

Towards a Southern African English Defining Vocabulary*

Lorna Hiles, *Department of Afrikaans and Dutch, Stellenbosch University, Stellenbosch, South Africa (lorna@dictionarian.co.za)*

Abstract: Controlled defining vocabularies have been used regularly in lexicography since the 1970s. They are mostly employed in learners' and school dictionaries and have been used to ensure that the definitions are easily understood by the dictionary user, and that the words used to define a lemma are not more difficult than the lemma itself.

In my experience, defining vocabularies compiled for English dictionaries for a British or American market are not entirely sufficient for southern African dictionaries. Words which would be well-known and used in southern Africa are not included in a foreign defining vocabulary.

With this article, it is my intention to find methods of adapting existing defining vocabularies and to compile one that is more useful for southern African school dictionaries.

A complete southern African English defining vocabulary is not part of the scope of this paper, but by providing some suggestions, I hope to make the reader realize that there is a need for a complete defining vocabulary that would cater to the needs of a southern African dictionary and its users. Further research would then be an extension of this paper into a full defining vocabulary (DV) that would be used for southern African dictionaries.

Keywords: CONTROLLED DEFINING VOCABULARY, DEFINING VOCABULARY, DEFINITIONS, LEARNERS' DICTIONARY, RESTRICTED DEFINING VOCABULARY, SCHOOL DICTIONARY, SOUTH AFRICAN DICTIONARIES, SOUTHERN AFRICAN DICTIONARIES

Opsomming: Op pad na 'n Suider-Afrikaanse Engelse definiëringswoorde-skat. 'n Gekontroleerde definiëringswoorde-skat word gereeld sedert die 1970's in die leksikografie gebruik. Dit word meesal in aanleerders- en skoolwoordeboeke benut en word gebruik om seker te maak dat die omskrywings geredelik verstaanbaar is vir die woordeboekgebruiker, en dat die woorde wat gebruik word om 'n lemma te omskryf, nie moeiliker is as die lemma self nie.

Die definiëringswoorde-skat wat vir Engelse woordeboeke vir Britse of Amerikaanse markte saamgestel word, is nie heeltemal voldoende vir Suider-Afrikaanse woordeboeke nie. Woorde wat in Suider-Afrika goed bekend is en gebruik word, word nie by 'n oorsese definiëringswoorde-skat ingesluit nie.

Met hierdie artikel is dit die voorneme om metodes te vind om die bestaande definiëringswoorde-skat aan te pas en om een saam te stel wat nuttiger is vir Suider-Afrikaanse skoolwoordeboeke.

'n Omvattende Suider-Afrikaanse definiëringswoorde-skat is nie deel van die bestek van hier-

* This article was presented as a paper at the Nineteenth Annual International Conference of the African Association for Lexicography (AFRILEX), which was hosted by the Research Unit for Language and Literature in the SA Context, North-West University, Potchefstroom Campus, Potchefstroom, South Africa, 1–3 July 2014.

die artikel nie. Deur 'n paar voorstelle te maak, behoort die leser daarvan bewus gemaak te word dat daar 'n omvattende definiëringswoordeboek nodig is wat in die behoeftes van 'n Suider-Afrikaanse woordeboek en die gebruikers daarvan voorsien. Hierdie artikel sou as wegspringplek kon dien vir voortgesette navorsing na 'n definiëringswoordeboek wat vir Suider-Afrikaanse woordeboeke van nut kan wees.

Slutelwoorde: AANLEERDERSWOORDEBOEK, BEPERKTE DEFINIËRINGSWOORDESKAT, DEFINIËRINGSWOORDESKAT, GEKONTROLEERDE DEFINIËRINGSWOORDESKAT, OMSKRYWINGS, SKOOLWOORDEBOEK, SUID-AFRIKAANSE WOORDEBOEKE, SUIDER-AFRIKAANSE WOORDEBOEKE

Introduction

Restricted defining vocabularies are used in pedagogical dictionaries — these being school dictionaries or learners' dictionaries. They are used to ensure that learners are able to understand the dictionary definitions with their vocabulary level. According to Herbst (1996: 324) a "controlled defining vocabulary is, quite obviously, used to avoid using words in the definitions which are less accessible to the learner than the word being defined". Defining vocabularies are mostly used for definitions, but also offer guidelines for examples and notes.

Defining vocabularies are developed and owned, for the most part, by commercial publishers. They usually contain anything from 2 000 to 5 000 words. They are generally compiled using corpus frequency data, which assumes that the most frequent words in a language are the ones that learners are taught first. "The lexicographer needs words of high generality, and some of these, such as *house*, are also very frequent, but others, such as *vessel*, are not." (Whitcut 1988: 49) The compiler then systematically works through a set of parameters, such as avoiding synonyms and antonyms, to determine which words are necessary to write definitions in a concise and simple way.

It has been found that existing defining vocabularies lack certain words that would make definitions more accessible to southern African learners, and therefore there is a need for a defining vocabulary that would be more suitable for southern African dictionaries — with regard to both their definitions and their example sentences. So, for this paper, methods that can be used to build a Southern African English defining vocabulary have been established. Two methods of examining the existing defining vocabularies have been used to determine what words can be deleted and what words should be added. A method of compiling a southern African defining vocabulary from scratch has also been suggested.

Defining vocabularies

The following existing defining vocabularies were worked with:

- Oxford South Africa, which contains 2672 entries. It was adapted in 2008

for the *South African Oxford School Dictionary 3e* and has been used with permission.

- Oxford 3000, which contains 3 540 entries, and is available on the Internet.
- Macmillan Dictionary Defining Vocabulary, which contains 2450 entries and is available on the Internet.
- Longman American Defining Vocabulary, which contains 2203 entries and is available on the Internet.

In the information about these defining vocabularies, the following comments are made: "The Defining Vocabulary has been carefully chosen after a thorough study of all the well-known frequency lists of English words. Furthermore, only the most common and 'central' meanings of the words on the list have actually been used in definitions." (Longman Dictionaries USA, 2014)

The Macmillan defining vocabulary, "... includes a few simple phrasal verbs like *put on* and *take off* (for talking about clothes), and a few simple phrases like *in order to* and *pay attention*. If they are not listed, we do not use them." (Macmillan Dictionary, 2009)

This research attempts to determine how much needs to change to make these existing defining vocabularies more useful for southern African lexicographers. These changes could be applied globally, but this research focuses specifically on the needs of southern African dictionaries.

Methods of examining defining vocabularies

The first method used in examining these defining vocabularies involved taking a random stretch of words from each of the defining vocabularies and combining them to compile a "megalist" of all the words in that particular stretch. See Appendix A. This megalist was used to search through a South African school dictionary from a different publisher. This dictionary does not use a defining vocabulary, so it was used to establish which of the DV words in the megalist are used in the unrestricted definitions.

The stretch used was from *lightning* to *low*. The megalist contained 62 words, while Oxford 3000 contained 53, Oxford South Africa 42, Macmillan 42, and Longman 34.

The Longman stretch contained only words, no phrases. The Macmillan stretch contained words and two phrases, *look after*, *look for*.

The Oxford 3000 stretch contained more phrases, *look after*, *look at*, *look for*, *look forward to*; derivatives such as, *like*, *unlike*, *likely*, *unlikely*, *load*, *unload*; and gave the part of speech where necessary, for example, *like*: *preposition*, *verb*, *conjunction*. It also provided International Phonetic Alphabet (IPA) pronunciation guides where necessary, for example at *live* /laiv/ and *live* /lɪv/.

The Oxford South African stretch contained all of the above, except that pronunciation guides were given as phonetic respelling, not as phonetic tran-

scriptions, and the phrase *look something up* was included as well.

Of the 62 words in the megalist, 18 were not used in any of the definitions in the dictionary. Examples of these words are *lodging, lodgings, locate, local, loneliness*.

This was an interesting exercise in comparison, but it would not be very useful as the only method used to compile a southern African defining vocabulary. For a start, one would need to check whether these words are used in many pedagogical dictionaries, not just one. The range of dictionaries would need to include different user groups, different age groups and different levels of English proficiency. Bilingual dictionaries could also be used, with the example sentences interrogated instead of definitions.

This method would, however, make a useful clean-up exercise, allowing one to take out words that are not used in any of the dictionary definitions, or words that are out dated.

The second method used was less systematic, but it was more useful. A selection of words was checked against the definitions in the school dictionary without a defining vocabulary, and then against the defining vocabularies. Words that were in the school dictionary but not in all four of the defining vocabularies were collated.

Lexical sets were mostly adhered to — for example, animals, parts of the body, technical words, foods, school words and irregular plurals, but the selection of words was otherwise random, and according to the author's intuition.

This process presented some interesting results.

One thing that became apparent was a lack of consistency within lexical sets: for example, *herb, spice, spicy* and *flavour* were all in a different number of the defining vocabularies.

Moreover irregular noun forms were not included in all the defining vocabularies, for example, *mice, feet, teeth, children, men, and women*. *Mice, feet* and *teeth* are certainly different enough from the singular that they cannot just be assumed. The convention for defining vocabularies is that "all normal forms of nouns, verbs, and adjectives" can be used in definitions (Macmillan Dictionary, 2009) so it allows for plurals, both regular and irregular, to be assumed. However, irregular plurals and verb forms are not predictable based on their base forms.

One needs to bear in mind that the defining vocabulary is a list of words that the user *should* know. Therefore irregular plurals should be considered for inclusion in a defining vocabulary.

In table 1, one can see that the irregular plurals mentioned above are in zero, one or three of the defining vocabularies. The third column shows how many times they appear in definitions.

School-related terms (Table 2) and technical terms (Table 3) that do not appear in all four defining vocabularies suggest that these defining vocabularies are out of date. The inclusion of modern school-related terms is an indication of the relevance of the defining vocabulary. In South Africa, the curriculum was given an overhaul in the early 2000s and terms such as "pupil" and

"standard" were dropped in favour of "learner" and "grade". It has not been possible to establish when two of the three foreign DVs were most recently updated. The Oxford 3000 was updated at the time of compilation of the Oxford Advanced Learner's Dictionary 6th edition, which was published in 2000.

	number of DV	number of definitions
mice	0	1
men	0	23
women	0	37
children	3	90
feet	0	43
teeth	1	38

Table 1

	number of DV	number of definitions
classroom	2	2
learner	1	3

Table 2

	number of DV	number of definitions
email	3	8
Internet	3	40
software	2	5

Table 3

This process would also be more useful if one checked the words against more southern African dictionaries. Again, the range of dictionaries used would be important. This method is more useful at providing or suggesting words to be added to a southern African defining vocabulary. See Appendix B for a list of words that are used in definitions in one South African school dictionary but do not appear in all four defining vocabularies.

The third method one can use to build a southern African defining vocabulary would be to examine a corpus of southern African English, and then compile the defining vocabulary from scratch. One could then use the two methods detailed above to delete unnecessary words, or add words that are not

necessarily frequent but that are required.

Conclusion

To conclude, existing defining vocabularies can be cleaned up a lot in order to compile an up-to-date southern African English defining vocabulary, and using a range of dictionaries as checks would be a valuable exercise. As mentioned above, both methods used to adapt existing defining vocabularies can be used more extensively and checked against more dictionaries to make them more accurate tools. One could also extend the search to example sentences and try to establish whether the words used in the examples are words that the typical user of a particular dictionary is likely to know.

Bibliography

- Herbst, T.** 1996. On the Way to the Perfect Learners' Dictionary: A First Comparison of OALD5, LDOCE3, COBUILD2 and CIDE. *International Journal of Lexicography* 9(4): 321-357.
- Longman Dictionaries USA. 2014. *The Longman American Defining Vocabulary* [Online] Available from: http://www.longmandictionariesusa.com/longman/defining_vocabulary [Accessed 30 June 2014].
- Macmillan Dictionary. 2009. *Macmillan Dictionary: Clear definitions* [Online] Available from: <http://www.macmillandictionary.com/learn/clear-definitions.html> [Accessed 30 June 2014].
- Whitcut, J.** 1988. Lexicography in Simple Language. *International Journal of Lexicography* 1(1): 49-55.

Appendix A: Megalist of words in one stretch taken from four existing defining vocabularies

lightning	literary	local	long	lost
like	literature	locally adv.	look	lot
likely	litre	locate v.	look after	loud
limb	little	located adj.	look at	loudly
limit	a little det., pron.	location n.	look for	loudness
limited	live /la?v/ adj., adv.	lock v., n.	look forward to	love n., v.
line	live /l?v/ v.	lodging	look sth up	lovely adj.
link	living adj.	lodgings	loose adj.	lover n.
lion	lively	log	loosely adv.	low adj., adv.
lip	liver	logic n.	lord n.	
liquid	load	logical adj.	lorry n. (BrE)	
list	loaf	loneliness	lose	
listen	loan	lonely	loss	

Appendix B: Words that are used in definitions but do not appear in all four defining vocabularies

	Number of DV	number of definitions
adventure	3	2
bucket	1	2
children	3	90
classroom	2	2
container	3	115
disease	3	66
email	3	8
feet	0	43
flavour	3	4
frog	0	2
goat	3	9
herb	0	7
Internet	3	40

learner	1	3
men	0	23
mice	0	1
permanent	3	8
rabbit	2	4
rat	2	1
sauce	3	12
shellfish	1	3
shiny	3	29
sickness	0	3
software	2	5
specific	2	1
spice	2	2
spicy	1	9
spider	2	4
teeth	1	38
temporary	3	5
thunder	1	2
tribe	2	6
women	0	37
worm	1	1
zoo	0	2