

TREATMENT OF MULTI-WORD LEXICAL ITEMS IN THE DICTIONARY: THE CURRENT SITUATION AND THE POTENTIAL PROBLEMS FACING DICTIONARY USERS

Alenka Vrbinc, Marjeta Vrbinc

Abstract. In this article, we first discuss the inclusion of idioms in the latest printed editions of monolingual learner's dictionaries; we then consider and evaluate the success of look-up operations in connection with the recognition of the correct part of speech. The central part of the article discusses and justifies the macro- and microstructural treatment of multi-word lexical items. Finally, we address the issue of including idioms in the microstructure as special senses of the lemma, as usage examples and/or in the idioms section. The inclusion and treatment of multi-word lexical items depends on the needs of potential dictionary users. Lexicographers should therefore strive to adopt a pragmatic approach that would help the user to locate a multi-word lexical item with minimal effort. When compiling a bilingual dictionary, contrastive differences between source and target languages should also be taken into consideration.

Keywords: monolingual dictionaries, bilingual dictionaries, idiomatic expressions, macrostructure, microstructure, dictionary users, English, Slovene

1. Introduction

Before starting work on any mono- or bilingual dictionary, lexicographers must reach a number of macrostructural decisions (e.g., number of entries, selection of entries) and formulate a coherent policy on the handling of all the elements that constitute a dictionary entry.

One of the 'classic' problems of dictionary structure is the position of multi-word lexical items and the strategies employed by users to find them in the dictionary (cf.

Svensén 2009: 470). All dictionaries must make decisions about the placement of multi-word lexical items if they are to be included within the entry of one of their component words. There is no right and wrong about how to present multi-word lexical items in a dictionary, and dictionaries do differ in this respect. If we study general dictionaries written in the Anglo-American tradition, we find that they tend to be atheoretical, which means that they rarely categorize multi-word lexical items when listing them. Multi-word lexical items are often treated together under a label such as ‘idioms’ or ‘phrases’ (cf. Atkins, Rundell 2008: 253–255, Moon 1998: 17–18). Different approaches employed in different dictionaries give rise to the problem of accessibility, and a lexicographer can never be sure how the user will approach this challenge.

Phraseological material consists of units that differ in structure as well as semantics. They can be placed on a continuum or scale ranging from transparent, freely recombinable collocations at one end, to formally invariable, unmotivated idioms at the other (cf. Cowie 1994: 3168–3169). If we take a look at the existing monolingual dictionaries, we see that idioms are included both as lemmas as well as subheadwords (in the specific section of idioms), the latter of which is the traditional, prevailing method. Since we know that an idiom is simplex as far as meaning is concerned and complex as far as structure is concerned, the micro-structural treatment of idioms is contrary to the idiom’s nature as an independent lexical item (cf. Svensén 2009: 194). Presenting an idiom as a lemma is more in accordance with the idiom’s classification as an independent lexical item having an opaque meaning. Here, however, the lexicographer is faced with other difficulties: e.g., one must be able to draw a clear dividing line between idioms and other word combinations, or one must decide which of the idiom’s component words is to form the basis of positioning in the alphabetical sequence of the lemma list. The same problems are also encountered when including idioms in the microstructure of a dictionary.

In this article, only printed editions of monolingual learner’s dictionaries have been taken into consideration, leaving aside electronic versions, where there should be fewer or no problems concerning the retrieval of multi-word lexical items. In the case of bilingual English–Slovene dictionaries, electronic versions cannot be researched because CD-ROMs do not exist at all. Recently, a test version of the most comprehensive English-Slovene dictionary (*Veliki angleško-slovenski slovar Oxford*) has been launched online (<http://www.e-slovarji.si>, but access is by password only, not free of charge), but it still suffers from “teething problems” and consequently, it cannot be the subject of any serious research.

The aim of this article is to present how idioms are treated in English monolingual dictionaries and to propose potential ways of including them in bilingual decoding dictionaries, more precisely yet, in an English–Slovene dictionary. Idioms are either hard to recognize or difficult to understand; dictionary users need help with understanding idioms, then translating them from L2 to L1. We will concentrate mostly on English monolingual learner’s dictionaries, because, aside from bilingual dictionaries, this is the type of dictionary foreign learners most often consult. Bilingual encoding dictionaries have not been taken into consideration, since in production, idioms can be regarded as ‘luxury articles’; i.e. it remains

possible to produce perfectly satisfactory speech without using them (cf. Hausmann 2004: 312).

Some multi-word lexical items that used to be classified as idioms and that were consequently included and treated as such in previous editions of monolingual dictionaries have been included in the headword list in recent editions of the same dictionaries, thus acquiring the status of compounds (e.g. *red herring*, *off colour*, *off chance*, *sitting duck* and *lame duck*). Such items are also included in this article.

2. The ordering of idioms in the latest editions of English monolingual learner's dictionaries

When trying to find idioms, especially long idioms, users are often faced with the problem of locating them: users are puzzled about the right entry under which to find the idiom in question. Again, since monolingual dictionaries differ in their policies, we should first touch upon the principles observed in existing dictionaries.

In OALD8, LDOCE5, CALD3 and MED2, the idioms are included and defined in the entry for the first lexical (or content) word (a noun, a verb, an adjective, or an adverb) of the idiom. This means that any grammatical words, such as articles and prepositions, are ignored. On the other hand, if one content word of an idiom can be replaced by another, as in *be a bag of nerves*, where *bag* can be replaced by *bundle*, this is shown by the use of a slash *be a bag/bundle of nerves* and the idiom is defined under the first full fixed word, i.e. *nerve*. In some idioms, many alternatives are possible. In the expression *disappear into thin air*, it is possible to replace *disappear* with *vanish*, *melt* or *evaporate*. In the dictionary, this is shown as *disappear, vanish, etc. into thin air*, indicating that we can use other words with a meaning similar to *disappear* in the idiom. Since the first content word of the idiom is not fixed, the expression is defined under *thin* with a cross-reference only under *air*. However, the use of 'etc.' is problematic for a non-native speaker, especially when encoding; simply including 'etc.' is not sufficient to indicate the whole range of possibilities. There are few very common verbs, especially delexical verbs, and adjectives, such as *bad* and *good*, that have too many idioms to all be listed in the entry. Instead, there is a note in some dictionaries (e.g. in OALD8) telling the user to look at the entry for the next noun, verb or adjective in the idiom (e.g. *take the biscuit* is included under *biscuit* rather than under *take*). What about phraseological units that contain only grammatical words, such as *one, it, or in*? These idioms are defined under the first word of the idiom, e.g. *one up on sb* is defined under the entry for *one*.

In COBUILD5, however, idioms are included and treated under the first noun they contain, which was the traditional approach of including idioms. If the idiom being sought contains one or more nouns, it will be found under the first noun. This means that, for example, *cost the earth* is found under *earth*. If there is no noun in the idiom, the idiom is listed under the first significant word, e.g. *thick and fast* is listed under *thick*.

A feature unique to CALD3 is the Idiom Finder, which is located at the end of the dictionary (pp. EH59–EH98). The Idiom Finder lists all the long idioms included in this dictionary under each lexical word in the idiom. This means that if a user looks

in the Idiom Finder for *be like a bear with a sore head*, he/she will find it listed under *like*, *bear*, *sore* and *head*, with the page number in the dictionary where he/she can find out what the idiom means. The word that the user is seeking is always printed in colour (blue) to make it easier to find. The 'Idiom Finder' contains only those idioms that have three or more lexical words. We welcome this feature, since it is of great help when trying to locate an idiomatic expression.

3. The recognition of the correct part of speech

Another difficulty in a successful look-up operation is recognizing parts of speech for words that are component elements of idioms. The problem is intensified by the fact that many of these are ungrammatical. Furthermore, the meaning of many phraseological units (especially idioms) cannot be guessed from the sum of meanings of individual elements, which makes the identification of parts of speech quite problematic. An interesting system of including idioms occurs in the *Oxford Dictionary of English* (a dictionary intended for native speakers of English). The lexicographers did not make a distinction between parts of speech of individual words within an idiom, but rather included all idioms after the last entry. For example, the word *hand* has two entries (noun, verb), but the idioms follow the second of the two entries, i.e. the verb, disregarding the part of speech of the word *hand* in individual idioms. This is by far the most user-friendly system of including idioms in dictionaries, because no demands whatsoever are imposed on the user regarding the recognition of parts of speech.

As is evident from a comparison of different monolingual English dictionaries, it can be established that determination of the part of speech of a particular component element of an idiom may pose problems to lexicographers; the same idiom may be included in entries for different parts of speech. For example, *look sharp* is listed in the entry for the adjective *sharp* in OALD8 and in the entry for the adverb *sharp* in LDOCE5. In the English–Slovene dictionary, it should be included under the adjective, since the verb *look* is a linking verb which is followed by *sharp* (= 'quick') functioning as a subject complement realized by an adjective:

sharp *prid*. ... **IDIOMI look sharp** 1. BRIT. pohiteti 2. AM. paziti ...

Figure 1. *Look sharp* in a bilingual English–Slovene dictionary

Similarly, to determine the right part of speech for numbers, one has to differentiate between a numeral and a noun, since in some idioms the headword is a numeral, whereas in others it is a noun (either in the singular or in the plural form). In the entries for numbers, monolingual learner's dictionaries often include just one entry and label it 'number' (e.g., the entries for *four* and *six* in LDOCE5). Sometimes, they include one entry and label some senses as a number and some as a noun, but make no distinction by parts of speech within idioms, thus listing all idioms in one entry. For example:

four ... **1 number** 4 ... **2 noun** a group of four people or things:
to make up a four at tennis ∅ a coach and four (= four horses)
3 noun (in CRICKET) a shot that scores four RUNS **4 noun** a team
of four people who ROW a long narrow boat in races; the boat
that they row **IDIOMS on all 'fours** (of a person) bent over with
hands and knees on the ground: *We were crawling around on all
fours.* **these four 'walls** used when you are talking about keep-
ing sth secret: *Don't let this go further than these four walls* (=
Don't tell anyone else who is not in the room now).
(OALD8)

Figure 2. All idioms listed in one section regardless of the part of speech of *four*

In some cases, the dictionary includes two entries (one for a numeral and one for a noun) but lists all the idioms under the numeral, although some of them should be listed under the noun. For example:

six ...

- ▶ **number** the number 6: *Look for a bus with a number six on the front of it.* • *"How many grandchildren do you have now?" "I've got six (grandchildren)."*
- ▶ **idioms at sixes and sevens** INFORMAL in a confused, badly organized or difficult situation: *We've been at sixes and sevens in the office this week.* • **be six feet under** HUMOROUS to be dead and buried: *There's no point worrying about it - we'll both be six feet under by then.* • **six of one and half a dozen of the other** INFORMAL SAYING said when you think that neither of two choices is better than the other: *"Shall we go by car or train?" "I don't know, it's six of one and half a dozen of the other."* • **six of the best** UK OLD-FASHIONED a beating, usually of six hits with a stick
- ▶ **noun** [C] in cricket, six points scored when the player hits the ball to the edge of the playing area without it touching the ground first: *Richards hit a six.*
(CALD3)

Figure 3. The treatment of idioms under one part of speech only

Such inconsistent treatment should be avoided, and the policy adopted by a particular dictionary should be pursued in all its entries. If a dictionary includes idioms under different parts of speech depending on the part of speech of the constituent elements of the idioms, there should be no deviations from or exceptions to this rule. Consequently, a recommended strategy for inclusion of idioms in any dictionary, be it monolingual or bilingual, would be to stick to the rule of part-of-speech differentiation (if this is the established policy). This can be illustrated by a sample entry from an English–Slovene dictionary, where *four* is treated as a numeral in the first entry and as a noun in the second.

four¹ *štev.* štiri, štirje **IDIOMI be scattered to the (four) winds** ⇒ SCATTER¹; **from the four corners of the earth/world** iz vseh koncev sveta; **these four walls** na štiri oči: *Don't let this go further than these four walls.* To naj ostane med nami.

four² *sam.* 1. štirica, štirka 2. četverica, četvorica 3. ŠPORT četverec 4. četveroprega 5. ŠPORT (kriket) met, ki šteje štiri teke **IDIOMI the big three/four** ⇒ BIG¹; **on all fours** po vseh štirih; **four on the floor** AM. (avto) štiri prestave

Figure 4. The treatment of idioms under appropriate parts of speech

If we want to adopt an even more user-friendly approach to the inclusion of idioms, we can follow the policy developed in the *Oxford Dictionary of English* and disregard the part of speech, thus treating all idioms together at the end of the final entry for a particular word.

four ...

- ▶ *štev.* štiri, štirje
- ▶ *sam.* 1. štirica, štirka 2. četverica, četvorica 3. ŠPORT četverec 4. četveroprega 5. ŠPORT (kriket) met, ki šteje štiri teke

IDIOMI be scattered to the (four) winds ⇒ SCATTER¹; **the big three/four** ⇒ BIG¹; **four on the floor** AM. (avto) štiri prestave; **from the four corners of the earth/world** iz vseh koncev sveta; **on all fours** po vseh štirih; **these four walls** na štiri oči: *Don't let this go further than these four walls.* To naj ostane med nami.

Figure 5. The treatment of idioms at the end of the final entry for a particular word

4. Multi-word lexical items in the macro- and microstructure

If we compare the inclusion of multi-word lexical items in different monolingual (English) dictionaries, we can observe that one and the same multi-word lexical item is listed in a special section in some dictionaries, and as a headword in others. If it is included as a lemma, the multi-word lexical item seems to have the status of a compound. Let us enumerate just a few examples of multi-word lexical items included as separate entries in at least some monolingual dictionaries: *big cheese*, *bread and butter*, *coffee-table book*, *upper crust*, *couch potato*, *salad days*, *knuckle sandwich* and *sugar daddy*. Interestingly, *white-bread* and *toffee-nosed* are included as keywords and as lemmas in the *Cambridge Idioms Dictionary*. Since they consist of two words and are hyphenated, they match the definition of a compound and are treated as such in monolingual dictionaries; this means that they are listed as separate headwords and not as microstructural elements, i.e. in the idioms section. *White-bread* and *toffee-nosed* are hyphenated compounds and as such do not merit inclusion in a phraseological dictionary.

Idioms such as *black and white*, *act of God* or *walk of life* can be included in a mono- as well as in a bilingual dictionary as headwords, but it is next to impossible to predict whether the users will look them up as a headword or simply look up one of the constituent elements of such a lexical item. This depends mainly on the user's ability to recognize a phraseological unit consisting of two or more words as a unit. It is therefore recommended that this problem be approached in a more user-friendly way, i.e. to include such items in two ways: as headwords and as idioms in the idioms section of the entries for all constituent elements, as in this example (of an English–Slovene dictionary):

black and white *prid.* 1. črno-bel 2. jasen, očiten **IDIOMI (in) black and white** črno-belo; **in black and white** črno na belem
black *prid.* ... **IDIOMI ... black and white** ⇨BLACK AND WHITE;
(in) black and white ⇨BLACK AND WHITE; **in black and white** ⇨BLACK AND WHITE
white *prid.* ... **IDIOMI black and white** ⇨BLACK AND WHITE; **(in) black and white** ⇨BLACK AND WHITE; **in black and white** ⇨BLACK AND WHITE

Figure 6. *Black and white* included as a lemma and as (parts of) idioms

Including phraseological units simultaneously as headwords and idioms in the idioms section with a cross reference is one possibility. However, there are cases where a phraseological unit can either be given full headword status, treated in the entry for one of its constituent elements separately or appear in the idioms section. For example, the multi-word lexical item *off day* can be included in an English–Slovene dictionary in the following ways:

- (a) as a headword:

off day *sam.* POG. slab dan

Figure 7. *Off day* as a headword

- (b) one of the senses of the adjective *off* is 'below the usual standard or rate' and 'dan, teden' (= day, week) can only function as an element of equivalent differentiation in the form of a collocator (sense 4 in the example below):

off *prid.* 1. (hrana) pokvarjen: *go off* pokvariti se 2. BRIT., POG. nevljuden, neprijazen, nesramen 3. BRIT., POG. nesprejemljiv 4. (dan, teden) slab 5. (sezona) mrtev

Figure 8. *Off day* as a sense of *off*

- (c) in the idioms section, although it cannot be classified as an idiom according to the phraseological criteria:

day *sam.* ... **IDIOMI ... off day** POG. slab dan ...

Figure 9. *Off day* as an idiom in the idioms section

Another problem deserving of mention is the inclusion of hyphenated and non-hyphenated items in mono- as well as bilingual dictionaries, since these are sometimes given full headword status (in this case, they are regarded as compounds), but at other times included as idioms in the idioms section. Hyphenated compounds, in particular, are treated as lemmas in most dictionaries, but non-hyphenated compounds can also be found as lemmas. It should be stressed that many such inconsistencies can be observed in one and the same dictionary. It is true that it has always been difficult to decide whether a compound should be entered in open, closed, or hyphenated form; therefore, the criterion of frequency comes in handy. Among the many advantages of using a corpus in lexicography, frequency counts may be the most important (cf. Landau 2001: 302–303). The relative frequency of variants in the spelling of a word can lead a lexicographer to a decision about what to regard as a lemma or preferred spelling. To illustrate how this criterion works in practice, we have taken compounds starting with *off* that are included as lemmas in OALD8. This dictionary includes compounds with *off* in the following ways:

(a) as hyphenated compounds:

off-air (adj), *off-Broadway* (adj), *off-centre* (adv, adj), *off-duty* (adj), *off-key* (adj), *off-licence* (n), *off-limits* (adj), *off-peak* (adj), *off-piste* (adj), *off-putting* (adj), *off-ramp* (n), *off-road* (adj), *off-roader* (n), *off-sale* (n), *off-screen* (adj), *off-street* (adj), *off-tackle* (adj), *off-the-cuff* (PoS not given), *off-the-grid* (adj), *off-the-shelf* (adj), *off-white* (adj);

(b) as non-hyphenated compounds:

off chance (n), *off colour* (adj), *off day* (n), *off season* (n), *off year* (n);

(c) as compounds spelt solid:

offbeat (adj), *offcut* (n), *offhand* (adj, adv), *offline* (adj), *offload* (v), *offprint* (n), *offset* (v, adj), *offshoot* (n), *offshore* (adj), *offshoring* (n), *offside* (adj, n), *offsider* (n), *offspring* (n), *offstage* (adj).

From the user's point of view, hyphenated compounds (item (a) above) and compounds spelt solid (item (c) above) are logically included as headwords; consequently, users will have no difficulty locating them. However, non-hyphenated compounds (item (b) above) may pose more problems, since it is likely that users will not recognize the two words as forming a unit. The resulting look-up operation will be problematic, as it can be presupposed that users will try to find them under the second element, i.e. under their nominal component. If such items are included as headwords, it is suggested that they also appear as idioms in the idioms section (sometimes with a cross reference to the place in the dictionary where they are included as headwords).

Another problem concerning the spelling and consequently also the manner of inclusion in a dictionary involves some idiomatic expressions that are hyphenated when used attributively and not hyphenated when used in other syntactic functions. In such cases, to ensure a successful look-up operation, the hyphenated item should appear in the macrostructure as an entry, whereas the non-hyphenated item should be included in the idioms section. The same policy should be adopted when compiling mono- and bilingual dictionaries. For example (in an English–Slovene dictionary):

cuff *sam.* ... **IDIOMI off the cuff** *iz rokava*
off-the-cuff *prid.* *iz rokava*

Figure 10. *Off the cuff* vs. *off-the-cuff* in the micro- and macrostructure, respectively

5. Idioms in the microstructure

The majority of existing mono- and bilingual dictionaries include idioms micro-structurally, i.e. within a special idioms section at the end of the entry. This section is variously called ‘idioms’ (OALD8, CALD3) or ‘phrase(s)’ (MED2).

In other dictionaries, phraseological units are listed together with the other senses of the word in frequency order (LDOCE5). This principle of arrangement of phraseological units has its advantages and disadvantages. Among the advantages is the inclusion of a commonly used phraseological unit at the beginning or at least in the first part of a dictionary entry. This is a good principle given the fact that an average dictionary user does not read the dictionary entry from beginning to end but rather picks things from the beginning. It is certainly more likely that dictionary users will come across a more frequent lexical item than a less frequent one. If, on the other hand, a dictionary user is faced with a less frequent phraseological unit, this principle is considerably less user-friendly. This non-alphabetical arrangement is especially problematic in long entries, because there is no single criterion for performing a look-up operation. In this case, the notion of frequency is vague and it is based only on data obtained from a corpus of texts that are not available to the general dictionary user.

The third method for including phraseological units is used in COBUILD5, where one idiom in the entry is shown separately with its own definition, whereas if there are two or more idioms within one entry, they are listed in the section entitled ‘phrase(s)’. This method produces inconsistency in the treatment of idioms, and dictionary users may be puzzled as to why they find an idioms section in some entries but not in others. In the second edition of CALD, idioms are included separately with their own definitions, but here a dictionary user should know that depending on the meaning, one word can appear as a lemma more than once. Consequently, phraseological units are included under several lemmas and are also grouped according to semantic similarities.

The above methods of including idioms can be designated as the direct method of treating idioms. However, some idiomatic expressions are dealt with in a much more indirect way, for instance as examples illustrating the use of the lemma. In OALD8, for example, *the cream of the crop* is used to contextualize the noun *cream* in sense 5:

cream ...
noun ... **5 the ~ of sth** the best people or things in a particular group: *the cream of New York society* ◊ **the cream of the crop** of *this season's movies* ...
(OALD8)

Figure 11. Idiom included as a usage example

The same treatment of this idiom can be observed in LDOCE5. In MED2, the idioms section contains *the cream of* and *the cream of the crop* as two separate idioms, whereas in CALD3, only *the cream of sth* is included as an idiom and *the cream of the crop* appears in a usage example to illustrate the use of the idiom *the cream of sth*. If the status of one and the same idiom differs to such a great extent in monolingual dictionaries, this prompts the question of how to treat it in a bilingual decoding dictionary. Inclusion and treatment largely depend on the contrastive differences observed in a specific language pair. *The cream of* can be followed by various nouns that have one semantic component in common, i.e. ‘the best people or things from a group’. If *the cream of* is translated with one and the same translational equivalent into the target language, regardless of the noun that follows the preposition *of*, *the cream of* can be treated as one of the senses of the noun *cream*. The word combination *the cream of the crop* should be included as an idiom because its meaning is unpredictable from the point of view of a non-native speaker.

cream¹ sam. ... **4. the cream of** (pomembni ljudje) smetana česa
IDIOM the cream of the crop smetana

Figure 12. Possible treatment of *the cream of* and *the cream of the crop* in an English–Slovene dictionary

On the other hand, if the translation differs depending on the noun in the *of*-phrase, *the cream of* can be included either as one of the senses of the noun *cream* or as an idiom. In both cases, illustrative examples of usage along with appropriate translations are a necessity. The following example is taken from *Muret-Sanders Langenscheidt, Großwörterbuch Englisch-Deutsch*, where *the cream of* is treated as one sense of the noun *cream*, with usage examples translated differently into German (the idiom *the cream of the crop* is omitted from this dictionary):

cream ... 1 s ... **6. fig.** Creme *f*, Auslese *f*, Elite *f*: **the cream of society**
 die Creme der Gesellschaft; **the cream of the cream** crème de la
 crème; **the cream of the players** die besten Spieler ...
 (Langenscheidt)

Figure 13. Different translations of examples

In one and the same monolingual dictionary, the same phraseological unit may be listed in the idioms section in one entry and as a usage example in another entry. For example, the phraseological unit *cut/pare sth to the bone* appears in OALD8 in the idioms section at the end of the entry for the noun *bone*:

bone ...
 noun ...
IDIOMS ... **cut, pare, etc. sth to the 'bone** to reduce sth, such
 as costs, as much as you possibly can ...
 (OALD8)

Figure 14. Idiom within the idioms section

However, it appears just as a usage example in the entry for the verb *pare*, sense 2, even though its meaning is explained by means of a gloss in brackets, i.e. ‘reduced to the lowest possible level’.

pare ... verb... **2 pare sth (back/down)** to gradually reduce the size or amount: *The training budget has been pared back to a minimum.* ◊ *The workforce has been **pared to the bone** (= reduced to the lowest possible level).* ...
(OALD8)

Figure 15. Idiom as a usage example

This treatment of the idiom *cut/pare, etc. sth to the bone*, as can be observed in the entry for the verb *pare*, is not to be recommended for two reasons: (a) if an idiom is included in the idioms section in one entry in the dictionary, one would expect to find it listed in the idioms section in the entries for all the constituent elements with cross-references to the entry where it is treated; (b) if the meaning is not transparent, which is evident from the gloss provided in brackets in the above example (see Figure 15), its inclusion as a usage example does not seem sensible. In a bilingual context, the opacity of the meaning of *cut/pare, etc. sth to the bone* clearly justifies its inclusion as an idiom.

Interestingly, OALD8 includes *go to sleep* in the entry for the noun *sleep* as an idiom in the idioms section, but the same string of words as a free combination illustrates the use of the noun *sleep* in sense 1.

sleep ...
noun 1 [U] the natural state of rest in which your eyes are closed, your body is not active, and your mind is not conscious: *I need to **get some sleep**.* ◊ *I didn't **get much sleep** last night.* ◊ *Can you give me something to help me **get to sleep** (= start sleeping)?* ◊ ***Go to sleep**—it's late.* ◊ *He cried out **in his sleep**.* ◊ *Anxiety can be caused by **lack of sleep**.* ◊ *His talk nearly **sent me to sleep** (= it was boring).* ◊ *Try to go **back to sleep**.* ... **IDIOMS** ... ***go to 'sleep** (informal)* if part of your body **goes to sleep**, you lose the sense of feeling in it, usually because it has been in the same position for too long ...
(OALD8)

Figure 16. *Go to sleep* as a usage example and as an idiom

This is not the only place in OALD8, where *go to sleep* can be found. Surprisingly, we come across the same word combination in the entry for the verb *go*, sense 15.

go ... verb, noun
...
► **STATE/CONDITION 15** [I] used in many expressions to show that sb/sth has reached a particular state/is no longer in a particular state: ***go to/into sth** She went to sleep.* ◊ ***go out of sth** That colour has gone out of fashion.* ...
(OALD8)

Figure 17. *Go to sleep* as a usage example in the entry for the verb *go*

LDOCE5, on the other hand, lists *go to sleep* as a polysemous idiom. From the users' point of view, this way of including it is much more user-friendly, because it requires no linguistic knowledge from the users, nor do they need to make a distinction between an idiom and a free combination (see Figure 18). Apart from this, LDOCE5 also lists *go to sleep* as an example illustrating the use of the phrasal verb *go to sb/sth* (see Figure 19).

sleep² *n* ... **3 go to sleep** **a**) to start sleeping: *I went to sleep at 9 o'clock and woke up at 6.* **b**) *informal* if a part of your body goes to sleep, you cannot feel it for a short time because it has not been getting enough blood ...
(LDOCE5)

Figure 18. *Go to sleep* as a polysemous idiom

go¹ ... *v* ...
go to sb/sth *phr v* [not in passive] **1** to begin to experience or do something, or begin to be in a particular state: *I lay down and went to sleep.* | *Britain and Germany went to war in 1939.* ...
(LDOCE5)

Figure 19. *Go to sleep* as a usage example in the entry for the phrasal verb *go to sb/sth*

What method of including *go to sleep* should then be employed in a bilingual dictionary? As far as the English–Slovene language pair is concerned, it is suggested that *go to sleep* be included in two entries, i.e. the entry for the noun *sleep* as an idiom and the entry for the verb *go* in the phrasal verbs section under the phrasal verb *go to*. This would mean that in this case the treatment is the same as in LDOCE5. When *go to sleep* is included as an idiom, it is sensible to treat it as a polysemous word combination, although the translational equivalent in Slovene is the same (*zaspati*). In order to differentiate between both senses of the verb *zaspati*, equivalent differentiation by means of *roka*, *noga* (= ‘arm, leg’) should be used in sense 2.

sleep *sam.* ... **IDIOMI** ... **go to sleep** **1.** *zaspati* **2.** *POG. (roka, noga) zaspati* ...

Figure 20. *Go to sleep* as a polysemous idiom in a bilingual English–Slovene dictionary

What justifies the inclusion of *go to sleep* as an example illustrating the use of the phrasal verb *go to* is the fact that *go to sth* is translated into Slovene by *začeti kaj* (= ‘begin sth’), whereas *go to sleep* is translated by the perfective *zaspati* (= ‘fall asleep’), and *zaspati* semantically corresponds to *začeti spati* (= ‘start sleeping’); however, *začeti spati* is a paraphrase rather than a translational equivalent proper.

go gl. ...
SESTAVLJENI GLAGOLI ...
go to ... 2. go to sth začeti kaj: *go to sleep* zaspati • *go to war*
 stopiti v vojno • *My brain went to work on what I should say.*
 Začel sem premišljevati, kaj naj rečem. • *She hates to see food*
going to waste. Ne prenese, da se hrano meče stran. ...

Figure 21. *Go to sleep* as an example illustrating the use of the phrasal verb *go to* in a bilingual English–Slovene dictionary

6. Conclusion

When lexicographers make decisions on how to handle multi-word lexical items, it is advisable to study their treatment in other dictionaries, to evaluate the user profile and then choose the method that best fits the needs of potential users of the dictionary that is being compiled (Atkins, Rundell 2008: 255). After deciding on a policy for including and treating multi-word lexical items, this policy should be pursued in as consistent a manner as possible. At the same time, lexicographers should strive to adopt a pragmatic approach that would help the user to locate a multi-word lexical item with minimal effort. We should bear in mind that users may already experience difficulties identifying such items in texts, and if they fail to identify them in a text, they cannot successfully look them up in a dictionary. Therefore it is recommended that one and the same multi-word lexical item be included in two or even more places (e.g. as a lemma, as an idiom or as a usage example), so that the user can find it as quickly as possible, perhaps even in the first place he/she looks in the dictionary.

Another aspect that should not be neglected is the necessity of raising users' awareness of the principles followed by the lexicographers. This knowledge is essential if they want to become efficient dictionary users and to retrieve all the information included in a particular dictionary. Consequently, good dictionary skills also lead to a reduction in the number of look-up operations.

The methods for inclusion and treatment of multi-word lexical items in mono- and bilingual dictionaries should, in principle, be based on similar criteria, but contrastive differences between the source and the target languages should be taken into consideration when compiling a bilingual dictionary.

References

- Atkins, B. T. Sue; Rundell, Michael 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Cowie, Anthony P. 1994. *Phraseology*. – R. E. Asher (Ed.). *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon Press, 3168–3171.
- Hausmann, Franz Josef 2004. *Was sind eigentlich Kollokationen?* – Kathrin Steyer (Hr.). *Wortverbindungen – mehr oder weniger fest*. Institut für deutsche Sprache. Jahrbuch 2003. Berlin, New York: de Gruyter, 309–334.
- Landau, Sidney I. 2001. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.

- Moon, Rosamund 1998. *Fixed Expressions and Idioms in English. A Corpus-Based Approach*. Oxford: Oxford University Press.
- Svensén, Bo 2009. *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.

Dictionaries

- Krek, Simon (Ed.) 2005. *Veliki angleško-slovenski slovar* Oxford. Ljubljana: DZS.
- LDOCE5 = Mayor, Michael (Ed.) 2009. *Longman Dictionary of Contemporary English*. 5th ed. Harlow, Essex: Pearson Education Limited.
- MED2 = Rundell, Michael (Ed-in-Chief) 2007. *Macmillan English Dictionary for Advanced Learners*. 2nd ed. Oxford: Macmillan Education.
- COBUILD5 = Sinclair, John; Clari, Michela (Eds.) 2006. *Collins COBUILD Advanced Learner's English Dictionary*. 5th ed. London: HarperCollins Publishers.
- Soanes, Catherine; Stevenson, Angus 2003. *Oxford Dictionary of English*. 2nd ed. Oxford: Oxford University Press.
- OALD8 = Turnbull, Joanna (Ed.) 2010. *Oxford Advanced Learner's Dictionary of Current English*. 8th ed. Oxford: Oxford University Press.
- Walter, Elizabeth (Ed.) 2006. *Cambridge Idioms Dictionary*. 2nd ed. Cambridge: Cambridge University Press.
- CALD2 = Walter, Elizabeth; Woodford, Kate (Eds.) 2005. *Cambridge Advanced Learner's Dictionary*. 2nd ed. Cambridge: Cambridge University Press.
- CALD3 = Walter, Elizabeth; Woodford, Kate (Eds.) 2008. *Cambridge Advanced Learner's Dictionary*. 3rd ed. Cambridge: Cambridge University Press.
- Willmann, Helmut (Ed.) 2001. *Langenscheidt Muret-Sanders Großwörterbuch: Englisch-Deutsch*. Berlin, München, Wien, Zürich, New York: Langenscheidt.

Alenka Vrbinc (University of Ljubljana) publishes articles dealing with lexicographic issues, has edited a German–Slovene dictionary and co-authored an English–Slovene dictionary.
alenka.vrbinc@ef.uni-lj.si

Marjeta Vrbinc (University of Ljubljana) is interested in contrastive lexicology, practical lexicography, and learner's dictionary skills. She has edited a German–Slovene dictionary and co-authored an English–Slovene dictionary.
marjeta.vrbinc@ff.uni-lj.si

MITMESÕNALISTE ÜKSUSTE ESITUS SÕNARAAMATUS: HETKESEIS JA SÕNARAAMATU KASUTAJATE VÕIMALIKUD PROBLEEMID

Alenka Vrbinc, Marjeta Vrbinc

Ljubljana Ülikool

Artiklis analüüsitakse mitmesõnaliste leksikaalsete üksuste esitust ükskeelsetes (inglise keele) sõnaraamatutes eesmärgiga töötada välja nende esitusviisi kodeeriva kakskeelse (inglise-sloveeni) sõnastiku jaoks.

Esimeses osas antakse ülevaade sellest, kuidas on mitmesõnalisi üksusi käsitlenud inglise keele ükskeelsed õppesõnastikud, ning kirjeldatakse sõnaliigi määramise probleeme seoses mitmesõnalise üksuse leidmisega sõnaraamatus. Paljud fraseologismid on struktuurilt ebagrammatilised, mistõttu on nende komponentide sõnaliiki keeruline määrata. Asja teeb veel keerulisemaks idioomide semantiline läbipaistmatus. Ükskeelsete (inglise) sõnastike võrdluses selgus, et üks ja seesama idioom võib olla välja toodud mitmes eri (sõnaliigiga) artiklis. Artikli kesktes osas uuritakse, kas idioome on eri sõnaraamatutes esitatud makro- ja/või mikrostruktuuri üksusena. Selgub, et üks ja seesama mitmesõnaline leksikaalne üksus võib olla esitatud püsiühendite ploki, samal ajal kui teisel on ta välja toodud iseseisva märksõnana (lemmana). Omaette probleemi kujutavad endast sidekriipsuga (ja ilma sidekriipsuta) kirjutatud üksused, mille selgemaks esituseks antakse mõned soovitusel. Arutletakse ka püsiühendite esituse üle omaette tähendusüksusena või kasutusnäitena.

Uurimuse kokkuvõte on, et leksikograafid peavad fraseoloogiliste väljendite esitusviisi sobitama kasutaja vajadustega. Kasutajapoolse vähima pingutuse nimel tuleks rakendada pragmaatilist lähenemisviisi: üks ja seesama mitmesõnaline leksikaalne üksus peaks olema esitatud kahes või enam kohas, et kasutaja leiaks otsitud info võimalikult kiiresti.

Võtmesõnad: ükskeelsed sõnaraamatud, kakskeelsed sõnaraamatud, idiomaatilised väljendid, makrostruktuur, mikrostruktuur, sõnaraamatu kasutajad, inglise, sloveeni