the following six major categories:

a) Fixed phrases (Lewis's 1997 *polywords*)
    Sag *et al.* (2003: 4) describes fixed phrases as "... fully lexicalized and undergo neither morphosyntactic variation ... nor internal modification". They are pre-assembled, extremely sta ble language chunks that cover a fairly heterogeneous group of MWU's including binomials which can be defined as "two or more words or phrases belonging to the same grammatical category, having some semantic relationship and joined by some syntactic device such as 'and' or 'or'", such as *ladies and gentlemen, spick and span, day and night, pure and simple, here and now* (Bhatia 1994: 143). They also include conventionalized discourse formulae (*on the one hand, on the other hand, last but not least*) and Latin and Greek borrowings such as *ad hoc, ad infinitum, carpe diem* etc. (Lennon 1998).

b) Institutionalized utterances
    Unlike Lewis (1997), we restrict this designation to complete sentences or fragments thereof which have been lexicalized and serve as conversational routines or social formulae such as greetings, ending a telephone conversation, saying good-bye, etc: *nice to meet you, so long, have a nice weekend, take care now, come off it, (well) what do you know.*

iii) Lexicalized sentence stems (Lewis's 1997 "sentence frames and heads")
    A lexicalized sentence stem is "... a unit of clause length or longer whose grammatical form and lexical content is wholly or largely fixed; its fixed elements form a standard label for a culturally recognized concept, a term in

the language" (Pawley and Syder 1983: 192). These include sentence heads (*if I were you, would you mind if, that's ... for you*), sentence tails (*as it were, and what have you, and so on*) and sentence slots (*... -er, ... -er*) (Lindstromberg 2003).

d) Idioms
    There is a great deal of disagreement in the literature on what constitutes an idiom; some would include similes and proverbs, others might list single words (*blarney, ergo*) or even acronyms (*WASP, VIP, UFO*) and Latin phrases borrowed into English (e.g. *magnum opus, de facto*) (Hirsch, Kett and Trefil 2002). *Oxford Idioms Dictionary for Learners of English* (henceforth *OID*) (2001) considers as idioms almost all the categories listed here as MWU's. Despite this apparent confusion, there is general consensus amongst linguists that idioms are semantically opaque and syntactically fixed (or frozen or fossilized) MWU's (see for example Baker 1992; see also Schmitt 2000 and Moon 1998). Examples are *light at the end of the tunnel, ball and chain, hold your horses* and *clear the decks.*

To these, we would like to add two further categories that have been overlooked by most researchers though they tend to be largely institutionalized and lexicalized entities (Lennon 1998). These are:

e) Similes
    A simile is basically a comparison between two things where similarity, whether real or not, is perceived to exist (Abu-Ssaydeh 2003). As such, similes may either be lexicalized (*as drunk as a skunk, sly as a fox and easy as pie*) or created by the language user on the basis of actual similarity or culturally-conditioned perceptions (*behave/sweat/be reared/live like a pig, dead as a doornail* and *work like a Trojan/hell/an automaton/a madman/a beaver/a slave*) (Abu-Ssaydeh 2003).

f) Proverbs
    These are usually sentence-long encapsulations of popular wisdom in a given culture that are passed down from one generation to the next (Mieder 1985): *many hands make light work, a bird in the hand is worth two in the bush, make hay while the sun shines.*

2. Variation in MWU's: Types of variation

A classification of the different types of variation MWU's in general may undergo is, in my opinion, essential for lexicographers and would greatly benefit translators. For one thing, such a classification will correlate the MWU category and the type(s) of variation it is likely to permit and will pinpoint similarities

and differences that may exist amongst different categories. Secondly, it will separate types of variation that are recognized by current lexicographers from those that have thus far been bypassed in the general and specialized dictionaries. Thirdly, it will set apart variation types that are lexicographically manageable (e.g. where variation is limited to one or two alternative words) from those that are too awkward to be dealt with in lexicography (such as cases where variation may be unpredictable). Fourthly, areas in MWU's where native speaker's creativity may come into play can be monitored and recorded if frequency warrants it. Finally, it could provide a theoretical grounding that may, in part, guide lexicographers to fine-tune the presentation of variational aspects in MWU entries and to propose some practicable solutions for dealing with the more problematic units and, in the process, help translators perform their job more efficiently.

Despite the claims made by most linguists regarding the notion of fixedness in MWU's and the non-committal qualifications the notion may be associated with (e.g. *often, frequently, semi-fixed,* etc.), such units in fact display amazingly complex patterns of variation that affect entire categories as well as members within the same category. These patterns can be summed up as follows:

2.1. Nil variation, i.e. the total absence of lexical or syntactic variation, typically in fixed phrases (see 1.2.(a) above) but also in many other categories:

1) *bread and butter* (fixed phrase)
2) *on the one hand* (fixed phrase)
3) *what a pity* (institutionalized utterance)
4) *no kidding* (institutionalized utterance)
5) *easy does it* (idiom)
6) *as the crow flies* (idiom)
7) *as easy as pie* (simile)
8) *like a red rag to a bull* (simile)
9) *no news is good news* (proverb)

2.2. Slot-fillers

This type of variation is unique to sentence frames, but it does not affect the lexical constituents or the grammatical patterns of the frames themselves:

10) *what if...*
11) *how come...*
12) *what's the matter with...*
13) *how about...*

2.3. Substitution by synonymous or near-synonymous items from a restricted set of expressions which native speakers regularly use in the production of MWU's and which lexicographers recognize as legitimate alternative realizations of the MWU in question:

14) *for the moment/present/time-being*
15) *leave a bad/bitter/disagreeable/nasty/sour taste in the mouth*
16) *seize/catch/clutch/grab/grasp at straws*
17) *sober/grave as a judge*
18) *haul/rake/drag over the coals*

2.4. Lexical substitution from a restricted set of non-synonymous but meaning-preserving lexical items that may occur at a certain syntactic point in the structure of the MWU:

19) *drive/back/box/push/paint in a corner*
20) *work like a beaver/dog/horse/slave*
21) *as quick as the blink of an eye/a flash/lightning/a wink*
22) *up to snuff/par/scratch/the mark*
23) *run like hell/mad/the wind/a hare*

It should be noted that variation in (c) and (d) above can be initial (*come/roll to a stop*), medial (*live to fight/see another day*) or final (*close call/shave*).

2.5. Optional elements, where one element or more of an MWU may be suspended

This pattern does not apply to fixed phrases, lexicalized sentence stems or similes, but it is observed in spoken expressions, idioms and proverbs:

24) *I hate to (have to) say this but* (spoken expression)
25) *(have) no fear* (spoken expression)
26) *put (an optimistic/positive) gloss on it* (idiom)
27) *stand/go toe to toe (and swap blows)* (idiom)
28) *(every man for himself and) the devil take the hindmost* (proverb)
29) *it is an ill wind (that blows nobody any good)* (proverb)

As the following example demonstrates, optionality may operate at more than one position in a given MWU:

30) *up a/the (shit) creek/river (with no paddles/without a paddle)* (idiom)

Similar in nature are cases where part of an MWU may occur independently in the native speaker's language. The MWU *shed crocodile tears* might be reduced to *crocodile tears*. *Go/Rise from rags to riches* occurs in actual usage as *rags to*

riches (saga/story/tale). Hold out/offer/extend the olive branch appears in language as *the olive branch* (e.g. *accept, pick up the olive branch*). This type of variation is particularly significant to non-native speakers and, consequently, to lexicography as I will demonstrate later.

2.6. Alternative syntactic/lexical patterns for some or most of the MWU

In this pattern, the syntax or the lexical constituency of the MWU may vary over a large part of the unit, thus necessitating the listing of the two forms in order to capture actual language use:

31) *cut his cloth according to his pocket/cut his coat according to his cloth*
(idiom)
32) *have your heart's desire/have everything your heart could desire* (idom)
33) *be in the firing line/be in the line of fire* (idiom)
34) *I (can) tell you/I am telling you* (spoken expression)

2.7. Regional variation: Primarily British and American

a) The same MWU may have different meanings in different regions that speak English like the UK and the USA:

35) *give it a lick and a promise* (informal): (1) Br Eng: 'wash or clean sth quickly and carelessly'; (2) Am Eng: 'to do a job quickly and carelessly' (*Longman dictionary of contemporary English* 1995, henceforth *LDOCE*).

b) Different MWU's may have the same sense:

36) *cut in line* (Am)
37) *jump the queue* (Br)

The British idioms *rain cats and dogs, it is pouring, rain hard, be pissing down, raining stair-rods* are expressed in a variety of ways in American English which are unknown in British English:

38) *coming down like a cow pissing on a flat rock*
39) *hit a frog strangler*
40) *it's a real toad choker*
41) *it's raining like piss out of a gum boot*
42) *it's raining pitchforks and plowhandles*
43) *it's raining to beat sixty*

c) An MWU may display deletion or lexical variation in its constituents (substitution):

American	British
44) *a new lease on life*	*a new lease of life*
45) *in light of*	*in the light of*
46) *take the cake*	*in the light of*
47) *shoot the moon*	*throw a moon*
48) *like a cat on a hot tin roof*	*like a cat on hot bricks*
49) *not a happy camper*	*not a happy bunny*

2.8. Complex patterns of syntactic and lexical variation

Unlike other MWU categories, idioms are the translator's Achilles' heel. For one thing, the range of lexical and syntactic variation that is witnessed in idioms is confusingly and dauntingly vast. Secondly, variation in idioms is, for the most part, unpredictable. But before we discuss variation in this category, it should be noted that no idioms permit certain transformations such as topicalization (*the bucket he kicked last week*), conjunction reduction or pronominalization (*he kicked the bucket and she will kick it too*). Moreover, idioms consisting of Adjective + Noun combinations (*easy prey, a breathing space, a hot potato*) do not allow predicative usages (*the prey is easy; *the space is breathing; * the potato is hot*), nominalizations (*the easiness of the prey; *the breathing of the space; *the hotness of the potato*), the formation of comparatives or superlatives, or modification (*an easier prey; *a hotter potato; *the hottest potato*). Beyond these restrictions, however, the majority of idioms undergo a great deal of change which is effected through the manipulation of certain grammatical categories such as tense, number and aspect, the adjectival inflection; by internal modification (the addition of adjectives or adverbs) or by the application of certain transformations like passivization, nominalization and negation. In this type of variation, the extent of the MWU's flexibility is so idiosyncratic that, to the translator, it would seem to be idiom-specific. To illustrate, let us examine the grammatical changes that the following units may undergo:

50) *make up his mind*: does not undergo passivization though it allows past (*made up his mind*), past/present perfect forms (*had/has made up his mind*), and the addition of *-ing* (*while he is making up his mind*);
51) *spill the beans*: allows passivization (*the beans are/were spilled*), *-ing* addition (*spilling the beans*), tense/aspect change (*spilled/will spill the beans*), change of mood (*spill the beans*) and negation (*didn't spill the beans*);
52) *kick the bucket*: permits inflection for the simple past only (*kicked the bucket*);
53) *drop a hint*: can be in the progressive aspect (*dropping a hint*);

54) *bury the hatchet*: allows all tenses (*buried, buries/is burying/will bury the hatchet*);

55) *keep tabs on*: allows the -ing form (*keeping tabs on*), the simple past (*kept tabs on*), combines with the modals would and must (*must/would keep tabs on*) and permits the insertion of the adjective *close* (*keeping close tabs on*).

The picture can be even more complex, combining lexical modification and grammatical manipulation: in LDOCE the idiom *send shivers (up and) down your spine* is listed as fixed, but the wealth of variations that this particular idiom displays is fascinatingly complex:

56) *the two should send shivers of disgust up your spine*
57) *sending shivers of danger down her spine*
58) *sending shivers of trepidation down her spine*
59) *sent shivers of awareness down her spine*
60) *send cold shivers down the spine of*
61) *shivers ran up my spine*
62) *sent shivers up viewers' spines*
63) *sending shivers through her body*
64) *sent tiny shivers all over the heated surface of her body*
65) *sent shivers rippling down her spine*
66) *sent cold shivers rippling down her spine*
67) *sending strange shivers down her spine*
68) *sent unexpected shivers down her spine*
69) *sent shivers running through her*
70) *sent shivers along her spine*
71) *sent shivers quivering frantically up and down her spine*

(BNC).

Cases of nominalization (i.e. a process whereby a phrase is changed into a Noun Phrase) and adjective formation have also been observed in idioms:

72) *break the ice = an ice-breaker*
73) *open his eyes = an eye-opener*
74) *pass the buck = buck-passing*
75) *it boggles the mind = a mind-boggling thing*
76) *breaks new ground = ground-breaking*
77) *blaze a trail = trail-blazing*

Examples cited so far may give the impression that it is only idioms that display this feature. This is not the case since unpredictable syntactic and lexical variation is equally evident in proverbs as the following example demonstrates:

78) *Make hay while the sun shines.*

a) *Make hay while the sun fitfully shines.*
b) *I'll make hay with what you've got, every penny of it.*
c) *He might as well make hay while the sun shone, he told himself.*
d) *Rush out and make hay before you reach your sell-by-date?*
e) *They preferred to make hay while the setting Labor sun still shone.*

(BNC).

2.9. There is another class of variation which is difficult to assign to any of the above categories and is basically creative in nature. It reflects the ability of the native speaker to manipulate an MWU in a way that is intended to create special effect or express certain meanings, a feat that lies almost completely outside the capabilities of the translator whose mother tongue is not English. Note, for example, how the following proverb has been modified:

79) *A bird in the hand (is worth two in the bush).*

a) *A bird in the hand is never worth an infinite number of birds in the bush.*
b) *A bird in the hand is worth two votes for Bush.*
c) *A bird in the hand is still worth two in the bush.*
d) *I wanted a bird in the hand, but this one was practically in Shepherd's Bush.*
e) *Applicants simply prefer the bird in the hand to that in the bush.*
f) *It has been well said, seems to assert "that a bird in the hand is worth less than the same bird in the bush."*

2.10. Reversibility which can be defined as the ability of a certain MWU to reverse the syntactic order of its elements

We shall come back to this feature in the next paragraph.

3. Correlation between the MWU category and variation types

3.1. Having reviewed the MWU categories and the types of variation that usually take place in them, can we make any meaningful generalizations or establish any correlation between category type and class of variation? Perhaps the easiest category to isolate here is the fixed phrases. The completely fossilized members in this category include borrowed expressions such as the French phrases *joie de vivre, bon appetit, crème de la crème* and the Latin phrases *ipso facto, magna cum laude, person non grata* and *post mortem*. But not all fixed phrases are that frozen. The binomials are a case in point. A corpus-based study by Bastow (2003) demonstrates that the binomial *men and women* occurs 381 times in the corpus, but *women and men* never does. By comparison, *friends and*

allies shows variation: *friends and allies* occurs 67 times compared to 47 occurrences of *allies and friends. Knife and fork* and *day in day out* are non-reversible, but *day and night* and *pepper and salt* can be reversed quite often. Apart from that, fixed phrases are the most stable in terms of variation since their representation in the dictionary is identical to their manifestation in actual language use. This inflexibility is not peculiar to fixed phrases; some institutionalized utterances (*how do you do*), idioms (*lock, stock and barrel*), certain lexicalized similes (*dry as dust*) and some proverbs (*familiarity breeds contempt*) are just as fixed in their patterns.

3.2. Sentence stems consist of elements that are fairly stable though they permit the slots to be filled by any semantically and syntactically compatible elements. Interestingly, sentence stems share this feature with novel similes (as opposed to lexicalized similes, see 5.4. below). Institutionalized utterances, in turn, may allow lexical or even pattern variation on a larger scale, but the number of permissible combinations is limited and can be easily isolated and listed from any good lexical data. Again, this feature is shared also by lexicalized sentence stems and idioms.

3.3. Proverbs typically permit nil variation (*look before you leap, once bitten twice shy*) or allow certain elements to be optional, especially from the end of the proverb pattern (*don't count your chicken*), with occasional lexical and syntactic variation. On the face of it, it seems that the amount of variation noticed amongst proverbs is far less than that recorded amongst idioms. Such a statement, however, remains largely impressionistic in the absence of any documented statistical data for variation in both categories.

3.4. So far as similes are concerned, there are three types of variation that are permitted:

a) Nil variation, as in the case of some lexicalized similes (*as cute as a button, as clean as a whistle, stubborn as a mule, fit as a fiddle*);

b) variation by substitution with non-synonyms (*drunk as a judge/a lord/a skunk; light as air/a feather*);

c) some lexicalized similes occasionally substitute their stable lexical constituents with others that native speakers of English may produce on the spot. Look at the following examples from the *BNC*:

80) *drunk as a lord/a skunk: drunk as a newt/a fiddler/March hares/a vicar/a pig/bishop/any sot on May day*
81) *old as the hills: old as the hills/Noah/time/history*
82) *thin as a rake: thin as a rake/rail/a willow wand/a ruler/a bean-pole/a stick*

Unlike other categories of MWU's, similes are created all the time, the only restriction being the limits on the human imagination. Take at random almost any adjective and you are very likely to find examples of similes for it in the lexical corpus:

83) *dumb: dumb as a beetle/a box of rocks/a dog/a drum with a hole in it/a farmer/hell/an ox/a stump* etc.
84) *ugly: ugly as sin/a pug dog/truth*
85) *fat: fat as balloons/a pig*
86) *tall: tall as a tower/a house/a giant*
87) *cold: cold as a freezer/ice/charity/steel/debt/death/a Siberian winter/granite/stone/a glacier/a grave*

Similes given in (80-87) above are very significant to the researcher; they represent a parting of the ways between the lexicalized/institutionalized units and the open-ended nature of language. In other words, the members of this class bridge the gap between formulaic language with its varied degrees of fixedness and the unfettered potential of creative language that has nothing in common with MWU's except similarity in patterning with one of its categories (Abu-Ssaydeh 2003).

3.5. The idioms remain the most challenging to both the lexicographer and the translator; evidently, with the exception of slot-filling, every possible type of variation we have documented here is displayed in the idiom category, with numerous idioms combining more than one type of variation. This is possibly the case because idioms are numerically superior to any other category and the occurrence of a verb element that makes idioms more susceptible to a variety of verb-related manipulations – tense, mood, voice, etc. They, however, do not display a uniform behavior as the examples cited above demonstrate.

The following table summarizes the types of variation permissible in each category:

Table 1.

	fixed phrases	institutionalized utterances	lexicalized sentence stems	idioms	similes	proverbs
nil variation	+	+	+	+	+	+
slot-fillers			+			
substitution by synonyms		+	+	+	+	

substitution by non-synonyms	+	+	+	+	
optionality	+		+		+
pattern/ lexical constituents variation	+	+	+		
syntactic and lexical variation				+	+
creativity				+	

4. Variation: Any theories?

This survey of variation in MWU's leads us to three conclusions regarding this lexical area. The first is that variation is most evident in the idiom category, but the overall range and size of variation in this area as a whole should not be underestimated since a high percentage of MWU's belongs to this category. Secondly, sweeping claims made by many linguists that MWU's are fixed entities have to be modified in order to accommodate the prevalent and complex patterns of variation witnessed in this area. Thirdly, lexicographers have to develop methodologies that will reflect variations through a methodology that is both suited to the inherently restrictive nature of the dictionary and sufficiently clear and detailed to allow the translator to generate MWU's variants with nativelike competence.

Having said that, though, can we describe variation in MWU's as being systematic enough to permit the formulation of a theory that would allow the translator to predict the behavior of MWU's every time one of them is encountered? It has to be emphasized that, from a translator's point of view, any theory would be valuable to the extent that it can satisfy one or both of the following pre-requisites: (a) it must enable the translator to predict variation in an accurate manner and/or (b) it must enable the lexicographer to compile dictionaries that can explicitly account for variation exhaustively yet concisely.

A survey of the relevant literature shows that while variation has been tentatively recognized by researchers, no adequate explanations have thus far been proposed. In fact, the only three interpretations found limit themselves exclusively to idioms and binomials, leaving variation in other categories without explanation. The first of these was put forward by Ettlinger (2002) who argues that the degree of idiom fixedness is determined by the age of the idiom. "Idioms with varying degrees of syntactic flexibility", states Ettlinger, "are at various stages in the diachronic process of becoming 'dead' idioms; new idioms fresh from simply being metaphors, will have little restrictions on (their) flexibility, and as with regular language, lexical items can be swapped with synonyms. Near-dead idioms will become inflexible and immutable as their literal meanings disappear into the past" (Ettlinger 2002: 12).

The major difficulty with this theory is that, from the translator's perspective, the vast majority of English idioms are in fact non-compositional; their meanings are not self-evident whether they are described by researchers as opaque, transparent or somewhere in between. Secondly, if the time factor is to be a vital feature that determines the extent of variation the idiom may tolerate, where and how would such information be available to the translator? Thirdly, even if such information becomes accessible, does the theory really apply to all "dead" idioms? How can we explain the difference in behavior between *kick the bucket*, *spill the beans* and *bite the bullet*, all of which are usually described as "dead" idioms? *Kick the bucket* permits only the past tense; *spill the beans* appears in different highly complex realizations and *bite the bullet* appears in the past tense, the infinitive, -*ing* form, simple present, the imperative form and with the modals *should* and *ought to*. Moreover, what do we do with the scores of culture-specific idioms which will always be opaque to the translator regardless of their age such as *carry coals to Newcastle*, *beat about the bush*, *run with the hare* and *hunt with the hounds* and *bark up the wrong tree*?

The second theory was proposed by Jackendoff and others and is summarized in Ifill (2002). Briefly stated, the theory claims that the degree of fixedness in an idiom is determined by syntactic transparency or the relation of the idiom to its non-idiomatic counterpart; if the idioms syntactically resemble their meanings (i.e. they are syntactically transparent), then transformations such as passivization as well as internal modifications are permitted. An example is the idiom *keep tabs on* which can be mapped on its meaning: 'maintain surveillance of'. The same principle applies to the idiom *spill the beans* which permits the passive transformation for the same reason ('reveal a secret'). On the other hand, the idioms *kick the bucket*, *he bought the farm* and *bite the dust* all of which mean 'die' cannot be mapped on their meaning ('die') since *kick*, *buy* and *bite* are transitive verbs that take an object whereas *die* is an intransitive verb (Ifill 2002).

But, again, there are serious flaws with this theory from the translator's standpoint. To start with, the relationship obtaining between the idiom and its meaning is so complex that it may not be simply interpreted as a matter of syntactic mapping. Secondly, a cursory examination of any random list of idioms would reveal that the number of idioms in English exhibiting such syntactic correspondence with their meanings is very small. Thirdly, the theory cannot explain this huge difference in variation existing between individual idioms. In fact, an examination of a few idioms would show that the range and nature of

variation found in idioms is almost idiom-specific. And even if we accept the conclusions arrived at by both Ettlinger (2002) and Jackendoff (1997), it is doubtful if any of these conclusions would contribute significantly to lexicography or translator training.

The third theory seeks to explain why certain binomials permit a reversed order while others resist it. According to a hypothesis suggested by Golenbock (2000), it is " ... the frequency of the two words in spoken English (that) determines which word comes first when spoken (in particular, the more frequent word comes first in the expression)". To test this hypothesis, Golenbock made a list of binomial expressions and tested the frequency of the words comprising them in a lexical corpus. The patterns of two-thirds of the nominal binomials and all adjectival binomials confirmed his initial hypothesis. For example in the expression *skin and bones*, it was found that the word *skin* occurred 5353 times in the corpus while *bones* occurred 961 times. Thus, the phrase would be non-reversible, which is the case. Similarly, *fish and chips* is non-reversible because the word *fish* appeared in the corpus 5219 times compared to 863 citations of the word *chips*. But a very large number of noun and verb combinations did not support the hypothesis, for example *stop and shop*, *hit and run* and *come and go*. Interesting as this hypothesis might be, it does not constitute a principled approach to variation which lexicographers can systematically utilize, nor can it be of any immediate benefit to translator.

5. MWU's in dictionaries

Both general use and specialized dictionaries list MWU's, the first as one of the primary features and the second as an exclusive feature. But do dictionaries of either type provide the translator with a surefire way of using them correctly, especially in so far as variation is concerned? It can be safely said that certain MWU categories do receive a fairly substantial coverage in both kinds of dictionaries. Examples of those are fixed phrases, institutionalized utterances and some lexicalized sentence stems; these are easily identifiable categories where variation is, lexicographically speaking, manageable and the amount of variation permitted therein is fairly limited. They are, therefore, a safe bet for both the lexicographer and the translator.

Idioms, similes and proverbs, by contrast, are a different kettle of fish. The range of lexical and syntactic variation permitted in these three categories is so vast and so complex that little has been done in lexicography to account for it. Even dictionaries of idioms fail to list rudimentary variation such as alternative lexical items or optional elements in the MWU. For example, *OID* (2001) lists *put/lay his cards on the table*. Nothing is said about the fact that *cards on the table* and *all cards on the table* may occur as independent, and therefore dictionary-quotable, variants. No dictionary I am familiar with provides any thorough

treatment of the syntactic behavior of MWU's; the transformations they may undergo, tense change, mood and aspect changes they accept or the lexical items that may be inserted at different points in their syntax. Such knowledge may be taken for granted by the native speaker, but it is crucial for the translator if he is to use MWU's effectively and accurately. The problem is that the only two sources for such variation are the knowledge stored in the brains of the native speakers and lexical corpora which are material representations of such knowledge. And since access to the native speakers' brains is impracticable, the only remaining and readily accessible source is lexical corpora. This being the case, lexicographers are required to operate on three fronts: to fine-tune their current methodologies in order to account for variation; to thoroughly examine lexical corpora such as the British National Corpus and the Bank of English in order to identify possible variational patterns and to find ways of handling unpredictable variation, especially in the idiom category. This proposal can be detailed as in the following sections.

5.1. The current practice of citing MWU's partially or in a haphazard manner represents a false picture of English in actual use. This effort has to be based on a systematic analysis of lexical corpora and a careful documentation of the findings.

5.2. Regional variations with their different realizations should be established and cited side by side for the benefit of the translator. For example, the American idioms *hit a home run*, *home stretch* and *second stringers* are absent from *OID* (2001). On the other hand, *pipped at the post*, *Bob's ypur uncle* and *get his knickers in a twist*, are typically British, rather than, American MWU's.

5.3. Where parts of a certain MWU are allowed to occur independently, such independence should be recognized through separate listings:

88) *play second fiddle*
    second fiddle

89) *send shivers up and down his spine*
    shivers run up/through/all over your spine/body

90) *shed crocodile tears*
    crocodile tears

91) *lay/put his cards o the table*
    all cards on the table
    cards on the table

92) *give him/get the green light*
    the green light

5.4. Similes that are cited with a fairly high frequency in the corpus should be represented. For example, the simile *as happy as a lark* is cited in LDOCE whereas the similes *as happy as the day is long/as a clam/as Larry* are cited in *OID* (2001). Other possible similes with the adjective *happy* that recur in the corpus are:

93) *as happy as a pig in the mire/poop/s**t*
    as happy as a sandboy

Ugly as sin on the other hand, is not cited in either dictionary, though it appears repeatedly in the corpus. The same principle applies to *howl like a banshee/a dog/a wolf, fat as a pig* and *cold as stone* which are also frequent in the corpus and should be recognized as lexicalized ones (for more examples, see Abu-Ssaydeh 2003).

5.5. Where the MWU may have alternative syntactic patterns, both patterns should be cited:

94) *I can't rightly say/don't rightly know*
    by no means/not by any means

This category should also include cases where nominalization and adjective-formation are permitted, in which case each variant would be listed as a separate entry in the dictionary:

95) *break the ice*
    an ice-breaker

96) *open his eyes*
    an eye-opener

97) *settle a score*
    score-settling

98) *pass the buck*
    buck-passing

5.6. The typical lexicographical practice of using the phrase *et cetera* to indicate possible variants at any given position should be replaced by actual listing of as many as possible of the lexical variants permitted in the structure of the MWU:

99) *brake/clatter/come/crash/draw/grind/jerk/puff/roll/screech/*
    shudder/skid/skim/squeal/tremble to a stop

100) *haul/rake/take him over the coals*
101) *a bag/can of worms*
102) *tear/cut/break loose from*
103) *not give a tinker's cuss/a damn/a donkey's arse/a fig/a hoot/a toss*
104) *like a bull in a china shop/closet*
105) *be in full flood/flow/spate*
106) *good grief/god/lord/heavens/gracious*

5.7. Reversibility in binomials, a feature that is totally absent from English dictionaries, should also be indicated:

107) *day and night*: list also under *night and day*
    friends and allies: list also under *allies and friends*
    men and women
    come and go

5.8. Grammatical modifications and possible lexical insertions should be somehow indicated in their appropriate places within the structure of the MWU. For this purpose, one could tentatively propose two complementary forms of representation in the MWU entry; firstly, grammatical modifications are shown in terms of grammatical categories such as past, passive, modal, interrogative and negative. Additionally, actual lexical items (such as adjectives, adverbs, verbs, nouns, etc.) can be cited along with the original lexical constituents of the MWU to show language at work, as it were. The original elements of the MWU can, furthermore, be in bold face to distinguish them from those that belong to the permissible variations. For example, the idiom *not see the wood for the trees* can be represented as follows (brackets indicate optionality):

108) *not see the wood for the trees*:

miss the forest for the trees; (interrogative: modal/(past perfect) *fail to/difficult/(a little) tricky/impossible to/not)* **see the wood for the trees**

This formulation means that this MWU occurs in two forms which we separate by a semicolon. The detailed representation of the second form (*see the wood for the trees*) is important for the translator since it does capture the actual realizations of this idiom in real use as the following examples from the *BNC* demonstrate:

109)

a) *It was impossible to see the wood for the trees.*
b) *Can we see the wood for the trees?*

c) *He could see the wood for the trees.*
d) *It is often difficult to see the wood for the trees.*
e) *We miss the forest for the trees.*
f) *The Inland Revenue had failed to see the wood for the trees.*
g) *T's a little tricky to see the wood for the trees in this one.*

(Interestingly, the examples cited above show that *not* is not an integral element of this MWU and should, therefore, be listed as an optional element.)

Other representations would include:

110) *Send shivers (up and down) your spine*:

(*would*) **send**(*-ing*) (*cold/strange/tiny/unexpected*) **shivers** (*of disgust/ danger/trepidation/awareness; quivering frantically*) **along/up/down/up and down** (*rippling down/running through*) **your spine**

111) *play second fiddle to*:

play *a* (*muted/poor/soft*) **second fiddle to**

112) (*take*) *with a grain of salt*:

(*take*) **with a** (*certain/hefty/large*) **grain** (*/pinch/fistful*) **of salt**

113) *a hard/tough act to follow*:

a (*hell of/very/exceedingly/great/no/not an easy/almost impossible/*) **tough/hard** (*-er*) **act to follow**

114) *keep tabs on*:

(*past/passive/modal*) **keep**(*-ing*) (*careful/close*) **tabs on**

115) *lay/put your cards on the table*:

put/lay(*-ing*) (*all*) (*your*) **cards on the table**

116) *a/the fly in the ointment*

a (*/the main/only/one enormous*) **fly in the ointment**
(*there was* (*only*) *one/an other/there were no obvious*) **fly** (*s*) **in the ointment**

117) *hold on by a hair/thread*

(*past/present*) **hold on/hang**(*-ing*) (*on*) **by a thread/hair**

118) *the icing on the cake*

this/the (*final/spectacular*) **icing on the** (*scrummy*) **cake**

6. Conclusion

MWU's are crucial elements in the English lexis and their acquisition is important for the translator. They occupy a position between the two more stable and well-defined levels of the single word and the rule-governed level of syntax. The changes ranging from nil modification typical of fixed phrases to the open-ended simile prove that variation in MWU's cannot and must not be described in absolute terms. They also show that it is inaccurate to constantly seek any correlation between a specific kind of variation and a specific MWU category. To the dismay of both the lexicographer and the translator, no single coherent theory has been thus far advanced to systematically explain the nature of variation in a way that would permit predictability by translators or assist lexicographers in their work. This failure means that they have to be left to their own devices to map out the directions variation takes in MWU's.

As a result, the problem of variation can be approached from several angles: by modifying lexicographical methodology so that wherever syntactic and lexical variations are allowed in a given MWU, such variations can be documented in an economical and user-friendly way in the dictionary. To this end, MWU's must be closely monitored in lexical corpora in order to determine the scope of variation and the permissible alternative patterns as well as to document observable lexical and syntactic variations and optional elements in each unit. It is equally vital for the translator to appreciate the complexity of variational patterns in MWU's, the idiosyncratic nature numerous MWU's exhibit and the extent to which they can be syntactically and lexically manipulated, not on the basis of a theory-based or rule-based predictability but on the grounds of permissible and corpus-verifiable changes. Until this task is accomplished, translators have to avail themselves of the online lexical corpora in order to ensure that variations they might like to make are indeed compatible with those native speakers of English would recognize as nativelike.

REFERENCES

Abu-Ssaydeh, Abdul-Fattah
    2003   "Similes: The dictionary and beyond", *Journal of Applied Linguistics* 17/2: 81-107.
    (forthcom- "Multi-word units in English-Arabic dictionaries: Status and translation strategies",
        ing)   *Turjuman*.
Anaud, Pierre J. L. – Henri Bejoint (eds.)
    1992   *Vocabulary and applied linguistics*. London: Macmillan.

Baker, Mona
    1992   *In other words coursebook on translation.* London – New York: Routledge.

Bhatia, Vijay
    1994   "Cognitive structuring in legislative provisions", in: John Gibbons (ed.), 136-155.

Coady, James – Thomas Huckin (eds.)
    1997   *Second language vocabulary acquisition: A rationale for pedagogy.* Cambridge: Cambridge University Press.

Cowie, Anthony P.
    1992   "Multiword lexical units and communicative language teaching", in: Pierre J. L. Anaud – Henri Bejoint (eds.), 1-12.

Everaert, Martin Erik – Jan van der Linden – André Schenk – Rob Schreuder (eds.)
    1995   *Idioms: Structural and psychological perspectives.* Hillsdale: Lawrence Erlbaum Associates.

Fellbaum, Christiane D. (ed.)
    1998   *Wordnet: An electronic data base.* Cambridge, MA: MIT Press.

Fillmore, Charles J.
    1979   "On fluency", in: Charles J. Fillmore – Daniel Kempler – Willaim Wang (eds.), 85-102.

Fillmore, Charles J. – Daniel Kempler – William S.-Y. Wang (eds.)
    1979   *Individual differences in language ability and language behaviour.* New York: Academic Press.

Gazdar, Gerald – Ewan Klein – Geoffrey Pullum – Ivan Sag
    1985   *Generalized phrase structure grammar.* Cambridge, MA: Harvard University Press.

Gibbs, Raymond W.
    1995   "Idiomaticity and human cognition", in: Martin Everaert – Jan van der Linden – André Schenk – Rob Schreuder (eds.), 97-116.

Gibbons, John (ed.)
    1994   *Language and the law.* London: Longman.

Jackendoff, Ray
    1997   *The architecture of the language faculty.* Cambridge, MA: MIT Press.

LDOCE   see Summers, Della – Michael Rundell *et al.* (eds.)

Lennon, Paul
    1998   "Approaches to the teaching of idiomatic language", *International Review of Applied Linguistics* 36/1: 11-20.

Lewis, Michael
    1993   *The lexical approach.* Hove: Language Teaching Publications.
    1997a   *Implementing the lexical approach.* Hove: Language Teaching Publications.
    1997b   "Pedagogical implications of the lexical approach", in: James Coady – Thomas Huckin (eds.), 255-270.

Summers, Della - Michael Rundell *et al.* (eds.)
    1995   *Longman dictionary of contemporary English* (LDOCE). (3rd edition.) Harlow: Longman.

McCarthy, Michael
    1990   *Vocabulary.* Oxford: Oxford University Press.

Mieder, Wolfgang
    1985   "Popular views of the proverb", *Proverbium* 2: 109-143.

Moon, Rosamund
    1998   *Fixed expressions and idioms in English: A corpus-based approach.* London: Clarendon Press.

Nattinger, James R. – Jeanette S. DeCarrico (eds.)
    1992   *Lexical phrases and language teaching.* Oxford: Oxford University Press.
    2001   *Oxford idioms. Dictionary for learners of English.* Oxford: Oxford University Press.

OID   see Nattinger, James R. – Jeanette S. DeCarrico (eds.)

Pawley, Andrew K. – Frances H. Syder
    1983   "Two puzzles for linguistic theory: Nativelike selection and nativelike fluency", in: J. C. Richards – R. W. Schmidt (eds.), 191-226.

Richards, Jack C. – Richard W. Schmidt (eds.)
    1983   *Language and communication.* New York: Longman.

Riehemann, Sussane Z.
    2001   A constructional approach to idioms and word formation. [Unpublished Ph.D. dissertation, University of Stanford.]

Schmitt, Norbert
    2000   *Vocabulary in language teaching.* Cambridge: Cambridge University Press.

Sinclair, John M.
    1987   "Collocation: A progress report", in: Ross Steel – Terry Threadgold (eds.), 319-331.
    1991   *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Steel, Ross – Terry Threadgold (eds.)
    1987   *Language topics: Essays in honour of Michael Halliday.* Vol. 2. Amsterdam: John Benjamins.

Widdowson, Henry G.
    1989   "Knowledge of language and ability for use", *Applied Linguistics* 10: 128-137.

Willis, Dave
    1995   "Review of lexical phrases and language teaching", *ELTJ* 49/1: 87-90.

Yorio, Carlos
    1980   "Conventionalised language", *TESOL Quarterly* 14/4: 433-442.

INTERNET SOURCES

Arnold, Doug J.
    1995   "Multi-word units: Collocations and idioms", available at http://www.essex.ac.uk/linguistics/clmt/MTbook/HTML/node55.html (date of access: 23 March 2003).

Bastow, Tony
    2003   "'Friends and allies': Binomials in a corpus of US defence speeches", English Language Postgraduate Seminars, University of Birmingham, available at http://www.english.bham.ac.uk/research/archive/2003-2004/langpgseminars03-04/s1w10a.htm (date of access: 9 April 2005).

British National Corpus
    2004   http://thetis.bl.uk/lookup.html (date of access: June 2004).

Ettlinger, Marc
    2002   "What's the deal with idioms?", available at http://socrates.berkeley.edu/~marce/ling/Papers/220.pdf (date of access: 15 February 2003).

Fatihi, Ali R.
    2004   "Lexical Approach: Putting theory into practice", *Language in India* 4/1, available at http://www.languageinindia.com/jan2004/lexical.html (date of access: 5 January 2004).

Fernando, Chitra – Roger Flavell
    1981   *On idiom: Critical views and perspectives.* (Exeter Linguistic Studies 5.) Exeter: University of Exeter, available at http://ling.kgw.tu-berlin.de/lexicography/data/IDIOM.html (date of access: 12 January 2004).

Golenbock, Janice
    2000   "Binomial expressions – does frequency matter?", available at
            http://www.andrew.cmu.edu/course/76-451/golenbock.html
            (date of access: 24 March 2003).

Ifill, Tim
    2002   "Seeking the nature of idioms: A study of idiomatic structure", available at
            http://www.swarthmore.edu/SocSci/Linguistics/papers/2003/ifill.pdf
            (date of access: 19 December 2003).

Lavelle, Thomas
    2003   "Teaching lexical phrases", available at
            http://bogglesworld.com/askthomas.htm (date of access: 3 July 2004).

Lindstromberg, Seth
    2003   "My good-bye to the Lexical Approach", *Humanising Language Teaching* 5/2,
            available at http://www.hltmag.co.uk/mar03/mart1.htm (date of access: 29 December
            2003).

Porto, Melina
    1998   "Lexical phrases and language teaching", *Forum* 36/3, available at
            http://exchanges.state.gov/forum/vols/vol36/no3/p22.htm
            (date of access: mid-January 2004).

Sag, Ivan A. – Timothy Baldwin – Francis Bond – An Copestake – Dan Flickinger
    2003   "Multi-word expressions: A pain in the neck for NLP", available at
            http://lingo.stanford.edu/pubs/WP-2001-03.pdf
            (date of access: 7 September 2003).

Hirsch, E. D. – Joseph F. Kett – James Trefil
    2002   *The New Dictionary of Cultural Literacy*. (2nd edition.), available at
            http://www.bartleby.com/59/4/ (date of access: February 2004).

Willis, Dave
    1997   "Lexical phrases in syllabus and materials design", available at
            http://langue.hyper.chubu.ac.jp/jalt/pub/tlt/97/sep/willis.html
            (date of access: 17 December 2003).