

Multiple Labels Marking Connotative Values of Idioms in the Oxford Idioms Dictionary for Learners of English

MARJETA VRBINC

*Faculty of Arts,
University of Ljubljana, Slovenia
marjeta.vrbinc@ff.uni-lj.si*

ALENKA VRBINC

*Faculty of Economics,
University of Ljubljana, Slovenia*

ABSTRACT

This contribution is aimed at studying multiple labelling in the Oxford Idioms Dictionary for Learners of English (OIDLE2). We sought to establish whether labels belonging to one and the same category combine with one another or whether multiple labelling consists of labels from different categories of labels, the latter providing different types of information. The database used for the analysis was compiled by searching manually through the dictionary and keying in all the idioms with multiple labelling into our database. Altogether, 392 idioms or their senses with two or more labels were found in OIDLE2. The findings of the study are: labels expressing different types of diasystematic information are included; the three most frequent labels appearing in combination with other labels are informal, humorous and old-fashioned; the combination of four labels is used only once, ten idioms were identified with three labels, while the majority of label combinations consist of two labels. The significance of the findings lies in the issues related to multiple labels, combinations of labels expressing different types of diasystematic information and other issues related to labelling in general. The inclusion of diasystematic information largely depends on the type of dictionary and its intended users. This is especially true of dictionaries intended for non-native speakers of a language, where one of the main functions is to promote the active use of a foreign language, and where every single piece of information included in the dictionary counts.

Keywords: phraseological dictionary; diasystematic information; labels; connotation; dictionary users

INTRODUCTION

Phraseological units are used in a wide variety of situations and contexts. They are typical of spoken language but can also be encountered in written language, especially in literary texts and in journalism. Connotation is one of the crucial features of the stylistic properties of phraseological units. Moon (1998, pp. 274-275) stresses that texts should be explored in detail in terms of genre and intertextuality as well as in relation to other choices of lexis if we want to examine the stylistics of phraseological units. Gläser (1998, p. 128), however, points out that “connotations are additional semantic markers which are associated with the value judgements of a speech community [...] or of an individual speaker or writer”. Among the connotations most frequently encountered in phraseological units, Gläser (1998, p. 129) enumerates expressive connotations (derogatory, taboo, euphemistic, humorous), stylistic connotations (informal, slang, formal, literary, archaic, foreign) and register markers (astronomy, economics, judicial, medical) (cf. Fiedler 2007, pp. 24-25, Philip 2011, pp. 67–68, Cowie et al. 1983, pp. 39-40). Fiedler (2007, p. 23) agrees that phraseological units are used by a speaker/writer to make the text more expressive, since they are used to express evaluations of people or events, to attract attention, illustrate facts or organise texts.

THE INCLUSION OF INFORMATION ON CONNOTATION IN DICTIONARIES

In a dictionary, the connotative value of a lexical item is described by labels, and in many cases, more than one label is needed to describe all the connotations carried by a lexical item. In metalexigraphy, all the restrictions and constraints on the use of a lexical item are referred to as diasystematic information or diasystematic marking (Hausmann 1977, 1989, Svensén 2009, pp. 315-332). The term diasystematic marking is defined by Svensén as describing information implying that “a certain lexical item deviates in a certain respect from the main bulk of items described in a dictionary and that its use is subject to some kind of restriction” (Svensén 2009, p. 315). The existing literature offers various taxonomies of diasystematic information proposed by various researchers, but the most detailed classification can be found in Hausmann (1989, p. 651), who identifies as many as eleven types of labels. His classification was also adopted by various other scholars, such as Bergenholtz and Tarp (1995, pp. 131-134) and Svensén (2009, pp. 326-332) and is used as the theoretical background for our research:

1. diachronic information: this associates a word or one of its senses with a particular period in the history of language (*old-fashioned, obsolete, archaic, old use* or *dated*);
2. diatopic information: this associates a word or one of its senses with a particular regional dialect or national variety (*regional, dialect*);
3. diaintegrative information: this associates a word or one of its senses with the dimension of integration into the native stock of words of a language (*Latin, French, Italian*);
4. diamedial information: this associates a word or one of its senses with a particular medium of communication (*written, spoken*);
5. diastratic information: this associates a word or one of its senses with a particular social group, consequently referring to sociolects, such as slang and different kinds of jargon (*slang, vulgar, taboo*);
6. diaphasic information: this associates a word or one of its senses with a particular register of a language (*formal, informal*);
7. diatextual information: this associates a word or one of its senses with a particular discourse type or genre (*poetic, literary*);
8. diatechnical information: this associates a word or one of its senses with a particular subject field (*mathematics, business, law, medicine*);
9. diafrequent information: this associates a word or one of its senses with a particular frequency of occurrence (*less frequent, rare*);
10. diaevaluative information: this associates a word or one of its senses with a particular attitude or evaluation or the speaker’s mood (*appreciative, derogatory, offensive, humorous, ironic, euphemistic*);
11. dianormative information: this associates a word or one of its senses with a certain degree of deviation from a cultural standard (*non-standard, substandard, disputed*).

The way dictionaries employ labelling depends largely on the policy of each individual dictionary and is usually explained in the front or back matter. In practice, dictionaries use diverse methods for providing information on deviations of any kind. Apart from labels, another method for providing information about the connotative value of phraseological units is the use of definitions. *Longman Idioms Dictionary*, for instance, states in the front matter that information about the context and situation in which idioms are used (*Longman Idioms Dictionary* 1998, pp. viii–ix) is included in definitions “by noting that idiom is insulting or rude, or by explaining that a particular idiom is used when you are annoyed, angry, pleased, etc”. Connotation can also be expressed verbally in brackets at the

end of the entire entry, as is the case in the *Oxford Idioms Dictionary for Learners of English* (e.g., the idiom *to hell with sb/sth* is labelled *spoken* and accompanied by additional information at the end of the entry (*Some people find this use offensive.*)).

AIMS OF THE STUDY

This contribution is aimed at studying multiple labelling in a specialised phraseological dictionary by answering the following research questions.

1. Which labels are often used in combination with one another?
2. Do labels belonging to one and the same category combine with one another?
3. Does multiple labelling consist of labels from many categories of labels?

Multiple labelling was studied because one single label represents one piece of information which a user may find easier to process than information provided by a string of two or more labels. It should be stressed that not all the labels listed in the dictionary analysed for the purposes of our study were taken into consideration (a more detailed treatment of the labels studied and those omitted is provided in the Methodology section). Labelling was studied with the intended user in mind, since it is the user who should benefit from diastematic information. Therefore, suggestions were advanced for further improving labelling strategies in phraseological dictionaries, since labelling can still be considered an area of lexicography where more research work leading to better and more user-friendly practical solutions tailored to the needs of the users should be done (cf. Atkins, Rundell 2008, p. 496) if we want users to put their phraseological dictionary to good use. As for the terminology, we use the term ‘idiom’ as a superordinate and a hyponymic term for all phraseological units included in the *Oxford Idioms Dictionary for Learners of English*, regardless of the level of semantic opacity. It should be pointed out that dictionaries in the Anglo-American tradition mostly use ‘idiom’ without making any further typological classification of phraseological units (Moon 1998, p. 4).

METHODOLOGY

This study examines multiple labelling in a British phraseological dictionary, i.e., *Oxford Idioms Dictionary for Learners of English* (hereafter referred to as OI2). All the labels used in this dictionary are enumerated on the reverse of the front cover, and some, though not all, are explained. However, not all the labels included in the lists of labels were taken into consideration; therefore, some criteria had to be defined as to which labels to exclude from our research. The group of labels expressing diatopic information was disregarded, because diatopic information refers to regional or dialectal variation, but our intention was not to focus on inter- or intravarietal peculiarities of English. All the other labels were taken into consideration on condition that they appeared in combination with other labels. It is necessary to point out that not all the labels listed in the front matter were found in combination with another label (i.e., *figurative, sport, technical*). Consequently, these labels were not included in the database used for the analysis of multiple labelling in phraseological dictionaries.

Apart from the labels proper, the symbol exclamation mark within a triangle (hereafter represented by *!*) has the same force as a label, as is evident from the reverse of the front cover of the OI2, where the following explanation can be found: “*!* indicates a word or a phrase that is likely to be thought offensive or shocking by many people”. This symbol can be regarded as a label expressing diastatic information; therefore, we included it

in our research as one of the labels. Another label that appears in combination with other labels is *saying*, a label that does not belong to any of the above categories but is included in the front matter of the dictionary under the title *Labels Used in the Dictionary*. However, the typography used in the dictionary microstructure (i.e., italics) is the same as that used for all the labels in the dictionary. Regarding all the criteria that have to be met to interpret an element in a dictionary as a label, *saying* cannot be interpreted as a label, as it does not express any limitations or restrictions on usage. Apart from that, one would logically expect to find *idiom*, *proverb*, *catchphrase*, etc., as labels, but these do not appear among the labels in the OIDL2. Another reason for not excluding *saying* from our research was that a dictionary user is likely to interpret it as a label.

The database was compiled by searching manually through the dictionary. Since the size of the dictionary is manageable, it was not necessary to select one or several letters to represent the material for our analysis, neither was a random choice of a page range needed. The entire dictionary was analysed, which means that all the idioms with multiple labelling were keyed into our database. Altogether, 392 idioms or their senses with two or more labels were found in the OIDL2. In some cases, labels are separated by the disjunctive conjunction ‘or’, implying that either one or the other label marks the connotative features of the lemma (i.e., an idiom) or one of its senses. Labels joined by the conjunction ‘or’ were also included in the database, thus classifying them as multiple labels, since a user must decipher both of them. The database is assessed as being sufficiently representative for us to be able to extract multiple labels used in the dictionary under investigation. The aim of our research is to find combinations of labels rather than to perform a statistical analysis of the frequency of combinations as opposed to the use of a single label only.

RESULTS

LABELS EXPRESSING DIFFERENT TYPES OF DIASYSTEMATIC INFORMATION

The labels appearing in combinations belong to very different groups according to the classification of diasystematic information (see the inclusion of information on connotation in dictionaries). From eleven possible groups, the labels belonging to as many as nine groups can be found in our database. There are no labels belonging to the group of labels expressing diatopic information or dianormative information. All other groups are represented by at least one label, and the total number of labels (including *saying*) amounts to 22. Different types of diasystematic information expressed by labels appearing as multiple labels in the OIDL2 are as follows:

1. Diachronic information: *old-fashioned*, *old use*;
2. Diainegrative information: *French*, *Italian*, *Latin*;
3. Diamedial information: *spoken*, *written*;
4. Diastratic information: *slang*, *!*;
5. Diaphasic information: *formal*, *informal*;
6. Diatextual information: *literary*;
7. Diatechnical information: *business*, *law*, *politics*;
8. Diafrequent information: *less frequent*;
9. Diaevaluative information: *approving*, *disapproving*, *ironic*, *offensive*, *humorous*.

As can be seen from the above groups, diaevaluative information is expressed by as many as five labels. Diatechnical information is expressed by three labels, which also holds true for diainegrative information. Diachronic information, diamedial information, diastratic

information and diaphasic information are represented by two labels each, and diatextual information and diafrequent information are represented by one label each. The label *saying*, which belongs to none of the above categories, also appears in combination with other labels. The most frequently used labels are those expressing diaphasic information (261 occurrences), followed by the labels expressing diaevaluative information (197 occurrences), diachronic information (103 occurrences), diastratic information (95 occurrences) and diamedial information (90 occurrences). Labels expressing diaintegrative information are far less common (14 occurrences), whereas those expressing diatextual information (6 occurrences), diatechnical information (5 occurrences) and diafrequent information (1 occurrence) appear only rarely. For more detailed information on the frequency of label groups, see Table 1.

TABLE 1. Frequency of label groups.

Label group	Frequency
<i>Diaphasic</i>	261
<i>Diaevaluative</i>	197
<i>Diachronic</i>	103
<i>Diastratic</i>	95
<i>Diamedial</i>	90
<i>Diaintegrative</i>	14
<i>Diatextual</i>	6
<i>Diatechnical</i>	5
<i>Diafrequent</i>	1

FREQUENCY OF LABELS

On the one hand, the number of labels combining with one another is relatively high; on the other, their frequency of use differs greatly. Three labels, in particular, stand out in terms of frequency of occurrence in this dictionary:

- *informal* – appearing in combination with other labels in 234 cases,
- *humorous* – appearing in combination with other labels in 119 cases and
- *old-fashioned* – appearing in combination with other labels in 102 cases.

As opposed to these three very frequent labels, five labels are used only once: *old use*, *less frequent*, *politics*, *law* and *Italian*. The frequency of use of all other labels found in combinations differs greatly and ranges from 2 to 81 times: the labels *spoken*, *disapproving*, *slang* and *!* are used with a relatively high frequency rate, whereas the frequency of *offensive*, *business* and *approving* is negligible. For more detailed information on the frequency of use of labels, see Table 2.

TABLE 2. Varying degrees of frequency of labels

Label	Frequency	Label	Frequency
<i>informal</i>	234	<i>French</i>	7
<i>humorous</i>	119	<i>Latin</i>	6
<i>old-fashioned</i>	102	<i>literary</i>	6
<i>spoken</i>	81	<i>offensive</i>	3
<i>disapproving</i>	54	<i>business</i>	3
<i>slang</i>	48	<i>approving</i>	2
<i>!</i>	47	<i>old use</i>	1
<i>formal</i>	27	<i>less frequent</i>	1
<i>saying</i>	20	<i>politics</i>	1
<i>ironic</i>	19	<i>law</i>	1
<i>written</i>	9	<i>Italian</i>	1

COMBINATIONS OF TWO OR MORE LABELS

In the whole dictionary, the combination of four labels is used only once to mark one of the senses of the polysemous idiom *tough/bad luck*. Both senses of this idiom are labelled *informal* and sense two is additionally labelled *ironic*. In sense two, variant forms of the idiom are provided, i.e., *tough shit*, *tough titty*, which are labelled *!*, *slang* (further explanation of the problematic interpretation of these four labels can be found under Issues related to multiple labels).

The combination of three labels is only slightly more frequent than the combination of four labels, but still rather infrequent. In the entire dictionary, only ten idioms with three labels were identified. The combinations of labels used to mark restrictions on the use of certain idioms are as follows:

- *informal, old-fashioned, ironic (the best of luck (to sb))*;
- *informal, old-fashioned, humorous (the boys in blue and be mother)*;
- *informal, old-fashioned, business (talk through your hat)*;
- *humorous, formal, Latin (be compos mentis)*;
- *humorous, formal, French (the crème de la crème)*;
- *humorous, formal, disapproving (a/the prodigal son)*;
- *humorous, literary, Latin (in flagrante (delicto))*;
- *humorous, old use, less frequent (Heaven forfend (that))*;
- *humorous, spoken, law (I rest my case)*.

In the OIDE2, however, the majority of label combinations consist of two labels. The label *informal* is the most frequently used label (229 occurrences in double combinations), and it combines with nine other labels (*humorous, old-fashioned, disapproving, spoken, saying, ironic, business, !, written*). The frequency of combinations varies considerably: by far the most frequent combination is *informal, humorous*, with 68 occurrences; this is followed by the combination *informal, old-fashioned*, with 56 occurrences, *informal, disapproving*, with 48 occurrences and *informal, spoken*, with 39 occurrences. All other combinations with *informal* are far less common and range from 7 occurrences (*informal, saying*), to one occurrence (*informal, written*). For more details on the combinations with *informal*, see Table 3.

TABLE 3. The label *informal* used in a double combination with other labels

<i>informal</i> +	Frequency	<i>informal</i> +	Frequency
<i>humorous</i>	68	<i>ironic</i>	6
<i>old-fashioned</i>	56	<i>business</i>	2
<i>disapproving</i>	48	<i>!</i>	2
<i>spoken</i>	39	<i>written</i>	1
<i>saying</i>	7		

The second most frequent label *humorous* (114 occurrences in double combinations) combines with eleven labels (*informal, old-fashioned, formal, saying, spoken, literary, ironic, Latin, French, disapproving, slang*), with varying degrees of frequency of occurrence. The most frequent combination is that with *informal*, which has already been mentioned (see Table 3); this is followed by the combination with *old-fashioned* and *spoken* with many fewer occurrences. As many as five labels (*literary, disapproving, slang, Latin, French*) appear only once in combination with *humorous*. For the sake of consistency, the same combinations of labels are presented in more than one table depending on which label is used as the starting point for different combinations. For example, the combination *informal* and *humours* is

presented in Tables 3 and 4. In Table 3, the starting point is the label *informal*, and in Table 4, the starting point is the label *humorous*. For more details on the combinations with *humorous*, see Table 4.

TABLE 4. The label *humorous* used in a double combination with other labels

<i>humorous</i> +	Frequency	<i>humorous</i> +	Frequency
<i>informal</i>	68	<i>literary</i>	1
<i>old-fashioned</i>	14	<i>disapproving</i>	1
<i>spoken</i>	10	<i>slang</i>	1
<i>saying</i>	7	<i>Latin</i>	1
<i>formal</i>	6	<i>French</i>	1
<i>ironic</i>	4		

In terms of frequency, *old-fashioned* occupies the third place with 102 occurrences, 98 of them being double combinations. It combines with nine labels (*informal*, *spoken*, *humorous*, *formal*, *saying*, *written*, *disapproving*, *slang*, *literary*), most frequently with *informal*, a combination already mentioned in Table 3. Also worth mentioning are *spoken*, which appears in combination with *old-fashioned* in 17 cases, and *humorous*, which appears in 14 cases (see also Table 4). All other labels are used much less frequently in combination with *old-fashioned*, i.e., from 4 times (*formal*), to only once (*disapproving*, *slang*, *literary*). For more details on the combinations with *old-fashioned*, see Table 5.

TABLE 5. The label *old-fashioned* used in a double combination with other labels

<i>old-fashioned</i> +	Frequency	<i>old-fashioned</i> +	Frequency
<i>informal</i>	56	<i>written</i>	2
<i>spoken</i>	17	<i>disapproving</i>	1
<i>humorous</i>	14	<i>slang</i>	1
<i>formal</i>	4	<i>literary</i>	1
<i>saying</i>	2		

The label *spoken* combines with other labels in 81 cases, 69 of these being double combinations. The most frequent combinations are those with *informal* (39 occurrences; see also Table 3), *old-fashioned* (17 occurrences; see also Table 5) and *humorous* (10 occurrences; see also Table 4), while as many as five labels (*!*, *slang*, *offensive*, *approving*, *disapproving*) are used only once in combination with *spoken*. For more details on the combinations with *spoken*, see Table 6.

TABLE 6. The label *spoken* used in a double combination with other labels

<i>spoken</i> +	Frequency	<i>spoken</i> +	Frequency
<i>informal</i>	39	<i>!</i>	1
<i>old-fashioned</i>	17	<i>slang</i>	1
<i>humorous</i>	10	<i>offensive</i>	1
<i>ironic</i>	8	<i>approving</i>	1
<i>formal</i>	2	<i>disapproving</i>	1

The labels *informal*, *humorous*, *old-fashioned* and *spoken* are the ones most commonly used. All other labels are used much less frequently. Among the pairs of labels that appear with greater frequency is the combination *slang*, *!*, with 43 occurrences. All other combinations of two labels are rarely used:

- *formal*, *written* (4 occurrences);
- *formal*, *Latin* (3 occurrences);
- *formal*, *literary* (3 occurrences);
- *ironic*, *saying* (2 occurrences);
- *written*, *French* (2 occurrences).

Some combinations, however, appear only once in the entire dictionary. For further details, see Table 7, which does not include the labels *informal*, *humorous*, *old-fashioned* and *spoken* in the left-hand column.

TABLE 7. Infrequent pairs of labels used in a double combination

Label combination	Frequency	Label combination	Frequency		
	<i>informal</i>	48	<i>spoken</i>	8	
	<i>old-fashioned</i>	1	<i>ironic</i>	<i>informal</i>	6
	<i>humorous</i>	1		<i>saying</i>	2
<i>disapproving</i>	<i>spoken</i>	1		<i>humorous</i>	1
	<i>formal</i>	1		<i>formal</i>	4
	<i>offensive</i>	1	<i>written</i>	<i>old-fashioned</i>	2
	<i>Italian</i>	1		<i>French</i>	2
	!	43		<i>informal</i>	1
<i>slang</i>	<i>spoken</i>	1		<i>written</i>	2
	<i>old-fashioned</i>	1		<i>saying</i>	1
	<i>offensive</i>	1	<i>French</i>	<i>humorous</i>	1
	<i>humorous</i>	1		<i>approving</i>	1
	<i>slang</i>	43		<i>formal</i>	1
!	<i>informal</i>	2	<i>Latin</i>	<i>formal</i>	3
	<i>spoken</i>	1		<i>humorous</i>	1
	<i>humorous</i>	6		<i>formal</i>	3
	<i>written</i>	4	<i>literary</i>	<i>humorous</i>	1
<i>formal</i>	<i>old-fashioned</i>	4		<i>old-fashioned</i>	1
	<i>literary</i>	3		<i>spoken</i>	1
	<i>Latin</i>	3	<i>offensive</i>	<i>slang</i>	1
	<i>spoken</i>	2		<i>disapproving</i>	1
	<i>disapproving</i>	1	<i>business</i>	<i>informal</i>	2
	<i>French</i>	1			
	<i>informal</i>	7			
	<i>humorous</i>	7			
	<i>old-fashioned</i>	2			
<i>saying</i>	<i>ironic</i>	2			
	<i>politics</i>	1			
	<i>French</i>	1			

DISCUSSION

ISSUES RELATED TO MULTIPLE LABELS

As has been established in the Results section, the maximum number of labels in the OIDE2 is four, but the combination of four labels is extremely rare and can be found for only one idiom, i.e., *tough/bad luck* and its variants *tough shit* and *tough titty*. The treatment of this idiom is extremely problematic, because it is difficult to decipher which label refers to which specific form of the idiom listed within sense 2. A close observation shows two possible interpretations:

- sense 2 of *tough/bad luck* is labelled *informal*, *ironic*, while its variants *tough shit* and *tough titty* are labelled *!*, *slang*;
- *tough shit* and *tough titty* are labelled *informal*, *!*, *slang*, *ironic*, that means using four labels.

The question that can rightly be posed here is whether such complex labelling would be understood and correctly interpreted by a dictionary user. The problem begins with the label *informal*, which appears right after the lemma *tough/bad luck*. Theoretically, the scope of this label is clear: it refers to the entire dictionary entry, i.e., to both senses of this

polysemous idiom. However, in practice, one may wonder whether a dictionary user would be aware of this. A possible solution would be to have two separate dictionary entries: i.e., *tough/bad luck* and *tough shit/titty*. Although the issue is not within the scope of this article, one cannot but point out the way the variant of the lemma is given: *tough shit, tough titty*. The use of the comma is misleading, since the intended dictionary user might incorrectly interpret it as indicating one idiom (*tough shit, tough titty*), rather than as two variants (*tough shit* and *tough titty*). It should not be forgotten that the primary dictionary users are non-native speakers of English who should be given clear guidance about the exact structure of the idiom in question, rather than ambiguous variants that allow multiple interpretations as regards structure.

The combination of three labels is also infrequent, since only ten idioms with three labels were identified in the whole dictionary. With the exception of the idioms *the boys in blue* and *be mother*, which are both labelled *informal, old-fashioned, humorous*, all other combinations of three labels are unique, since a comparison of the triple combinations shows no repetition, which may be due to the lack of any template concerning triple labelling in the dictionary under investigation. In the idioms with a combination of three labels, *humorous* is the label found in as many as eight idioms, being followed by *informal* and *old-fashioned*, which are found in four cases, *formal* in three cases and *Latin* in two cases. All other labels found in a combination of three labels appear only once. This is in line with the general frequency of labels found in multiple combinations, since *informal, humorous* and *old-fashioned* occupy the first three places in terms of frequency. At the same time, it clearly shows that idioms are often used in informal and spoken language as opposed to more formal occasions and in written language. If we compare the frequency of the labels *informal* and *spoken*, on the one hand, and that of *formal* and *written*, on the other, we can see that the former far exceed the latter (315 and 36, respectively) in combinations of three or two labels.

An observation worthy of comment regarding the label *humorous* is that in as many as 54 out of 119 occurrences, the label is accompanied by *or* (29 instances), *often* (22 instances) or *usually* (3 instances). A more detailed examination of the cases where *humorous* is combined with the conjunction *or* shows that it most commonly appears in combination with the label *old-fashioned*. Only in four instances does *humorous* not combine with *or* if it is used in combination with *old-fashioned*. This suggests that an old-fashioned expression can very often be used humorously, which means that lexical items that are not often used nowadays but are used by older people or were used by people in the recent past can acquire additional, i.e., humorous connotation.

The analysis of double combinations shows that *slang* and *!* combine in as many as 43 instances, which means that *slang* is used in combination with another label in only 4 cases, whereas *!* is used in combination with another label in only 3 cases. The front matter explanation for *!* suggests that a lemma or one of its senses marked by *!* is thought to be very offensive in a way that many people find shocking, whereas there is no explication in the front matter of the OIDE2 for the label *slang*; however, this label normally marks items used in very informal spoken English by a small group of people. The reason may be sought in the type of vocabulary items treated in this specialised dictionary, since in very informal language, an idiom or one of its senses often has a very rude and/or offensive connotation.

COMBINATIONS OF LABELS EXPRESSING DIFFERENT TYPES OF DIASYSTEMATIC INFORMATION

As far as the combinations go, it is to be expected that labels expressing different types of diasytematic information be combined, since it seems logical to provide the user with

different types of information. This assumption is proved by studying the combinations of labels in the OIDL2, which are as follows:

Diaphasic information (*formal, informal*) is combined with:

- diaevaluative information (*disapproving, humorous, ironic*),
- diachronic information (*old-fashioned*),
- diamedial information (*written, spoken*),
- diaintegrative information (*Latin, French*),
- diatextual information (*literary*),
- diatechnical information (*business*),
- diastratic information (!),
- the “label” *saying*.

Diastratic information (*slang, !*) is combined with:

- diaevaluative information (*humorous*),
- diamedial information (*spoken*),
- diachronic information (*old-fashioned*).

Diaevaluative information (*humorous, disapproving, ironic, offensive*) is combined with:

- diachronic information (*old-fashioned, old use*),
- diatextual information (*literary*),
- diamedial information (*spoken*),
- diaintegrative information (*Italian, Latin, French*),
- the “label” *saying*.

Diachronic information (*old-fashioned, old use*) is combined with:

- diamedial information (*spoken, written*),
- diatextual information (*literary*),
- the “label” *saying*.

Diamedial information (*spoken, written*) is combined with:

- diaintegrative information (*French*).

The “label” *saying* is combined with:

- diatechnical information (*politics*),
- diaintegrative information (*French*).

Interestingly, different labels expressing diaevaluative information and therefore belonging to the same group of labels can also be combined. The use of the label *humorous*, for instance, which is the second most frequently used label marking 119 idioms in the OIDL2, suggests that it is diaevaluative information that idioms often convey. Apart from *humorous*, other labels expressing diaevaluative information (*approving, disapproving, ironic, offensive*) are used to label 78 idioms, which means that as many as 197 idioms assigned two or more labels tend to communicate diaevaluative information. Given the nature of idioms, i.e., expressing evaluation on the part of the speaker/writer, this finding is more than expected and corresponds to Moon’s findings (1998, p. 219, pp. 223-225). This is also confirmed by the fact that of all the labels expressing different types of diasystematic information, the ones expressing diaevaluative information are most numerous, since as many as five labels (*humorous, approving, disapproving, ironic, offensive*) from this group are represented in this dictionary. Similarly, labels expressing diastratic information, i.e., *slang*

and !, can also be combined. The groups of labels expressing diaevaluative and diastratic information are the only groups of labels whose members combine with one another (at least in our database). This is, however, not surprising, especially if we consider the labels expressing diaevaluative information, since this group comprises various labels, or more precisely, it is the most heterogeneous group, with labels expressing different types of evaluation. On the other hand, if we take the labels expressing diaphasic information, we can see that they are represented by the labels *formal* and *informal* and the diamedial information is marked by *spoken* and *written*. These labels are beyond doubt labels expressing opposing stylistic levels, thus being mutually incompatible. The same observation can be made in connection with the labels *approving* and *disapproving*, expressing diaevaluative information, which are also mutually exclusive, but can freely combine with other labels belonging to this group and still provide different types of information about the use of one particular lexical item labelled in this way. The same holds true for the labels *old-fashioned* and *old use*, marking diachronic lemmata or senses and belonging to the group of labels expressing diachronic information.

If we take a closer look at the labels expressing other types of information, we can see that there is a great difference in the frequency of occurrence between the individual groups. Diatechnical information is expressed by three labels, and the same holds true for diaintegrative information. Both types of information are infrequent in our database. As regards diatechnical information, this could be expected, since idioms are not frequent in technical and scientific vocabulary. It is true that some scientific disciplines are richer in idioms than others. In business English, for instance, one comes across more idioms than in the terminology of medicine. Generally speaking, terminology is not characterised by the use of idioms. It should, however, be stressed that popular scientific articles appearing in semi-specialist journals prefer a variety of stylistic devices to attract and retain the reader's attention and interest, whereas academic research articles are almost completely devoid of figures of speech (Gläser 1998, p. 132). As for the labels expressing diaintegrative information, it can be claimed that the number of idioms used in the language of origin is not particularly high. There is considerable calquing in different languages (calquing is probably quite a productive method of extending the stock of idioms in certain languages), but calques are not labelled in dictionaries using labels expressing diaintegrative information. Diachronic, diamedial, diastratic and diaphasic information categories are represented by two labels each, but despite the low number of labels representing each type of information, the labels appear with a high frequency of occurrence. Given the nature of these types of information and the nature of idioms, this finding is more than expected. Idioms are characterised by informal, spoken and slang connotations, which is why such labels are very common.

OTHER ISSUES RELATED TO LABELLING

While studying the dictionary entries in the OIDE2 to compile the database of idioms with multiple labels, it was observed that in some cases, the lexicographers explain the restrictions on the use of certain idioms rather than making use of labels proper. For instance, the idiom *to hell with sb/sth* is labelled *spoken*, but at the end of the dictionary entry, one finds the following explanation provided in brackets: "Some people find this use offensive". Similar treatment can also be found for *what the hell!, god/Heaven help sb* and in the case of some other idioms. In those cases, one can only wonder why the lexicographers did not include the information about the offensive connotation by means of the label *offensive*.

Another problematic use of labels can be found in the idiom *blow smoke (up sb's ass)*, where ! can be found at the end of the part in brackets, while the entire idiom is additionally marked with *AmE* and *slang*, i.e., *blow smoke (up sb's ass !)* (*AmE* and *slang*). This means

that ! refers to the addition in brackets, and the shorter as well as the longer version are marked *AmE* and *slang*. This labelling strategy can be regarded as highly complicated, and raises the question whether an average dictionary user can interpret this combination of labels appropriately. A possible solution would be to list this idiom twice and to label the variant appropriately: *blow smoke (AmE, slang)* and *blow smoke up sb's ass (AmE, slang, !)*.

In reference to the combination of two or more labels, it should be stressed that it is quite demanding for the dictionary user to comprehend and process this information if we take into consideration that it is probably not only labels and their comprehension that present a problem area for the user. The user does not consult the dictionary to check the labels but primarily to seek other information (e.g., meaning, grammatical information or context; cf. also Kim Hua & Woods 2008, Namvar 2012). This means that the user may regard labels as secondary (possibly inessential) information.

CONCLUSION

Multiple labelling can be regarded as being difficult for an average dictionary user to comprehend and interpret correctly. This especially holds true for those cases of multiple labelling with several labels listed one after the other. The problem is even more aggravated if some labels appear in brackets and refer to variants of the lemma, or when some are found outside the brackets and refer to the lemma and/or to the variant(s) given. Information provided in the form of labels is very important and is especially significant when it refers to idioms, which are demanding for non-native speakers from the aspect of decoding and even more so from the aspect of encoding. This is why labels should be included in a way that would enable a dictionary user to interpret them appropriately and consequently to use the lemma correctly.

Some suggestions on how to improve multiple labelling of idioms are as follows:

- It would be advisable not to list variants whose restrictions and connotations differ from those of a lemma as variant forms under the lemma but rather to include them in a dictionary as independent lemmata with their specific labelling needs. This should, however, be examined more exhaustively by conducting a user survey to gain better insight into users' comprehension and interpretation of labels.
- As for the label *saying*, it can be concluded that this label is either to be retained or discarded. If it is retained, other labels with a similar value, such as *proverb*, *idiom*, *catchphrase*, should be introduced to provide the entire set systematically and to mark different types of phraseological units more precisely. Another solution would be to discard such precise labelling of phraseological units, which seems to be a more sensible decision in phraseological dictionaries intended for non-native speakers and non-linguists. An average dictionary user is not interested in a theoretical classification of phraseological units; moreover, a dictionary user would not gain any practical benefits in terms of decoding and encoding needs from such a theoretical classification. On the other hand, the symbol !, which has the value of a label, can be quite effective; in fact, it can be more effective than the label proper (e.g., *taboo*). As such, its use in a phraseological dictionary is a welcome feature.
- All the labels or the symbols used instead of labels in a particular dictionary should be precisely explained in the introductory part of the dictionary to enable the dictionary user to get an idea about the correct meaning of each label.
- The introductory part of the dictionary should also include a detailed explanation of the scope of the label, which depends on the position of the label(s), i.e., either before the

first sense number (in this case, the label refers to all the senses that follow) or after the sense number (in this case, the label refers to that specific sense only).

In conclusion, the inclusion of diasystematic information largely depends on the type of dictionary and especially on its intended users. Therefore, lexicographers' decisions about whether to use a label and how to use it appropriately should be based on the user profile, a recommendation which is also proposed by Ptaszynski (2010, p. 437). This is especially true of dictionaries intended for non-native speakers of a language, where one of the main functions is to promote the active use of a foreign language, in which case every single piece of information included in the dictionary counts.

REFERENCES

- Atkins, B. T. Sue & Rundell, Michael (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Bergenholtz, H. & Tarp, S. (1995). *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. Amsterdam: John Benjamins Publishing Company.
- Cowie, Anthony P., Mackin, Ronald & McCaig, Isabel R. (1983). *Oxford Dictionary of Current Idiomatic English. Volume 2: Phrase, Clause & Sentence Idioms*. Oxford: Oxford University Press.
- Fiedler, Sabine (2007). *English Phraseology*. Tübingen: Gunter Narr Verlag.
- Gläser, Rosemarie (1998). The Stylistic Potential of Phraseological Units in the Light of Genre Analysis. In Anthony P. Cowie (Ed.), *Phraseology: Theory, Analysis, and Applications* (pp 125–144). Oxford: Oxford University Press.
- Hausmann, Franz Josef (1977). *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Tübingen: Niemeyer.
- Hausmann, Franz Josef (1989). Die Markierung in einem allgemeinen einsprachigen Wörterbuch: Eine Übersicht. In Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta (Eds.), *Wörterbücher /.../ Ein internationales Handbuch zur Lexikographie /.../. Erster Teilband* (pp 649–657). Berlin, New York: Walter de Gruyter.
- Kim Hua, Tan and Woods, Peter Charles (2008). Media-Related or Generic-Related Features in Electronic Dictionaries: Learner's Perception and Preferences. *GEMA Online[®] Journal of Language Studies*. Vol. 8(2), 1-17.
- Moon, Rosamund (1998). *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Clarendon Press.
- Namvar, Firooz (2012). The Relationship between Language Proficiency and Use of Collocation by Iranian EFL Students. *3L: The Southeast Asian Journal of English Language Studies*. Vol. 18(3), 41-52.
- Parkinson, Dilys and Francis, Ben (2006). *Oxford Idioms Dictionary for Learners of English*, 2nd Ed. Oxford: Oxford University Press. (OIDLE2)
- Philip, Gill (2011). *Colouring Meaning: Collocation and Connotation in Figurative Language*. Amsterdam, Philadelphia: John Benjamins Publishing.
- Ptaszynski, Marcin Overgaard (2010). Theoretical Considerations for the Improvement of Usage Labelling in Dictionaries: A Combined Formal-Functional Approach. *International Journal of Lexicography*. Vol. 23(4), 411-442.
- Stern, Karen and Engineer, Sue (eds.) (1998). *Longman Idioms Dictionary*. Harlow, Essex: Addison Wesley Longman Limited.
- Svensén, Bo (2009). *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.