Oxford Handbooks Online

Abstract and Keywords

This chapter provides an overview of multi-word items (in English). After discussing contested terms such as *idiom, formula*, and *multi-word item* itself, it then examines criteria for classification of a sequence as a multi-word item, and the phraseology continuum. Next it considers different descriptive and theoretical approaches to multi-word items. The following sections look in turn at idioms, proverbs, phrasal verbs, binomials, similes, then looser kinds of formulae and phraseological item including prefabs. The final section comments briefly on some cross-linguistic aspects.

Keywords: formula, idiom, idiomaticity, multi-word item, non-compositionality, phrasal verb, phraseology, prefab, proverb

7.1 Introduction

THIS chapter on phraseology deals specifically with the phenomenon of multi-word items: that is, lexical items which consist of more than one 'word' and have some kind of unitary meaning or pragmatic function. While the existence of idiomatic, irregular, or unpredictable combinations of words seems universal in natural languages, there are marked differences in morphological and semantic types amongst different languages, and so, to avoid generalization, this chapter focuses primarily on multi-word items in English, adopting a largely synchronic and descriptive approach.

The following subsection discusses terms, then section 7.2 discusses general criteria for multi-word items and section 7.3 gives a short overview of descriptive and theoretical approaches. Section 7.4 looks at idioms; section 7.5 at proverbs and some other types of

multi-word item; and section 7.6 at formulae, frames, and prefabs. Finally, section 7.7 comments briefly on cross-linguistic aspects.

7.1.1 Terminology

As many have pointed out, the study of multi-word items is complicated not just by the range of types of item and the unclear boundaries between types, but also by the range of terms in use, often in conflicting ways. The problem arises because different methodological approaches, theoretical departure points or end purposes, have led to different systems of classification, which, however internally consistent and satisfactory, do not fit well together. Consequently, new terms are introduced while existing terms are redefined and applied to categories with divergent characteristics. (For discussion of this problem, see Howarth 1996; Moon 1998; Wray 2002.) For the sake of clarity, I will keep to relatively standard categories, using central terms as follows.

'Multi-word item' (MWI) is used as a superordinate term for the more specific types described below: idioms, formulae, and so on. Other general terms in use include (p. 121) 'fixed expression',[1] 'fixed phrase', 'lexical phrase', 'multi-word unit', 'phrasal lexeme', 'phraseme', and 'phraseological unit'. Recurrent sequences of words, including complex forms of verbs (*has been standing, to be considered*) and freely-formed collocations (*hot weather, walk slowly*), cannot be considered MWIs, nor are quotations unless associated with additional meanings or functions (as with, for example, *an eye for an eye* and *thereby hangs a tale*). Also excluded are compound nouns, verbs, and adjectives, where multi-wordedness is a matter of orthography, with open forms often in free variation with hyphenated or solid forms; consider examples like *textbook, text book, text-book* and the slight but increasing tendency for prefixes to be written as separate words (*the anti war demonstrators, non existent*).

The term 'idiom' is particularly confusing since there are multiple established uses. I will adopt a middle position here, and use 'idiom' to designate MWIs which are problematic because of their semantics: potentially ambiguous, often figurative, and also often evaluative or connotative. Canonical examples are *bury the hatchet, follow in someone's footsteps, move heaven and earth, out on a limb, rain cats and dogs, a skeleton in the cupboard/closet, spill the beans, with all guns blazing*. Broader uses in the phraseology literature include non-transparent compounds (*nosedive, paperback*), and some grammatically anomalous items (*at all, by and large*) or other fossilized sequences (*thank you, needless to say*), which I will instead classify as formulae. *Idiom* is sometimes extended still more broadly to constructions such as *neither ... nor* or *prefer X to Y*: I exclude these altogether. In contrast, the narrowest uses limit 'idiom' to opaque polysemous phrases with a viable literal meaning: thus *bury the hatchet* and *spill the beans* are considered idioms, but transparent items like *follow in someone's footsteps* and hyperboles like *move heaven and earth* are not. This does not seem a helpful distinction to

me, as literal interpretations are usually precluded by context, and all need some kind of explanation.

Formula is used here to label a subset of MWIs characterized by their pragmatic functions: greetings, discourse markers, and other situation- or text-bound items. Some are agrammatical, some well-formed, and examples include *how do you do?, many happy returns, thank you, mind the gap!, by the way, in fact, needless to say, and so on*, etc. See further in section 7.6.

7.2 Criteria: the phraseology continuum

There are no hard-and-fast rules for distinguishing MWIs from other recurring word sequences, nor are there clear-cut divisions between MWIs and non-MWIs. Instead, it is generally agreed that there is a phraseology continuum and that several factors need to (p. 122) be taken into account: in particular, the institutionalization of a sequence; its fixedness or frozenness; and its degree of semantic or pragmatic compositionality.

'Institutionalization' as a criterion refers to the extent to which a string of words recurs. This can be measured by inspecting, say, corpus evidence, though some MWIs are comparatively infrequent, or restricted to certain kinds of language.[2] Some would add that for a string to be considered an MWI it must be recognized as a holistic sequence in the lexicon in order to exclude very frequent freely-formed strings such as *in the middle of* or *it is possible to*. As a criterion, this is more subjective and imprecise, though it can be partially assessed through informant testing or checking dictionary inclusion.

'Fixedness' refers to the extent to which a string of words is fossilized, either paradigmatically (other words cannot be substituted or added, component words cannot be omitted) or syntagmatically (with restrictions on sequencing and/or regular grammatical operations such as inflection, passivization, and negation). This too can be checked with corpus evidence. Variability is discussed further in section 7.2.1.

'Non-compositionality' refers to the extent to which a string of words has a unitary meaning that cannot be derived by decoding, literally, each component word. Sometimes unitary meanings are obscure and impossible to retrieve synchronically (as with *rain cats and dogs* = 'rain heavily'[3]). Others are more amenable to interpretation (*alarm bells ring*) or have specialized pragmatic functions (*happy birthday*).

These three criteria are variables: idioms like *kick the bucket* are institutionalized (at least in the folk lexicon), fixed, and non-compositional, as are anomalous formulae like the emphasizer *at all* following a negative or *any. Live and let live, change horses in midstream* are institutionalized and relatively fixed, but partly compositional. *Miss the bus/boat* varies without significant change of meaning; *break the ice* permits a wide range

of grammatical operations. Sequences such as *at the same time* are semantically compositional but fill discourse roles in argumentation.

Further criteria used to distinguish English MWIs from ordinary sequences include some kind of syntactic unity (e.g. functioning as a prepositional phrase, noun phrase, verb+complement, clause, etc.) and, in spoken discourse, phonology and intonation. To these can be added some notion of non-productivity, since there can be no rules for MWI formation beyond the usual grammatical and semantic resources of a language and, sometimes, a process of analogy. At the same time, some MWIs are particularly susceptible to creative manipulations, and there are also some very productive phraseological frames (see section 7.6.3).

(p. 123) 7.2.1 A note on variation

Statements about the incidence and nature of MWI fixedness reflect different approaches to lexical variability. One approach—a lexical approach[4]—is to consider each lexically unique item as independent: thus *throw in the towel, chuck in the towel*, and *burn one's boats, burn one's bridges* constitute pairs of idioms rather than variant forms. A second or lexicographical approach would consider each pair as a single idiom with institutionalized variations: in a conventional dictionary, these are likely to be defined together under a stable lexical element (e.g. *towel, burn*—if necessary, with comments where variations represent British/American distinctions, as with *flog/beat a dead horse*).[5] Neither approach deals well with the phenomenon of extreme and multiple variations, as with these clusters:

> *shake/quake/quiver in one's shoes*
> *shake/quake/quiver in one's boots*
> *shake/quake/quiver in one's loafers/sandals/wellies/size nines* (etc.)

> *fan the fire/fires/flames (of something)*
> *add fuel to the fire/flame/flames*
> *fuel the fire/fires/flame/flames (of something)*

A single metaphor underlies each cluster, with the component lexis selected from restricted sets, but nothing except *in* and *the* is fixed. A conceptual approach might therefore consider them as single idioms. In my terms, they are 'idiom schemas' (1998: 161ff.). See also Nunberg et al. (1994: 504ff.), who term them 'idiom families', and discussion of sets of variants in Taylor (2012: 77ff.).

7.3 Multi-word items in theory and practice

Multi-word Items

This section looks briefly at different kinds of approach to the study and description of MWIs, and provides a short selection of references. As already mentioned, it is this diversity of approach which has led to conflicts in terminology and indeed in taxonomic description, irrespective of other differences arising from the different languages being examined.

At the most descriptive end of the range are lexicographical approaches, where MWIs are identified as fixed and semi-fixed items needing to be defined or explained in dictionaries (this is essentially the approach I adopted in Moon 1998, from which I draw heavily in the following discussions of types of MWI). Mainstream lexicography generally seeks to be consistent and clear, rather than conforming tightly to theoretical models, and uses a restricted range of typological labels—sometimes only *phrase* and *phrasal verb, proverb, saying*. But lexicography also has the discipline of dealing non-selectively with a whole lexicon, pair of lexicons, or substratum of a lexicon, thus treating several thousand items from across the phraseology continuum, not just best-case exemplars or the much-discussed few such as *bury the hatchet, spill the beans, kick the bucket, kith and kin*.[6] Sometimes linked to the lexicographical are lexicalist approaches, which segment the phraseology continuum into major categories reflecting degrees of collocational frozenness as well as semantic opacity: see, for example, the work of the Russian lexicalists, including Melčuk, Vinogradov, and Amosova as in Cowie's account (1998: 4ff., 209ff.). See too discussion by Howarth (1996), where he describes 'four major phraseological categories': 'free collocations' (*blow a trumpet*), 'restricted collocations' (*blow a fuse* in the literal sense), 'figurative idioms' (*blow your own trumpet*), and 'pure idioms' (*blow the gaff*) (1996: 33 and *passim*). Similar segmentations of the phraseology continuum have been widely adopted.

Corpus approaches, also descriptive, are driven by a concern with the phraseological patterning observed in the target language. Since many subtypes of MWI are relatively infrequent, the focus of corpus approaches tends to be on collocation and recurrent strings, which may be entirely compositional. Sinclair's work with collocation led to his model of language structure as composed of both an Open-Choice Principle and Idiom Principle,[7] with phraseologies often of 'indeterminate extent' (his example is *set eyes on*, which tends to co-occur with *never* or temporal expressions) and allowing internal variation: see Sinclair (1991: 109ff. and *passim*); later, his corpus-driven work on meaning led to his view that 'units of meaning are expected to be largely phrasal' (2004: 30). Biber et al.'s corpus work on grammar led to identification of 'lexical bundles', which they explain as 'recurrent expressions, regardless of their idiomaticity, and regardless of their structural status' that are three or more words in length (1999: 990ff.): amongst their four-word examples are *I don't know (what), have a look at, the nature of the, as a result of, are likely to be*. These two approaches typify a very productive field: special interests include co-textual constraints; what amount to prefabricated sequences; and implications for the study and teaching of MWIs.

With pedagogically oriented approaches, concerns relate to teaching language learners about MWIs in their target language, including recognition of multi-word sequences as idiomatic units and asymmetries between first and foreign language: for English, phrasal verbs are a particular problem area. However, these approaches are also concerned more broadly with idiomaticity and phraseology, since these are essential for fluency in speech and writing. Issues include, in particular, collocations (e.g. Howarth 1996; Nesselhauf 2005), formulae, and general phraseological phenomena, as with (p. 125) prefabricated sequences, routines, lexical phrases (e.g. Nattinger and De Carrico 1992), or lexicalized sentence stems (Pawley and Syder 1983). See also discussions in collections edited by Schmitt (2004) and Meunier and Granger (2008). Cognitive approaches, for example Kövecses and Szábo (1996) and Boers (2011), have contributed insights to the teaching of more figurative MWIs by drawing attention to the conceptual metaphors and other patterns which underlie sets of idioms and other MWIs; Wulff (2008) combines cognitive and corpus approaches in her exploration of the nature of idiomaticity.

Partly linked to these are psycholinguistic approaches, including the pioneering work of Wray on the formulaic nature of language (2002, 2008): her 'heteromorphic lexicon model … permits multiple part-mappings of the same information in the lexicon' such that phraseologies can be stored both holistically and decomposed into 'subparts' (2008: 14–15), thus allowing for routine, idiomaticity, and creativity. A further important strand of psycholinguistic research relates to the processing of MWIs, in particular whether idioms and similar non-compositional items are processed as gestalts or first processed word-by-word (reaction times support the first): classic studies here include many by Gibbs (1985, 1986, 1995, 2007) and work reported in Cacciari and Tabossi (1993).

Lastly, a significant and dominant strand in MWI theory has developed from syntactic work: since MWIs are gestalts and often ill-formed grammatically, they cannot be accommodated within standard generative or rule-governed models of language, and have to be accounted for elsewhere, for example via a look-up list of exceptions. Important discussions of this include those by Katz and Postal (1963), Katz (1973), Weinreich (1969), and, partly revisionist, Jackendoff (1995) and Nunberg et al. (1994); see also Fraser's work (1970) on degrees of fixedness and grammatical defectiveness in MWIs.

7.4 Idioms

Even adopting a constrained definition of 'idiom', this is a very diffuse category in English in terms of grammatical behaviour and degrees of fixedness, institutionalization, and non-compositionality (and frequency or infrequency of occurrence). First and foremost, however, these items are characterized by semantic peculiarities; examples include:

between a rock and a hard place

bury the hatchet
cut corners
get along like a house on fire
give someone the cold shoulder
once in a blue moon
a piece of cake
pull someone's leg
smell a rat
pigs might fly/when pigs fly

(p. 126) Amongst these, *cut corners* is relatively transparent, as is *smell a rat*, where its meaning, 'be suspicious about something', could just about be retrieved from the conventional negative connotations of *rat* in English (untrustworthiness, dishonesty, etc.), leaving aside the more obvious interpretation of a dead and decomposing rat. More opaque is *get along like a house on fire*, where a common-sense interpretation via literal meaning may not produce the appropriate idiomatic meaning, but instead suggest 'destructively' rather than 'extremely well'. For items like *a piece of cake*, the idiomatic meaning might be guessed from context, but might not; it could be interpreted as 'something pleasant or indulgent' rather than 'something easy'. Thus non-compositionality is subjective, depending on individuals' linguistic and metaphorical competence and their decoding of component words.

Since idioms have non-literal meanings they are potentially ambiguous, and this ambiguity is interesting from theoretical perspectives. In practice, however, ambiguity is less of an issue, partly because context excludes literal readings, and partly because data suggests that idiomatic meanings are privileged over literal. For example, the 450-million word Bank of English corpus (BoE)[8] has 145 examples of *smell a rat*, 226 of *bury the hatchet*, and 50 of *kick the bucket*: in each case just one example is literal. This tendency is further supported by psycholinguistic evidence, which suggests that informants process idiomatic meanings faster than literal uses: that is, the collocational sequence predisposes receivers to expect the idiomatic meaning. There are, however, a few genuinely ambiguous idioms, including some body language idioms, such as *shake hands* (on a deal, in reconciliation) or *raise one's eyebrows* (expressing surprise or doubt), which may or may not imply a physical action alongside agreement or reaction. Even where literal equivalents of idioms almost never occur, literal meanings may still be partly retained and exploited in puns or acknowledged in some other way, as in ... *there are times when I fancy I smell if not a rat, then a pretty sinister mouse* (from BoE); some further examples are given below.

While the origins of some idioms are uncertain or idiosyncratic, there are others which correspond to conceptual metaphors (Lakoff and Johnson 1980), and have been used in the literature to explore those metaphors. For example, idioms such as *hot under the collar, hot and bothered, see red, one's blood boils* realize the metaphor ANGER IS HEAT (Lakoff 1983: 380ff., drawing on joint work with Kövecses). Similarly with clusters of metonymic idioms, such as those containing *hand* which refer to actions carried out by those hands (*lend a hand, turn one's hand to, get one's hands dirty*), or *eye*, referring to

understanding or observation (*see eye to eye, keep one's eye(s) on, turn a blind eye*). For further discussion and examples, see Kövecses (2000, 2002) and Goatly (1997, 2007); also Moon (1998: 202ff.).

An important aspect of idiom meaning is evaluative orientation. While, for example, *a piece of cake* and *get along like a house on fire* evaluate a task or friendship positively, (p. 127) *between a rock and a hard place* and *give someone the cold shoulder* evaluate negatively. Evaluations are sometimes predictable, as with *smell a rat*, and sometimes not, as with *cut corners*, which evaluates negatively but where the metaphor might imply efficiency and so evaluate positively. See Moon (1998: 215–308) for discussion of the discoursal roles of idioms, including evaluative and cohesive functions; Fernando (1996) also gives special focus to the functions and evaluations of idioms in text.

7.4.1 Grammatical aspects

Idioms are often noted for grammatical peculiarities, such as transformational deficiencies or inflectional defectiveness. It is difficult to establish these fully except by considering items individually, and the following draws on BoE for information (see also discussion in Moon 1998: 94ff.). Inflectional issues relate particularly to whether component nouns inflect. For example, *leg* in *pull someone's leg* mostly pluralizes consistently with the number of persons being teased. However, *shoulder* in *give someone the cold shoulder* usually remains singular; with *lend a hand, turn a blind eye to*, the component nouns conventionally but illogically remain singular (cf. Sinclair 2004: 30ff., who comments that that singular *eye* is typically associated with figurative and idiomatic usage rather than literal). Occasionally, items such as *collective* are inserted to mitigate singular/plural mismatches. Some examples of these:

> *I am quite interested in what Damien Hirst does, even though a good deal of it **is pulling our legs***.

> *They say **banks gave them the cold shoulder** because their only real collateral was books, which are hard to liquidate at a good price.*

> *… **we** need to all pitch in, **lend a hand** and do our part to help forge a brighter future for this country.*

> ***They have chosen to turn a blind eye to** the fact that most of their pledges are undeliverable at Westminster.*

> *… its parliamentarians seemed ready **to toss the Commission out on its collective ear** for failing to deal with long-running fraud and corruption.*

Where idioms include transitive verbs they should in principle be passivizable; however, some such idioms passivize freely, some only infrequently, and some not at all. *Pull someone's leg* occurs both as *someone's leg was pulled* and *have one's leg pulled*, though active uses are more common. While *give someone the cold shoulder* sometimes passivizes, with the indirect object becoming subject (BoE: *She was given the cold shoulder by the staff there*), more frequent is the lexicalized alternative *get the cold shoulder*. Very few ditransitive idioms have double passivization: a rare example is *pay lip service to someone/something*, where *lip service is paid (to)* and *be paid lip service* are both found. *Kick the bucket* does not passivize at all—logically, since its meaning 'die' is essentially intransitive. *Smell a rat* also seems not to passivize, though this is less easily explained.

(p. 128) Some idioms are canonically negative, as in *not lift a finger, not put a foot wrong*; idioms such as *pull one's weight* and *up to scratch* are more often found with negatives than without. Polarity may also be expressed lexically through antonymous pairs of idioms, as in *keep track of something, lose track of something; in one's element, out of one's element; go up in the world, come down in the world*. Noun groups in idioms are

rarely pronominalized, since the idiomatic meaning produced through the noun group would be lost or concealed, though compare institutionalized *on the one hand ... on the other*; also *pull the other one*, used independently of *pull someone's leg*. A more frequent phenomenon is the number of idioms which generate nominal or adjectival derivatives, often involving inversion of elements: for example, *leg-pull, corner-cutting*; from *blaze a trail, trail-blazing, trail-blazer*; from *break the ice, ice-breaking, ice-breaker*. Compare too the verb *to cold-shoulder* and adjectival forms as in *a once-in-a-blue-moon treat*.

The above comments relate to the internal grammatical workings of idioms. In context, many idioms are further associated with particular structures or selection restrictions. For example, *rock the boat* typically follows a modal or negative expression; *pour cold water on something* needs a human subject, and 'something' is typically realized by an idea or suggestion.

> *We are just going to have to accept their decision because we **don't want to rock the boat.***

> *Strategists spent the conference **pouring cold water on** optimistic forecasts.*

7.4.2 Variation

A number of idioms have regular lexical variations, with little or no change in idiomatic meaning, as in the following:

> *lower/drop one's guard, let down one's guard*
> *bend/stretch the rules*
> *make one's blood run cold, make one's blood freeze*
> *a skeleton in the closet/cupboard*
> *from head to foot/toe*
> *throw someone to the wolves/lions*
> *a bad/rotten apple*
> *bleed someone dry/white*
> *by/in leaps and bounds*
> *out of/from thin air*

At the same time, a few pairs of idioms which might look like simple variations actually have distinct meanings, as in

> *get one's hands dirty* (get involved)
> *have dirty hands* (be guilty)
> *fill one's boots* (get something valuable)
> *fill someone's boots/shoes* (replace someone)

(p. 129) In another type of variation, forms represent different grammatical relationships, as with possession, causation, or aspect:

have/get/develop cold feet
have no/an axe to grind, with/without an axe to grind
cross one's fingers, keep one's fingers crossed
come to a head, bring something to a head
get/have the wind up, put the wind up someone
one's hackles rise, raise someone's hackles
let the cat out of the bag, the cat is out of the bag
open the floodgates, the floodgates are open

In lexical manipulations of idioms, items are inserted to emphasize, add focus or humour, or signal idiomaticity or metaphoricity, as in these from BoE:

*But later, when the Goss Government decided on corporatisation, the Queensland Opposition **smelt an ideological rat**.*

*Beating arch rivals Dublin was **the icing on the league cake**.*

*Hong Kong banks are **weathering the economic storm** better than their regional counterparts.*

*There have been numerous new electronic gadgets which claim to convert the front room into the front stalls–and it need not **cost a sofa-arm and chair-leg**. All home cinema systems are based on* [etc.].

*Some of the world's best wine regions are **literally on your doorstep** and close enough for you to visit on a one to two day trip.*

*Indeed they **got along like the proverbial house on fire**.*

Compare too items such as *hang up one's boots* ('retire'), which has a large number of variations with *boots* replaced by another contextually appropriate, metonymic noun:

*But he was forced to **hang up his boots** last month because of a knee injury, after making only 30 Saints appearances.*

*The housewives' favourite is **hanging up his microphone** after 26 years as Ireland's most popular broadcaster.*

*Now Dowell is **hanging up his dancing shoes**.*

Lastly, exploitations: these are idiosyncratic but interesting stylistically and for ways in which idioms remain sufficiently intact for meanings to be retrieved:

She who pays the undertaker calls the tune. [sc. *pay the piper*]

Will the prisoner be about to spill the enchiladas on Joe when he is suddenly spirited away and will Joe barely escape with his life? [sc. *spill the beans*]

Journalism's job is to tell truth to power. That happens once in an avocado moon. [sc. *once in a blue moon*]

But although the merry Duffield looks like Santa in disguise (if the hood fits, wear it), his generosity may not always be what it seems. [sc. *if the cap/shoe fits*]

(p. 130) 7.5 Proverbs, phrasal verbs, binomials, and similes

This section looks in turn at some types of MWI which have traditionally been picked out as special categories of item for formal or pragmatic reasons.

7.5.1 Proverbs and proverbial sayings

Proverbs are generally characterized as traditional maxims or sayings with didactic meanings, typically offering advice or reflections on life and human nature, social relations, or the natural world. Because many are old and have cognates in other languages and because they express social practices and/or beliefs, they are historically and culturally very interesting, generating a substantial paremiological literature: Mieder (2004) provides an extended discussion. Of the following English examples, some have parallels or sources in classical writers and the Bible:

> *absence makes the heart grow fonder*
> *a bird in the hand is worth two in the bush*
> *first come, first served*
> *a friend in need is a friend indeed*
> *if in doubt, do nothing/nowt*
> *a leopard does not change its spots*
> *live and let live*
> *a penny saved is a penny earned*
> *one swallow does not make a summer* (sometimes as *does not a summer make*)
> *they that sow the wind shall reap the whirlwind*
> *you can take a horse to water but you cannot make him drink*
> *don't judge a book by its cover*

Some proverbs are literal, some figurative; however, the non-compositionality of proverbs lies particularly in their illocutionary and perlocutionary roles.

Canonically, proverbs take the form of complete utterances—simple-present statements presenting universal truths, imperatives, quasi-conditionals—though structures may be adapted to a grammatical context (cf. *you cannot have your cake and eat it* and 'he's trying to have his cake and eat it'). Formally, proverbs may pattern internally, for example with paired noun groups, clauses, repeated items, and alliteration, reinforcing their rhetorical effect. Many proverbs contain archaisms; others have changed over time, as with *don't spoil the ship for a ha'porth of tar*, where *ship* was previously *sheep*. Some proverbs have standard lexical variations (*many hands make light work/labour*, or antonymous *familiarity breeds content/contempt*). In context, proverbs may be truncated and used allusively or manipulated: (p. 131)

> *In fact every time she saw Flo she threw her out. Muttered something about **two's company**—there's no point in quoting the rest.*

> *Well, the compromise was found and it is a very traditional compromise. It's what you call **half a loaf**.*

> *The problem comes, of course, with that large area in which **one person's** experimental **meat is another person's** moral **poison**, where so much is a matter of personal opinion.*

Occasionally reduced forms of proverbs exist as independent MWIs, as with these:

> *call the tune; he who pays the piper calls the tune*
> *last/final straw; it's the (last) straw that breaks the camel's back*
> *make hay; make hay while the sun shines*
> *rolling stone, gather no moss; it's a rolling stone that gathers no moss*
> *silver lining; every cloud has a silver lining*

(The last also has a recurrent variation/reversal, *every silver lining has a cloud*.)

7.5.2 Phrasal verbs

Phrasal verbs are combinations of verbs and adverbial or prepositional particles (e.g. *in, on, up*), which are associated with meanings that cannot necessarily be derived or predicted from ordinary meanings of those component verbs and particles. Examples include:

> *break down*
> *chicken out*
> *get back*
> *give up*
> *go off*
> *let on*
> *make up*
> *put down*
> *set in*
> *take out*

Many combinations are highly polysemous and frequent. As an MWI subtype, they are particularly associated with English and northern Germanic languages such as Dutch, Danish, and Swedish: cf. the phenomenon of separable verbs in German and Old English. Amongst the most prolific contributing verbs in English are *come, get, give, go, hold, lay, put, run, set, take*, etc., while the most prolific particles are *down, in, off, on, out*, and especially *up* (of the 60 possible combinations of these items alone, no fewer than 58 are found as phrasal verbs). Phrasal verbs are sometimes marginalized as informal and most have single-word synonyms, often Latinate; however, phrasal verbs are often less marked as choices. Compare *give up, stick out, draw up* vs. more formal *desist, protrude,* *formulate.* See Bolinger (1971) for a classic study of phrasal verbs; Biber et al. (1999: 407 ff.) for distributional statistics and corpus-informed discussion.

One reason why English phrasal verbs attract attention is that they are particularly problematic for language learners, who not only need to recognize them as special combinations in text but understand how to manipulate them grammatically: whether they passivize, and where particles are placed in relation to grammatical objects. This requires identifying whether the particle is adverbial or prepositional. For transitive uses of adverb combinations, pronouns and shorter objects are typically placed between active verb and particle, but longer or foregrounded objects are placed after the particle: hence *they set up a new organization, they set it up, they set the organization up*. When passive, the adverb follows: *the organization was set up*. In active prepositional combinations, the prepositional object/complement must follow the particle: *they went over the problem, they referred (us) to the rules, they put thought into the matter*. Some combinations of apparently intransitive verbs and prepositions passivize, others do not: *the problem was looked at* but not normally **they were sided against*. By the same token, a few

combinations of apparently transitive verbs and adverbs do not passivize (*she was got down by her problems*).

Phrasal verbs are sometimes described as arbitrary combinations, and many are semantically non-compositional, not least where verbs are non-literal or semantically depleted: qualities of takeness, getness, giveness, may be difficult to explain. However, there are some distinct patterns in combinations, where particles are associated with recurrent uses or meanings, and this leads to a measure of productivity. For example, *up, out*, and *away* have completive functions, where combinations describe actions done fully, to the maximum extent or point of disappearance (*burn up, fold up, die out, stretch out, fade away*, etc.); *off* occurs in combinations describing separation or blocking (*cordon off, divide off, wall off*); *on* in combinations indicating perseverance or continuation (*go on, push on, soldier on*). *Up* is associated with intensification or improvement, as in *brighten up, speed up, whip up* and combinations such as *wise up, glam up, man up*; a completive use of *out*, to the point of exhaustion, occurs not just in *tire(d) out, wear/worn out* but in *shopped out, partied out*, even ad hoc *business-conferenced out*. While not all combinations are susceptible to rationalization, many fall neatly into semantic sets linked by some regularity of meaning, and this is particularly interesting for cognitive linguists investigating the conceptual metaphors inherent in notions of up and down, in and out (Lindner 1981), also for teachers of English as a foreign/second language—see Kövecses and Szábo (1996), Boers (2000), and Lindstromberg (2010)—and lexicographers producing phrasal verb dictionaries for learners of English.[9]

A number of phrasal verbs have derived adjectives and nouns, sometimes with inversion of the verb/particle sequence: for example, *rundown, throwaway, downcast, uplifting; leftover(s), takeup, outpouring, overspill*. Compare too single-word verbs such as *input, outgrow, overturn* which correspond with phrasal verb senses for *put in, grow out (of), turn over*.

(p. 133) 7.5.3 Binomials and trinomials

English binomials and trinomials are institutionalized pairs or triples of words, often linked by *and, or, but*, or *to* and occurring in a fixed order: for example,

> *back and forth*
> *bloody but unbowed*
> *come and go*
> *head to foot*
> *high and dry*
> *hit and/or miss*
> *lock, stock, and barrel*
> *loud and clear*
> *sink or swim*
> *whys and wherefores*

wrack/rack and ruin

Some are alliterative or assonantal; some fossilize (near-)obsolete words, which survive as unique items in these MWIs (*wrack* and *wherefore* above; also *to and fro, spic(k) and span*—phraseological fossils are discussed in section 7.6.2). While most pairs or triples are irreversible, there are some exceptions, such as *day and night, night and day*, also institutionalized exploitations, such as *poacher turned gamekeeper* and *gamekeeper turned poacher*. Most of the items listed above emphasize or indicate repetition or completeness, and there also seem to be other patterns underlying the frozen sequences, for example logicality (*born and bred, cut and dried*) or speaker orientation (*come and go, now and then*).[10]

7.5.4 Similes

These conventionalized comparisons also emphasize. Items with *like* include a range of grammatical types and degrees of compositionality, some of which can equally be classified as idioms:

> *know something like the back of one's hand*
> *like a bolt from the blue*
> *like clockwork*
> *get along like a house on fire*

(p. 134) Much more restricted syntagmatically is the *as*-simile frame, for example

> *as clean as a whistle*
> *as cool as a cucumber*
> *as dry as a bone*
> *as fit as a fiddle*
> *as hard as nails*
> *as quick as a flash*
> *as solid as a rock*
> *as white as a sheet*

Initial *as* is frequently omitted. All emphasize the adjective, which may be non-literal: *cool* above refers to behaviour, *white* to pallor caused by illness or shock. Some similes have single-word adjectival derivatives (*bone-dry, rock-solid*) and the frame itself is very productive: candidates for MWIs are simply those that recur. A variation on the frame, *about as ADJECTIVE as NOUN GROUP*, is especially creative and used ironically (cf. Sinclair 2003):

> *about as alluring as a wet sock*
> *about as reliable as an election promise*
> *about as useful as a chocolate parasol*

See Moon (2008) for discussion of *as*-similes.

7.6 Formulae and formulaic items

This section groups together recurrent idiomatic sequences of different kinds: some are compositional, some not, and many are associated with pragmatic functions. While from cultural or logophiliac perspectives, they can seem less interesting than idioms, proverbs, and the like, from the perspectives of psycholinguistics, applied linguistics, and general phraseology, they are not only interesting but crucially important in relation to the production of language.

7.6.1 Formulae

Many MWIs could be placed under this vague heading: what links them is their association with specific functions or meanings. One subgroup consists of ritual or conventionalized formulae used in greeting, well-wishing, thanking, responses, conversational gambits, text structuring, modalizing, and so on:

> *come to think of it*
> *excuse me*
> *for example/instance*
> *good luck*
> *good morning*
> (p. 135)
> *happy birthday*
> (*it's*) *needless to say*
> *no comment*
> *see you (later)*, (*I'll*) *be seeing you*
> *thank you*
> *well done*
> *you're welcome*

There are many other items with such strong pragmatic functions or force that they must be considered MWIs, even if compositional:

> *above all*
> *as a matter of fact*
> *at best, at worst*
> *for a start*
> *in effect*
> *in fact*
> *in other words*
> *no doubt*
> *not exactly*
> *talking of —*
> *to all intents and purposes*

to say the least

To these can be added conversational fillers such as *you know, you never know, I see, I mean, I guess*: perfectly understandable, yet so frequent that they may receive treatment as MWIs in reference and teaching materials for language learners. There is no fixed or stable set of such items, nor can there be, since their MWI-ness is a matter of interpretation and convention. Moreover, the process whereby phraseologies acquire pragmatic functions is a productive one; some items are long institutionalized, others are transient.

Other kinds of institutionalized formulae include catchphrases, slogans, and some quotations, for example *the best thing since sliced bread, you can't take it with you, I'll have what she's having, diamonds are a girl's best friend, because you're worth it, — doesn't get any better than this, tomorrow is another day*. These are sometimes labelled as sayings, but while 'saying' is widely used to refer to some kind of colloquialism, recurrent formula, or platitude, it is scarcely a definable linguistic category. In the following BoE examples, it is simply a metalinguistic comment on a choice of words:

> *He was, as the saying went, 'a good mixer.'*
> *But as the saying goes, 'You see what you want to see.'*
> *He is a great believer in the saying that too many cooks spoil the broth.*

7.6.2 Defective phraseologies

A subset of formulae consists of anomalous items, uninterpretable either because it is difficult to work out the syntactic relations between elements, or else because (p. 136) component items are too semantically depleted for holistic meanings to be retrieved (in any case, many are associated with pragmatic functions). There are many degrees and types of anomaly, and examples include:

> *at all*
> *by and by*
> *by and large*
> *hard done by*
> *how do you do?*
> *in case*
> *of course*
> *so long*
> *to do with*
> *to hand*

Some are partially decomposable (in the cases of *by and large* and *of course* only through reference to their historical origins). In the following, component words are used in atypical grammatical roles:

> *all of a sudden*
> *go for broke*
> *in the know*
> *on the make*
> *on the up and up*
> *the dos and don'ts*

Items like these result from diachronic processes, though are not productive in the usual sense: some are fossilized relics of more coherent structures, some contain non-standard grammatical uses (e.g. *no can do, come what may*). Another type of phraseological fossil is where MWIs contain words which are otherwise obsolete, thus making the sequence non-compositional: Makkai (1972) refers to such words as 'cranberry morphs'. Examples include:

> *by dint of*
> (*days*) *of yore*
> *in cahoots (with)*
> *kith and kin*
> *out of/off kilter*
> *put the kibosh (on)*[11]
> *run/go amok*
> *short shrift*

In a few cases, the unique item is more or less restricted to the MWI, though occasionally occurs in other, freer, structures, perhaps with allusion to the MWI: examples here (p. 137) are *dudgeon, fettle, grist, umbrage, wend* which are mainly associated with *in high dudgeon, in fine/good fettle, grist to one's mill, take umbrage, wend one's way*. In other cases, unique items are homonymic with words in ordinary distribution (*hue and cry, leave someone in the lurch, to boot*), while some MWIs include obsolete or peculiar senses of polysemous words (*beg the question, forlorn hope, toe the line*).

7.6.3 Phraseological frames

Following on from these is the phenomenon of productive phraseological frames, where a slot can be filled by any of a set of items and where there is some consistency of meaning. Several such frames consist of prepositions and nouns, as in

> *beyond + belief, description, doubt, question, recognition, repair*
>
> *in + a fix, hole, mess, spot*
>
> *on + the alert, boil, bubble, fly, hoof, hop, march, run*
>
> *under + consideration, discussion, examination, observation, scrutiny*

While some of these combinations are non-compositional, there are clearly conventionalized patterns of meaning and usage, not least in the metaphorical meanings of the prepositions (at the same time, some would consider these simple collocations). Frames can be compared to clusters of idiom-like phraseologies, such as these metonymic items formed with the verb *hit* (cf. discussion in Ruhl 1979; Moon 1998: 148):

> *hit the books/bottle/deck/road/sack* (etc.)

They seem related to semantically depleted uses of *take, make, give*, etc. in strings such as *take a look, make a comment, give a smile*, but that model of periphrastic construction is really too broad and general to be considered an MWI frame. Compare too the existence of what Fillmore et al. term 'formal idioms' (1988): that is, idiomatic syntactic frames, sometimes clausal or above the level of the clause, such as *the X-er, the Y-er* (e.g. *the more the merrier*) and *X, let alone Y*. See also Taylor's discussion of such frames (2012: 84ff.).

A few phraseological frames are particularly associated with creative uses, typically humorous: cf. the ironic *about* subtype of simile discussed in section 7.5.4. Some have recurrent realizations, but in many cases the point of using them is to be inventive and entertain, as with these:

> *a few beers short of a six-pack* (= inadequate, stupid)
> *a few ice floes short of a Titanic*
> *two sandwiches short of a picnic*
> *two thumb-screws short of a torture chamber*
>
> *busier than a three-legged horse in a dog-food factory* (= very busy)
> *busier than a one-armed juggler*
> *busier than a two-dollar hooker*
> _(p. 138) *not the sharpest pencil in the box* (= stupid)
> *not the sharpest tool in the shed*
> *not the sharpest knife in the drawer*
> *not the brightest crayon in the box*
> *not the brightest bulb in the chandelier*

and drawing from Veale (2012: 65f.):

> *Tungsten is the Cleopatra of the elements*
> *The Manhattan is the Cary Grant of cocktails*
> *Pac Man is the King Lear of the 1980's 8-bit videogame* revolution

See Veale (2012) for a book-length account of creativity and the patterns underlying creative formulations like these.

7.6.4 Prefabs

Last are 'prefabs', or prefabricated sequences: a useful label to encompass a range of idiomatic phraseological phenomena, not necessarily MWIs, and so scarcely fitting into the remit for this chapter: they lie at the very edge of MWI-ness. What characterizes them is that they recur and seem preferred ways of construction and/or articulation: they may be two-word sequences, complete sentences, or something in between (or longer). The following extract of unscripted dialogue from the BoE contains a variety of phraseological items:

> **A** and erm I was almost hoping the children would do something a little bit sort of <laughs> not bad

> **B** Yeah that's right.

> **A** but they were so good and everyone was saying such nice things about them I and I was always a bit of a goody two shoes and I did all the things that mums and dads want you to do you know went to university all this stayed at home with my < ... > children I've always done the right things and erm

> **B** Paragons of virtue.

> **A** you know all these children < ... > the < ... > the children I mean cos they're obviously < ... > n+ < ... > not better than not brilliant at home but I mean they did pull the stops out last weekend I'm happy to say.

> **B** Oh that's good.

At the idiom end of the spectrum are *goody two shoes* and *pull (all) the stops out*; binomial-like *mums and dads* is over 60 times as frequent as its reversal; *paragons of virtue* is a restricted or recurrent collocation which accounts for 10 per cent of BoE occurrences of *paragon*. Further units include *a little bit, a bit of, at home*, and speech formulae/fillers *you know, I mean, sort of*. But there is evidence here of plenty more chunks of language—not just pragmatically oriented *that's right/good* and *I'm happy to* (p. 139) say, but prefabricated sequences such as *saying (such) nice things about, went to university, stay at home (with my children), done the right things*; also *I was (almost) hoping* and *I was always a*, which have narrative-building functions. While there are different ways of considering such strings descriptively and theoretically, they are further evidence of the phraseological nature of lexis and point towards phraseological models for the lexicon rather than atomistic ones.

7.7 Cross-linguistic aspects

This chapter has looked only at English MWIs, their structures and types. There are closely parallel phenomena in many other languages, though equally some languages differ markedly from English, for example in the morphology of non-compositional units

(Chinese is a case in point), or in the anisomorphism of items which may be multi-word sequences in one language but simplex words in another. This presents problems for language teaching and learning, for the creation of bilingual dictionaries, and for translation: also, of course, for language engineering.

But cross-linguistic comparisons foreground many interesting features. Some proverbs and other items recur in almost identical forms in different languages, showing cross-cultural influences or common cultural origins. The first of the following can be traced back to medieval Latin, the second to Greek (which has 'spring', like the French equivalent, rather than 'summer'):

> *all roads lead to Rome*
> French: *tous les chemins mènent à Rome*
> German: *viele Wege führen nach Rom*

> *one swallow does not make a summer*
> French: *une hirondelle ne fait pas le printemps*
> German: *eine Schwalbe macht noch keinen Sommer*

Proverbs with loose parallels in non-European languages/cultures reflect the universality of their messages:

> *once bitten, twice shy*
> Chinese: *'once bitten by a snake, he is scared all his life at the sight of a rope'*
> *birds of a feather flock together*
> Malay: *'hornbills are like hornbills, sparrows like sparrows'*
> *don't teach your grandmother to suck eggs*
> Congolese: *'You do not teach the paths of the forest to an old gorilla'*

There are also individual conceptual metaphors which recur in MWIs in different languages. For example, Kövecses (2002: 163ff. and 95ff.) discusses metaphors for anger in Hungarian, Japanese, Chinese, Polish, and isiZulu: these parallel English MWIs where (p. 140) anger is conceptualized in terms of pressure in a container (*blow one's top/stack, blow a fuse/gasket, blow up, bottle up (one's anger), go through the roof, hit the roof/ ceiling, be steamed up*), and heat (see examples in section 7.4). In other cases, idioms etc. have similar meanings or even structures in different languages, but are realized through different metaphors:

English: *kill two birds with one stone*

Dutch: *twee vliegen in één klap slaan* ('kill two flies with one blow')

French: *faire d'une pierre deux coups* ('make two shots with one stone')

Italian: *prendere due piccioni con una fava* ('catch two pigeons with one bean')

Portuguese: *matar dois coelhos de uma cajadada só* ('kill two rabbits with a single stick')

See Piirainen (2012) for an extensive cross-linguistic study of idioms, including 74 European languages and 17 non-European.

Notes:

([1]) Including by myself, as in Moon (1998).

([2]) To give some idea of idiom etc. (in)frequencies: according to evidence in the Bank of English corpus (see footnote 8), *bury the hatchet* and *spill the beans* occur roughly 5–6 times per 10 million words; *kick the bucket, kith and kin, rain cats and dogs* occur roughly once per 10 million words; *bag and baggage, cupboard love, one man's meat is another man's poison* occur less than once per 20 million words.

([3]) There are several theories about the historical meaning of *rain cats and dogs*, including the drowning of small animals in floods following torrential rain, and a possible derivation from Greek *katádoupoi* 'cataracts'.

([4]) See Moon (2008) for discussion of these approaches in relation to *as*-similes.

([5]) Different dictionaries adopt different policies for their listing of multi-word items, particularly idioms and other phrases; some list them under the first lexical element, some under a component noun.

([6]) I exclude logophiliac dictionaries of lexical curiosities, which are driven by a very different set of non-linguistic principles.

([7]) 'Idiom' approximates here with idiomaticity.

([8]) The Bank of English corpus was created by COBUILD at the University of Birmingham. All extended/contextualized examples are drawn from this corpus.

([9]) For example, dictionaries of phrasal verbs edited by Sinclair and Moon (1989) for Collins Cobuild, and by Rundell (2005) for Macmillan, both draw attention to regularities in particle meaning.

([10]) Cf. Lakoff and Johnson's discussion of a 'me-first' orientation in such sequences (1980: 132f), reporting Cooper and Ross's (1975) observations. Other factors in sequencing include the relative length of words (shorter is often first) and as John Taylor points out (personal communication), the relative frequencies of words (the more frequent word tends to come first: Fenk-Oczlon 1989) and the sequencing of the stressed vowels (vowels with a lower second formant tend to follow vowels with a higher second formant: Oakeshott-Taylor 1984).

([11]) *Kibosh* is occasionally found as an independent verb (BoE: 'The net effect is invariably to kibosh the Conservative Party').

Multi-word Items

Rosamund Moon

Rosamund Moon is a Senior Lecturer in the Department of English at the University of Birmingham. She was previously a lexicographer, working on the Cobuild Dictionary Project at Birmingham (1981-1990, 1993-1999), and also at Oxford University Press (1979-1981, 1990-1993). Her main research areas are lexis and phraseology, lexicography, figurative language, and corpus linguistics; her publications include Fixed Expressions and Idioms in English: a Corpus-based Approach (1998: Oxford University Press), and (with Murray Knowles) Introducing Metaphor (2006: Routledge).