Diasystematic Information in the "Big Five": A Comparison of Print Dictionaries, CD-ROMS/DVD-ROMS and Online Dictionaries

Marjeta Vrbinc (marjeta.vrbinc@ff.uni-lj.si), Faculty of Arts, University of Ljubljana, Ljubljana, Slovenia and Alenka Vrbinc (alenka.vrbinc@ef.uni-lj.si), Faculty of Economics, University of Ljubljana, Ljubljana, Slovenia

Abstract: The information provided by labels is called diasystematic information, which gives restrictions and limitations concerning the use of a lexical item. The focus of the study, the findings of which are presented in this contribution, is five British monolingual learner's dictionaries (OALD9, LDOCE6, COBUILD7, CALD4, MED2), which are often referred to as the "Big Five". The aims of the study are to compare the print edition and the electronic versions (CD-ROM/DVD-ROM and online dictionaries) of the same dictionary to see whether the lists of labels used in one particular dictionary coincide across versions of one and the same dictionary. Parallels are then drawn between dictionaries to determine similarities and differences in the use of labels providing different types of diasystematic information. Some of the most important findings of the study are that lists of labels differ in all three versions of one and the same dictionary and that some labels enumerated in the lists either are not used in the A–Z section at all or are used in a different form. Apart from that, some labels used in the dictionaries are so close in interpretation that the intended user will probably experience difficulty in distinguishing between them.

Keywords: DIASYSTEMATIC INFORMATION, TAXONOMIES OF DIASYSTEMATIC INFORMATION, LABELS, MONOLINGUAL LEARNER’S DICTIONARIES, PRINT DICTIONARIES, DICTIONARIES ON CD-ROMS/DVD-ROMS, ONLINE DICTIONARIES, RESTRICTIONS AND LIMITATIONS CONCERNING USE, LISTS OF LABELS IN FRONT MATTER, ACTUAL USE OF LABELS

Opsomming: Diasystematiese inligting in die "Groot Vyf": 'n Vergelyking van gedrukte woordeboeke, CD-ROMS/DVD-ROMS en aanlyn woordeboeke.

Die inligting wat deur etikette verskaf word, word diasystematiese inligting genoem. Dit dui die beperkings en begrensings rakende die gebruik van 'n leksikale item aan. In die studie, waarvan die bevindings in hierdie bydrae aangebied word, word gefokus op vyf Britse eetalige aanleerderswoordeboeke (OALD9, LDOCE6, COBUILD7, CALD4, MED2) wat dikwels die "Groot Vyf" genoem word. Die doel van die studie is om die gedrukte uitgawe en die elektroniese weergawes (CD-ROM/DVD-ROM en aanlyn woordeboeke) daarvan te vergelyk om vas te stel of die lyste etikette wat gebruik word in een spesifieke woordeboek ooreenstem met ander weergawes van die-

As one of the information categories in the dictionary entry, labels have been present in dictionaries for a very long time. The information provided by labels is called diasystematic information, which gives restrictions and limitations concerning the use of a lexical item. A comparison of different dictionaries reveals certain similarities and differences in the inclusion and treatment of diasystematic information that will be addressed in more detail in this contribution. Since labels give information on the connotative value of lexical items, they are of particular importance to non-native speakers of a language; thus, the focus of the study is five British monolingual learner’s dictionaries (MLDs), i.e., *Oxford Advanced Learner’s Dictionary of Current English (OALD9)*, *Longman Dictionary of Contemporary English (LDOCE6)*, *Collins COBUILD Advanced Dictionary of English (COBUILD7)*, *Cambridge Advanced Learner’s Dictionary (CALD4)* and *Macmillan English Dictionary for Advanced Learners (MED2)*, which are often referred to as the “Big Five”. All these dictionaries are either accompanied by a CD-ROM or DVD-ROM or they have an online version (free access and/or access by a unique PIN code). This study examines whether a user of different versions of one and the same dictionary gets the same information as regards the meaning and the use of labels, or whether one version provides information, whereas the other one lacks it. Before taking a closer look at the labels used in the “Big Five”, it is necessary to outline the theoretical background to diasystematic information provided in the form of labels.

### 1.1 Taxonomies of diasystematic information

The existing literature offers various taxonomies of diasystematic information proposed by different researchers. Mostly, they vary in degree of complexity, but there is also a great deal of overlapping. Jackson (2002: 109-115), for instance, enumerates seven types of usage labels, i.e., dialect; formality; status;
effect; history; topic or field and disputed usage. Landau (2001: 217-272), however, classifies eight common kinds of usage information, i.e., 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 and status or cultural level. Atkins and Rundell (2008: 182-186) distinguish nine types of linguistic labels: domains; region; dialect; register; style; time; slang and jargon; attitude and offensive terms. The most detailed classification can be found in Hausmann (1989: 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: 131-134) and Svensén (2009: 326-332) and is also taken as a theoretical background in our research. Below, the classification proposed by Hausmann (1989: 651) is presented, and parallels with other classifications enumerated above are drawn:

1. diachronic information: this 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. The most common temporal labels found in contemporary dictionaries are old-fashioned, obsolete, archaic, old use or dated. This group is referred to as history by Jackson (2002), currency or temporality by Landau (2001) and as time by Atkins and Rundell (2008).

2. diatopic information: this associates a word or one of its senses with a particular regional dialect or national variety. Most British dictionaries nowadays include words or senses typically used in different varieties of English. Sometimes regional areas within a country are specified; thus, regional or dialect is used as a label. Jackson (2002) refers to this group as dialect, but says that "dialect labels refer to geographical restrictions, and we can take this to include both national varieties and regional dialects within a national variety" (ibid: 110). Landau (2001) calls diatopic information regional or geographic variation, while Atkins and Rundell (2008) divide it into two separate groups: region and 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. Monolingual dictionaries usually provide information on the language of origin, mostly for words that have retained their original form (e.g., from Latin, from French, from Italian). Interestingly, diaintegrative information is not included in any of the other classifications and can also be regarded as belonging to etymology as a different information category.

4. diamedial information: this associates a word or one of its senses with a particular medium of communication. The most common labels are written and spoken. Apart from Hausmann, Landau (2001) is the only scholar who includes diamedial information in his classification of diasystematic
information and refers to it as style, functional variety, or register.

5. diastatric 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. The most common labels are slang, vulgar and taboo. Diastratic information is included in all the taxonomies: Jackson (2002) calls it status, Landau (2001) divides it into two groups, i.e., restricted or taboo and scatological usage and slang, and Atkins and Rundell (2008) also have two groups for diastratic information, i.e., slang and jargon and offensive terms.

6. diaphasic information: this associates a word or one of its senses with a particular register of a language, the most common labels being formal and informal. Diaphasic information is also included in all the taxonomies: Jackson (2002) calls it formality, Landau (2001) refers to it as style, functional variety, or register, and Atkins and Rundell (2008) call it register.

7. diatextual information: this associates a word or one of its senses with a particular discourse type or genre. The most common labels are poetic and literary. Diatextual information is disregarded by Jackson (2002) but is referred to as style, functional level, or register by Landau (2001) and as style by Atkins and Rundell (2008).

8. diatechnical information: this 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 present 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/specialized/specialist instead of giving detailed information on specific subject fields. Diatechnical information is included in all classifications: Jackson (2002) calls it topic or field, Landau (2001) technical or specialized terminology and Atkins and Rundell (2008) refer to it as domains.

9. diafrequent information: this associates a word or one of its senses with a particular frequency of occurrence. Labels used to indicate frequency are less frequent and rare. It should be stressed that dictionaries label only less frequent items, which means that unmarked items are more common. Hausmann’s classification is the only classification that includes diafrequent information — all other scholars disregard it altogether.

10. diaevaluative information: this 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 and euphemistic. Diaevaluative information is one of the categories included in all classifications studied: Jackson (2002) calls it
effect, Atkins and Rundell (2008) attitude, while Landau (2001) divides it into two groups depending on the value a particular label expresses. Labels denoting positive connotation (approving, humorous) belong to style, functional variety, or register, while those expressing negative connotation (offensive, derogatory, disapproving) belong to the group referred to as insult.

11. dianormative information: this 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 and disputed. In other words, the acceptability of items marked with one of these labels is questionable as regards linguistic correctness. Disputed usage is the term used by Jackson (2002) to describe dianormative information, while Landau (2001) refers to it as status or cultural level. Atkins and Rundell do not include this group in the classification of diasystematic information.

Table 1 below summarizes these four classifications.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>diachronic</td>
<td>history</td>
<td>currency or temporality</td>
<td>time</td>
</tr>
<tr>
<td>diatopic</td>
<td>dialect</td>
<td>regional or geographic variation</td>
<td>— region; — dialect</td>
</tr>
<tr>
<td>diaintegrative</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>diamedial</td>
<td>/</td>
<td>style, functional variety, or register</td>
<td>/</td>
</tr>
<tr>
<td>diastatic</td>
<td>status</td>
<td>— restricted or taboo and scatological usage; — slang</td>
<td>— slang and jargon; — offensive terms</td>
</tr>
<tr>
<td>diaphasic</td>
<td>formality</td>
<td>style, functional variety, or register</td>
<td>register</td>
</tr>
<tr>
<td>diatextual</td>
<td>/</td>
<td>style, functional variety, or register</td>
<td>style</td>
</tr>
<tr>
<td>diatechnical</td>
<td>topic or field</td>
<td>technical or specialized terminology</td>
<td>domains</td>
</tr>
<tr>
<td>diafrequent</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>diaevaluative</td>
<td>effect</td>
<td>— insult;</td>
<td>attitude</td>
</tr>
</tbody>
</table>
Table 1: Summary of classifications.

<table>
<thead>
<tr>
<th>dianormative</th>
<th>disputed usage</th>
<th>status or cultural level</th>
<th>— style, functional variety, or register</th>
</tr>
</thead>
</table>

The label *figurative* is used as a label in many dictionaries, but it does not fit 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 for a word to acquire semantic features that might not have been present in the original meaning (Atkins and Rundell 2008: 289). This label typically marks examples illustrating the use of the lemma or one of its senses and can thus be said to denote secondary senses that have the status of conventional metaphors (Hanks 2006: 28). 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.

The label *trademark* is also used quite frequently. According to Landau (2001: 218), this label provides diatechnical information, which he calls technical or specialized terminology, but such a classification of this label can be disputed, since it expresses the origin of the lexical item without any connection whatsoever with technical or specialized terminology, for example, Kleenex, Levi’s, Lycra, Polaroid, Rolex, Sellotape, Skype, Tupperware, Teflon, Thermos, etc. On the other hand, lexical items marked as trademarks can also belong to terminology typical of a specific subject field, for example, AZT or Prozac (medical or pharmaceutical terms). In such cases, the user would profit from getting the information concerning the subject field rather than getting the information on the origin of the lexical item.

Against this theoretical background, the aims of this study are: (1) to present a more detailed view of current practices employed in the British monolingual learner’s dictionaries under investigation, in which labels of various kinds are abundantly provided; (2) to compare the print edition and the electronic versions (CD-ROM/DVD-ROM and online dictionaries) of the same dictionary to see whether the lists of labels used in one particular dictionary coincide across versions of one and the same dictionary; (3) to draw parallels between dictionaries to determine similarities and differences in the use of labels providing different types of diasystematic information.

2. Labels in print, CD-ROM/DVD-ROM and online versions of the "Big Five"

2.1 Labels in OALD9

In the print edition of OALD9, the labels are listed on the inside front cover
under the title "Labels used in the dictionary". They can be found in two columns: the first one contains twelve labels "used with words that express a particular attitude or are appropriate in a particular situation" (OALD9: inside front cover), i.e., approving, disapproving, figurative, formal, humorous, informal, ironic, literary, offensive, slang, specialist and taboo, whereas the second one includes five labels that "show other restrictions on the use of words" (ibid), i.e., dialect, old-fashioned, old use, saying and TM. A closer look at the first twelve labels reveals that they belong to different classes of diasystematic information. The labels approving, disapproving, humorous, ironic and offensive express diaevaluative information; formal and informal provide diaphasic information; literary, diatextual information; slang and taboo, diastratic information and specialist expresses diatechnical information, whereas figurative belongs to none of the categories mentioned in Section 1.1. The labels enumerated in the OALD9’s second column again provide different types of diasystematic information: the labels old-fashioned and old use express diachronic information and dialect, diatopic information, whereas saying expresses the type of lexical item rather than its connotative value; it is thus not taken into account in any of the taxonomies discussed in Section 1.1. The last label listed in the second column is TM, which gives information on the origin of the lexical item (cf. also Section 1.1). Apart from that, fifteen labels expressing diatopic information are listed under “Abbreviations used in the dictionary”: AustralE, BrE, CanE, E AfrE, IndE, IrishE, NAmE, NEngE, NZE, S AfrE, ScotE, SE AsianE, US, WAfrE, WelshE. As has been mentioned, the label dialect, which also belongs to the group of labels providing diatopic information, is, however, listed under "Labels used in the dictionary” and not under "Abbreviations used in the dictionary”.

Contrary to the CD-ROM of the previous, i.e., the 8th edition, the DVD of the 9th edition does not contain the Advanced Search option, which makes it impossible to compare the list of labels on the DVD with those given in the print and online editions. Among the labels expressing diatechnical information, the print dictionary lists only one very general label, i.e., specialist. The label specialist has been introduced into the 9th edition, while the label technical was used in previous editions. The list of labels in the online edition, however, still includes technical instead of specialist, although the label specialist is used in the A–Z section, an error that should be corrected in the online edition. A closer observation of the A–Z section, however, shows that besides the very general label specialist, numerous other subject-field labels are abundantly provided, though not listed in the front matter. For instance:

antilogarithm — mathematics
polygon — geometry
antimatter — physics
melanoma — medical
anus — anatomy
antivirus — computing
meltdown, sense 2 — economics
niches, sense 2 — business
polymer — chemistry
polyphony — music
polysemous — linguistics
magnitude, sense 2 — astronomy
magnitude, sense 3 — geology
chromosome — biology
fiduciary — law

It has to be stressed that the online version "Oxford Learner's Dictionaries" is the only online dictionary among the dictionaries studied that includes information on the labels. This can be accessed via a tab entitled "About", where the "Guide to Symbols and Labels" section can be selected and the section "Labels used in Oxford Learner's Dictionaries" provides information on the labels. Interestingly, the lists of labels as well as explanations of each individual label are identical and the examples illustrating each label are almost identical in the print dictionary and in the online dictionary. The online dictionary, however, does not include the section Abbreviations, which can be found on the inside front cover of the print dictionary, which means that this information is completely lacking in this version of OALD. Although these labels are not provided in the "Guide to Symbols and Labels", they are used in the online dictionary in exactly the same way as in the print dictionary.

2.2 Labels in LDOCE6

Similar to OALD9, LDOCE6 also lists labels on the very first page of the dictionary. Here, the labels are subdivided into three categories: Words which are used only or mainly in one region or country (BrE, AmE, AusE), Words which are used in a particular situation, or show a particular attitude (approving, disapproving, formal, informal, humorous) and Words which are used in a particular context or type of language (biblical, law, literary, medical, not polite, old-fashioned, old use, spoken, taboo, technical, trademark, written). The first category of labels is homogeneous, since all the labels express diatopic information; this cannot be claimed for the second and third categories. The second category includes labels providing diaphasic (formal, informal) and diaevaluative information (approving, disapproving, humorous). The third category is the most heterogeneous one, since the labels law, medical and technical provide diatechnical information; literary and biblical, diatextual information; not polite, diaevaluative information; old-fashioned and old use, diachronic information; spoken and written, diamedial information; and taboo, diastatic information. For the classification of trademark, see Section 1.1.

The total number of labels included in LDOCE6 is twenty, which makes it the dictionary with the fewest labels among all the five dictionaries under consideration. Among the labels expressing diatechnical information, the generic
Marjeta Vrbinc and Alenka Vrbinc

label technical and two more specific labels, i.e., law and medical, are used. In the online version offered by the access code, on the other hand, all labels are listed in the Advanced Search option under Register. The labels that are listed in the print dictionary under “Words which are used in a particular situation, or show a particular attitude” and “Words which are used in a particular context or type of language” coincide with the labels under Register in the online version except for the labels approving and disapproving, the labels which were added in the print edition of LDOCE6 but were not included in the previous (i.e., the 5th) edition. The labels expressing diatopic information provided in the print edition under “Words which are used only or mainly in one region or country” cannot be found among the labels in the online version, which does not, however, imply that they are not used in this version. On the contrary, they are used in the same way as in the print dictionary. The free online dictionary seems promising at first sight, since it offers the tab “How to use the LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH ONLINE”. Here, the user would most certainly expect to find instructions as to the use of all features found in the dictionary, including labels. Disappointingly, the labels cannot be found, nor are they explained, but they are used in the dictionary in the same way as in the print dictionary and in the online version offered by the access code.

2.3 Labels in COBUILD7

The first thing one notices when searching for metalinguistic information in COBUILD7 is that the online dictionary accessed by a unique code provides no information on diasystematic information or any other kind of metalinguistic information. The consequence is that no comparison can be made, but at the same time, it can be claimed with a high degree of certainty that labelling is done in much the same way in both versions of the COBUILD dictionary. The labelling described in this contribution is therefore based on the print edition.

If compared to the other four MLDs, COBUILD7 is clearly the only dictionary with quite extensive front matter. In the other four dictionaries, front matter is reduced to a list of labels and/or other abbreviations used in the dictionary and a graphic presentation of the dictionary entry taken from each individual dictionary. COBUILD7, however, differs greatly in this respect. On pages xiii-xv of the front matter, the user gets precise information on Style and Usage (p. xiii-xiv) and Pragmatics (p. xiv-xv). The Style and Usage section has two subsections: Geographical labels and Style labels. The Geographical labels subsection lists two labels, i.e., Brit and Am, and provides brief explanations of them. Then follows a list of other geographical labels used to mark a lexical item typically used in other varieties of English, such as Australian, Irish, Northern English and Scottish. Interestingly, these labels are introduced by ‘e.g.’, implying that not all the labels used in the A–Z section are listed here. Nevertheless, all these labels provide diatopic information. The Style labels section gives twenty labels in total (business, computing, dialect, formal, humorous, infor-
mal, journalism, legal, literary, medical, military, offensive, old-fashioned, rude, spoken, technical, trademark, very offensive, very rude, written), but a close inspection of these labels shows that not all the labels listed belong to "style labels" as they are referred to in the title of this section. Among these labels, we can find business, computing, legal, medical, military and technical, which are obviously subject-field labels, i.e., labels expressing diatechnical information. The label dialect belongs to the same group as Brit and Am mentioned above; the labels formal and informal express diaphasic information; humorous, offensive and very offensive, diaevaluative information; journalism and literary, diatextual information; rude and very rude, diastatic information; spoken and written, diamedial information and old-fashioned, diachronic information. For the classification of trademark, see Section 1.1.

The Pragmatics section starts by explaining what pragmatics is and expands on how pragmatic information is included in the dictionary. The "pragmatics labels" include approval, disapproval, emphasis, feelings, formulae, politeness and vagueness, and it is obvious that some of these coincide with "style labels" as used in other dictionaries, i.e., approval with approving, disapproving with disapproving, polite word/expression with polite.

### 2.4 Labels in CALD4

In the print edition of CALD4, the labels are found under the title "Style and usage labels used in the dictionary". The list includes 31 labels (abbreviation, approving, Australian English, child’s word/expression, disapproving, female, figurative, formal, humorous, informal, Indian English, Irish English, legal, literary, male, Northern English, not standard, offensive, old-fashioned, old use, polite word/expression, saying, Scottish English, slang, South African English, specialized, trademark, UK, US, written abbreviation, A1, A2, B1, B2, C1, C2). The labels express different types of diasystematic information: diatopic information (Australian English, Indian English, Irish English, Northern English, Scottish English, South African English, UK, US); diaevaluative information (approving, disapproving, humorous, offensive, polite word/expression); diachronic information (old-fashioned, old use); diaphasic information (formal, informal, child’s word/expression); diastatic information (slang); diatextual information (literary); diatechnical information (legal, specialized); dianormative information (not standard), and diafrequent information (A1, A2, B1, B2, C1, C2). For the classification of trademark and figurative, see Section 1.1, and for the classification of saying, see Section 2.1.

Several labels stand out as they are not labels proper, meaning that they do not fall into any of the categories of labels expressing diasystematic information. Two of the labels peculiar to CALD4 are female and male, which are listed but not explained in the front matter. This means that the user cannot find any information about what they mark. It should be stressed that female and male are actually not labels proper, since they introduce the feminine or the masculine form of the lemma: for example, female lioness in sense 1 of the lemma lion. On the CD-ROM, we are faced with the opposite situation. The
two labels are not enumerated in the list of labels, but if the user comes across them in the A–Z section, s/he gets a brief explanation about their meaning by simply placing the cursor over the label: for example, female is explained as ‘only applies to women’. The labels abbreviation and written abbreviation should also not go unmentioned, since they indicate the type of lemma rather than functioning as labels proper and are usually treated as part-of-speech indicators.

As in OALD9, there are discrepancies between the print edition and the CD-ROM version of CALD4 in that the print edition gives two labels not included in the list of labels on the CD-ROM, i.e., female and male. An obvious discrepancy concerns the labels expressing diatopic information. In the print dictionary, this category is represented by the eight labels mentioned above, while the CD-ROM version offers two labels (British English only and American English only) and a third choice ‘other regions’. The search results for British English only show that the lemmata or their senses marked with the label UK are obtained, and similarly the search for American English only yields those lexical items labelled US. As is to be expected, the possibility ‘other regions’ yields all other regional varieties, which are more precisely listed in the print dictionary. With respect to labels expressing diatopic information, it can be established that the A–Z section shows certain inconsistencies when compared with the lists of labels: labels not found in any of the lists are used in the A–Z section. If we compare the following two lemmata

anyroad NORTHERN FOR anyway
bairn SCOTTISH ENGLISH OR NORTHERN ENGLISH a child

we can see that "anyroad" is marked Northern, which is a label not found in any of the lists, while “bairn” is marked Northern English, which is a label included in the list in the print edition; besides that, it is self-explanatory; thus any explanation seems to be redundant.

2.5 Labels in MED2

The print edition of MED2 lists the labels under three titles:
— Style and attitude labels (twelve labels: formal, humorous, impolite, informal, literary, offensive, old-fashioned, showing approval, showing disapproval, spoken, very formal, very informal);

— Subject labels (eighteen labels: art, astronomy, biology, business, chemistry, cinema, computing, economics, legal, linguistics, literature, maths, medical, music, physics, science, theatre, tourism);

— Regional labels (three labels, i.e., British, American, mainly American, are enumerated and explained; in contrast, eleven labels, i.e., Australian, Irish, Scottish, Canadian, Caribbean, East African, Indian, New Zealand, South African, Welsh, West African, are listed but no explanation is provided).

These titles are to be found on the inside front cover, as is the case in most of its competitor dictionaries. The labels formal, informal, very formal and very informal express diaphasic information; showing approval, showing disapproval, offensive, humorous and impolite, diaevaluative information; old-fashioned, diachronic information; spoken, diamedial information; literary, diatextual information, whereas diatechnical information is provided by all the labels listed under Subject labels and diatopic information by all the labels enumerated under Regional labels. In the print edition, the labels under Style and attitude labels are explained, whereas all others are considered to be obvious and self-explanatory. The same holds true of the treatment of labels on the CD-ROM accompanying the print edition, where the labels can be found under Study pages, About the dictionary, Style and attitude labels. The online version, on the other hand, lacks lists of labels as well as explanations of labels.

It is important to point out that this is the only dictionary which gives a list of subject-field labels and does not use the generic label technical/specialized/specialist. A comparison of the lists in the print dictionary and the lists in the Super Search under the Advanced Search option on the CD-ROM shows only minor divergences. Apart from the labels listed in the print edition under Style and attitude labels, the electronic version offers the following additional labels: often humorous, journalism, mainly journalism and mainly literary. As regards labels expressing diatopic information, two labels enumerated in the print dictionary are absent from the list on the CD-ROM, i.e., East African and Welsh. Apart from that, one and the same label is given as Indian English on the CD-ROM and as Indian in the print dictionary. The variant given on the CD-ROM, i.e., Indian English, is also used in the A–Z section of the print dictionary. The subject-field labels listed in the print dictionary and on the CD-ROM mostly coincide, the only difference being that the label trademark is listed only on the CD-ROM but is lacking in the print edition.

The online dictionary does not provide any guidance whatsoever as to the labels used in the dictionary.
3. Labels across the dictionaries studied

In this section, we would like to make a survey of labels included in our research in relation to whether they can be found in all five dictionaries under investigation or whether they appear in one or two dictionaries, but are not used by the lexicographers of other dictionaries. It seems logical to start with labels that can be found in all five dictionaries. There are seven such labels: approving, disapproving, formal, humorous, informal, literary and old-fashioned. In MED2, approving and disapproving are expressed as showing approval and showing disapproval and in COBUILD7 as approval and disapproval. One label appears in all dictionaries except LDOCE6: offensive. Five labels can be found in three dictionaries: law/legal (legal in CALD4 and COBUILD7, law in LCODE6), old use (OALD9, LDOCE6, CALD4), spoken (LDOCE6, COBUILD7, MED2), technical/specialized/specialist (technical in LDOCE6, specialized in CALD4, specialist in OALD9) and trademark (OALD9, LDOCE6, CALD4). As many as ten labels are used in two dictionaries: figurative (OALD9, CALD4), not polite/impolite (not polite in LDOCE6, impolite in MED2), polite/politeness (polite in CALD4, politeness in COBUILD7) medical (LDOCE6, COBUILD7), saying and slang (OALD9, CALD4), taboo (OALD9, LDOCE6), written (LDOCE6, COBUILD7), dialect (OALD9, COBUILD7) and journalism (CALD4, MED2). If we disregard labels typical of MED2 and COBUILD7 which only express intensification or frequency of one and the same label (MED2: very formal, very informal, often humorous, mainly journalism, mainly literary, mainly spoken; COBUILD7: very offensive, very rude), all other labels listed in the front matter of the dictionaries under consideration are used in one dictionary only:

OALD9: ironic;
LDOCE6: biblical;
COBUILD7: business, computing, military, rude;
CALD4: not standard, abbreviation, written abbreviation, child’s word/expression, female, male, A1, A2, B1, B2, C1, C2.

As far as the subject-field labels are concerned, a more extensive list can be found only in MED2 (see Section 2.5), whereas other dictionaries give just a few subject-field labels, a situation which does not facilitate an apt comparison:

OALD9: specialist (the A–Z section, however, includes several subject-field labels, see Section 2.1);
LDOCE6: law, medical, technical;
COBUILD7: business, computing, legal, medical, military, technical;
CALD4: legal, specialized.

If we compare the labels expressing diatopic information, we can see that there is a great deal of overlapping, since the labels British English, American English and Australian English are used in all five dictionaries; Irish English and Scottish English in four (OALD9, COBUILD7, CALD4, MED2); South African English
Diasystematic Information in the "Big Five"

(OALD9, MED2, CALD4) as well as Northern English (OALD9, COBUILD7, CALD4) appear in three dictionaries; whereas New Zealand English, West African English and Welsh English are used in two dictionaries (OALD9, MED2) and the same holds true for Canadian English and East African English (OALD9, MED2); South-East Asian English and English from the United States (OALD9), Indian English (CALD4) as well as Caribbean English (MED2) are used in only one dictionary under consideration.

4. Discussion

4.1 Interpretation and intelligibility of labels easily confused

If we refer back to Section 3, we can see that five labels (formal, informal, literary, humorous and old-fashioned) are used to mark lexical items of various kinds 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, especially if they are used together, which is often the case. This 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, 2nd edition (Dictionary of Standard Slovene 2014: §133) explains that the label knjižno is assigned to 'words, senses or phraseological units used particularly in fiction or in scientific texts'. As a matter of fact, 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 CALD4, 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 sufficiently clear for a foreign learner to distinguish accurately between these two labels, because literary is explained as labelling formal language. The addition of 'in literature' to 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 meta-
phorical 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 the 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: 229).

Another label used only in LDOCE6 and COBUILD7 is written. The explanation provided in the print edition of LDOCE6 says that it refers to "a word or phrase that is used only, or nearly always, in written English". 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", we can see that they are both explained in a similar way. Therefore, the question can be posed whether a learner of English would spot the difference between these two labels. Apart from the difference between written and formal, another pair is used in LDOCE6 and COBUILD7 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 explanations 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 sufficiently clear are the explanations in LDOCE6:

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"

The same holds true for the explanations found in COBUILD7:

informal — "used mainly in informal situations, conversations, and personal letters, e.g. pep talk"
spoken — "used mainly in speech rather than in writing, e.g. pardon"

Another problem concerns the interpretation of the labels old-fashioned (used in all five dictionaries) and old use (used in OALD9, LDOCE6 and CALD4). 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. COBUILD7 and MED2 use only the label old-fashioned, and considering the fact that the intended user may not be able to understand the difference between the two labels, this seems a sensible decision. The labels old-fashioned and old use are definitely labels that should be used with care in dictionaries if both labels are used in one and the same dictionary.

4.2 Labels expressing opposing connotative values

Sets of labels expressing opposing connotative values (formal and informal, polite and impolite, spoken and written) are also 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 LDOCE6 and MED2, but interestingly, neither of these dictionaries has the opposing label polite. Polite, on the other hand, is used in CALD4 as well as in COBUILD7 (here, the label politeness is used), but neither of them has the label impolite. This means that the entire pair polite/impolite is not used in any of the dictionaries included in our study. From the point of view of dictionary users, this may indicate inconsistency or a failure to mark the opposing value of the 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: 93). Similarly, the label spoken can be found in MED2, and one would rightly expect the label written to be used in the same dictionary, but this is not the case. LDOCE6 and COBUILD7, on the other hand, also use the label spoken, but the label written can also be found. It seems just as illogical as in the case of the pair polite/impolite to find spoken without its opposing label written, which is in line with Fedorova (2004: 269). What is more, MED2, which lacks the label written, has a very precise labelling of spoken connotation, since it uses two labels to mark that a lexical item or one of its senses belongs to spoken language: spoken and mainly spoken. Obviously, the question can be posed why it is necessary to label one concept so precisely while omitting the other concept completely.

4.3 Labels expressing diatechnical information

Another interesting feature is that MED2 is the only dictionary that gives a list of subject-field labels (cf. Section 2.5). In OALD9, for example, the only subject-field label enumerated in the list of labels is the generic label specialist, but as mentioned in Section 2.1, other subject-field labels are also used in the diction-
ary. Here, we are faced with a problem of higher-level domain markers (technical/specialized/specialist), on the one hand, and lower-level domain markers (mathematics, physics, biology, business, phonetics, etc.), on the other. This means that some lexical items are marked with a generic label without giving further details as to the specific subject-field of a particular lexical item, while others are labelled with very specific subject-field labels giving the user precise information about the subject-field where the lexical item functions as a term (cf. Vrbinc and Vrbinc 2013: 449, 454). The question is firstly, what is the criterion for deciding which label to use (a generic or a specific one) and secondly, what level of precision would benefit the end user. Is it enough to simply use the label technical/specialized/specialist just to indicate that the lemma or one of its senses thus labelled belongs to terminology? Or would the user expect to find the information about the precise subject-field or sub-field? These are the questions that cannot be answered without carrying out a user survey, which is beyond the scope of this study.

The opposite extreme is the use of labels denoting sub-fields (e.g., business vs. economy, medical vs. anatomy, mathematics vs. geometry). A general dictionary user cannot be expected to recognize the subtle differences between such closely related labels. If such labels are used, one would expect an explanation of the distinction between them, but taking account of the type of dictionary and the target audience, it can be claimed with a high degree of certainty that this is an unnecessary complication. In MLDs, one would expect that the subject-field labels would refer to fields of science only and would disregard the sub-fields. That means it is recommendable to use generic subject-field labels rather than more specific subject-field labels.

### 4.4 Labels expressing diatopic information

The labels indicating diatopic information are numerous, which is to be expected, given that the dictionaries claim to cover the vocabulary of the entire English-speaking world. The print edition of OALD9 gives three labels that need to be discussed in more detail: NAmE (North American English), US (English from the United States) and Canadian English. Although the difference between them might be obvious to a linguist or a native speaker of English, it is questionable whether a learner makes any real difference between these labels. The labels US and Canadian English may be understandable, but the label NAmE is most certainly not because it can be interpreted as a label encompassing both US and Canadian English. We dare to assume that an average user of a learner’s dictionary would equate the label NAmE with the label US or AmE (as used in other dictionaries); thus, the label NAmE would not do the job it intends to do. Similarly, the CD-ROM version of CALD4 offers two labels (British English only and American English only) and a third choice ‘other regions’. The search results for British English only show that the lemmata or their senses marked with the label UK are obtained; similarly, the search for American English only gives
those lexical items labelled US. The use of 'other regions' instead of giving a full range of labels found in the dictionary is also a feature typical of the CD-ROM accompanying the print edition of CALD4. A similar situation can be observed in MED2, which lists two variants of the same label, one in the print dictionary, i.e., Indian, and one on the CD-ROM, i.e., Indian English. In the A–Z section, however, the label Indian English is used in both versions of the dictionary. Such a discrepancy between the labels used in any of the lists (either in the front matter of the print dictionary or in the list provided within the advanced search options on the CD-ROM) and those actually used in the A–Z section should by all means be avoided. Interestingly, COBUILD7 does not list all the labels indicating the varieties of English but uses 'eg' to indicate that the list is incomplete. This policy is far from ideal, since a user cannot be supposed to know which other labels apart from those listed s/he can encounter in the A–Z section of the dictionary.

4.5 Register or style labels?

A general observation is that some MLDs list very different labels among what they call register or style labels, although they are far from being register or style labels, e.g., biblical, law, medical, technical, trademark (LDOCE6); abbreviation, female, male, saying, specialized, trademark (CALD4); business, computing, dialect, legal, medical, military (COBUILD7). This practice should be avoided and two possible solutions can be put forward. The first one is that labels should be grouped according to the connotative value they express, where the theoretical considerations proposed by metalexicographers within the framework of the classification of diasystematic labelling could be made use of. It is, however, true that considering the very specific target audience of MLDs, lexicographers try to simplify all dictionary components in order to make them more approachable and user-friendly. Taking this into consideration, all the labels used in a particular dictionary could be listed together in alphabetical order without any further division, and the title could simply be "Labels used in the dictionary".

4.6 Labels or not?

In COBUILD7, some "pragmatics labels" coincide with labels proper as used in other dictionaries, while some of them show no parallel with labels proper. Among the latter, three labels should be highlighted: emphasis, feelings and vagueness. The question is whether these convey sufficient information for the dictionary user to make full use of them. The print edition offers explanations for them, explanations which are lacking in the online dictionary. The following are the explications found in the front matter of COBUILD7 (xv):

emphais — "The label EMPHASIS indicates that you use the word or expres-
tion to show that you think something is particularly important or true, or to draw attention to it. An example of a word with this label is absolutely.”

feelings — “The label FEELINGS indicates that you use the word or expression to show how you feel about a situation, a person, or a thing. An example of a word with this label is unfortunately.”

vagueness — “The label VAGUENESS indicates that you use the word or expression to show lack of certainty. People often use vague language to make statements ‘softer’, so that what they say does not appear too direct or too strongly stated. Examples of vague language are presumably ..., Do you know what I mean?, kind of ..., and sort of ...”

The labels saying, TM/trademark, abbreviation and written abbreviation also deserve attention, since they cannot be considered labels proper; consequently, they do not fall into any of the categories for the classification of diasystematic information. These labels give information on the type of lexical item (saying, abbreviation, written abbreviation) or the origin of the lexical item (TM/trademark), but they themselves do not give any information on the connotative aspect of the lexical item they are used to mark. The question can therefore be posed why some dictionaries (print dictionaries or their CD-ROMs) list these labels under style or register labels. It is misleading for the user to find labels that cannot be regarded as style or register labels among other labels that address the connotative aspect of lexical items. Apart from that, phraseological units classified as sayings are included in the special idioms sections in CALD4. The usability of the label saying, however, remains questionable, since it is doubtful whether the user needs the information about the type of phraseological units. Besides that, it seems somehow illogical to label sayings and omit the labelling of formulae, catchphrases, collocations and other types of phraseological units. The precise classification of phraseological units affects neither the decoding nor the encoding process; it can thus be regarded as superfluous for the learner who is the intended user of a learner’s dictionary.

4.7 Too precise labelling of the same connotation

If we address the issue concerning the labels used in one dictionary only, we can see that MED2 and COBUILD7 stand out in that they both include as many as six labels that occur in no other dictionary (MED2: often humorous, very informal, journalism, mainly journalism, mainly literary and mainly spoken; COBUILD7: very offensive, very rude, emphasis, feelings, vagueness and formulae). 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, offensive vs. very offensive, rude vs. very rude. 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
Diasystematic Information in the “Big Five”

4.8 Labels included in the dictionary front matter but absent from the A–Z section

In LDOCE6, the labels approving and disapproving are listed in the print version but are lacking in the online version. This means that the Advanced search option yields no results as regards these two labels. A manual search in the print edition was rather unsuccessful, which means that we were unable to locate a lemma labelled approving or disapproving. Several lemmata, however, were found where the definition of a lemma or one of its senses said: "use this to show approval" (e.g., the lemma succinct is defined in the following way: clearly expressed in a few words — use this to show approval) or "used to show disapproval" (e.g., the lemma agitator is defined in the following way: someone who encourages people to work towards changing something in society — used to show disapproval). This means that the labels proper are not used, since labelling is integrated into the definition part. The question can therefore be raised why the dictionary front matter includes the labels even though they are most probably not used to label the (dis)approving connotative values of the lemmata or their senses.

5. Conclusion

As opposed to print dictionaries, online dictionaries (the only exceptions being OALD and CALD) are the most deficient as regards information on diasystematic labelling as well as other metalexicographic information, which is a finding that cannot be viewed with much optimism. Nowadays, in the age of modern technology, one can logically expect an increasing number of users for online dictionaries. It is true that currently, not everyone in the world enjoys good internet connectivity, but it is also true that access to the internet is becoming increasingly widespread, and this is expected to improve in the future. Also, the future of print dictionaries seems to be questionable (Macmillan, for example, announced in November 2012 that, in future, only online dictionaries will be available; http://www.macmillandictionaryblog.com/bye-print-dictionary), which means that the future of lexicography probably lies in online dictionaries. Consequently, users of online dictionaries need guidance as to certain aspects of dictionary use, and diasystematic information is doubt-
lessly an information category of the dictionary entry that needs to be explained in a comprehensive, clearly formulated, understandable and consistent way. Not all labels are self-explanatory; thus, it is essential for every type of dictionary, regardless of the medium, to provide its users with the necessary information. This is also in line with Norri (2000: 93), who says that "[i]n many cases greater care could be taken when explaining the method of labelling to the reader. All too often, the preface provides scant guidance in this important matter".

Another very important aspect is that all labels listed in the front matter of print dictionaries or in the CD-ROM or online versions should be used in the same form in the A–Z section. To have one form of a label in a list of labels and another form in the A–Z section is undesirable and most user-unfriendly and should certainly be avoided by compilers of all dictionaries, regardless of the target audience. Apart from that, all labels should be followed by a detailed explanation — one that the intended user will understand. In cases, where two different labels express similar connotative values, special care should be taken to explain the subtle differences as precisely as possible. The explanations should be short, concise and to the point, since a user should understand them immediately. Lexicographers should also avoid listing labels that are not used in the A–Z section (e.g., the labels approving and disapproving in LDOCE6), or omit labels that are used only sparsely in the entire A–Z section (e.g., the label biblical is used only seven times in LDOCE6).

Information about the restrictions and limitations provided by diasytematic information is vital for decoding and even more so for encoding. It should therefore be tailored to the needs and skills of the intended dictionary users, so that they can use this type of information correctly and efficiently. We should bear in mind that the intended user is a foreign learner who should receive clear guidance as to the use of lexical items s/he looks up in a dictionary. Consequently, dictionaries would benefit from theoretical classification of labels for their presentation in the front matter or in the electronic versions. To conclude, the possible problems that deficient (treatment of) diasytematic information could have for users should be regarded as a set of hypotheses that still have to be tested empirically, which should be a logical follow-up to this study.

**Bibliography**

**Dictionaries**

Diasystematic Information in the "Big Five"


Other literature


Internet Reference