



## Deseret Language and Linguistic Society Symposium

---

Volume 15 | Issue 1

Article 11

---

3-14-1989

### Supermarket Lexicography or Finding New Terms to Put in the Dictionary

Robert L. Good

Follow this and additional works at: <https://scholarsarchive.byu.edu/dlls>

---

#### BYU ScholarsArchive Citation

Good, Robert L. (1989) "Supermarket Lexicography or Finding New Terms to Put in the Dictionary," *Deseret Language and Linguistic Society Symposium*: Vol. 15 : Iss. 1 , Article 11.

Available at: <https://scholarsarchive.byu.edu/dlls/vol15/iss1/11>

This Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in Deseret Language and Linguistic Society Symposium by an authorized editor of BYU ScholarsArchive. For more information, please contact [scholarsarchive@byu.edu](mailto:scholarsarchive@byu.edu), [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

SUPERMARKET LEXICOGRAPHY  
OR  
FINDING NEW TERMS TO PUT IN THE DICTIONARY

Robert L. Good  
ALPNET

For the past five years or so I have been the chief lexicographer at a local company called ALPNET that produces computer-assisted translation software as well as other natural language processing software. Part of the work I have been doing has involved the creation of monolingual English linguistic databases and bilingual dictionaries to support our software applications.

Creating dictionaries from scratch is not an easy task. The first obvious question that arises is "What am I going to put into this dictionary?" The answer, of course, depends on a multitude of other questions: What is the dictionary going to be used for? Who will use the dictionary? What features will allow the users easy access to the contents of the dictionary? What medium will it end up in? What are the size limitations? What sources should be consulted, and how, in order to create the dictionary? What format will the physical entries take? Will outside experts be required? What is the budget for this dictionary? What are the deadlines involved and how will they affect everything else?

These are questions every editor must answer to produce the best dictionary possible given the constraints as represented by the answers to these questions. If the dictionary is a wholly new creation, then identifying the list of headwords to be treated is a significant undertaking. Even if one is just making a revision of a pre-existing dictionary, it is still not trivial. It is likely that it will be necessary to add new terms to the dictionary to increase the desired coverage. Where do these lists of terms for consideration come from?

Traditionally, publishers of dictionaries have access to individuals whose job or hobby it is to read through copious amounts of text--newspapers, magazines, books of all sorts, professional journals, etc.--looking for examples of new terms or old terms used in novel ways. With the advent of massive amounts of electronic corpora and corpus analysis tools for analyzing them, it is now possible to have the computer do much of the work, at least in the step of identifying new terms. Much of my work at ALPNET has

centered around the design of computer-based corpus analysis tools for identifying new terms, both single-word and multiple-word.

The multiple-word terms have been of particular interest because, when translating from one language to another, word-for-word translations are rarely ever adequate. At ALPNET we have explored syntax parsing algorithms and implemented a system for identifying multiple-word terms. This has made it possible to identify potentially useful strings of words for automatic inclusion into our translation dictionaries.

But, in addition to manually reviewing texts or analyzing them using a computer, there is perhaps another source of new terms for the dictionary that does not appear to have been tapped yet. And that is the local supermarket. Most people probably are not aware that they cannot buy toilet paper, soda crackers, lunch meat, or dishwashing soap at any of their local supermarkets. They might get what they think is toilet paper, but this is not the label that appears on the packaging. Of course, once we are told that the advertiser's name for the product is not 'toilet paper' we instinctively go through some mental gymnastics to conjure up a suitable euphemism. The one that is used is 'bathroom tissue'.

Similarly, 'lunch meat' never appears on the package; the specific meat is always indicated. 'Dishwashing soap' is really 'dishwashing liquid' and 'soda crackers' are typically 'saltine crackers', at least as long as they have salt on them. One manufacturer called its regular and low salt crackers 'saltine crackers', but its unsalted crackers were merely 'crackers'.

None of these products claims to be toilet paper, lunch meat, dishwashing soap, or soda crackers. What we see here is a discrepancy between at least one common name and the generic name manufacturers have chosen to label these products. We have all seen this phenomenon with brand names that have become the common name for the product, like linoleum. We have also seen other cases where many of us use brand names as if they were generic names, even though the brand name is not yet in the public domain. Examples include Kleenex, Band-Aids, and Xerox machines, which are more generically referred to as facial tissues, bandages, and photocopying machines. These generic names are actually quite descriptive. But this is not always the case. Occasionally a manufacturer will pick a name that is really not descriptive at all. For example, if someone sent you to the store to buy some 'plastic strips' what would you bring back?

The term 'plastic strip' happens to be Curity's name for what I always call 'band-aids'--regardless of the brand--and what Johnson & Johnson insists on calling 'plastic bandages' so that they can retain a proprietary right to the name 'Band-Aid'. Curity's name is far from transparent.

Dictionaries seem to do a more thorough job of identifying terms that are proprietary but may also have a life of their own as generic terms in popular usage. This attention to detail is not so evident for non-proprietary names like 'bathroom tissue'. In fact, there appears to be quite a gap between the general monolingual desk dictionary's content with regard to these terms and the products on the shelves at the supermarket.

Appendix B lists the products and terms I checked the shelves for at a large local supermarket and then looked for in the dictionaries listed in Appendix A. I was sometimes surprised by the terms underrepresented in the dictionaries. Terms like 'eye drops', 'light' (meaning low in calories), 'ice-cream sandwich', 'lunch meat', 'Q-Tips', and several others seemed to be worthwhile terms.

On the other hand I was somewhat surprised at some of the generic terms chosen by manufacturers to call their products. Everyone knows what an ice-cream cone is, even though none of the dictionaries explicitly state that it does not have to be in the shape of a cone. I call the flat-bottomed ones cones, as I do the truly cone-shaped sugar cones. But the manufacturers of the flat-bottomed ones prefer the term 'cup' to describe their product. That seemed a little strange to me.

After looking at the various products in Appendix B and their commercial names, what are we to conclude? Should any of these terms appear in a dictionary of English? This is a difficult question. The mere existence of a term does not mean that it ought to be included in a dictionary. And if the term is composed of more than one word, it is even more difficult to decide if it warrants individual attention. The size, scope, and budget for the dictionary naturally restrict the number of terms that can be included and the depth of their treatment. The frequency and distribution and durability of each term must also be considered. Nonce words do not make good entries in a general dictionary. Neither do terms so specialized that no one is likely to ever come across them. But a good selling point for a dictionary can be the number of "new terms" that have been added since the last edition.

Simon and Schuster has recently published the third college edition of their New World Dictionary. The dust jacket mentions that 5,000 new terms have been included. I did not notice in the introductory material any discussion of these

terms, where they came from, or by what criteria they were admitted into the printed lexicon, but there are 5,000 of them and some of them are probably ones whose exclusion from earlier editions has puzzled us for years. The factors mentioned before concerning frequency, distribution, and durability, as well as the publishing requirements no doubt influenced the decision to include or exclude each term. The publishers very likely had many more new terms than they had room for or thought it wise to include without additional citations.

Some of the terms that I found browsing the shelves at the supermarket may not seem very important, but they have great exposure and distribution. Some of them are possibly the kinds of terms one would like to be able to look up in a dictionary. For example, when speaking of ground beef, we know that there are at least three kinds, differentiated by fat content: regular, lean, and extra lean. I don't know if there is a standard throughout the beef industry, but at one large local supermarket it states right on the package what the upper limits for the percentages are. I thought it was interesting that not only are the grades of hamburger missing from the six dictionaries I consulted, but also the word 'ground beef'.

Another specific set of terms that might be included are the ones describing what I call sliced cheese products. These are the individually wrapped "cheese" slices, some of which the supermarkets do not even refrigerate, which makes one wonder what is in them. The range of names for these are varied and some are so ominous sounding that I would really like to know what the differences are. Some of the names of these products include: pasteurized process (American/Swiss) cheese; pasteurized process cheese product; pasteurized process cheese spread; pasteurized process cheese food; pasteurized process cheese food substitute. The first one sounds like real cheese. I have no idea what I should think about the last one. Perhaps these names are dictated by law because of the content. If so, are they sufficiently standardized that they could be treated in a dictionary?

Dictionaries with more precise descriptions of consumer products could provide a useful service to the average consumer. Such descriptions might specify the different grades of ground beef, the general contents of cheese products, and percentages of butter fat in ice cream, ice milk, and light ice milk, etc.

I recognize that the editorial process requires that some terms be excluded in preference to others. The desk dictionaries that I consulted were certainly not unabridged dictionaries and there is only so much time, space, and money that can be devoted to making a dictionary. Even unabridged dictionaries have their limitations. I am not

claiming that any of these particular terms demands treatment to the exclusion of some other term. I am merely recommending one more place, and a very common one at that, to look for possible new words to put into the dictionary.

## Appendix A

### Dictionaries Consulted for this Study

#### Desk dictionaries

American Heritage Dictionary, 2nd college edition. Houghton Mifflin, 1982. 200,000 precise definitions; 25,000 new words and meanings.

Random House College Dictionary, revised edition. Random House, 1982. 170,000 entries.

Webster's New World Dictionary, 3rd college edition. Simon and Schuster, 1988. 170,000 entries.

Webster's Ninth New Collegiate Dictionary. Merriam-Webster, 1984. 160,000 entries and 200,000 definitions.

#### Unabridged dictionaries

The Random House Dictionary of the English Language, 2nd edition unabridged. Random House, 1987. 315,000 entries; 50,000 new words; 25,000 new meanings.

Webster's Third New International Dictionary, Unabridged. Merriam-Webster, 1981. 460,000 entries.

## Appendix B

### Supermarket Product Names and their Distribution Among the Six Reference Dictionaries

Parenttheses indicate possible generic names for the products that are evidently not used on packaging or they contain explanatory notes. The terms marked "general" often appeared in the dictionary with only their more general definitions rather than with the word sense desired for the specific context given here.

Term	Number of Dictionaries
Band-Aids (Trademark)	6
band-aids	2
bandages	6
plastic strips	0
pasteurized process (American/Swiss) cheese	0

pasteurized process cheese product	0	
pasteurized process cheese spread	0	
cheese spread	1	
pasteurized process cheese food	0	
cheese food	1	
pasteurized process cheese food substitute	0	
cotton balls	0	
cosmetic puffs	0	
puffs	5	
dishwashing liquid	0	
dish detergent	0	
dishwashing detergent	0	
liquid detergent	0	
eye drops	2	
eyedropper	6	
ground beef	0	
regular	0	general
lean	0	general
extra lean	0	
hot dog	6	
frank	6	
wiener	6	
ice cream	6	
ice milk	6	
light ice milk	0	
light	2	general
sorbet	5	
ice cream sandwich	0	
(ice-cream cone)	5	
cups	0	general
ice cream cups (flat bottom cones)	0	
sugar cones (pointed cones)	0	
cone	3	general
facial tissues	2	
tissues	6	
imitation krab flakes	0	
imitation krab salad	0	
krab	0	
(lunch meat)	1	
cold cuts	6	
luncheon meat	3	
paper towels	0	
towels	6	
cotton swabs	0	
Q-Tips	2	
swabs	6	
soda cracker	6	
saltine cracker	0	
cracker	6	
saltines	6	
bathroom tissue	1	
bath tissue	1	
(toilet paper)	6	
(toilet tissue)	3	