

Knud Troels Thomsen
The Danish Dictionary

On the Hardships Involved in Creating a Labelling System for Usage Restrictions in a Comprehensive Monolingual Dictionary

Abstract

This paper deals with the process of making a two-component system for labelling language usage in The Danish Dictionary (to be published 1999). It is assumed that some basic language restrictions are dependent on the inherent property of *metaphoricality/figurativeness*. Some labelling approaches in the German lexicographic tradition are discussed.

0. Introduction

Among meta-lexicographers it has been discussed whether usage restrictions are part of the meaning of a word¹, or whether usage restrictions and meaning description are complementary to each other. For a discussion see (Ludwig 1991:5ff and 227). We partly agree that usage restrictions are complementary to the meaning description. Partly, because we think it is crucial to distinguish between restrictions that are general, which means that they may restrict any word, and restrictions that are specific, which means that they are not general. In the case of specific restrictions, the relation between word meaning and restriction is that of *dependency*, whereas the general restrictions are complementary to word meaning, which means that there is no direct relation between word meaning and usage restriction. Furthermore, in consequence of the dependency relation, we cannot regard usage restrictions as a more or less linear sequence of different, but equivalent criteria as stated e.g. by Hausmann (in HSK 5.1:650f). Those are the starting points for our two-component language restriction labelling system.

1. The two-component labelling system

The most important precondition for labelling usage restrictions depends on the presence or absence of figurativeness or metaphoricality, which is an inherent property of a word. The reason for this division lies in the way words function, depending on whether they are metaphorical or not. The non-metaphorical words can form ontological taxonomies based on conceptual relations like 'superordinate-subordinate', 'part-whole', 'before-after', &c, whereas the metaphorical words can form only semantic

taxonomies, e.g. in the form of word fields, and they therefore cannot be used for structuring ‘world knowledge’. There are cases in which this rigid criterion does not hold. Especially in LSP texts one finds many more or less figurative words like the *shoulders* of a bottle, the *arms* of a river, &c. In those cases the ontological conceptual taxonomy criterion wins, which means that if the word fits into a taxonomy, where it can be defined through its relations to other words (concepts), it should be listed as a non-metaphor. This division also has implications for the meaning description. The word *idiot* is used in two senses: 1. A mentally retarded person; this meaning is defined through its objective relations to other words in a taxonomy of *degrees of retardation*, whereas in the second, metaphorical meaning: 2. A foolish or stupid person, *idiot* can be used subjectively to characterize any person. – Since the above mentioned distinction between plus or minus metaphorical has implications first of all for the specific restriction labels, we bring the initially made system for labelling specific restrictions in Figure 1.

Specific restriction labels

| Class-formers (non-metaphorical) | Non-class-formers (metaphorical) | |
|-------------------------------------|-------------------------------------|--|
| LSP Words | LSP Slang (jargon) | Subjectively marked language categories |
| General language | Argot | |
| | Taboo | |
| | Slang | |
| | (Slang) | |
| | Attitudinal words | |
| Grammatical words | | |

Figure 1

The categories in this figure are discrete, and cannot combine with each other.

1.1 Class-forming words

All the words in the left hand column can be described by means of their *genus proximum* and *differentiae specifica*e. Examples are for instance *table, bicycle, valve, universal joint, gear box, cartilage, carbine, &c.* Not only nouns can be categorized as class-forming words, verbs, adjectives and adverbs, too, may fit into some sort of ontological class hierarchies headed by typologizing expressions. On the horizontal level the non-class-forming words are divided into LSP words and general language.

1.1.1 Class-forming subcategories

LSP words are ascribed to a domain, which is indicated by a domain-label. We use a fixed list of domains. Following Ludwig (1991: 249ff) every LSP word should be labelled for domain. We use only a domain label, however, if the user cannot infer the domain a word belongs to from the meaning description. In the opposite case, we regard the label as redundant, and consequently it will not contribute further to the description. General language is used unmarked, since it is used in a **non-specialized** way, and therefore it remains unlabelled. Ludwig (1991:232) states that non-markedness of a word, means that the word can be used in any context without restrictions. We don't agree with him, because 1. in general the actual use of language depends on the communication situation, and derived from that 2. LSP texts very often require a high degree of specificity, which it is not possible to obtain by means of general language expressions. Consequently, it might be very inadequate to use a general language expression in an LSP text.

1.1.2 Non-class-forming subcategories

In contrast to the class-forming words, the categories of the non-class-forming words do not have the same ability to form ontological taxonomies, like class-forming words, however, they can form semantic taxonomies, preferably hierarchies and/or word fields. The word classes that fit into this category are the same as the above mentioned. The relation between class-forming-words and non-class-forming words is that of lexical **synonymy**. – In this main category the following subcategories are placed: **LSP slang** (=Jargon), **argot** (of social groupings), **slang, (slang)**, **attitudinal words** and **taboo words**. By **(slang)** is meant words that were originally conceived of as slang words, but which are (now) in common use. **Grammatical words** like prepositions, con- and subjunctions, some adverbs, interjections, &c, do **not** form any hierarchies at all, but they can form syntactic or semantic classes.

2. General restriction labels

General restriction labels can combine freely with each other and with the specific restriction labels.

2.1.1 Geography: regionalisms (not dialects!)

2.1.2 Chronology: *childrens language, young people's (teenagers') language, old people's language, archaic.*

2.1.3 Frequency (absolute or relative): *rare.*

2.1.4 Medium: (frequency in) *spoken or written language.*

3. Why changes were necessary

Looking back, the above mentioned labelling system contained several evident inconsistencies. Thus, it was quite difficult to distinguish between *taboo* and *slang* words, because the relation between the two categories is unidirectional: all *taboo* words belong to the *slang* category, whereas *slang* words are not necessarily *taboo* words. Therefore the category *taboo* words was removed and included in *slang*. Likewise, it was difficult to distinguish between (*slang*) and *slang*. Not because the categories were included in each other, but because according to the definition (*slang*) did not belong to marked language. Consequently the (*slang*) category was combined with *slang* and put into the same line bridging the categories *general language* and *slang*. Furthermore, *argot* did not meet the needs for registering language usages, which are bound to certain situations, so we made a new category of marked language, called *situational jargon*. This category should also meet the needs for classifying words as *solemn, poetic*, because with Hausmann (HSK 5.1:650) we believe that not the words themselves, but the **situations** or the **texts** are *solemn, poetic* &c. Another problem turned up. In very many cases we were reluctant to categorize as *slang* some metaphorical usages which did not fit into the other *marked language* categories. A word like 'flash' in connection with 'eyes' is metaphorical without being slang. The first major revision resulted in the making of a new general usage restriction category called *Value* with the subcategories: *taboo, informal, formal, euphemistic/pejorative*, a new specific category: *situational jargon*, and with the deletion of *taboo, (slang)* and *attitudinal words* as subcategories equal to the other non-class-forming subcategories.

4. Why further changes were necessary

This new system, too, had several basic problems: 1. the delimitation of the various marked language categories was too difficult, 2. there was no clear distinction between for instance *metaphorical* & *informal* language use as opposed to *slang*. This meant that the rôles of the general and the specific markers respectively were inconsistent. There were severe problems with

defining not only the situations, but also labels for the category *situational jargon*. It turned out that most of the labels were more or less equivalent to text genres. We therefore decided to make a new general category called *Genre*, which is an open list of defined text types (ex: lyric, newspaper article, after-dinner speech, funeral speech, &c). This step allowed us to label all sorts of expressions according to their Genre-boundness. This also seems to be an advantage with regard to e.g. grammatical words occurring (almost) exclusively in LSP texts (instructions, forms, law texts &c). Furthermore, there was no consensus among the editors as to when a word should be categorized as *slang* or not. We think this might be due to the fact that some people regard *slang* as very chic or smart, whereas other people think of slang as 'low level language', see also Ludwig (1991:244). Consequently, we left the category *slang*. Instead of *slang* we combine the inherent property of *figurativeness* with the value-subcategory *informal* or *taboo*. If a word happens not to be figurative, the subcategories *taboo* and *informal* can be used alone, of course. The subcategories of *Jargon* and *Argot* will be suffixed by hyphen-'slang', e.g. *rocker-slang*. The final system for specific restrictions is shown in Figure 2:

| | | |
|-------------------------------------|-------------------------------------|--------------------|
| Class-formers (non-metaphorical) | Non-class-formers (metaphorical) | |
| LSP Words | Jargon | Marked language |
| General language | Argot | |
| Grammatical Words | | |

Figure 2

To this system for labelling specific restrictions the complementary system for labelling general restrictions must be added: *Frequency*, *Geography*, *Chronology*, *Medium*, *Value*, *Genre*.

Compared to lexicographic practice a large number of distinctions cannot be found within this framework. For some of the missing categories this is due to the fact that the restriction is implied by the meaning description, and they

therefore seem redundant. Ludwig (1991:255f) mentions e.g. some derogatory words in German which are all labelled “spött.”(spöttisch=derogatory):

“Apostel spött. 'jmdn, der sich mit (allzu) großem Eifer für etw. ... einsetzt' ... Hinterwäldler spött. jmd., der völlig weltfremd, hinter der allgemeinen Entwicklung zurückgeblieben ist ...”, &c (Ludwig 1991:255f)

In the above – mentioned examples, the meaning description includes the derogatory meaning, and so it seems unnecessary to add a restriction to the entry.

5. Summing up

As some sort of concluding remark, I would like to compare the above presented labelling system with Hausmanns “Makromodell der Markierung im Wörterbuch” (HSK 1989: 650ff). The similarities of the two approaches lie in the linearity of the categories: *Chronology/Zeitlichkeit*, *Geography/Räumlichkeit*, *Medium/Medialität*, *Genre/Textsorte*, *Frequency/Frequenz* and partly *Value/Formalität*. So what happened to the rest of the categories respectively? Hausmanns category “Nationalität” we regard as an etymological kind of information, the “sozio-kulturelle Gruppe” is partly comparable to our *Argot*-category, partly to the *Chronology*-category, e.g. as far as “Kinder” is concerned, since this occurs to be more a matter of age than a matter of social grouping. “Technizität” is in some respects comparable to our *LSP language*, with the exception that this category has an other status in our system, since it is defined primarily through conceptual non-metaphorical relations. “Attitüde” is part of our *Value*. “Normativität” (normativity) is dealt with in another way, since we do not consider this a usage restriction in the narrow sense of the word. But the crucial difference, I think, lies in the fact that due to the linearity Hausmann’s system does not foretell which categories combine and which do not.

Notes

- 1 By *word* we refer to any language sign that might be thought of as an entry.

References

- Coseriu, E. 1971: *Sprache. Strukturen und Funktionen. XII Aufsätze zur allgemeinen und romanischen Sprachwissenschaft*. Hrsg.: Uwe Petersen. Tübinger Beiträge zur Linguistik 2, 53–72.
- HSK 5.1= *Handbücher zur Sprach- und Kommunikationswissenschaft. Wörterbücher: ein internationales Handbuch zur Lexikographie*= Dictionaries/Hrsg. F.J.Hausmann, Berlin; New York; de Gruyter, Band 5, Teilband 1.
- Ludwig, K.–D. 1991: *Markierungen im allgemeinen einsprachigen Wörterbuch des Deutschen. Ein Beitrag zur Metalexikographie*. Lexicographica Series Maior 38, Niemeyer, Tübingen.