Metalinguistic labels in online English MLDs – theoretical and practical considerations

Magdalena Tomaszewska ✉
Uniwersytet w Siedlcach/University of Siedlce
https://orcid.org/0000-0003-2710-7396
magdalena.tomaszewska@uws.edu.pl

Abstract

The article presents theoretical and practical considerations on metalinguistic labels in online English Monolingual Learner's Dictionaries (MLDs). Initially, it discusses traditional rationale for employing labeling conventions in such sources. Further, it considers the application of metalinguistic labels in five prevailing online MLDs. The labels in relevant noun-related headwords are presented in general, the ones accompanying nominal lemmata are analyzed in detail on the basis of headword entries in the online version of the *Cambridge Dictionary*. Finally, the present contribution places labeling strategies in the context of teaching or developing skills in dictionary use, and argues in favor of more in-depth cooperation on unifying labeling strategies and presenting them explicitly to the user.

Keywords: metalinguistic dictionary labels, online Monolingual Learner's Dictionaries, MLDs, nouns, *Cambridge Dictionary*

1. Introduction

English Monolingual Learner's Dictionaries (MLDs) are offered as either traditional commercial print works that can be supplemented with CD- or DVD-ROMs, or open-access sources available on the Internet. Contemporarily, learners of English have a wide range of dictionaries at their disposal, including online

versions of the well-established print editions, e.g., the *Cambridge Dictionary* (*CD*), the *Collins English Dictionary* (*CED*), the *Merriam-Webster English Dictionary* (*MWED*), the *Longman Dictionary of Contemporary English* (*LDOCE*), and the *Oxford Advanced Learners' Dictionary* (*OALD*)[1]. Although online MLDs have been under scrutiny in general (Heuberger, 2020; Lew, 2011), the present contribution focuses on various labeling strategies and discusses their application in the selected dictionaries. A broader perspective is taken on nominal lemmata in the online version of the *Cambridge Dictionary*. The discussion of theoretical and practical aspects is accompanied by considerations about the end users – learners of English. Thus, the issue of teaching or developing dictionary skills is also briefly referred to.

The article is intended to complement previous research into dictionary headword entries. The study makes use of the comparative and empirical methods. Alongside discussing the concepts underlying dictionary making, the present analysis compares and contrasts representative noun-related headwords in the five dictionaries. The data presented in the article are based on the author's own observations and experience as a language learner and teacher.

2. Metalinguistic labels in digital English MLDs

It is common for dictionaries to make use of their own specific conventions of labeling. Introducing labels is motivated by space saving and lexicographic tradition (Lew, 2015, p. 1). Although "[i]n all kinds of monolingual dictionary, basic grammar is supplied in the form of wordclass markers" (Atkins & Rundell, 2008, p. 400), it has been a common practice to develop individual dictionary-specific metalanguage (labels, abbreviations and symbols). Only very few distinctions are marked by means of the same symbols (e.g., *C* for 'countable use of a noun' and *U* for 'uncountable use of a noun'). Furthermore, labels coding diasystematic information (term according to Vrbinc & Vrbinc (2015)) such as: *formal*, *informal*, *humorous*, *literary* and *old-fashioned* are used in the best known MLDs on CD-ROMs (Vrbinc & Vrbinc, 2015, p. 116), with *literary*, *old-fashioned* and *old use* being preferred as indicators of obsolescence (own calculations based on Norri, 2022, p. 404)[2]. Commonly accepted uniformity, though, is non-existent. "Coding

---

[1] Besides (a) being MLDs available online and (b) having recognised book counterparts, the five mentioned dictionaries provide the reader with extensive dictionary entries including (c) usage examples. The online dictionaries that do not fulfill the three criteria (e.g. *Chambers 21st Century Dictionary*, see https://chambers.co.uk/, or *Dictionary.com*, see http://www.dictionary.com, etc.) were not considered in the study.

[2] The three labels may, however, be used within a single dictionary, which is the case of *OALD*. The labels there are claimed not to be mutually interchangeable as (a) *literary* refers

systems are among the features where the dictionaries on the market are most distinct, and arguably one in which the greatest improvements have been made in the past few decades" (Heuberger, 2016, p. 36).

Since the lexicon is analyzable not only linguistically (by means of graphic, phonetic, morphological, syntactic and semantic analysis), but also extra-linguistically (by means of cognitive, pragmatic, psychological and sociological analysis), "[i]nformation concerning each of these levels can and must be codified (according to different theories) in a lexicon, especially in the case of computerized lexicons" (Calzolari, Picchi, & Zampolli, 1987, p. 68). In other words, dictionary makers have been adopting different labeling strategies according to various motivations (e.g., to save space and organize different types of information concisely, to differentiate their product, to unify analyses within a theoretical framework, etc.). The underlying principles of labeling conventions in MLDs are simplicity and selectiveness, because such sources tend to be learner-oriented, even at the cost of theoretical accuracy. Dziemianko (2006, p. 5) points out that "the ease of accessibility is difficult to reconcile with the accuracy of description, since, as a rule, the more detailed the information, the more elaborate the system of presenting it".

Although encoding is of merit to lexicographers, decoding surface data can pose problems to dictionary users. On the one hand, learners "only reluctantly read the prefatory matter and are not prepared to study the coding systems adequately" (Heuberger, 2016, p. 35). On the other hand, Dziemianko (2008) suggests that there is a strong tendency for intermediate and advanced students of English to use and appreciate (especially verbal) codes even if the codes are more complex than the mainstream ones. In Dziemianko's study, the students were presented with substitutes (less known lexical items of English), accompanied by semantic and grammatical information on their more regular counterparts. What was of importance in the study was that (a) students could not rely on their *a priori* knowledge of the items, (b) successful completion of the task depended on understanding the codes that the students were exposed to.

Furthermore, the factor that significantly impacts the learners' reluctance to consult and understand the coding systems, at least in the case of online dictionaries, is that the users are not explicitly exposed to them (the lists of labels and codes are either hidden or simply lacking). It seems that publishing either a list of the labels, or a hyperlink to such a list, in a visible place, for example on the dictionaries' homepages, could raise the users' awareness of the dictionary-specific conventions. Additionally, the explanations accompanying labels must

---

to items that are "used mainly in literature and imaginative writing", (b) *old-fashioned* labels items that "are passing out of current use", and (c) *old use* is restricted to items "that are no longer in current use" (see https://www.oxfordlearnersdictionaries.com/about/english/labels).

be consistent and cohesive for the end user. All in all, fewer abbreviations mean greater accessibility (especially from the point of view of less advanced students), but codes are beneficial, on the condition they are understandable, applied consistently, and the learner is aware of their existence and motivated to study them.

3. Metalinguistic labels in noun-related and nominal headwords in online MLDs

Dictionaries use a limited number of formal and functional labels or codes, which generally fall into two categories. The newer approach results in the use of lexical descriptors (e.g., *noun*, *verb*, *plural*, *transitive*, *often attributive*, etc.), the older approach is preserved in abbreviated items (e.g., *T*, *obj*, *OPP*). The five online dictionaries selected for the present analysis make use of both strategies, and their coding conventions are discussed on the basis of the labels applied in the headword entry for the adjective *uncountable* and the related entries. Such an approach guarantees uniformity and comparability.

LDOCE and *OALD* do not explain metalinguistic terms in a separate help section, but incorporate them as headwords[3]. The dictionaries recognize only the grammatical sense of *uncountable* and, consequently, their definitions of the adjective refer to uncountable nouns. *OALD*, however, has the separate headwords *uncount noun* and *mass noun*. Although *LDOCE* has the separate headword *noncount*, which cross-references to *uncountable*, the reverse direction of referencing is not provided. In *OALD*, the separate headwords *non-count* and *uncountable* cross-refer to each other. In fact, the definition in both entries is analogous and is reiterated in the headword *mass noun*. Both sources additionally inform that the opposite of *uncountable* is the adjective *countable*, and *OALD* further encourages the user to check the difference between *uncountable* and *countless*. Although *LDOCE* informs that uncountable nouns in the dictionary are marked with the acronym [U], in fact, they are not. In the headword entries for *furniture*, *happiness*, etc., *LDOCE* uses the lexical label [uncountable].

MWED, *CED* and *CD* explain linguistic terms in their headword entries, and in the supplemental help or grammar sections. Among the five discussed sources,

---

[3] Additionally, in *LDOCE* each headword entry for a grammar term shows the link "Related topics: Grammar", which redirects to the headword *grammar*. In this particular entry, the user is given a possibility to get a first glimpse of several linguistic terms generated collectively in a word cloud. The link "Show all entries from Topic: Grammar" leads to a word cloud with a greater number of such terms. Similarly, in *OALD*, the list of grammatical terms is available under the button *Language*. It is hyperlinked to a sub-page, where the results can be filtered by topics, e.g. *Grammar*, *Features of language*, etc., and the level of language abilities of students at different levels of learning, i.e., CEFR A1-C2.

MWED is exceptional in that it generally does not use the grammatical distinction between countable and uncountable[4], and provides only one non-linguistic sense for the adjective *uncountable*, that is 'unable to be counted'. Although *MWED* has the separate headword entry for *countable* (only in the non-linguistic sense 'capable of being counted'), there is no mention of the mutual relationship between the two adjectives. It is partly because antonyms are not cross-referred in the dictionary, but the headword *noncount noun* has a cross-reference to the headword *mass noun*, which, surprisingly, cross-references to its antonym *count noun*. Also the reverse direction is provided in the case of the mentioned antonyms, but neither of the two redirects to *noncount noun*.

CED has four headwords related to uncountable nouns: *uncountable*, *uncountable noun*, *uncount noun*, and *mass noun*. Interestingly, the dictionary organizes the same type of information in various parts of the entries. Thus, the relation between uncountable nouns and mass nouns is hinted at through the cross-referential *see* placed under the definition (*CED*, *uncountable*), but the relation between uncountable nouns and uncount nouns is provided in the definition (*CED*, *uncountable noun*). The headword *uncount noun* does not redirect to any of the other three entries, although it seems to have the clearest definition and most specific information on usage (i.e., "a noun such as 'gold', 'information', or 'furniture' which has only one form and can be used without a determiner" *CED*, *uncount noun*). Otherwise, the definitions are intuitively over-complicated (i.e., "denoting a noun that does not refer to an isolable object" *CED*, *uncountable*), or circular if not tautological (i.e., "a noun such as 'wine' which is usually uncount but is used with 'a' or 'an' or used in the plural when it refers to types of that substance, as in 'a range of Australian wines'" *CED*, *mass noun*). One of the compensations provided by *CED*, however, is a component with simple descriptions and examples of the morpho-syntactic structures in Modern English - *Easy Learning Grammar*[5].

CD has two metalinguistic headwords related to uncountable nouns. The entry for *uncountable noun* is less developed than the one for *uncountable*. Even

---

[4] Entries for typically countable (regular) nouns and typically uncountable nouns are treated as neutral and are not signaled in the dictionary. Consequently, they are indistinguishable. Otherwise, *MWED* provides the plural forms under either the lemma/pronunciation, or individual senses. The plural forms are provided in the case of (a) nouns whose plural stem differs from the one in the singular, e.g. *child - children*, *wolf - wolves* (and even if the plural form is nonexistent, e.g. *pity - pities*), (b) nouns that usually have uncountable uses, but in some senses developed plural forms, e.g. *novelty - novelties*, *water - waters*.

[5] It can be accessed through the bookmark *Grammar* available in each headword entry either above the headword, or higher, on the main menu bar. The latter button expands into a drop down list with more elements, i.e., *Easy Learning Grammar*, *Grammar Patterns*, *English Usage*, *Teaching Resources*, and *Video Guides*.

though the latter provides only the linguistic sense of the term[6], the two head-words are not cross-referred. Of the five online dictionaries discussed in the present analysis, the headword *uncountable* in *CD* seems to be the most exhaustive from the point of view of the learner. The entry contains a simple definition, a few sentence examples, and cross-references to related grammar topics. What is particularly useful is that the user is able to preview the headlines and the initial lines of grammatical descriptions, e.g., "Nouns: countable and uncountable" "Countable and uncountable nouns with different meanings" or "Uncountable nouns used countably" (conveniently, the hyperlinks redirect to one sub-page where the discussion of the grammar points is presented collectively). *CD* (*uncountable noun*) informs: "uncountable nouns are marked [U] in this dictionary" and indeed they are thus labeled.

4. Nominal metalinguistic labels in *CD*

CD developed relatively complex codes, and is the only dictionary (of the five chosen for the study) that published the list of the employed labels explicitly. The ones accompanying nouns are reproduced in Figure 1.

Nouns

noun	A word that refers to a person, place or thing.
[C]	Countable noun: a noun that has a plural.
[U]	Uncountable or singular noun: a noun that has no plural.
[S]	A singular noun.
plural	The plural form of a noun.
noun [plural]	A noun that can only be used in the plural.
[usually plural]	A noun usually used in the plural.
[usually singular]	A countable noun usually used in the singular.
[+ sing/pl verb]	A noun that refers to a group of people acting collectively. When used in the singular it can be followed by either a singular or a plural verb in British English. In American English a singular verb is preferred.

Figure 1 Nominal metalinguistic labels (part of the list of codes used in *CD*). Source: https://dictionary.cambridge.org/help/codes.html#nouns

---

[6] However, the majority of the examples from the *Cambridge English Corpus* (available midway between the top and the bottom of the entry) show the uses of *uncountable* in the non-linguistic sense, e.g. *among uncountable families, uncountable number of symmetries, uncountable gallons of water*, etc.

The abundance of data and potential analyses inevitably results in different approaches to labeling, because "the information derived from traditional data collection always has in fact been supplemented heavily by the lexicographer's interpretation of the words concerned - most obviously regarding their meanings, but also regarding their part of speech and so on" (Hudson, 1988, p. 289). In a similar vein, but more specifically in relation to nouns, it has been observed that "countability is a property of words in context. Very few words are inherently and absolutely categorized as countable vs. non-countable in the lexicon, though some may be. Rather, concepts are *presented as* countable or non-countable in particular contexts" (Payne, 2011, p. 115).

Since dictionary making is a tedious and cooperative process, it is natural that labels happen to be misassigned, and some labeling choices may be burdened with errors. In *CD*, for example, the information about the codes and what they represent is not entirely accurate or consistent with how the labels function in the dictionary. Problems arise when (a) headword entries combine conventions from various dictionary components that are not mutually cohesive, and (b) lemmata that belong to the same category are tagged with different combinations of labels.

Instances of (a) occur when *CD* displays results from various dictionaries collectively, e.g., under the section *American Dictionary* (from the *Cambridge Academic Content Dictionary*), or *Business English* (from the *Cambridge Business English Dictionary*), see Figure 2. In the entry *inception*, in which the provided definitions are homogenous and synonymous (i.e., the presented item has the same sense), the noun is labeled as: [S], [U] and [C, usually singular, or U]. Similarly, the headword *goods* is labeled in a threefold way, as either: *phrase* [plural], *plural noun*, or *noun* [plural], of which the two former are not formally represented in the list of labels, see Figure 1. On the other hand, compounds with *goods* are labeled consistently as *noun* [plural], see the lemmata e.g., *dry goods*, *soft goods*, *white goods*.

Instances of (b) occur because the labels seem not to have been properly conceptualized, especially the labels [S] and *noun* [plural]. For example, it is questionable why both [S] and [U] can represent "a singular noun", see Figure 1. Qualitatively, the headwords tagged with [S] correspond to *singularia tantum* of which many are singular terms[7], marked with:

- the definite article, e.g., *the ark*, *the federal minimum wage*, *the gold standard*, *the onus*, *the underworld*,

---

[7] In philosophy of language, singular terms are expressions that signify one particular individual thing (i.e., person, place, object, etc.). Examples of singular terms are: proper names (e.g. *Santa*), definite descriptions (e.g. "the imaginary elderly man who is believed to bring presents to good children at Christmas"), and indexicals such as personal pronouns (e.g. *I*) and demonstratives (e.g. *this*).

- an initial capital letter, e.g., *GSM*, *Lloyd's List*, *Orion*, *Jupiter*, *Mercury*, *Saturn*, *Uranus*[8],
- both, e.g: *the City*, *the Decalogue*, *the Enlightenment*, *the S&P 500*, *the White House*.

Figure 2 CD, inception

---

[8] Similar lemmata, e.g. *Mars*, *Neptune*, *Venus*, are not labeled with [S].

Other headword nouns tagged with [S] lack a visual marker of this type. Their distribution, however, can be deduced on the basis of sample sentences. Such items are used in the singular, which is signaled with the indefinite article, e.g., *a frazzle*[9], *a no-brainer*, *a people person*, *a renaissance*, *a thaw*, etc., or are ambiguous, because the structures in the sample sentences/sample phrases are typical of countable and uncountable uses. For example, some of such nouns are:

- used as premodifiers, e.g., *fourth market activity*, *world pole vault champion*,
- preceded by a possessive noun phrase, e.g., *China's booming luxury market*, *her father's say-so*,
- preceded by the definite article, e.g., *the luxury market*, *the pole vault*, *the shot put*[10].

On the whole, [S] is applied to nouns that (a) are usually used in the singular, or (b) have uncountable uses. Taking into consideration that the label is to single out singular nouns, it is counter-intuitive that it is also applied to (c) collective nouns. Within a single entry, the distribution/patterns of such items can be deduced on the basis of the examples adduced either in sections for individual senses, or collectively at the bottom of the entry. It is worth mentioning, however, that examples of the latter type come from various sources, and are lacking in some headwords.

The application of the label [S] to individual senses can be exemplified with *footing*, which denote 'situation' in *a better financial footing*, or 'track' in *better footing*, e.g.,

- *New leadership is needed to return the company to a better financial footing*,
- *They will try to keep the field dry and ensure players have better footing.*

In some cases, the noun marked with [S] can be used countably or uncountably, but semantic nuances are not signaled explicitly, e.g.,

- *A vast bureaucracy with a red pencil and a blue pencil, giving the answers that the pricing system has failed to give?*
- *Evidently, virtually the entire movement had to be cut, and is crossed through with blue pencil, leaving intact a mere 12 bars towards the end of the piece,*

It seems that *blue pencil* in *a vast bureaucracy with (...) a blue pencil* should have a more literal reading than in *the entire movement (...) is crossed through with blue pencil*, because the phrase denotes a pencil with a blue lead used in marking

---

[9] In this and the following in-text examples, the items in bold correspond to the headwords.
[10] Other headwords defined as a sports competition are either not labeled, e.g. *the javelin*, labeled with [C], e.g. *biathlon, decathlon, hill climb, local derby, time trial, track meet*, or with [C usually singular or U], e.g. *pentathlon*.

corrections or cuts, but can also be interpreted as a synonym of censorship. The label is also employed in the case of collective nouns, which refer to either groups of people, e.g., *staff*, *the upper crust*, *the politburo*, *the populace*, or events, e.g., *the World Series*. Such nouns have only one form (similarly to typically uncountable nouns), and can be followed by a singular or plural verb (similarly to typically countable nouns). The case of *the World Series* deserves a brief comment. Although it is a singular term (see f. 7), it can be singular, e.g., *The World Series is a series of baseball games*, or plural, e.g., *three consecutive World Series*. Possibly, the plural form is due to analogy to the noun *series*, which has isomorphic singular and plural forms.

The label [S] can combine with other codes. For example, *staff* in the sense 'the group of people who work for an organization', *the upper crust*, *the politburo* and *the populace* are tagged with [S, + sing/pl verb]. [S] can also combine with labels related to countability, cf.:

- [S or C] in: *black bloc*, *Bonfire Night*, *Ebola virus*, *golden ratio*, etc.
- [S or U] in: *earth* 'planet' and 'ground', *sun*, *camaraderie*, *consensus*, *diversity*, *bear's breeches*, *susurrus*, etc.

Conversely, some headwords that name concepts that seem to conveniently fall under [S] are tagged by means of other codes, cf.:

1. [C usually singular] in *S-bend*,
2. [C, + sing/pl verb] in *government* 'the group of people who officially control a country'.

Occasionally, the combination of labels seems to be self-contradictory, as in the entry *wage*, where [S] is accompanied by the information "also *wages* [plural]", see Figure 3, and similarly, *the ruling class*, where the label is accompanied by "also *ruling classes* [plural]". Moreover, there are cases of the label [S] assigned incorrectly. For example, *pneumococcus* has the attested plural form *pneumococci*, which makes it a countable noun.

Taking into consideration the conventions adopted in *CD*, one can formulate a marginal objection as to why the nouns that "can only be used in the plural", which are to some extent, symmetrical to the nouns marked with [S], are labeled with *noun* [plural] instead of a single-letter symbol, see Figure 1.

Qualitatively, the headwords tagged with *noun* [plural] correspond to (a) *pluralia tantum*, including proper names that are never singular in form, e.g., *the Alps*, *the Andes*, *the Everglades*. However, names of countries that are always plural and intuitively belong to the same category are analyzed differently, cf.:

- [plural, + sing/pl verb] in *the Netherlands*,
- noun [+ sing/pl verb] in: *the Philippines*, *the Seychelles*, *the USA*.

wage

noun [S]

UK /weɪdʒ/ US /weɪdʒ/

(also **wages** [plural])

Add to word list

B1

a particular amount of money that is paid, usually every week, to an employee, especially one who does work that needs physical skills or strength, rather than a job needing a college education:

• *a very low/high wage*

• *an hourly/daily/weekly/annual wage*

• *He gets/earns/is paid a good wage, because he works for a fair employer.*

• *The job pays very low wages.*

Figure 3 CD, wage

The label *noun* [plural] is also applied to (b) nouns denoting singular bipartite items ending in -*s*, e.g., *pyjamas*, *scissors*, *trousers*, (c) plural nouns, including collective nouns, e.g., *alms*, *clothes*, *furnishings*, *odds*, *premises*, *suds*, *territorial waters*, *thanks*, *the suburbs*, *surroundings*. What deserves a brief comment in this subgroup is the case of the nouns borrowed in the plural form, e.g., *heroics*, *basics*, *paparazzi*. From the point of view of contemporary speakers of English, nouns containing Latin plural markers, e.g., -*ica* or -*a*, are no longer recognizable as having the same origin. Similarly, in *CD*, they are assigned various labels, cf.:

1. *plural noun* in *statistics*,
2. *phrase* [plural] in: *tactics*, *the tropics*,
3. *noun* [U] in: *linguistics*, *mathematics*, *gymnastics*, etc.
4. *phrase* [U] in: *ethics*,
5. noun [U, + sing/pl verb] in *data*,
6. noun [plural, + sing/pl verb] in *the mass media*.

unemployed

noun [plural]

UK /ˌʌn.ɪmˈplɔɪd/ US /ˌʌn.ɪmˈplɔɪd/

C1

people who do not have a job that provides money:

• *There are now over four million unemployed in this country.*

• *Our department provides services for **the** unemployed.*

+ SMART Vocabulary: related words and phrases

rich

noun [plural]

UK /rɪtʃ/ US /rɪtʃ/

▸ **the rich**

C1

rich people considered together as a group:

• *The resort is frequented by the rich and famous.*

Opposite
the poor

Figure 4 CD, unemployed, rich

The label *noun* [plural] is further used in the case of (d) noun phrases containing *the + adjective*, e.g., *the unemployed*, *the rich*, *the young*, etc. Since the mere concept of the discussed structure is plural, e.g., *the rich*, *the British are*, *CD* labels such items appropriately. Graphically, however, the information is not displayed in a consistent way, see Figure 4.

The analysis of *the + adjective* depends on the interpretation of the lexical element. In some approaches, *unemployed*, *rich*, etc. are adjectives that function as nouns (Brinton & Brinton, 2010, p. 140-141) and "serve as the head of a noun phrase" (Biber et al., 2021, p. 517). It is argued that even though the whole phrase is nominal, *unemployed*, *rich*, etc. are adjectives because they lack the plural marker. However, the same feature is characteristic of nouns that have one fixed form. In terms of subject-verb concord, *the rich*, *the British are* is similar to zero plurals, e.g., *deer*, *salmon*, *sheep*, etc. *are*, and collective *s*-less plural nouns. e.g., *gentry*, *police*, *staff*, *miscellanea*, etc. *are*. Another interpretation assumes that the nominal head (a dummy noun), e.g., *people*, is understood or ellipted (Brown & Miller, 2016, p. 24). If it is the case, *the* + adjective (+ noun) is similar to exocentric compound nouns of the structure adjective + noun (+ noun), e.g., *redhead* 'someone who has red hair', *highbrow* 'someone who is interested in serious subjects', *lazybones* 'someone who is lazy', *high-flyer* 'someone who has the desire and the ability to be very successful in their job or their studies', *paleface* 'someone who has white skin', etc. Exocentric compounds are sometimes called pseudo-compounds (exactly because their nominal element is not the head, which is to say that a redhead is not a head, a highbrow is not a brow, etc.), or *bahuvrihi*-compounds (the name is taken over from ancient Indian grammarians). Jespersen notes that "[b]ahuvrihi-compunds nearly always denote living beings (or personified inanimate things), the final member being generally the name of some part of the body or dress. They are frequently used as (nick)names: *Bluebeard*, *Edmund Ironside*, like other (nick)names indicating a single conspicuous characteristic of the person or thing to be named" (Jespersen et al., 1954, p. 149).

Similarly to [S], the label *noun* [plural] can be part of combinatory labeling, which has been shown on the basis of the nouns in *-ics* and *-a*. Additional examples are:

- *phrase* [plural] in: *compasses*, *groceries*, *trimmings*,
- *plural noun* in *bowels*,
- [C usually plural] in *fitting* 'a small part or thing'.

Analyzing an item is not always straightforward. One word-form can function in a variety of morphosyntactic and semantic contexts, e.g., *like*, *round* (see Brinton & Brinton, 2010, p. 140). Moreover, a word can develop new properties and undergo recategorization, either within a class (e.g., countable/uncountable

shift), or between classes (e.g., in the case of functional shift, i.e., conversion). Under such circumstances, the question arises to what extent labeling can be accurate and consistent. With regard to [S] and *noun* [plural], another question is if these particular labels are indispensable from the point of view of an average user when they are eagerly modified or replaced with other labels. In a broader perspective, however, even a defective way of organizing information is better than nothing, because it can help dictionary users to see similarities and understand the system of a language, even if to a certain degree.

5. Dictionary labels in the context of dictionary skills

Alongside concerns about the content, lexicographers have developed an interest in the end user. "Gradually, dictionary makers have begun to recognize that dictionary users do not necessarily understand all the conventions implicated in the presentation of lexicographic data. (...) With time, it has become increasingly clear that most people possess limited skills when it comes to using dictionaries" (Lew, 2015, p. 1). Consequently, it has become crucial to research on and improve the dictionary competence of language users, both native speakers and language learners. One study pointed out that if students are unfamiliar with basic grammatical concepts in their mother tongue, they can have problems in understanding foreign language data: "German students of English - that is language students - have a very poor knowledge of grammar. They are not familiar with basic grammatical terms, not even the most traditional ones, and they cannot analyze simple sentences of the types *Sie ist in der Küche*, *Er ist Lehrer*. One has the impression that students have never acquired the rudiments of grammar, either in the foreign language they are being taught, or in the mother tongue they learnt much earlier" (Herbst & Stein, 1987, p. 123). The authors rightly conclude that dictionary skills should be viewed in a broader perspective as being part of foreign language skills: "Thus what is needed to improve the dictionary competence of learners is not so much the development of special methods of teaching dictionary use, as the recognition that[,] for foreign learners[,] competence in dictionary use is part of their wider competence as speakers, readers or writers of the foreign language" (Herbst & Stein, 1987, p. 127).

In another study (performed circa three decades later), dictionary skills were analyzed in the context of improving and revising texts in mother tongue. The results were that the majority of students can use some of the dictionary information successfully, but "(...) the presence of lexicographic resources alone does not automatically lead to better revision results" (Wolfer et al., 2018, p. 22). Herbst and Stein's conclusion seems to apply equally to native speakers'

linguistic competence. Language users in general need to be competent enough to extract and incorporate the relevant linguistic information in their oral and written texts. If they are not, dictionary skills can be taught.

In a similar vein, Frankenberg-Garcia (2020, p. 33) states that existing research recognizes the need to train users in dictionary-consultation skills, but at the same time shows that little progress has been made in this respect. In other words, the shift to e-lexicography impacts old and new dictionary skills: "Some traditional skills are becoming largely obsolete, such as those related to paper page navigation or reducing a word form to its citation form. However, new skills arise from the numerous new search techniques afforded by electronic dictionaries" (Lew, 2013, p. 29).

Dictionaries, either paper or digital, are reference systems and "any reference system requires its users to apply specific reference skills" (Herbst & Stein, 1987, p. 116). Ideally, the users should be aware of what their dictionaries have to offer (the structure, the arrangement of the material, the labels, the scope of vocabulary control, the function of cross-references and hyperlinks, and much more, see Lew & Galas, 2008, Lew, 2013). All the skills needed to use a dictionary efficiently constitute the user's information literacy, and can be trained as part of the curriculum: either in the form of automated online courses - as tutorials (Ranalli, 2013[11]), or in the classical classroom context, where the teacher's role is to "[h]elp learners to use dictionary entries to look around words they know already in order to enrich what they can do with apparently familiar items" (Scrivener, 2014, p. 307). However, as Chi (1998, p. 566) points out "English teachers may themselves lack expertise and knowledge in using dictionaries", especially online ones, because "in many countries teachers tend to be left behind in the digital revolution" (Lew, 2013, p. 18).

Another aspect of information literacy is that language users need to use their critical assessment to be able to review their own knowledge and diagnose what type of information is relevant. "Users have to employ their cognitive faculties and discriminating skills to single out from the information offered the items they need" (Herbst & Stein, 1987, p. 124). Thus, more dictionary work is needed in the form of guided teaching in the classroom and tutorials, but equally important is self-study, or self-education. According to Chen (2017, p. 246), "teachers should provide instructions to improve learners' dictionary use

---

[11] Ranalli (2012, 2013) describes the *Virtual Vocabulary Trainer* project in which video presentations and text-based practice activities were designed to help learners improve their dictionary skills, and information and computer literacy. The participants were taught, for instance, how to enhance online dictionary searches by means of (combinations of) keyboard shortcuts. Thus, it seems justified to claim that courses on information and computer literacy in the context of dictionary use should present even simple tools and their practical application in linguistic queries. In other words, the users' prior knowledge of the basics should not be taken for granted.

skills", but more importantly, learners "should acquire good habits of dictionary use such as being attentive, careful, and involved during dictionary lookups".

It seems that still there is another party involved in facilitating the process of teaching or developing skills in dictionary use, i.e., dictionary providers. They should make attempts to unify labeling strategies, and present their labeling conventions explicitly to the user. From the point of view of the learner, a homogeneous system of labeling would simplify the comparison of different theoretical interpretations. What is more, lexicographers should not be urged to make ultimate decisions. They may equally well present antagonistic analyses and let more advanced learners decide which arguments are more persuasive. It seems that a more in-depth cooperation could be of benefit to all.

6. Conclusions

Employing dictionary labels is grounded in lexicographic tradition and is motivated by numerous advantages. Abbreviations and symbols: save space, code various elements of the linguistic and extra-linguistic reality (i.e., can provide graphic, phonetic, morphological, syntactic, semantic, cognitive, pragmatic, psychological and sociological information), and help to concisely organize different types of information, including reconciling various theoretical approaches. However, labels can be used as a hallmark to distinguish a dictionary as a product on the market. Consequently, very few distinctions are marked by means of the same metalanguage.

Taking into consideration the point of view of the learner, lexicographers are guided by simplicity and selectiveness (which are the underlying principles of labeling conventions), but at the same time learner-orientedness and accessibility underpin the approach to use lexical descriptors, e.g., *countable*, instead of abbreviated items, e.g., *C*. One reason is that such descriptors are more straightforward. Another is that decoding symbols may be problematic to dictionary users. It seems, though, that what follows from Dziemianko's (2008) study is that if students are aware of the existence of the applied labeling conventions and understand the dictionary-specific notation system, such codes are beneficial. Consistent and cohesive labels can help learners understand and draw similarities between metalinguistic concepts, and even misassigned labels can be of value as they represent a lexicographic interpretation, and can be a starting point for further research.

Dictionary skills, seen as part of foreign language skills or information literacy, can be developed or taught. The parties involved are: the learner, the teacher, and the lexicographer. With respect to the latter, the present contribution argues in favor of (a) providing the user with more explicit information about the adopted conventions (for instance, by publishing relevant hyperlinks

on the dictionary's homepage in the case of online MLDs), and more in-depth cooperation on (b) unifying labeling strategies, and (c) describing and illustrating them accurately, which would enable the learner to compare data from various sources, and allow more advanced students to become acquainted with arguments for different theoretical approaches and interpretations.

References

A. Dictionaries

Cambridge Dictionary (*CD*). Access 20.04.2024. https://dictionary.cambridge.org/

Collins English Dictionary (*CED*). Access 20.04.2024. https://www.collinsdictionary.com/dictionary/english

Longman Dictionary of Contemporary English (*LDOCE*). Access 20.04.2024. https://www.ldoceonline.com/

Merriam-Webster English Dictionary (MWED). Access 20.04.2024. https://www.merriam-webster.com/

Oxford Learner's Dictionaries (*OALD*). Access 20.04.2024. https://www.oxfordlearnersdictionaries.com/

B. Other sources

Atkins, B. T. S., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford University Press. https://doi.org/10.1093/oso/9780199277704.001.0001

Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). *Grammar of spoken and written English.* John Benjamins. https://doi.org/10.1075/z.232

Brinton, L. J., & Brinton, D. (2010). *The linguistic structure of modern English*. John Benjamins. https://doi.org/10.1075/z.156

Brown, K., & Miller, J. (2016). *A critical account of English syntax: Grammar, meaning, text.* Edinburgh University Press. https://doi.org/10.1515/9780748696093

Calzolari, N., Picchi, E., & Zampolli, A. (1987). The use of computers in lexicography and lexicology. In Cowie, A. P. (Ed.), *The dictionary and the language learner: papers from the EURALEX seminar at the University of Leeds, 1-3 April 1985*. (European Association for Lexicography. *Lexicographica*. Series Maior 17), (pp. 55-77). Max Niemeyer Verlag. https://doi.org/10.1515/9783111340500-007

Chen, Y. (2017). Dictionary use for collocation production and retention: a CALL-based study. *International Journal of Lexicography*, 30(2), pp. 225-251. https://doi.org/10.1093/ijl/ecw005

Chi, M. L. A. (1998). Teaching dictionary skills in the classroom. In T. Fontenelle, P. Hiligsmann, A. Michiels, A. Moulin, A., & S. Theissen (Eds.), *Euralex '98 proceedings* (pp. 565-577). Université Départements d'Anglais et de Néerlandais. https://euralex.org/wp-content/themes/euralex/proceedings/Euralex%201998%20Part%202/Man%20Lai%20Amy%20CHI%20Teaching%20dictionary%20skills%20in%20the%20classroom.pdf

Dziemianko, A. (2006). *User-friendliness of verb syntax in pedagogical dictionaries of English*. (*Lexicographica*. Series Maior 130). Max Niemeyer Verlag.

Dziemianko, A. (2008). Noun and verb codes in pedagogical dictionaries of English: user-friendliness revisited. In E. Bernal & J. DeCesaris (Eds.), *Proceedings of the XIII EURALEX International Congress, 15-19 July 2008, Barcelona* (pp. 1243-1250). Universitat Pompeu Fabra. https://www.euralex.org/elx_proceedings/Euralex2008/124_Euralex_2008_Anna%20Dziemianko_Noun%20and%20Verb%20Codes%20in%20Pedagogical%20Dictionaries%20of%20English_User-friendliness%20Revisited.pdf

Frankenberg-Garcia, A. (2020). Combining user needs, lexicographic data and digital writing environments. *Language Teaching*, *53*(1), pp. 29-43. https://doi.org/10.1017/S0261444818000277

Herbst, T., & Stein, G. (1987). Dictionary-using skills: A plea for a new orientation in language teaching. In A. P. Cowie (Ed.), *The dictionary and the language learner: Papers from the EURALEX seminar at the University of Leeds, 1-3 April 1985*. (European Association for Lexicography. *Lexicographica*. Series Maior 17), (pp. 115-127). Max Niemeyer Verlag. https://doi.org/10.1515/9783111340500-010

Heuberger, R. (2016). Learner's dictionaries. History and development; current issues. In P. Durkin (Ed.), *The Oxford handbook of lexicography* (pp. 25-43). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199691630.013.3

Heuberger, R. (2020). Monolingual online dictionaries for learners of English and the opportunities of the electronic medium: A critical survey. *International Journal of Lexicography*, *33*(4), pp. 404-416. https://doi.org/10.1093/ijl/ecaa018

Hudson, R. (1988). The linguistic foundations for lexical research and dictionary design. *International Journal of Lexicography*, *1*(4), 287-312. https://doi.org/10.1093/ijl/1.4.287

Jespersen, O., Christophersen P., Haislund N., & Schibsbye K. (1954). *A modern English grammar on historical principles*. Part VI: *Morphology*. Routledge.

Lew, R. (2011). Online dictionaries of English. In P. A. Fuertes-Olivera & H. Bergenholtz (Eds.), *E-lexicography: The Internet, digital initiatives and lexicography* (pp. 230-250). Continuum.

Lew, R. (2013). Online dictionary skills. In I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets, & M. Tuulik (Eds.), *Electronic lexicography in the 21st century: Thinking outside the paper. Proceedings of the eLex 2013 Conference, 17-19 October 2013, Tallinn, Estonia* (pp. 16-31). Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut. http://hdl.handle.net/10593/8138

Lew, R. (2015). Dictionaries and their users. In Hanks, P., & de Schryver, G.-M. (Eds.), *International handbook of modern lexis and lexicography* (pp. 1-9). Springer-Verlag. http://hdl.handle.net/10593/13849

Lew, R., & Galas, K. (2008). Can dictionary skills be taught? The effectiveness of lexicographic training for primary-school-level Polish learners of English.

In E. Bernal & J. DeCesaris (Eds.), *Proceedings of the XIII EURALEX International Congress, 15-19 July 2008, Barcelona* (pp. 1273-1285). Universitat Pompeu Fabra. http://hdl.handle.net/10593/342

Norri, J. (2022). Temporal labels and specifications in monolingual English dictionaries. *International Journal of Lexicography*, *35*(4), pp. 397-426. https://doi.org/10.1093/ijl/ecac013

Payne, T. E. (2011). *Understanding English grammar*: *A linguistic introduction*. Cambridge University Press.

Ranalli, J. (2012). *The VVT Project: A web-based platform for strategy instruction and research into self-regulated learning of L2 vocabulary*. https://dr.lib.iastate.edu/server/api/core/bitstreams/239e8b6d-034e-4181-bfce-a931a30582fe/content

Ranalli, J. (2013). Online strategy instruction for integrating dictionary skills and language awareness. *Language Learning & Technology*, *17*(2), pp. 75-99. http://hdl.handle.net/10125/44325

Scrivener, J. (2014). *Learning teaching: The essential guide to English language teaching*. Macmillan Education.

Vrbinc, M., & Vrbinc, A. (2015). Diasystematic information in learner's dictionaries: The usability of multiple labels. *GEMA Online Journal of Language Studies, 15*(1), pp. 111-128. https://doi.org/10.17576/GEMA-2015-1501-07

Wolfer, S., Bartz, T., Weber, T., Abel, A., Meyer, C. M., Müller-Spitzer, C., & Storrer, A. (2018). The effectiveness of lexicographic tools for optimising written L1-texts. *International Journal of Lexicography, 31*(1), pp. 1-28. https://doi.org/10.1093/ijl/ecw038