Diasystematic Information in Learner's Dictionaries: The Usability of Multiple Labels

Marjeta Vrbinc
marjeta.vrbinc@ff.uni-lj.si
University of Ljubljana, Slovenia

Alenka Vrbinc
alenka.vrbinc@ef.uni-lj.si
University of Ljubljana, Slovenia

ABSTRACT

Labels in dictionaries provide information about restrictions and constraints on the use of certain words or senses in the contexts in which they occur or in relation to other words described in a dictionary. These restrictions are referred to as diasystematic information. This contribution deals with diasystematic information in four British monolingual learner's dictionaries (OALD8, LDOCE5, CALD3, and MED2) with the emphasis on multiple labels. A detailed classification of labels is presented and an in-depth overview of labels used in combination with one another in the dictionaries under investigation is given. Then follows a discussion of labels which are often used in combination, the aim being 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 findings of this study show that labels belonging to different groups are mostly combined, and apart from that, labels expressing diaevaluative information and those expressing diastratic information can be combined within the group. Possible reasons for this are discussed in the contribution. 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. This is especially true of monolingual learner's dictionaries, where one of the main functions is to promote the active use of a foreign language where every single piece of information included in the dictionary counts.

Keywords: mulitple labels; diasystematic information; monolingual learner's dictionaries; foreign learners; comprehension of labels

INTRODUCTION

The tradition of using labels in dictionaries has existed for a long time, but lexicographers as well as dictionary users are still faced with many problems related to the system of labelling (Ptaszynski, 2010, pp. 411-412), one of the main reasons being that "many labels are umbrella terms that conceal a good deal of variation" (Atkins & Rundell, 2008, p. 496). Labels in dictionaries are aimed at providing information about restrictions and constraints on the use of certain words or senses in the contexts in which they occur or in relation to other words described in a dictionary. These restrictions and constraints are referred to in lexicographic theory as diasystematic information or diasystematic marking (Hausmann, 1977, 1989; Svensén, 2009, pp. 315-332), a term that is defined by Svensén (ibid, p. 315) as 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". Diasystematic information is, no doubt, the type of information that changes over time. That

is why labels used in individual dictionaries change from edition to edition. In order to reflect language change, labels need to be updated, which is in line with Norri (2000, p. 93) as well as Bauer (1994), who carried out research into changes in the labelling of abusive nouns in five different editions of *The Concise Oxford Dictionary* from the period between 1911 and 1990.

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 different methods for providing information on deviations of any kind; two of these methods should be foregrounded: definitions and labels. If information usually expressed by a label is incorporated within the definition, it can be found either in the introductory part of the definition (e.g., *exec* is defined as 'an <u>informal</u> word for executive' in LDOCE5; *CAT scan* as 'a <u>medical</u> examination that uses a computer to produce an image of the inside of sb's body from x-ray or ultrasound pictures' in OALD8; underlined by the authors) or in any other part of the definition (e.g., *catapult* is defined as 'a weapon used <u>in the past</u> to throw heavy stones' in OALD8; *princely* as 'a princely amount of money is very large – often used <u>humorously</u> to mean a very small amount of money' in LDOCE5; underlined by the authors). As this research focuses on labels only, diasystematic information within definitions falls beyond the scope of this contribution.

The existing literature offers various taxonomies of diasystematic information proposed by different researchers. Mostly, these vary in the degree of elaboration, but there is also a great deal of overlap. Jackson (2002, pp. 109-115), for instance, enumerates seven types of usage labels: dialect; formality; status; effect; history; topic or field and disputed usage. Landau (2001, pp. 217-272), however, classifies eight common kinds of usage information: currency or temporality; regional or geographic variation; technical or specialized terminology; restricted or taboo sexual and scatological usage; insult; slang; style, functional variety, or register; status or cultural level. Atkins and Rundell (2008, pp. 182-186) distinguish nine types of linguistic labels: domains; region; dialect; register; style; time; slang and jargon; attitude; offensive terms. 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 presented below:

1. diachronic information (criterion: time) - a feature which associates a word or one of its senses with a particular period in the history of language. This dimension includes a range of labels that can be arranged chronologically from archaic, via obsolete to contemporary words or senses and neologisms. In practice, lexical items that do not represent old use lack a label, which means that neologisms are usually unmarked. The most common temporal labels found in contemporary dictionaries are *old-fashioned, obsolete, archaic, old use* or *dated*[i].

2. diatopic information (criterion: place) – a feature which associates a word or one of its senses with a particular regional dialect or national variety. Every language community has certain conventions as to what is standard, thus unlabelled in a dictionary. Most British dictionaries nowadays include words or senses typically used in different varieties of English, from North America to South Africa, from Australia and New Zealand to the Caribbean. Sometimes regional areas within a country are specified; thus, *regional* or *dialect* is used as a label.

3. diaintegrative information (criterion: nationality) – a feature which associates a word or one of its senses with the dimension of integration into the native stock of words of a language. Monolingual dictionaries usually provide information on the language of origin, mostly with words that have retained their original form (e.g., from *Latin,* from *French,* from *Italian*).

4. diamedial information[ii] (criterion: medium) – a feature which associates a word or one of its senses with a particular medium of communication. The most common labels are *written* and *spoken*.
5. diastratic information (criterion: socio-cultural specificity) – a feature which 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. The most common labels are *slang, vulgar* and *taboo*.
6. diaphasic information (criterion: formality) – a feature which associates a word or one of its senses with a particular register of a language, the most common labels being *formal* and *informal*.
7. diatextual information (criterion: text type) – a feature which associates a word or one of its senses with a particular discourse type or genre. The most common labels are *poetic* and *literary*.
8. diatechnical information (criterion: technicality) – a feature which associates a word or one of its senses with a particular subject field. In monolingual dictionaries, subject-field labels, field labels or domain labels indicate that a certain word or one of its senses belongs to technical or scientific vocabulary. The large number of sublanguages typical of different subject fields presents a problem even for educated native speakers, since each subject field has its own vocabulary. Consequently, some dictionaries, especially learner's dictionaries, often use the general label *technical* or *science* (cf. also Vrbinc & Vrbinc, 2013, p. 449) instead of giving detailed information on specific subject fields.
9. diafrequent information (criterion: frequency) – a feature which associates a word or one of its senses with a particular frequency of occurrence. Labels used to indicate frequency are *less frequent, rare*.
10. diaevaluative information (criterion: attitude) – a feature which associates a word or one of its senses with a particular attitude or evaluation or the speaker's mood. Labels used to denote diaevaluative information are *appreciative, derogatory, offensive, humorous, ironic, euphemistic*.
11. dianormative information (criterion: normativity) – a feature which associates a word or one of its senses with a certain degree of deviation from a cultural standard. Labels expressing dianormative information are *non-standard, substandard, disputed*. In other words, the acceptability of items marked with one of these labels is questionable as regards linguistic correctness.

The label *figurative* is used as a label in many dictionaries, but it does not fit in any of the above-mentioned categories of labels, since it refers to the meaning extension of a lexical item rather than expressing any restriction on usage. In the process of figurative extension, it is possible that a word acquires semantic features that might not be present in the original meaning (Atkins & Rundell, 2008, p. 289).

This detailed typology shows that it is essential to define and classify restrictions and constraints in order to be able to incorporate diasystematic information in dictionaries; as Ptaszynski (2010, p. 415) claims, "identifying particular types of restrictions and discussing their nature are important to the understanding of the notion of diasystematic information".

The scope of a label is a convention which lexicographers should follow to the greatest possible extent, and the position of the label determines its scope: a label preceding sense division in a polysemous entry refers to all senses, whereas a label following the number indicating the sense division applies to that particular sense only. Thus, it can be said that, if the label precedes the sense number, it is located at a higher level in the hierarchy of the dictionary entry than the label within the same dictionary entry that follows the sense

number. The relation between the superordinate and subordinate labels can be regarded as cumulative (Svensén, 2009, p. 321).

The aim of this contribution is to give an overview of labels used in British monolingual learner's dictionaries (MLDs) and to discuss which labels are often used in combination. We also wanted 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. Multiple labelling was studied because one single label represents one piece of information which a user may find easier to process than a sequence of two or even three pieces of information in the form of multiple labels. It has to be stressed that not all the labels listed in the dictionaries were taken into consideration (a more detailed treatment of the labels studied and those omitted is provided in the Methodology section). Labelling was studied having the intended user in mind, since it is the user who should benefit from diasystematic information. The theoretical considerations presented in this article are aimed at improving usage labelling in MLDs, since we cannot but agree with Atkins & Rundell (2008, p. 496), who claim that "labelling is an area of lexicography where there is more work to be done".

METHODOLOGY

This study examines multiple labelling in four British MLDs, i.e., *Oxford Advance Learner's Dictionary*, 8th edition (OALD8), *Longman Dictionary of Contemporary English*, 5th edition (LDOCE5), *Cambridge Advanced Learner's Dictionary*, 3rd edition (CALD3) and *Macmillan English Dictionary*, 2nd edition (MED2). Among the 'big five', only *Collins Cobuild Advanced Dictionary of English* (COBUILD) was left out, the reason being that the CD-ROM of this dictionary does not allow for a label search, since it does not offer the Advanced Search option. 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 are excluded from our research. Four groups of labels expressing diasystematic information were disregarded: diatopic, diaintegrative, diatechnical and diafrequent for the following reasons:

- The diatopic information refers to regional or dialectal variation, but our intention was not to focus on inter- or intravarietal peculiarities of English.
- Labels giving diaintegrative information provide only information on the language of origin of a certain lexical item rather than on restrictions and constraints on the use of certain words or senses, which is typical of diasystematic information in general.
- Diafrequent information is actually not indicated by means of labels in MLDs. For instance, LDOCE5 marks the most frequent lexical items in spoken and written language by alpha-numerical codes such as S3 (one of the 3,000 most frequent words in spoken English), W2 (one of the 2,000 most frequent words in written English); CALD3 marks frequency by E = essential, I = improver, A = advanced; MED2 uses the star symbols (three stars mark the 2,500 most common and basic English words; two stars mark very common words, whereas one star marks fairly common words). OALD8 uses the key symbol to mark the items used most frequently, but at the same time it is the only dictionary under scrutiny that has two labels expressing diafrequent information, i.e., *less frequent* and *rare*, which are included in the list of labels on the CD-ROM, but are not given in the print edition.
- Diatechnical information refers to a specialist field in which a lexical item is used; consequently, the subject-field or domain labels represent a special group of labels indicating terminology, but our intention was not to focus on terms belonging to different specialist fields or even subfields.

All the other labels (i.e., diachronic, diamedial, diastratic, diaphasic, diatextual, diaevaluative, dianormative) were taken into consideration and form the database used for the analysis of multiple labelling in MLDs.

The database was compiled using CD-ROMs of the above-mentioned dictionaries. The letter S was chosen as being sufficiently representative, since it is one of the 'big' letters, or more precisely, it is the letter with the greatest number of entries in English dictionaries. The database was created using the Advanced Search option (in OALD8, LDOCE5 and CALD3) and the Super Search option (in MED2). In OALD8, the dropdown menu in Advanced Search offers 19 labels under "Register", from which 14 were selected as appropriate for our research. The dropdown menu in Advanced Search of LDOCE5 lists 15 labels under "Register", 10 of which were included in our database, while in CALD3, 13 out of 21 labels listed under "Usage" were analysed. In the Super Search option of MED2, the Advanced Search offers 17 labels under "Style", and all of these were taken into consideration in our research.

Each label was selected separately, and each search yielded a list of words marked with this particular label. All these words in the stretch beginning with the letter S were manually checked to see whether multiple labels were used. If a lemma, one of its senses or a phraseological unit was marked with one label only, the item was not entered into our database. If, on the other hand, two or more labels were used, the lemma, one of its senses or a phraseological unit was included in the database. 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, one of its senses or a phraseological unit. 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. Such a search yielded 187 lemmata or senses in OALD8, 35 lemmata in LDOCE5, 202 lemmata, senses or phraseological units in CALD3 and 58 lemmata, senses or phraseological units in MED2. Altogether, 482 lemmata, senses or phraseological units with multiple labels were studied. The database is assessed as being sufficiently representative for us to be able to extract multiple labels used in each particular 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.

Here, some weaknesses of the CD-ROMs should be mentioned. In OALD8, for instance, the Advanced Search option is far from being ideal for linguistic research, since a label search yields only (one- or multi-word) lemmata, which means that phraseological units are disregarded. The Advanced Search option in LDOCE5 searches lemmata marked with a certain label. This means that, if the entire lemma (in the case of polysemous lemmata) is marked with a label, such a lemma is found, but, if a certain sense is labelled, the Advanced Search does not include it in the Search results list, and the same holds true for all phraseological units included as separate senses of the lemma in frequency order. The CD-ROM of CALD3 gives the most comprehensive picture of labels, since certain senses or even illustrative examples that are labelled are also included in the Results list. Apart from that, the list also includes multi-word lexical items treated as lemmata as well as phraseological units. Multi-word lexical items with labels are also not found by using the Super Search option in MED2, a problem that has already been mentioned in OALD8 and LDOCE5. The CD-ROM of MED2, however, searches among phraseological units.

RESULTS

Before taking a closer look at multiple labelling by dictionaries, we should start by making a survey of the labels included in our research in relation to whether they can be found in all

four dictionaries under investigation or whether they appear in one or two dictionaries but are not used by the lexicographers in other dictionaries.

It seems logical to start with labels that can be found in all four dictionaries. According to our database, there are five such labels: *formal, humorous, informal, literary* and *old-fashioned*. These labels belong to four different groups of labels expressing different types of diasystematic information: diaphasic (*formal, informal*), diaevaluative (*humorous*), diatextual (*literary*) and diachronic (*old-fashioned*). Four labels appear in three dictionaries: *approving, disapproving, offensive* (OALD8, CALD3, MED2; in MED2, *approving* and *disapproving* are expressed as *showing approval* and *showing disapproval*) and *old use* (OALD8, LDOCE5, CALD3); they belong to the groups of labels providing diaevaluative (*approving, disapproving*) and diachronic (*old use*) information. Six labels are used in two dictionaries: *figurative, slang, non-standard* (OALD8, CALD3; in CALD3, *not standard* is used instead of *non-standard*), *spoken, impolite* (LDOCE5, MED2; in LDOCE5, *not polite* is used instead of *impolite*) and *taboo* (OALD8, LDOCE5). The latter labels are classified as giving dianormative (*non-standard*), diastratic (*slang, taboo*), diamedial (*spoken*) and diaevaluative (*impolite*) information. As has already been mentioned in the Introduction, the label *figurative* cannot be grouped into any of the label categories. That is why it must be treated separately as a label expressing certain shades of semantic meaning as well as a certain degree of stylistic level. If we disregard the labels typical of MED2 which only express intensification or frequency of one and the same label (*very formal, very informal, often humorous, mainly journalism, mainly literary, mainly spoken*), only four labels included in our research are used in one dictionary: *ironic* (OALD8), *journalism* (MED2), *written* (LDOCE5) and *child's word* (CALD3). These labels belong to the labels expressing diaevaluative (*ironic*), diatextual (*journalism*), diamedial (*written*) and diaphasic (*child's word*) information.

Below, combinations of labels in each individual dictionary are presented to see whether different groups of labels combine or whether combinations of labels belonging to one and the same group are also used to mark a lemma, sense or phraseological unit.

MULTIPLE LABELS IN OALD8

Our database shows that a combination of a maximum of three labels can be found in OALD8, although the two-label combination is the most frequent. The labels that are found in a triple combination are the following:

- *taboo, slang, disapproving* (shit, noun – sense 4);
- *old-fashioned, slang, disapproving* (spiv);
- *old-fashioned, informal, disapproving* (swank).
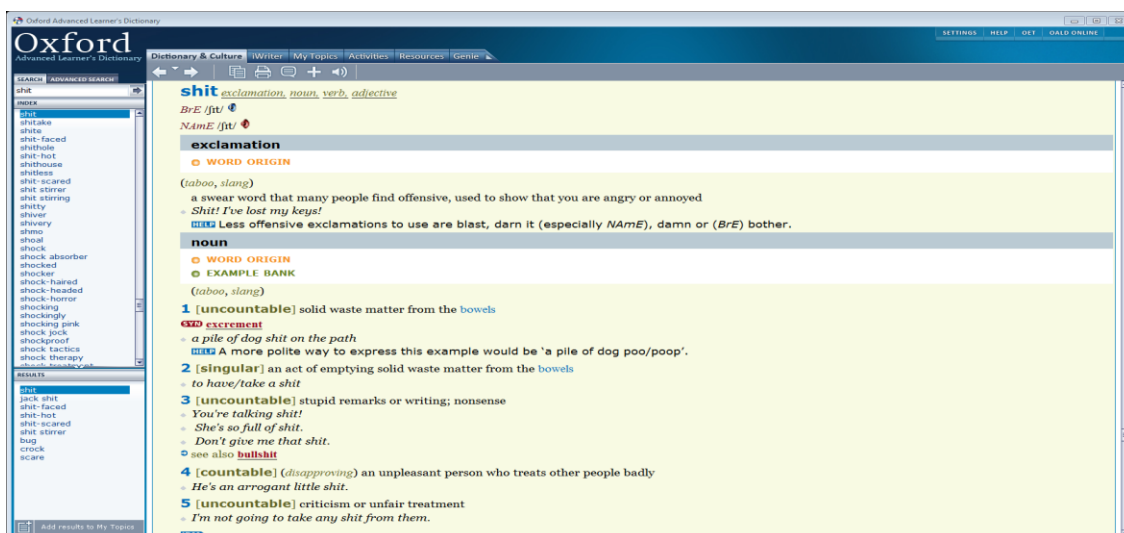
An example of an entire dictionary entry:

FIGURE 1. An example of a dictionary entry for *shit* from OALD8.

The most usual combination contains two labels, the combinations being quite varied. The label *informal* combines with other labels by far the most frequently. In our database, as many as 82 out of 187 items are labelled *informal* (43.9%). The label *informal* is found in combination with the following labels:

- *disapproving*: 44 items: scab;
- *old-fashioned*: 26 items: super (adjective);
- *humorous*: 7 items: skedaddle;
- *approving*: 2 items: swanky;
- *ironic*: 2 items: Sherlock;
- *figurative*: 1 item, i.e., shove (verb).

The next most common label is *formal*, but in comparison with *informal*, it appears much less frequently, since 36 out of 187 items in our database are marked *formal* (19.3%). The label *formal* combines with the following labels:

- *literary*: 11 items: scion (noun);
- *old-fashioned*: 11 items: slacks;
- *figurative*: 6 items: salivate;
- *disapproving*: 5 items: supine;
- *old use*: 2 items: save (preposition);
- *humorous*: 1 item: superannuated.

The labels *taboo* and *slang* are used in an almost identical number of items, 27 and 29, respectively. In the great majority of cases (26 in all), they are used in combination with one another; there is only one case where *slang* combines with *offensive* (scumbag), one case where it combines with *old-fashioned* (strewth) and one case where it combines with *humorous* (shrink (noun)). The same holds true for *taboo*, which combines with *old-fashioned* only in the lemma savage (adjective).

       Most of the labels that combine with *informal* and *formal* also appear in combination with other labels. *Approving*, for instance, combines with *literary* in one case only (steadfast), and the same holds true for the combination with *figurative* (slim – sense 1). *Disapproving*, on the other hand, is used in combination with *humorous* in three items in our database (scribbler), with *old-fashioned* in two cases (spinster) and in one case with *offensive* (slut),

old-use (strumpet) and *literary* (sullen). The label *literary* combines with *figurative* in three cases (sentinel), twice with *old use* (smite) and once with *old-fashioned* (slay – sense 1). *Humorous* is combined with *old-fashioned* in 8 cases (skulduggery), with *literary* in two cases (steed), and in one case it is used in combination with *slang* (shrink) and *old use* (stripling). With the label *humorous*, one peculiarity captures our attention, i.e., the frequent use of the conjunction 'or', implying that the items marked can be interpreted either as *informal/old-fashioned*, etc., or as *humorous*.

The label *non-standard* does not appear in combination with any other label in the letter S, which means that it is the only label studied that appears in isolation rather than in combination, at least in the stretch studied.

MULTIPLE LABELS IN LDOCE5

The database consisting of labels found in LDOCE5 is much smaller than the database with labels in OALD8; nevertheless, a number of multiple labels were detected. There is only one case of triple labelling: the noun scumbag is labelled *spoken, informal, not polite*.

Two labels stand out in frequency: *not polite* (19 out of 35 items, i.e., 54.3%) and *spoken* (17 items, i.e., 48.6%). Both of these labels appear in the triple combination mentioned above, but besides that, they also combine with other labels. *Not polite*, for instance, appears in combination with the following labels:

- *spoken*: 12 items: shit (noun);
- *informal*: 6 items: sissy;
- *old-fashioned*: 1 item: savage (noun).
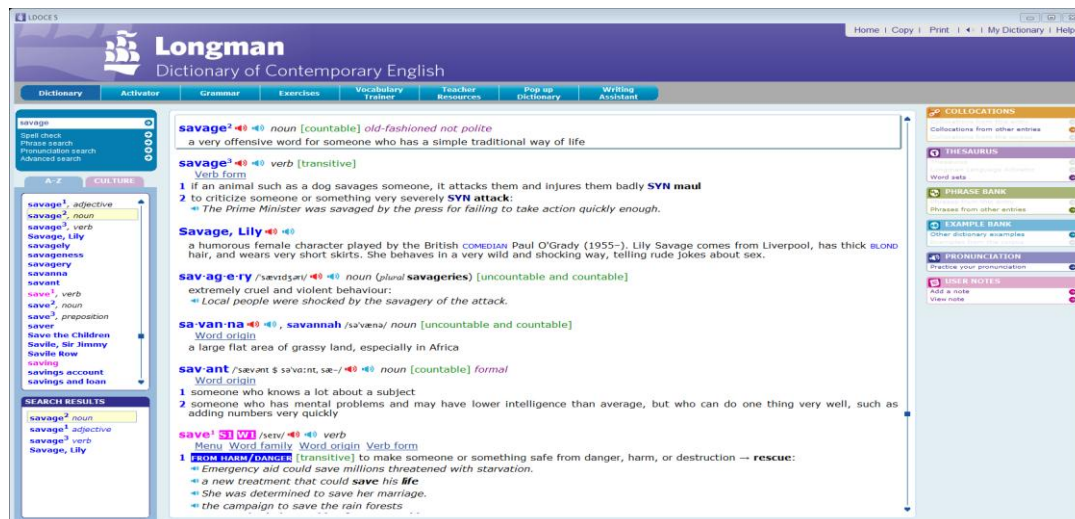
An example of an entire dictionary entry:

FIGURE 2. An example of a dictionary entry for *savage* (noun) from LDOCE5.

Spoken, as the second most frequent label in LDOCE5, is used in combination with the following labels:

- *not polite*: see above at *not polite*;
- *informal*: 4 items: sorted;
- *old-fashioned*: 1 item: sonny.

Another label whose frequency is comparable to that of *not polite* and *spoken* is *informal*, which is used to label 15 items in our database (i.e., 42.9%). It combines with *spoken* and *not*

polite (mentioned above) as well as with *old-fashioned*, which is used to label four items (swinger), and with *taboo* in one case (slut). The label *formal* appears in combination with other labels in three cases in our database, namely in combination with *literary* (sully), *written* (sic (adverb)) and *old-fashioned* (seafarer). In our database, *literary* combines in one case with *biblical* (spake), *formal* (sully) and with *old-fashioned* (sainted). The last label that needs to be mentioned is *taboo*, which combines with *informal* mentioned above and with *slang* (Sambo). Interestingly, the label *slang* is to be found neither in the list of labels in the print edition of LDOCE5 nor in the Register list of the Advanced Search option on the CD-ROM, but it is obviously used to mark slang expressions.

MULTIPLE LABELS IN CALD3

In CALD3, three items in the range studied are marked with three labels: spanking is labelled *approving, informal, old-fashioned*, spiv is marked *disapproving, informal, old-fashioned*, while sense 3 in seizure is marked *old use*, and the example illustrating the use of this sense is additionally labelled *figurative, humorous*.

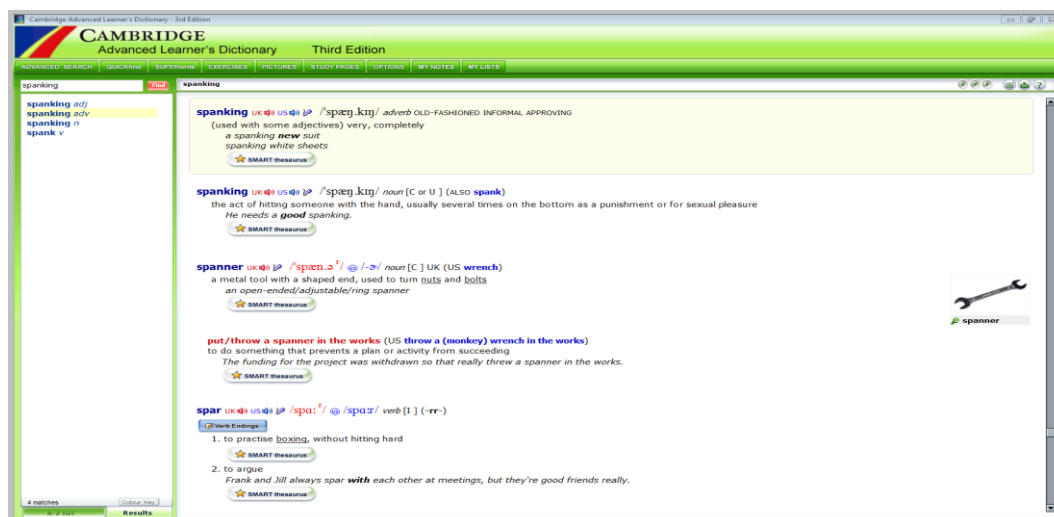An example of an entire dictionary entry:

FIGURE 3. An example of a dictionary entry for *spanking* (adverb) from CALD3.

All other items, however, are marked with two labels, of which three should be emphasized in terms of frequency: *informal* (133 out of 202 items, i.e., 65.8%), *disapproving* (104 items, i.e., 51.5%) and *old-fashioned* (60 items, i.e., 29.7%).

    Informal combines with five different labels:
- *disapproving*: 71 items: schizo;
- *old-fashioned*: 40 items: screwy;
- *humorous*: 15 items: scallywag;
- *approving*: 4 items: snazzy;
- *figurative*: 1 item: shut (sb) up.

The second most frequent label is *disapproving*, which is used in combination with the following labels:
- *informal*: see above at *informal*;
- *slang*: 13 items: sluttish;

- *formal*: 12 items: sentimentalism;
- *humorous*: 3 items: scrounge;
- *old-fashioned*: 2 items: shrew – sense 2;
- *old use*: 2 items: scarlet woman;
- *literary*: 1 item: spawn – sense 2.

Another label that is fairly frequent in CALD3 is *old-fashioned*, which combines with these other labels:

- *informal*: see above at *informal*;
- *humorous*: 5 items: scoundrel;
- *slang*: 4 items: slewed;
- *literary*: 3 items: stouthearted;
- *disapproving*: see above at *disapproving*;
- *formal*: 2 items: sovereign remedy;
- *offensive*: 1 item: spade – sense 4.

Apart from the combinations presented above, CALD3 includes only very few other combinations. Three additional combinations were found with the label *humorous*, which combines with *old use* in five items in our database (spake,) and with *slang* in two items (shekel – sense 2) and *figurative* (species – illustrative example) and *literary* in only one item (sage (noun) – sense 2). The label *approving* combines with *literary* in two cases (seraphic) and *figurative* in one case only (slim). One more combination should be mentioned that appears four times in our database, namely the combination *literary, old use* (silvan). It is interesting to note that *literary* and *old use*s are separated by the disjunctive 'or' in all four cases. Last but not least, the label *child's word* appears in combination with *informal* in two cases only (swot – noun, verb) and in both cases, the two labels are separated by the disjunctive 'or'.

MULTIPLE LABELS IN MED2

As far as the number of combinations is concerned, MED2 differs greatly from the other three dictionaries. In MED2, two labels stand out as regards frequency, i.e., the labels *informal* and *old-fashioned*, but a closer examination reveals that *old-fashioned* combines only with *informal* and not with any other label. The same applies to the label *disapproving*, which is found in combination with *informal* only. Fifty-one out of 58 items in our database are marked *informal* (87.9%), while the combination *informal, old-fashioned* appears in 39 items (67.2%) and the combination with *disapproving* in 11 items (19.0%). Regarding the data in our database, by far the most usual combination is *informal, old-fashioned*, which is used to label items such as smasher. The second most frequent combination is *informal, disapproving*, which appears in items such as showoff. Only one item in our database is labelled with the combination *informal, approving*, namely stonker.

An example of an entire dictionary entry:

FIGURE 4. An example of a dictionary entry for *smasher* from MED2.

Only four other combinations can be found in our database with a negligible number of appearances:

- *formal, spoken*: 4 items: shall I/we say*;*
- *humorous, literary*: 1 item: sans;
- *humorous, mainly spoken*: 1 item: make someone sick;
- *impolite, spoken*: 1 item: (the) same to you.

As many as five labels used in MED2 do not appear in combination with other labels: *very informal, journalism, mainly journalism, literary* and *mainly literary*. These labels can be used with labels expressing diatopic information, especially *American* and *British*, but these combinations are beyond the scope of this contribution.

DISCUSSION

LABELS USED IN THE MAJORITY OF DICTIONARIES UNDER REVIEW AND THEIR INTERPRETATION

Five labels (*formal, informal, literary, humorous* and *old-fashioned*) are used to mark lexical items in all dictionaries under consideration. Among these labels, the labels *formal* and *literary* may pose some problems for the intended dictionary users, i.e., advanced learners of English. These two labels are often used together, but interestingly, in OALD8, they are always separated by the disjunctive 'or' (e.g., sibilant; sully is an example of a lemma which is marked *formal* or *literary* in OALD8 as well as in LDOCE5), which only adds to the complexity of the problem, since the connotation they mark may not be obvious to every user. The reason can also be sought in the users' mother tongue or more precisely, in the method of labelling that is familiar to the user from monolingual dictionaries written in his/her native language. In Slovene lexicography, for example, the label *knjižno* at least roughly corresponds to the English label *formal*. The front matter of the *Slovar slovenskega knjižnega jezika* (Dictionary of Standard Slovene, 1997, p. xxvi) explains that the label *knjižno* is assigned to 'words, senses or phraseological units used particularly in fiction or in scientific texts; in colloquial speech, they sound unusual'. *Knjižno* can also be used to mark the same connotation as the English label *literary*, which means that dictionary users familiar with a system of labelling similar to that for the native speakers of Slovene may be faced with the problem of the correct interpretation of the difference between *formal* and *literary*. A

comparison of the explications of these two labels in the front matter of MLDs reveals that these are very simple and sometimes even overlap. In CALD3, for instance, the distinction between these two labels is explained as follows:

> *formal* – "used in serious or official language or when trying to impress other people"
> *literary* – "formal and descriptive language used in literature"

It is questionable whether these two explications are clear enough for a foreign learner to distinguish accurately between these two labels, because *literary* is explained as labelling formal language. The addition of 'in literature' in the explication of *literary* may also puzzle the user who may not know the connotation of this restriction. Does this label mark expressions that the user is not supposed to use without sounding poetic? Do such expressions express metaphorical and metonymic transfer? Another problem that should be mentioned in connection with *literary* is that in some dictionaries (though not in the MLDs under investigation), there is a seemingly similar label *literature*, which is a domain rather than a style label. Users familiar with the label *literature* are likely to confuse these two labels, thus interpreting the label *literary* as being a domain label indicating terms from the field of literature instead of words found in literary and poetic texts. This confusion is also touched upon by Atkins and Rundell (2008, p. 229), but this only applies to a dictionary user who would use two or more different dictionaries using either *literature* or *literary*.

Another problem concerns the interpretation of the labels *old-fashioned* and *old use*. Both labels are used in OALD8, LDOCE5 and CALD3. The difference is that *old-fashioned* marks lexical items that are not often used nowadays but are used by older people or have been used by people in the recent past, while *old use* implies that the lexical item marked with this label is no longer in current use. This distinction may be recognized by a native speaker of English, but it may not be so obvious to a foreign learner, who is the target user of MLDs. MED2, the most recent newcomer to the field of learner lexicography, is the only dictionary under consideration that does not include *old use* in the list of labels, and considering the fact that the intended user may not be able to understand the difference between the two labels, this may seem a sensible decision.

OPPOSING LABELS

Besides the labels *old-fashioned* and *old use*, which should be used in dictionaries with care, the labels *formal* and *informal*, *polite* and *impolite*, *spoken* and *written* are also pairs worth considering. The labels *formal* and *informal* are used in all dictionaries, whereas *polite* and *impolite*, *spoken* and *written* are not. *Impolite* (or *not polite* as used in MED2) is a label found in LDOCE5 and MED2, but interestingly, neither of these dictionaries has the opposing label *polite*. *Polite*, on the other hand, is used in CALD3, which, however, does not have the label *impolite*. This means that the entire pair *polite/impolite* is not used in any of the dictionaries included in our study. It is true that it seems unlikely that foreign learners will need a warning that a given usage is polite, as politeness is generally the unmarked register but from the point of view of dictionary users, this may indicate inconsistency or lack of marking the opposing value of a lexical item in question. It seems sensible to label lexical items that imply politeness as well as those that imply impoliteness. This is especially true of dictionaries for foreign learners, who need precise guidance on lexical items expressing a certain degree of politeness/impoliteness, since labelling such words is aimed at warning those unfamiliar with them that they may be either polite or impolite. Doubtlessly, pragmatic information supplied by labels is as vital as the semantic information supplied by the definition (cf. also Norri, 2000, p. 93).

LABELS UNIQUE TO ONE DICTIONARY ONLY AND THEIR INTERPRETATION

If we address the issue concerning the labels used in one dictionary only, we can see that MED2 stands out in that as many as six labels (*often humorous, very informal, journalism, mainly journalism, mainly literary* and *mainly spoken*) can be found that occur in no other dictionary. It should, however, be emphasized that the majority of these labels appear in pairs: *humorous* vs. *often humorous, informal* vs. *very informal, literary* vs. *mainly literary* and *spoken* vs. *mainly spoken* and *journalism* vs. *mainly journalism* (the labels *journalism* and *mainly journalism* are used in MED2 only). *Mainly*, which precedes *journalism, literary* and *spoken*, implies that the lexical item thus marked is chiefly used in journalistic language, in literary works or in spoken language. *Very*, modifying *informal*, expresses intensification and is close to the label *slang*, which is not used in this dictionary. *Often*, however, restricts the meaning of *humorous*, implying that the lexical item with this label is not necessarily used with humorous connotation in all contexts. Such labels, therefore, provide more detailed information on the connotative value of the lexical item in question, but do not really contribute to clearer labelling and easier interpretation by the intended dictionary user.

Another label used in only one dictionary, i.e., LDOCE5, is *written*. The explanation provided in the print edition of this dictionary says that it refers to "a word or phrase that is used only, or nearly always, in *written* English" (authors' italics). If we compare this with the explanation given for the label *formal*, which states that it refers to "a word that is suitable for formal speech or *writing*, but would not normally be used in ordinary conversation" (authors' italics), we can see that they are both explained in a similar way. Therefore, the question can be posed whether a learner of English would really spot the difference between these two labels. Apart from the difference between *written* and *formal*, another pair is used in LDOCE5 as well as in MED2, which presents the same difficulty in interpretation as *written* and *formal*, namely, *informal* and *spoken*. A comparison of the description of these two labels offers no solution to this problem, since in MED2, the explications provided are extremely close:

informal – "more common in speech than in writing and not used on a formal occasion"
spoken – "used in speech rather than writing"

Slightly more precise, but still not clear enough are the explications in LDOCE5:

informal – "a word or phrase that is used in normal conversation, but may not be suitable for use in more formal contexts, for example in writing essays or business letters"
spoken – "a word or phrase used only, or nearly always, in conversation"

Although the explications of labels found in the front matter of the print editions of MLDs may not always be sufficiently precise and clear, the user has at least a reference to consult when in doubt about the exact meaning of a particular label. The problem is that nowadays, an increasing number of dictionary users do not use a print edition of a dictionary, since users prefer to consult the online dictionaries that are freely available to users with an internet connection or the CD-ROM versions accompanying the MLDs. These electronic dictionaries, however, do not necessarily offer the user the explanation of labels (the explanation of labels can be found in the online version of OALD and the CD-ROM of MED2).

POSSIBLE COMBINATIONS OF LABELS

As far as the combination of different labels is concerned, it has already been mentioned in the Results section that the maximum number of labels found in this study is three, but often the diatopic information, mostly *British English, American English*, can also accompany

other labels, which means that the number of labels used to mark one lexical item or one sense can amount to as many as four (in OALD8, for example, the noun spiv is labelled *old-fashioned, BrE, slang, disapproving*; the verb swank is labelled *old-fashioned, BrE, informal, disapproving*). Each combination of labels is unique, since a comparison of the triple combinations shows no repetitions, which may be due either to the database not being sufficiently large to indicate such repetitions or to the lack of any template concerning the triple labelling in the dictionaries under investigation. In OALD8, the only label that appears in all three triple combinations is *disapproving*, while *old-fashioned* and *slang* appear in two combinations. The labels *taboo* and *informal* are used once in the triple combinations found. In LDOCE5, only one triple combination was found, which does not allow a comparison with other similar labels, while in CALD3, no label is used in all three triple combinations. The labels *old-fashioned* and *informal* are used twice, while all other labels (*approving, disapproving, old use, figurative* and *humorous*) are used once only. This implies that even the inter-dictionary comparison yields no result. All that can be said in reference to a triple combination of labels is 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 probably does not consult the dictionary only to check the labels but also to seek other information (e.g., meaning, grammatical information, collocations or context; cf. also Hua & Woods, 2008; Namvar, 2012). This means that the user may regard labels as secondary (possibly inessential) information.

Another observation worthy of comment is the label *figurative*, which also appears in multiple labels. In all cases in our database, this label is used to mark examples illustrating the use of the lemma or one of its senses. This feature typical of the label *figurative* may also be why this label does not fit into any of the categories of labels enumerated in the Introduction. The label *figurative* can thus be said to denote secondary senses that have the status of conventional metaphors (cf. Hanks, 2006, 28).

As far as the combinations go, it is to be expected that labels belonging to different groups 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 found in all four dictionaries under investigation. Diaphasic information (*formal, informal*) is combined with diaevaluative information (*approving, disapproving, humorous*), diachronic information (*old-fashioned, old use*) and diatextual information (*literary*), as well as with the label *figurative*. Diastratic information (*slang, taboo*) combines with diaevaluative (*offensive, humorous*) and diachronic information (*old-fashioned*). Diaevaluative information is found in combination with diachronic (*old-fashioned, old use*), diatextual (*literary*) and diastratic information (*slang*) as well as with the label *figurative*. Interestingly, labels expressing diaevaluative information and therefore belonging to the same group of labels can also be combined; thus, *disapproving* is combined with *humorous* and *offensive*, and *approving* with *humorous*. Similarly, labels expressing diastratic information, i.e., *slang* and *taboo*, can also be combined. These two groups of labels 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 diaevalutive information, since this group is comprised of various labels, or more precisely, it is the most heterogeneous group, with labels expressing different types of evaluation. 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 incompatible with each other. 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. On the other hand, the labels *old-fashioned* and *old use*, which belong to the group of labels expressing diachronic information, also do not combine, since they both mark diachronic lemmata or senses.

OTHER METHODS OF PROVIDING DIASYSTEMATIC INFORMATION

When studying multiple labelling in MLDs, it can be seen that different pieces of information are not always given in the form of labels themselves. In many cases, the use of navigational devices replaces the label proper. In OALD8, for instance, the shortcut preceding senses 8 and 9 of the lemma safe (adjective) says APPROVING, which is identical with the label *approving* also used in this dictionary. Additionally, both senses are labelled *BrE, informal*, which means that the shortcut in combination with the label is an example of an indirect use of multiple labelling. Sense 8 of the above-mentioned adjective is defined as 'used by young people to show that they approve of sb/sth', the definition itself suggesting that safe is used approvingly in this sense. For example:
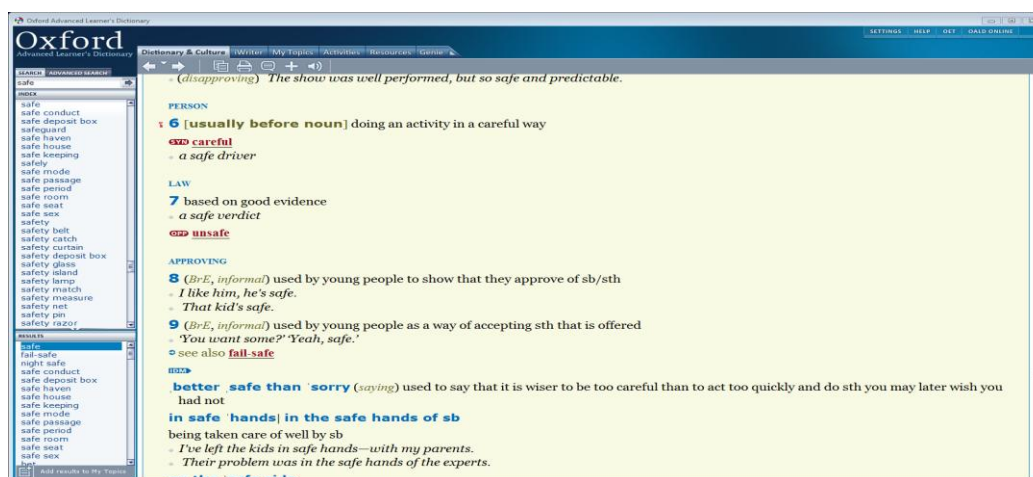

FIGURE 5. Senses 8 and 9 in the entry for *safe* from OALD8.

This corresponds to the method mentioned in the Introduction, by which deviations and restrictions of any kind can be integrated into different parts of definitions.

LABELLING THAT SHOULD BE AVOIDED

Some confusing examples of diasystematic data can be found in our database. An example is the treatment of the undefined run-on entry shamelessly in OALD8, which is labelled 'usually *disapproving* but sometimes *approving*'. Here, the opposing labels belonging to the same group (i.e., expressing diaevaluative information) are used to mark one and the same lemma, and at least the first of the two illustrative examples (i.e., The whole film is shamelessly romantic and glamorous; He shamelessly admits his part in the crime) does not provide further details on the connotative use of this adverb, thus leaving the user to make his/her own interpretation regarding the use of this adverb. Since the target user of the MLDs is a non-native speaker of English, s/he should get more precise and less confusing information about what makes this particular word approving or disapproving in specific contexts.

CONCLUSION

At first sight, labels used in MLDs seem to be so clear that no explanation of their value is needed for the user (a foreign learner) to interpret them correctly. A more careful study, however, shows that this is only wishful thinking and that, in certain cases, the user needs guidance as to the exact meaning of some labels. The print versions of the latest editions of MLDs lack precise front matter giving the user the possibility to get as much information as possible from the dictionary s/he consults; instead, they provide a graphic presentation of the entries with very basic information. The print versions, however, all include a list of labels, briefly explaining the meanings of labels, but the electronic versions of the same dictionaries, be they the CD-ROM versions or the online versions, only exceptionally provide explanations of the labels used. The user is therefore left to his/her own devices, not knowing, in many cases, how to interpret certain labels s/he encounters when consulting a dictionary. Nowadays, dictionary users prefer the electronic versions to print editions because of the numerous advantages these dictionaries provide; this preference obviously results in the necessity for explaining metalinguistic information of various types, labels being, most certainly, one of these. The absence of this explanation in electronic dictionaries is a serious shortcoming of electronic dictionaries, a shortcoming that first needs to be identified to be addressed later on. An additional tab providing explanations of different types of metalinguistic information or the information provided when the label is clicked on would be of great help to the user.

In pairs of labels that are close to each other as regards the restrictions or limitations they express, thus presenting problems of interpretation, special care needs to be taken to select only such labels as are easily interpretable by the intended users, i.e., foreign learners of English. Apart from that, two labels that express opposing values should by no means be combined to form a multiple label, since this confuses the user rather than gives him/her the piece of information s/he needs, especially when encoding. Different combinations of labels used to mark lemmata or their senses may also sometimes act as deterrents to clarity, as users are faced with the issue of comprehension of several different labels listed one after the other. This is in line with the findings reported by Namatende Sakwa (2011, p. 312). Another point that should be highlighted is that the position of the label should be carefully chosen, since it determines the scope of the label. Otherwise, the dictionary user may interpret the scope of the label incorrectly, which may lead to inappropriate use of a particular lexical items or one of its senses. Repeating one and the same label to mark the same lemma or the same sense should also be avoided if lexicographers want to give intended users the correct guidance as to the use of the lemma or one of its senses.

When studying labels in LDOCE5, it was noted that there is a discrepancy between the list of labels and the actual use of labels in the dictionary (cf. *slang* is used in LDOCE5, although it is not included in any list of labels, as has been mentioned in the section Multiple labels in LDOCE5). Such discrepancies are also observed by Namatende Sakwa (2011, p. 308), who gives examples of usage labels not listed in the front matter but used to mark the words in the dictionaries she studied. Needless to say, all the labels used in the A–Z sections should be included and explained in the list of labels provided. Otherwise, diasystematic information can be regarded as inadequate and incomplete.

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 MLDs, where one of the main functions is to promote the active use of a foreign language, where every single piece of information included in the dictionary counts. It should, however,

be stressed that MLDs are complex reference tools that are not easy to use if one wants to extract all the information they contain. Dictionary using skills are, therefore, very important at all stages of learning a foreign language, since users should be well aware of what they can find in their dictionaries, and they should know how to interpret and access different items of information included. Consequently, in the course of learning a foreign language, they should be trained in dictionary use. The role of the teacher in this respect is crucial, since the ability of learners to exploit as large a proportion as possible of the total information available (the labels being no exception) in learner's dictionaries depends almost solely on the teacher of those learners. School curricula still do not put sufficient emphasis on the various dictionary using skills, although constant practice in these skills in language classes seems essential. Teaching should be supported by materials designed to increase dictionary awareness and to develop reference skills, such as the various workbooks that accompany monolingual dictionaries.

REFERENCES

Atkins, B. T. S. & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.

Bajec, A. et al. eds. (1997). *Slovar Slovenskega Knjižnega Jezika*. Ljubljana: DZS.

Bauer, L. (1994). *Watching English Change: An Introduction to the Study of Linguistic Change in Standard Englishes in the Twentieth Century*. London, New York: Longman.

Bergenholtz, H. & Tarp, S. (1995). *Manual of Specialised Lexicography. The Preparation of Specialised Dictionaries*. Amsterdam: John Benjamins Publishing Company.

Hanks, P. (2006). Metaphoricity is Gradable. In A. Stefanowitsch & S. T. Gries, (Eds.). *Corpus-Based Approaches to Metaphor and Metonymy* (pp. 17-36). The Hague: Mouton de Gruyter.

Hausmann, F. J. (1977). *Einführung in die Benutzung der neufranzösischen Wörterbücher*. Tübingen: Niemeyer.

Hausmann, F. J. (1989). Die Markierung in Einem Allgemeinen Einsprachigen Wörterbuch: Eine Übersicht. In F. J. Hausmann, O. Reichmann, H. E. Wiegand, & L. Zgusta, (Eds.). *Wörterbücher /.../ Ein internationals Handbuch zur Lexikographie /.../. Erster Teilband* (pp. 649-657). Berlin, New York: Walter de Gruyter.

Hua, T. K. & Woods, P. C. (2008). Media-Related or Generic-Related Features in Electronic Dictionaries: Learner's Perception and Preferences. *GEMA Online® Journal of Language Studies. 8*(2), 1-17.

Jackson, H. (2002). *Lexicography: An Introduction.* London, New York: Routledge.

Landau, S. I. (2001). *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.

Mayor, M. ed. (2009). *Longman Dictionary of Contemporary English*. 5th ed. Harlow, Essex: Pearson Education Limited. (LDOCE5)

Namatende Sakwa, L. (2011). Problems of Usage Labelling in English Lexicography. *Lexikos. 21*, 305-315.

Namvar, F. (2012). The Relationship Between Language Proficiency and Use of Collocation by Iranian EFL Students. *3L: Language Linguistics Literature®, Southeast Asian Journal of English Language Studies. 18*(3), 41-52.

Norri, J. (2000). Labelling of Derogatory Words in Some British and American Dictionaries. *International Journal of Lexicography. 13*(2), 71-106.

Ptaszynski, M. O. (2010). Theoretical Considerations for the Improvement of Usage Labelling in Dictionaries: A Combined Formal-functional Approach. *International Journal of Lexicography*. *23*(4), 411-442.

Rundell, M. ed. (2007). *Macmillan English Dictionary for Advanced Learners*. 2nd ed. Oxford: Macmillan Education. (MED2)

Svensén, B. (2009). *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.

Turnbull, J. ed. (2010). *Oxford Advanced Learner's Dictionary of Current English*. 8th ed. Oxford: Oxford University Press. (OALD8)

Vrbinc, M. & Vrbinc, A. (2013). Differences in the inclusion and treatment of terminology in OALD3, OALD4 and OALD8. *Lexikos. 23*, 440-455.

Walter, E. & Woodford, K. eds. (2008). *Cambridge Advanced Learner's Dictionary*. 3rd ed. Cambridge: Cambridge University Press. (CALD3)

---

[i] Some dictionaries make the following distinction between *obsolete* and *archaic*: *obsolete* means that there is no evidence of use since 1755 (i.e., the date of publication of Samuel Johnson's *A Dictionary of the English Language*), whereas the label *archaic* is used to label a lexical item or a sense that used to be in common use but is now found only rarely and is suggestive of an earlier style or period. The labels *dated* or *old-fashioned* indicate that the word is still in use, but not by younger people, since it is associated with an earlier time.

[ii] Diamedial, diastratic, diaphasic and diatextual information all relate to style or stylistic level. The division into four separate groups is a more detailed presentation of style labels, but as a matter of fact, they could also be regarded as one uniform group.

ABOUT THE AUTHORS

Dr Marjeta Vrbinc is Associate Professor at the Faculty of Arts, University of Ljubljana, Slovenia. Her main research interests are contrastive lexicology and phraseology, practical lexicography, with the emphasis on English monolingual learner's dictionaries, and learners' dictionary skills. She has edited a German-Slovene dictionary and co-authored an English-Slovene dictionary.

Dr Alenka Vrbinc is Assistant Professor at the Faculty of Economics, University of Ljubljana, Slovenia. She has been active in the field of lexicography, primarily learner lexicography, but also in the field of phraseology. She has edited a German-Slovene dictionary and co-authored an English-Slovene dictionary.