Problems of Usage Labelling in English Lexicography*

Lydia Namatende Sakwa, Department of Languages, Faculty of Humanities and Social Sciences, Uganda Martyrs University, Nkozi, Uganda (lydiasakwa@yahoo.com)

Abstract: Landau (1991: 217) stipulates that ‘usage refers to any or all uses of language’. It is the study of good, correct, or standard uses of language as distinguished from bad, incorrect, and non-standard uses of language. Usage may also include the study of any limitations on the method of use, whether geographic, social or temporal. Basically it alerts users that certain terms should not be uncritically employed in communication. This article discusses the treatment of usage in English lexicography. It analyses the labelling practices in six monolingual English dictionaries namely: the Oxford Advanced Learner’s Dictionary (OALD), the Macmillan English Dictionary (MED), the Longman Dictionary of Contemporary English (LDOCE), the Cambridge International Dictionary of English (CIDE), the World Book Dictionary (WBD) and the New Shorter Oxford English Dictionary (NSOED). Discrepancies in the contextual usage labelling in the dictionaries were established and are discussed.

Keywords: LEXICOGRAPHY, LINGUISTICS, USAGE, LABELLING, DICTIONARY, CORPUS/CORPORA, DESCRIPTION, PRESCRIPTION, PROSCRIPTION


Sleutelwoorde: LEKSIKOGRAFIE, LINGUISTIEK, GEBRUIK, ETIKETTERING, WOORDEBOEK, KORPUS/KORPORA, BESKRYWING, PRESKRIPSIE, PROSKRIPSIE

* This article is a revised version of a paper presented at the Thirteenth International Conference of the African Association for Lexicography, organized by the Bureau of the Woordenboek van die Afrikaanse Taal, Stellenbosch, Republic of South Africa, 30 June–3 July 2008.

1. Introduction

Comments on usage are included in dictionaries as a guide to speakers of a language in their decisions on how to use words appropriately. They may be unsure, for example, whether wireless is an old-fashioned (and chiefly British) word now almost entirely replaced by radio, or whether to call a woman petite, slim or slender implies an approving attitude towards her in contrast to skinny, which suggests disapproval. To help learners with these difficulties, a number of labels are used in dictionaries to denote the stylistic values of words or the technical fields in which they are used; these are called ‘usage labels’.

Although some readers may interpret all dictionary data as if they were ‘authoritative guides’ (Abecassis 2008: 1) and/or ‘sources of knowledge’ (Tarp and Gouws 2008: 236) on how to spell, pronounce, interpret meaning or check usage, the question surrounding the appropriateness of usage labelling as found in most dictionaries has raised concern among many linguists (Ptaszynski 2010: 411).

This article intends to show the dilemma surrounding usage labels, which makes them confusing and sometimes misleading from the point of view of second language learners. The focus will be on demonstrating the problems with regard to the labels assigned to words from the perspective of the second language speaker. The article therefore attempts to guide future lexicographers to make more informed usage labelling decisions by expounding the inadequacies of usage labelling in English lexicography.

2. Inconsistencies

Although usage labels are widely employed, their use is not at all consistent (Abecassis 2008: 3) and the same word will not represent the same label from one dictionary to the next. Very frequently, one dictionary labels one sense of a word as slang while the next labels the same sense informal. For instance, while MED uses the label British English (BrE), impolite for the noun arse in the sense of ‘the part of the body that you sit on’, LDOCE uses the label BrE; CIDE BrE & AustE, slightly taboo, slang; the WBD slang; the OALD BrE, taboo, slang; and the NSOED uses the label now coarse slang.

It follows then that if dictionaries can agree that bonk is a verb that is spelled with ‘o’, and that one of its senses is ‘to have sex with somebody’, why is there discrepancy on the kinds of situations in which it may be used? If a noun like coolie in the sense of ‘a worker in the Eastern countries with no special skills of training’ is assigned the labels old-fashioned taboo (OALD), offensive (MED), old-fashioned (LDOCE), dated (WBD), and now rare or obsolete (NSOED), which usage label should users perceive as the correct one since all the dictionaries are in disagreement on the usage of the word?

These inconsistencies have been found to be prevalent in English dictionaries and, although lexicographers acknowledge the difficulty of labelling
words consistently (Crenn 1996), as a result of these inconsistencies, the authority of dictionaries in guiding users on contextual usage becomes questionable. Furthermore, these discrepancies can confuse users especially in the case of non-native speakers who may not have acquired a good command of the English language.

It has also been observed that lexicographers do not share the same view on the neutral register (the unmarked/unlabelled words) since in some instances words are labelled in some dictionaries while they are not labelled in others. For example, the adjective *antiquated* meaning 'old-fashioned and no longer suitable for modern conditions' is labelled *usually disapproving* in the OALD, *archaic* in the NSOED; while no label is assigned to it in the MED, LDOCE, CIDE or in the WBD; *anthropoid* is also labelled *technical* in the OALD while it is not labelled in the other five dictionaries. These discrepancies are very prevalent in English dictionaries.

It is remarkable how certain words that need to be labelled are not assigned usage labels in English dictionaries. For example, the word *cretin* meaning 'a very stupid person' is neither labelled by the NSOED nor by the WBD although the other three dictionaries label it; *dago* 'a very offensive word for a person from Italy, Spain or Portugal' is also neither labelled by the LDOCE nor by the CIDE; *slut* and *trollop*, which both mean 'a woman who has many sexual partners' are not labelled by the NSOED. Since usage labels are assigned to words to caution users about the potential danger of using some words, it may be a disservice to users to be deprived of such usage information for words like these.

3. **Criteria for Usage Labelling**

The question as to how lexicographers determine usage labelling is still an enigma. Ptaszynski (2010: 411) explains that 'Lexicographers have been searching in vain for an exhaustive and precise answer to the questions of which words to label in what kind of dictionaries and how to do it'. This position is reiterated by Crenn (1996) asserting that 'no particular monolingual dictionary is chosen as a basis for labelling by the lexicographers'. It is evident that usage labels are assigned to words according to the lexicographer's discretion. This probably accounts for the inconsistencies portrayed in section 2. Ptaszynski (2010: 411) explains that such problems 'stem from the lack of a firm theoretical basis for the application of disystematic information (i.e. information about restrictions on usage) in dictionaries'. It is no wonder then, that words transcribed the same way, from the same word class and with the same sense are labelled differently depending on the choice of dictionary one makes. The problem can affect both native and non-native speakers of English who, given a variety of labels for a single word, no doubt experience problems inferring the usage of such a word.
Furthermore, because there is no agreed-on criteria for usage labelling, the issue of subjectivity cannot be ruled out, for the interpretation of labels in other dictionaries and the corpus context is itself dependent on the individual lexicographer. Crenn (1996) suggests that 'although scholars [...] may point out the necessity of descriptive dictionaries, a completely objective system of register labelling has not yet been developed because lexicographers cannot feasibly let go of their subjective viewpoints'. Subjectivity is therefore another problem that may hamper good usage judgements in English lexicography.

Ptaszynski (2010) investigates the causes of the unsatisfactory theoretical treatment of information on usage, and outlines some general, theoretical guidelines for making adequate judgments concerning usage labelling. Research into a guide on usage labelling should impact positively in reducing some of the discrepancies in the usage labelling practices in lexicography. This is reiterated by Crenn (1996: 181) who asserts that although the issue of consistency in usage labelling might not be solvable in the near future, suggestions can be made to 'help lexicographers decide on register labels and to improve the consistency of register labelling in dictionaries'. Consistency has generally been linked to reliability of the dictionary (Swanepoel 2008: 216; Tomaszczyk 1988: 2), which according to Swanepoel (2008: 216) 'is increased if lexicographers adopt a set of clearly defined principles and adhere to them consistently'.

4. The Front Matter Information of Dictionaries

The front matter information of dictionaries was analyzed to show how practising lexicographers view and classify their comments on usage.

It was observed that few dictionaries provide justification for their usage labels. Questions that may arise from examining usage labels in English dictionaries can rarely be answered by the consultation of the front matter of the dictionaries. In fact, most dictionaries do not even explain what they mean by the terms 'usage labels' or 'usage labelling', taking it for granted that dictionary users know the meaning of the terms.

For instance, the LDOCE, the MED, and the CIDE explain the types of labels they select but there is no explanation of the terms 'usage labels' or 'usage labelling'. This leaves room for users not only to misinterpret what the lexicographers may mean but also to remain ignorant of the meaning of labelling, especially in the case of non-native speakers of English.

Lexicographers also have the tendency of assigning usage labels not listed in the front matter to words in the dictionary. This shows that inadequate information is given, which is evidence that the front matter notes in question are incomplete and therefore unreliable. Examples of usage labels assigned to words but not listed in the front matter are: vulgar in the WBD, nonstandard in the LDOCE and not standard in the CIDE.

Since some lexicographers do not explain or define the usage label terms in the front matter notes, it demonstrates their assumption that users of the
dictionaries will automatically understand what these labels mean. This assumption may be wrong, so that users may thus be misled in understanding the usage labels in question; this is substantiated by Hartmann and James (1998) who explain that the major hindrance to using dictionaries emanates from a lack of necessary skills to utilize them efficiently as well as the insufficient guidance offered by the dictionary preface to guide users to exploit its possibilities.

Crenn (1996: 49) implies that since there is great variation in the way dictionaries deal with usage, each dictionary should explain its usage labelling practices by explaining how it groups its usage labels and its explanations of these labels. This information is usually included in the dictionary front matter.

5. **The Choice and Classification of Usage Labels**

There seems to be little or no consensus on how to categorize the usage labels and as a result, the groups of labels as chosen by the lexicographers concerned vary from one dictionary to the next. There is also no uniformity in the choice or array of usage labels as presented in English dictionaries. The different dictionaries exhibit a diverse range of labels and categorize them differently depending on the lexicographer.

The LDOCE divides the usage labels into four categories namely: regional, foreign, attitudinal (formal, informal, humorous, approving) and style labels (biblical, dialect, law, literary, non technical, old-fashioned, old use, poetic, slang, spoken, taboo, technical and trademark).

The OALD does not divide its labels into broad groups, it simply presents the following list: approving, disapproving, figurative, formal, humorous, informal, ironic, literary, offensive, rare, slang, spoken, technical, AmE, BrE, dialect, old-fashioned, written, taboo, old use, saying and trademark.

The array of labels in the MED is divided into three categories namely: style and attitude labels (formal, humorous, impolite, literary, offensive, old-fashioned, spoken, very formal, and very informal), subject labels (for example business, computing and journalism) and regional labels (for example Australian, Canadian and American).

In the WBD there are labels like archaic, dialect, English speaking variation labels, foreign language labels, informal, obsolete, poetic, professional terms, slang, substandard, trademark, unfriendly use and figurative usage.

The NSOED divides its labels into four categories, namely: those showing restriction to geographical area, those referring to style or register, those indicating branch of knowledge or field of activity, and lastly, labels indicating frequency or extent of use.

The CIDE has the following array of labels: approving, dated, disapproving, not standard, old use, poetic, regional, slang, female, figurative, formal, humorous, informal, law, literary, male, medical, specialized, taboo and trademark.
It is evident that each dictionary makes its choice of the labels to be assigned to words. What is not clear is the criteria the lexicographers use in selecting certain labels over others.

Certain labels are not listed in some dictionaries, while they occur in others. For example, the OALD has labels mainly atypical to the other five dictionaries such as *ironic* and *saying*, the labels *very formal* and *very informal* are commonly used by the MED and rarely in the other five dictionaries. The WBD lists and explains the term *combining form* as a usage label but the other dictionaries do not have this as a usage label. The CIDE also employs *female* and *male* as labels, the two being alien to the other five dictionaries studied.

It is therefore apparent that there is no agreement on the number of dimensions in usage. It is evident that each dictionary makes its choice of the number, type and classification of usage labels to be assigned to words. Quirion (1995: 347) states that ‘the reduction in the number of labels could lead to some consensus in labelling’.

6. Definitions

As already mentioned in Section 4, some of the English dictionaries do not try to explain the terms ‘usage label’ or ‘usage labelling’; they simply embark on listing the labels assuming that users can discern the meaning by looking at the list of examples of usage labels or otherwise by looking at the dictionary articles, particularly at the entries assigned usage labels and then automatically comprehending what is meant by usage, which may not be the case. This has been established in three of the six dictionaries consulted for this study, namely the LDOCE, the MED and the CIDE.

Some of the definitions of the labels are imprecise, leaving room for questions. For example, the OALD defines the label *spoken* as indicating ‘expressions used mainly in informal conversation’. This is a vague definition because of the use of the word *mainly*. Should it then be concluded that an adjective like *kaput* in the sense of ‘not working properly’ or an idiom like *no probs* with the meaning ‘there is no problem’ and labelled *spoken* can sometimes be used in say formal situations? This type of definition can therefore be misleading.

As a result of the inadequacies in the definitions accorded, some of the labels in English dictionaries seem to overlap and in the end, it appears as if there is a replication. In other words, labels, which are apparently synonymous, seem to be assigned to unconnected words. The inadequacy in the way these labels are defined can be misleading.

This can be illustrated by the way the labels *archaic* and *obsolete* are defined in the WBD; *archaic* is described as referring to ‘a word that is no longer in general use. It may be found in special contexts, such as law or special styles of writing such as the Scriptures. Sometimes modern writers use archaic words to give their writings an old-fashioned flavor.’ This definition has the same implications as the label *obsolete*, which is defined in the same dictionary as indica-
tion for ‘a word only found in writings of an earlier time or in modern writings that imitate the style of earlier writing.’ These definitions of the two apparently different labels need to be revised if the difference is to be realized, otherwise they can be misleading as the definitions do not clearly show the difference.

Additionally, when the difference between the way OALD and LDOCE define the labels old use and old-fashioned is analyzed, a discrepancy in the LDOCE’s definition becomes evident. The LDOCE defines old use as indicating ‘a word used in earlier centuries’ and it defines old-fashioned as indicating ‘a word that was used earlier in this century but would sound old-fashioned today’. On the other hand, the OALD defines old use as referring to ‘expressions that are no longer in current use’ while it defines old-fashioned as referring to ‘expressions that are passing out of current use’. On analyzing the way the two dictionaries define the two different labels, it is evident that the OALD shows a clear distinction between the labels leaving no doubt that there is indeed no duplication of the labels. The LDOCE on the other hand does not evince this difference. Consequently, it could be argued that old use and old-fashioned have the same meaning and therefore are synonyms since even a word used in the earlier century (old use) would certainly sound old-fashioned today. This illustrates the deficiencies of some of the definitions of labels used in English dictionaries.

The CIDE does not even try to define the labels selected. It simply gives a list of the labels without even briefly explaining them.

Undefined or inadequately defined terms coupled with the fact that some label definitions vary from one dictionary to the other shows the inaptness of usage labelling in English lexicography.

7. Ambiguity

The editors do not spell out what specific areas are covered by labels like informal, nonstandard, slang, vulgar, disparaging, offensive … To give a few examples: Does the label literary mark only the effect on the reader or listener since it is defined in the OALD as indicating ‘language used mainly in literature and imaginative writing’, although logically, the use of literary language depends also on the setting and the kind of communication intended?

Similarly, nonstandard seems to indicate only the kind of speaker who might use it since it is defined as ‘not conforming to the speech or grammar of educated persons and often regarded as a marker of low social status’. However, can someone not adopt nonstandard words in certain settings to be witty/comical? Or perhaps in an attempt to satirize a situation?

Further still, a label like slang as explained by Crenn (1996: 186) ‘is understood differently by each dictionary (and presumably) every user and unless dictionaries specify to which group a particular slang word corresponds …’, for example slang military, slang law … then labels showing degree of formality (such as very formal, informal) could be more appropriate in achieving clarity.
In addition, one group may read one particular usage as informal or non-standard whereas another may use it as entirely natural for any conceivable circumstance. According to Landau (1991: 255-258), dictionaries cannot in good faith mark out some usages as informal without saying to whom they are informal. Informal may actually mean informal to those of the higher social classes, especially older, well-educated authors and professors in the humanities.

The combinations of usage labels assigned to words also sometimes pose as deterrents to clarity. It can be difficult to comprehend what exactly the lexicographer means by assigning two or more labels to a word. This problem may be illustrated by the following examples:

The noun vale meaning ‘a valley’ is assigned the label now archaic or poetic in the NSOED. Likewise, the noun vagrant meaning ‘a person who has no job or home especially one who begs’ is labelled slightly dated or law in the CIDE, the adjective wee meaning ‘very small in size’ is labelled colloquial also chiefly Scottish in the NSOED, the noun wank in the sense of ‘an act of masturbation’ is assigned the label British and Australian taboo slang in the CIDE, the adjective woebegone meaning ‘looking very sad’ is assigned the label mainly literary in the MED, and the noun gloom in the sense ‘total darkness’ is labelled especially literary in the LDOCE.

The use of conjunctions like ‘or’, ‘and’, and adverbs like ‘slightly’ and ‘usually’ can be puzzling. The entry recuperate is labelled slightly formal in the CIDE. The question then arises whether it should be used in formal situations. If it is slightly formal, then exactly in what situations it may be used? Outlandish, an adjective assigned the label usually disapproving in the OALD and politicking labelled especially disapproving in the CIDE are other examples.

The use of indications like ‘mainly’, ‘especially’, ‘chiefly’ in representing usage sometimes cause ambiguity as dictionary users may be uncertain as to what exactly the lexicographer means.

8. Dictionary Tone

The approach in a dictionary can be descriptive, prescriptive, or proscriptive (Tarp and Gouws 2008: 235; Van der Merwe 2008: 346). When it is descriptive, ‘it avoids classifying occurring forms as either recommended or not permissible. It reflects the spectrum of actual language in use’ (Tarp and Gouws 2008: 237). It shows the user what by illustration, it says, ‘it is like this’ thereby leaving the responsibility to the user to decide whether a word is adequate for use in a particular situation. When it is prescriptive, ‘lexicographers inform the user how to use language’ (Tarp and Gouws 2008: 236); they say ‘do it like this’. A prescriptive approach imposes the lexicographer’s ‘point of view on the dictionary and the target users of the dictionary’ (Tarp and Gouws 2008: 236). They explain the term ‘proscription’ in lexicography which deviates from the English word ‘proscribe’ since it does not refer to the state of being forbidden.
A proscriptive approach informs 'the user not only about language use but also about the form recommended by the lexicographer' (Bergenholtz 2003: 13).

Landau (1989: 194) suggests that until the latter half of the 20th century, dictionaries felt duty-bound to prescribe 'correct usage'. However, there is claim that 'the attitude to lexicography has changed since the 1970s with dictionaries being increasingly descriptive rather than prescriptive' (Abecassis 2008: 2). This study exposes the contrary, with lexicographers being implicitly prescriptive as illustrated below.

The label taboo is defined in the OALD as 'a word that should not be used because it is very rude or offensive'. It is evident that the lexicographer is telling his users not to use this word because of the consequences that would follow. The lexicographer in this case uses a prescriptive tone. Examples of words labeled taboo are: a verb like fuck which means 'to have sex with somebody' as labeled in the LDOCE, the NSOED and CIDE. The noun prick meaning 'penis' is assigned the label taboo in the LDOCE, the CIDE, the OALD, and coarse slang which is synonymous with taboo in the NSOED, the noun pussy meaning 'the female sexual organs especially the vulva' is also labeled taboo in the OALD, the CIDE, the LDOCE and the NSOED.

Offensive is defined by the MED as 'extremely rude and likely to cause offense'. This is another typically prescriptive label where the lexicographer implicitly warns the dictionary user against using the word in question. Examples of nouns labeled accordingly are queen to mean 'a male homosexual who behaves more like a woman than a man', this is found in the MED, shit 'an unpleasant person who treats other people badly' is another example from MED, squaw 'a native American woman' is an example from the NSOED.

The standard/substandard dichotomy implies a right and a wrong in a language. Abecassis (2008: 2) states that dictionaries 'still strongly emphasize the standard language'. According to Landau (1989: 175), 'variations based on register have no right or wrong; in certain situations, a word may be appropriate while in others, it may be inappropriate'. An example of a word labeled substandard is ain't to mean 'is it not?' as found in the WBD.

Other labels used in English dictionaries that are indicators of prescriptiveness are vulgar and obscene. Vulgar is prescriptive because it is assigned to a word (s) to show that it is generally uncouth, crude or coarse. The implication here is that such a word should never be used because it is basically in bad taste. Obscene as a label is prescriptive because it suggests that a word it has been assigned to is abominable and therefore should not be used. Prescriptiveness limits the use of a given word to a specific situation of communication and yet, in certain situations, a word may be adequate and inadequate in others.

Tarp and Gouws (2008) and Bergenholz (2001, 2003 as cited in Tarp and Gouws 2008: 239) and Van der Merwe (2008: 345-346) explore the issue of dictionary tone/approach and make recommendations that are complemented by this study, to guide lexicographers in the area of dictionary tone.
9. Conclusion

There is no agreed-on criteria for making usage decisions in English lexicography. This is probably the reason for most of the problems concerning usage labelling in English dictionaries. It causes the inconsistencies which characterize usage labelling, and the problems, which emerge from such inconsistencies. As Stein (2002: 14) asserts, ‘it is admittedly very difficult to make objective assessments on the social status of words but it seems … that we need much more research in this area’. Leech and Nesi (1999: 295) recognize that dictionaries ‘fall well short of perfection’; I contend that lexicographers, in their quest for perfection, can benefit invaluabley from research from lexicographers as well as different perspectives of users, like the second language user perspective provided by the current study.

References


