

The Codification of Usage by Labels

Maarten Janssen, Frank Jansen, Henk Verkuyl

1. What is a label?

In order to get a clear picture of the notion of label let us consider an example of a lexical entry taken from a well-known English-English dictionary:¹

diffuse (difiu.z), *v.* 1526. [-diffus-, pa. ppl. stem of L. *diffundere*; see prec.] †1. To pour out as a fluid with wide dispersion; to shed -1734. 2. To pour or send forth as from a centre of dispersion; to spread widely, shed abroad, disperse, disseminate 1526. *fig.* to dissipate 1608. 3. to extend or spread out (the body, etc.) freely (*arch.* and *poet.*) 1671. 4. *intr.* (for *refl.*) To be or to become diffused, to spread abroad (*lit.* and *fig.*) 1653. 5. *Physics.* To intermingle, or (*trans.*) cause to intermingle, by diffusion 1808. 6. to distract. Lear I,iv. 2.
1. *Temp.* iv, I 79. 2. D. thy riches among thy friends, JOHNSON. To d. geniality around one MASSON. 3. See how he lies at random, carelessly diffused MILT. *Sams*, 118. ...

The Shorter Oxford English Dictionary on Historical Principles. 1972

The entry contains pieces of information such as (v.), (1526), (fig.), (intr.), (poet.), (arch.), (Physics), (†), (refl.), (lit.), (adv.) and (trans.), mostly abbreviated. In most dictionaries we find similar indications: (dial.), (inf.), (coll.), (loc.), (vulg.), (Am. Eng.), (Art), etc.² They are generally called *labels*.

Some of them are connected to formal aspects of the word, some of them to its meaning. To make this more precise, if one considers an entry as a form-meaning pair <f,m>, then labels like (v.), (pl.), (refl.), etc. are generally considered as belonging to the f-side. They concern a specific form or subform of the headword to which a certain meaning is given. Labels like (trans.) and (intr.) are formal from the grammatical point of view: they concern information about the format in which a certain verb needs a direct object or not to express

¹ We will use the term *entry* for the whole article (= headword or catchword + definition) and the term *headword* for the word form described in the entry.

² This footnote will contain all the labels used in the present chapter. We will order them alphabetically: † = obsolete, 1526 = first occurrence in 1526, adv. = adverbial, Am.Eng = American English, Angl. = Anglicism, arch. = archaic, coll. = colloquial, fig. = figurative, Gall. = Gallicism, Germ. = Germanism, inf. = informal, intr. = intransitive, lit. = literary, loc. = local, off. = offensive, pej. = pejorative, poet. = poetic, refl. = reflexive, reg. = regional, trans. = transitive, v. = verb, n. = noun, vulg. = vulgar. In the text we will write all labels x that we are going to discuss uniformly as (x).

a certain meaning. In the above entry the fourth sense can be analysed as <**diffuse**_{intr.}, to be or to become diffused; to spread abroad (lit. and fig.) 1653.>, which says that the verb *diffuse* in its intransitive form has the meaning 'to be or to become diffused, to spread abroad'. So, formal indicators, like (trans.), (pl.), and (n.) will not count as usage labels, even though in many dictionaries they occur at places where also usage labels appear.³

This distinguishes (intr.) clearly from labels like (lit.) and (fig.). Grammarians do not distinguish between literal and figurative forms. The same applies to labels such as (off.) or (vulg.) among others. In the remainder of the present chapter we will restrict the application of the term *label* to what are called usage labels; that is, to labels at the m-side of a form-meaning pair.

It is important to observe that a label may be argued to be simply an artefact of the traditional format of a dictionary: a book in print with a very restricted amount of space. In that sense, (inf.) could really be taken as just a shorthand for "an informal way of saying" saving 19 space units in a definition. Given modern technology allowing more space to electronic dictionaries, this offers the possibility of working away labels in longer definitions. However, this does not imply that labels will disappear. It only means that we have to make the notion of label independent of the specific medium in which a dictionary is presented. So, even if we read somewhere in a dictionary: <**buck**, an informal way of saying *dollar*> the entry in fact contains a usage label.

In general, usage labels provide specific information about the domain of application of the definition. In the more abstract sense just given, a usage label is to be taken as a higher-level instruction, as a meta-linguistic device. This means that it cannot be equated with the definition itself: it restricts the definition to a certain context. The definition of a word given by a dictionary entry is intended for a group of users belonging to those who speak or want to speak the standard form of the language of the dictionary in question⁴. It is

³ One reason to make this distinction is the following. Suppose that (trans.) is taken as a usage label on the m-side to distinguish the transitive use of a verb from its intransitive use. Then by the same reasoning one could take (n.) and (v.) as usage labels too. In the case of *delegate*, for example, they would distinguish the substantive use from the verbal use.

⁴ The same applies to bilingual dictionaries that are assumed to give translations from one standard to another standard.

with respect to the standard use of a language that usage labels find their justification:

Dollar and **buck** have the same meaning, but differ in another way. **Buck** is informal in style, so it would not be a suitable word to use in a business letter. Information about the style of the word, or the kind of situation in which it is normally used, is provided in the dictionary. (LDOCE, p. F27)

In this example two words are asymmetrically related to a norm: *buck* is marked as informal, whereas *dollar* has a default value. It would not make sense to provide a label (inf.) for the Dutch word *huppelen* (= to frolic) saying that it is an informal word, since there is no alternative word available⁵. Usage labels like (inf.) or (vulg.) find their justification in helping to choose appropriately between alternative words applicable to the same situation. Sometimes there are entire ranges of alternatives, as in the domain of sexual words providing a host of (near-) synonyms ranging from the extremely formal to the utterly vulgar.

Someone reading a business letter containing the word *buck* will generally not consult a dictionary to see whether the word is appropriate or not in that context. This suggests that the incorporation of labels in dictionary entry is mainly justified for the purpose of language production. When writing a text, one often has to make choices that may be made dependent on the public that is supposed to read the text. Labels are supposed to guide the writer through a set of alternative options.

In general, it is the marked alternative that is labelled because an unlabelled part of the definitions is considered to have the default value. The reason for marking a certain use has traditionally been to warn users about the possible social consequences of a word. Until recently, most European dictionaries have given labels a prescriptive, normative force, the current tendency being to give them a more descriptive load. For example, for the Dutch Van Dale it was customary to label words adopted from foreign lan-

⁵ This does not hold for all kinds of labels. Some group labels (to be introduced) can occur without alternative: that there is no common Dutch alternative for the dialectal word *onk* (odd as in 'missing its pair') still makes it a valid dialectal word.

guages as (Germ.), (Gall.), (Angl.), etc., explicitly indicating a negative opinion about their use as loan words. The current label (< German) is supposed to leave it to the users to decide whether or not they want to use it. Whether users accept the transition from prescription to description is an open question.

With these basic considerations concerning labels in place, the rest of this chapter will in turn focus on the two major questions regarding labels. Section 2 will consider the question of what kinds of labels there are, whereas section 3 will focus on the function of labels.

2. Classification of Labels

The usage label was described above as a restriction on the domain of application of a word. We will distinguish between two sorts of domain by speaking about *group labels* and *register labels*⁶. This distinction reflects a difference between characterizing a word as used by a group of speakers in a specific domain, and guiding an individual language user in making an appropriate choice between alternatives.

2.1. Group Labels

Group labels are labels indicating that a word (or word meaning) is restricted in its use. Following common practice, four kinds of group labels are distinguished: geographical, temporal, frequency and field labels. The restrictions marked by these labels concern regional, professional or social domains or a temporal restriction on the application of the word or word sense (Cf. Zgusta 1971). Each of these four classes will now be discussed in more detail.

Geographical labels indicate that a certain word is marked as not belonging to the standard language because it is only used in a certain region. The clearest of those are labels like (AmE.) or (SAE.). They warn the user that the labeled word is not Standard English, but only used in America or South Africa. Differences between alternatives can be related to either side of the <f,m>-pair. Consider British and American English: the two entries <**behav-**

⁶ This distinction is an adaptation of the use/user distinction discussed in Milroy & Milroy (1990), and further developed in Crenn (1996). Many of the traditional classifications, like Hausmann (1989) do not use this bipartition.

ior, the way a person behaves...> and <**behaviour**, the way a person behaves...> have a different spelling. Including the second pair in an English dictionary provides an f-difference that can be taken as a higher-level instruction not to use the headword in question in British English. At the other side, the different British and American English meaning of *pissed* in the pairs <**pissed**, very drunk> and <**pissed**, very annoyed>, illustrate an m-difference.

A British dictionary will have to mark the first members of each pair as (AmE.), but it could have decided to not include the information about the American counterparts. In that case, the lexicographer could have made an excellent British dictionary. Why is it then that the labels are given? The answer is that, although speakers of British English are aware of the considerable overlap between American and British English, they turn out to be inclined to overestimate the overlap by taking their own variant as standard. So, in this sense (AmE.) is a warning sign.

In the case of the American and British dictionaries, the problem is that English has developed two standards. The *Grote Van Dale* faces a different situation: it contains Dutch words used in The Netherlands as well as in Belgium. In all previous editions it presupposed a standard with respect to which many Flemish words are to be marked. As a result the label (Flem.) is put on a par with labels for words occurring in Dutch regiolects or dialects. Eventually this might conflict with current cultural, political and social tendencies to create an own language norm in Belgium itself.

Labels like (reg.) and (dial.) prevent that words occurring in regiolects and dialects are considered to belong to the standard language. They mark the peripheral status of the word in the standard use. In the Dutch tradition their occurrence in dictionaries is often quite arbitrary. Mostly they are there because some previous lexicographer attested the word in one of the works of an admired literary writer. For example, Van Dale has included the regional word *rild* to provide a possibility for a Dutch reader to find its meaning just in case they happen to read the now obsolete author Maurice Roelants who used it in the phrase *rilde naakte knapen* (slender naked boys).

Obviously, the justification of including word with labels as (reg.) or (dial.) in a standard dictionary is to be made dependent on the chance for the word in question to become naturalized as a (near-) synonymous alternative to an existing standard word. In that case the label should be on its way out. This

means that eventually a word like *rilde* will have to be removed from the dictionary if it fails to penetrate into the standard language.

Temporal labels can be divided in those signalling the first occurrence of a word in the sources of the dictionary and those indicating its last occurrence. In the entry of *diffuse* above we see dates like 1526 and 1608 and at the bottom of the entry the relevant texts are given.⁷ Some dictionaries restrict themselves to cite authors whose dates are known. Lexicographers never had the instruments to reliably register the occurrences of a word in domains outside literary sources. It follows that temporal labels marking the first occurrence of a word or a new sense are not well founded in empirical investigation and also that they are rather rough-grained. The same applies to last occurrences.

Given the existence of a word, the last occurrence of it may be accidental: it may be re-used for some unpredictable reason. This even holds in such cases where the label is seemingly well defined, as in the case of Webster, which uses a label (obs.) to indicate that no appearance has been registered since 1756.

There are two ways for a word to lose its firm position in the standard language: (a) the word has been pushed aside by another word; or (b) people do no longer speak about its referent. The archaic use of *diffuse* in its third sense labelled as (arch.) is a special case of (a): the sense 'to spread out freely' has disappeared according to the lexicographer. As to (b), if a word pertains to something that no longer exists, a label like (hist.) can be used to express that fact. There may be alternatives for these labels. For example, Van Dale uses the past tense to define *schout* (bailiff) in its historical sense implying that this meaning no longer applies. In the same way, it uses (Ind.) to mark the outdated reference to the function of police commissioner in one of the former Dutch colonies (now Indonesia).

Modern technology will most probably change the picture considerably: it is possible now or in the near future to follow the lifetime of individual words in detail. For example, it is an empirical fact that many words known to peo-

⁷ There are reasons to be skeptical about what is suggested by the entry: it is difficult to believe that the figurative use of the verb had to wait about eighty years in order to come into existence (cf. section 2.3).

ple in their sixties and seventies are no longer used by the generations of their children, let alone their grandchildren. This could be made visible in a dictionary by investigating the use of language of say three different generations and marking first and last occurrences of words and word senses (perhaps by defining thresholds for acceptance in a statistical way). This would break away from the paper dictionaries which still present a language as a constant, practically unchanging extended object, as the newest edition of Van Dale does by covering the period 1850 - 1999. In practice, however, one in five words changes every generation.

Frequency labels are generally not used in printed dictionaries, as we just noticed, although sometimes dictionaries indicate which one out of two forms is used most frequently.⁸ Nowadays, the frequency of word forms can be established with respect to large corpora and so each word can (in principle) be marked as to the number of times it occurs in a certain year or in a certain period that is interesting to... To whom? Not to the average dictionary users, who simply hope to find the word they need leaving the decision to include it to the lexicographer. It follows that frequency labels may be part of the databases underlying a paper dictionary rather than being included in the dictionary itself.

Field labels mark words as belonging to a certain professional or social domain. Even though speakers of certain professions use a general non-dialectic language, they use either word forms that are not part of the standard language or they use word forms that are part of it but have a very specific meaning within the field.

From the point of view of the dictionary user it is often interesting (and necessary) to know what the use of a word is in various domains. The Dutch word *wissel* can be translated into the English *bill of exchange*, *change*, *exchange*, *track*, *changeover* and *switch* dependent on whether it is used for money trade, hunting, sports or railway. The translation into apparently unrelated mean-

⁸ For example, Van Dale indicates by the label (w.g.) in *gaarbak* w.g. *vergaarbak* that the first form is used infrequently.

ings indicates that the Dutch word has a heavy load to fulfil. Its senses are determined by the fact that it is used in a specific professional field.

It is virtually impossible for lexicographers to include the terminology of the entire professional domain. For example, the number of terms used in the shipping industry is formidable. One cannot expect a dictionary publisher to include all the terms available, certainly not because there are so many other domains in which technical terms are necessary. Apart from that, professional domains undergo rapid technical changes. Ordinary dictionaries do not have the function to record the lifetime of a professional term.

In this respect, the transition from printed dictionaries to electronic databases will change the picture. It will be certainly possible to provide more or less coherent sub-domains expressing terms used to communicate about things going on in a profession. Due to the improved possibilities to record the occurrences of individual words, it will also be possible to have more information about the lifetime of a professional label as well as about its penetration into the language for which the dictionary is made. This latter process may either mean that the professional meaning will lose its labelled status or that the lexicographer will add a new more general sense without a label.

2.2. Register Labels

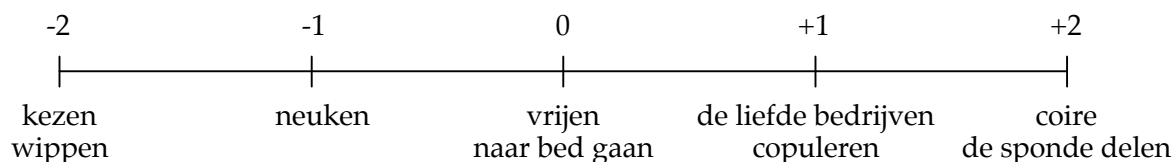
Language users generally operate in different social domains (family, employment, bureaucracy, church, social class, etc.) which are characterized by having a set of behavioural rules determining what can or cannot properly be said. They use different style registers to master this problem. What can be said when addressing an audience in a political meeting is quite different from what can be said privately. One is really not supposed to say that a journalist is a major league asshole if the microphone is open. Dictionaries want to protect people from using the wrong words in the wrong contexts.

A register label is therefore intended as a device to guide individual language users in their use of language with respect to a social group judging its appropriateness. A dictionary is often used by speakers with the need to address a certain audience, by poets, by writers and by journalists to avoid a word that they do not want to use or to find a more appropriate alternative for the specific purpose of the occasion. Some dictionaries provide synonyms

and near-synonyms marking differences among them by labels like (form.), (vulg.), (poet.), (bibl.), etc.⁹

There has been a long tradition in which the dictionary is considered as exemplifying the use of language in higher social classes. In that sense, the direction is one-way: lower class people are supposed to learn the words of the proper language rather than that higher-class people are supposed to learn vulgar words. The strategy of lexicographers is to consider unmarked words as having the default value in a median bandwidth. Below that area words are to be marked in order to warn people, above that area words are marked to indicate that these words only function in certain formal situations. In this sense *kick the bucket—die—expire* form a clear triplet. *To kick the bucket* is an informal way of describing the meaning of *die*, whereas *expire* is a very formal way of doing that.

At the end of the seventies one of the authors of the present chapter was involved in making a dictionary that could compete with the leading Dutch Dictionary De Grote Van Dale.¹⁰ The writers of the blueprint solved the problem of the formality register label by having a five point scale with the neutral value in the middle of the scale: -2 -1 0 1 2. That is, one can construct a scale as exemplified in the following table.



To identify this neutral value it was attached to the Dutch spoken by people presenting the NOS-journaal, the major Dutch news show known to every Dutch speaker: lexical variants observed in this show were neutral by defini-

⁹ The picture can be made more complex by labels like (Spoken) or (Written). The distinction between spoken and written language is quite important in French, where heavy restrictions exist on the application of words belonging to the spoken language to the realm of written language. Most French speakers use *bouffer* when they talk about eating and manger when they write about it, although *bouffer* nowadays occurs in newspapers like *Libération*. The strategy seems to be to marked the spoken usage by (pop.) or (fam.) rather than by (Written).

¹⁰ Due to the economic crisis at the time the enterprise ended before the blueprint for the dictionary could be made concrete.

tion. And in addition to this, all variants that could be imagined being used in the show without the raising eyebrows in the audience. The neutral value for the written variety was connected to two of the leading Dutch newspapers. An advantage of this approach is clearly that it provides a sense of asymmetry giving a negative direction and a positive direction. Everyone has access to a common neutral point of departure. This means that changes in the language of the anchoring point can be registered as a shift from 0 into the direction of -1, which then may become a new neutral value. The scale can be used to also fix a neutral value for the written language. For Dutch its zero point will be located a little bit more to the right of the zero point on the oral scale and thus it can be fixed by its relation to the spoken use of language.

In this way, dictionary makers can in principle follow the history register values in certain social domains. For example, the word *neuken* (fuck) was generally considered as -2, until the posh speaking and formally dressed well-known journalist Joop van Tijn used it for the first time publicly on TV in the early seventies. It has now gradually moved to -1, at some occasions even appearing in the original 0-zone of the seventies. For foreigners improving their Dutch with the help of a monolingual dictionary it should still be marked in order to warn for the social consequences if the word is used in certain social domains.

The downward values (informal) and (very informal) label the area of restrictions on the use of words in public, the upward values (formal) and (very formal) mark the area of juridical, scientific, pedantic, highbrow, posh language use. Sometimes the words marked like that are used appropriately if the social setting is scientific or highbrow, but often they tend to fall out of tune and that is exactly the reason why lexicographers use a warning sign. The five scale values can in principle be trimmed down to a one-, two-, three- or four-valued scale dependent on the number of alternatives of the word itself.

2.3. Figurative use

The idea of a scale also can be applied to literal and figurative use. This can be done by giving the literal use of the word a zero, and the figurative use a -1, for example: <**Sinterklaas**, Mythological Saint who gives all children presents on

his name day> receives a zero and <**Sinterklaas**, generous person> a -1. The negative value expresses the fact that in an analytical semantic theory that uses markers, the figurative use is always characterised by stripping one or more markers that are relevant in the literal meaning.

The application of a scale has two additional advantages. The first advantage is, that it enables the lexicographer to use the plus side of the scale as well. And in fact, this might be needed in the case of semantic specialisation, like in *situation* 'set of circumstances' (nog natrekken) which has developed into 'set of circumstances of a critical nature'. Along the same lines we could use the scale for words which enter the standard language from a specific field. Examples are psychiatric terms like Dutch *hysterisch* 'hysterical' and *neurotisch* 'neurotical'. which had a precise definition in the beginning of the 20th century, but developed rather vague meanings in the course of this century, as 'expressing exalted emotional behaviour' and 'expressing unreasonable, nervous behaviour' respectively.

The second advantage of implementing scales in the field of literal and figurative meaning is that it is flexible enough to describe all kinds of polysemous extension, which are to be considered as similar to the literal-figurative distinction. An example of what we have in mind is the 'picture of' extension, which we see in *nude* 'naked person' and 'portrait of a naked person'. This can be put on a par with the 'mock/play model of' extension, which we see in *garage* 'real building for real cars' and 'playground model of a garage for children to play with'. Both will receive a -1 mark, because (in spite of their long definitions) they lack several characteristics of the original objects. The lexicographer has another option as well: to omit the non-literal meanings altogether, because dictionary users simply need no help in dealing with polysemous senses of a word. They are aware that, given a zero value for a word, technical specialisation can add a +1-sense; this is the case with *water*, which was used in its daily zero sense long before its H₂O-sense was added. They also know that a word can start its life at +1 and then get a more general but less accurate default sense that has the zero value, as in the case of *schizofreen*. And they also know that a zero valued word can get a specific figurative meaning at -1. As said the number of values on the scale are not important: it

is the zero value and the two directions that are crucial for determining the relation between senses.

At this point one could try to solve the problem of how to present the different senses on a scale. The solution to this problem may result in doing away with labels like (fig.), because it is possible to simply make the figurative sense part of the enumeration of senses. One argument in favour of this would be the simple observation that lexicographers do not have a well-established theory about literal and figurative meaning enabling them to indicate precisely the figurative use of a word. This is no such semantic theory around.

The question is therefore what a dictionary user will miss if the label (fig.) is not given. Take it away in the entry at the beginning of the present chapter and see what difference it makes. Under sense 4 the figurative use of *diffuse* is said to be 'to spread abroad'. What would be the loss of information if the lexicographer had not given the two labels? Nothing, because dictionary users learn that *diffuse* can be used to indicate something that can be spread abroad.

2.4. The offensive use of words

The social pressure to promote so-called political correctness in the use of language has increased in the past decennia, in particular in the United States. It is obvious that this trend affects the use of labels in dictionaries. In general, European dictionaries turn out to be quite reluctant to give up their habit to put clearly offensive uses of a word under headings like (fig.) or (not literal). Most European countries have a colonial past and also a past in which minority groups have been treated in a way, which is nowadays considered very questionable. Traces of this past can be found in the language. There has been a tradition to consider these traces from the point of view of the majority, to which lexicographers in general belong. This has often resulted in the application of labels used for the non-literal application to the domains of racism, sexual offences, cultural differences, and so on. In particular, clearly racist

senses of words were and are still treated like that.¹¹ In the literature on labels one can find a host of labels trying to express that one should not use a word on the penalty of offending people among which (derogatory), (offensive), (disparaging), (sexist), (coarse), (rude), etc. etc.¹²

In the lexicological literature on labels there are attempts to distinguish between them on the ground that some of them can be said to concern the role of the speaker whereas other labels take the side of the listener or reader. Someone using *frog* for a Frenchmen is considered to have a bad intention to offend the French. The label (derogatory) is then said to warn dictionary users to evade this term even though many French are not at all offended by it. They even might consider it a honorary nickname, e.g. in sport circles. A label (off.) is then said to function as a warning sign that people in the extension of the term might be offended. That is, <**jew**, (off.) impostor> can be said to warn those who use the word *jew* in the sense of impostor¹³.

We do not consider this a productive way of looking at the function of these labels: to call impostors *jew* will generally not be taken as an offence by impostors. So, this label seems to be counterproductive and too superficial. What the lexicographer should say about the semantic connection between the pair <**jew**, impostor> is that it should not be treated on a par with the pair <**robin**, bird>. The latter format expresses that all robins are birds, whereas what the lexicographers want to express by <**jew**, (off.) impostor> is certainly not that all jews are impostors. This means that a systematic semantic pairing of the two words is simply wrong and its inclusion in the dictionary itself can therefore be argued to be offensive. One could, of course argue, that <**jew**,

¹¹ One of the present authors has analyzed in some detail entries like <**jew**, (fig) impostor> in Van Dale showing that this label is ill-chosen. It is easy to find European dictionaries in (university) libraries in which offensive use of language is labelled in terms of figurative use: [].

¹² See in particular Juhani Norri (2000).

¹³ We guess that the use of labels will never be precise enough to give an adequate description of the feelings of the population, as not only the word form itself and its extension are relevant, but also the characteristics of the people who use the questionable word: the in-group has more rights to use the term than the out-group, cf. the use of *nigger* by blacks. Furthermore, the context and situation in which the word is used is important: supporters attending a soccer match may be heard to use words (and get away with it), which are not accepted outside the stadium. These and other aspects cannot be accounted for in a label, and even not in this paper.

(stereot.) impostor> can be defined as the appropriate format to treat this type of socially unacceptable use of language, but it can be argued as well that dictionaries never should contain pairs in which a <**robin**, bird>-relation of inclusion is imposed on pairs of words whose referents do not stand in that relation.

Returning to the scale discussed above, one can easily see that the idea of a scale also applies to this particular domain: to use the word *frog* for a French person can be seen along the lines of figurative use discussed above: it receives a -1 value. Along the dimension of offensive use it also receives a -1 value. What the two values have in common is that in both cases one speaks about the non-standard use of the word *frog*. In other terms, offensive and figurative use have in common that they bring about polysemous extensions of a word having a neutral value.

Class	Subclass	Oxford labels
Group labels Indicate word as belonging to group of speakers	Geographical	Afr, dial, north, Amer, <i>etc</i>
	Temporal	arch. mod. obs.
	Frequency	freq.
	Field	Aer. Alch. poet. techn. <i>etc.</i>
Register labels Guide user in choosing between alternatives		colloquial, slang, jocular, derogative, vulgar, archaic, literary, euphemistic, figurative, pejorative suggested: very informal, informal, \emptyset , formal very formal

Table 1. Classification and examples of labels

In table 1, an overview is given of the classification of labels suggested in this section, with classes and subclasses. As an illustration, the labels as found in the Oxford dictionary are given. No explicit sub-classification of register labels is given, since though one can observe various classes like formality, offensive use, figurativeness, and mode of text, giving a list of subclasses would give an inappropriate sense of exhaustiveness and independence. Also, next to the Oxford labels, an example of a scalar labelling is given.

3. The functions of labels

With a classification of labels in place, we can now turn to the question: what is the use of labels? Why do labels appear in dictionaries and what purpose

are they meant to fulfil? The general conclusion will be that labels are less useful than they are commonly taken to be.

There are a number of conceivable uses of labels by people writing or preparing a speech in their mother tongue. These can be shown by some examples: consider a Dutch writer who has written the adjective *onk* (odd) in an article for the general public and who wishes to check whether this word is indeed standard Dutch. The label (gew.) (= regional) is a warning sign and he might therefore replace his *onk* for another word, or explain the meaning in the text. Or consider someone who wants to write an open letter about a politician who in his opinion is corrupt. He may feel the need to find the most derogatory name for the politician without the risk that his letter will not be published. So our writer will look up *flessentrekker* (crook) in his dictionary. If he finds no label attached to it, he could consider *flessentrekker* a name that is neutral enough to use.

We pursue the line of thought somewhat further. What level of specification is needed for the labels to make them useful? Does it really help someone to learn that *onk* is a word in the North-Hollandic dialects? The answer is negative, except for the very uncommon case in which a dictionary is used to obtain information about dialects. The same applies to the latter example: the newspaper may still find the use of the word politically incorrect without there being any label in the dictionary. This means that for productive use the labels in several categories could be trimmed down drastically. For example, instead of specifying all kinds of professional groups, one label would suffice: (technical). And in electronic databases one could provide a word with information about its occurrences in all sort of possible periods suppressing the signal function of temporal labels (more on databases in the next chapter).

Given these conceivable uses, is there any empirical evidence that monolingual writers in fact *do* use labels to adapt their lexical choices? The answer is a plain no. There are several reasons for this. The regrettably few empirical investigations of dictionary use indicate that there is little chance users will even read labels or read about them in the front matter. Wolf (1992) found in her survey of the use of monolingual dictionaries by DDR users that in more than 90% of the cases the users were interested in fixing formal, most orthographical problems. This percentage is considerably higher than the percentage for looking up some aspects of meaning (75%). Only a minority of 23%

claimed to use a dictionary to decide on a stylistic problem. Labels were never mentioned as helpful in this respect.

The same conclusion can be drawn from more recent and excellent presented surveys on the use of bilingual dictionaries. An analysis of the authoritative investigation by Atkins and Varantola (1997) – tapping the dictionary use of several hundreds of users, from experts to novices in English as a second language – reveals that the information provided by labels belongs to a wastebasket category of *Other types of information* comprising less than 5% of the total looking up activities. In the same vein, neither Scholfield (1999) nor Rundell (1999) do mention the use of label information. From a slightly different perspective, Höhne (1991) comes to the same conclusions: the dictionary is used for language advice on the stylistic level in only 3.2% of the cases, which is very low compared to orthography (58%), syntax (18%), and morphology (12%).

Confronted with this negative result, there might still be other situations in which register labels are more useful. Are they useful for productive users confronted with a task urging them to decide on stylistic choices? We did some small scale “working aloud” experiments to find out whether experienced dictionary users use labels in the following rewriting tasks: make an informal text more formal, and make a formal text less formal. The results were unequivocal: even in this situation, where people could reflect on the nature of the information given in an entry, the subjects made minimal use of its labels. There are several reasons for this, each of which is putting the usefulness of labels into perspective.

In the first place, users do not feel the need for assistance by label information. It is easy to understand why: experienced dictionary users are also experienced language users, capable of making a register choice on their own. And if they need the information provided by the label, there is a big chance that it will be overlooked. This is caused by its position in an entry, on a par with the grammatical labels that laymen usually do not understand. Also, the way in which usage labels are graphically presented is considered a signal of their minor importance for the content, in other words, as an implicit hint to skip them.

If a user has noticed the label, there is a chance that it will not be recognized correctly. Many of the labels are abbreviated, and some of those abbrevi-

viations (but not all) are difficult to complete. A few examples, from another small-scale experiment, should suffice. The majority of university students in Dutch language and literature did not manage to complete (bel.) to *beledigend* (offensive) and (min.) to *minachtend* (contemptuous). They thought (bel.) to be an abbreviation of *Belgisch* (Belgian), or *belangrijk* (important), and (min.) of *minder vaak* (less often).

Even if users are able to complete the abbreviated label some of them turn out to have problems interpreting its meaning. The students just mentioned were generally able to extend the label (gem.) to *gemeenzaam* (common) and (volkst.) tot *volkstaal* (slang), but they could not give a meaningful definition to these words. In fact, they were obliged to use the dictionary itself to see what the labels meant. If they happen to use the newest edition of Van Dale they do not find the meaning of the label only in a sense opened by the label (veroud.), which means 'outdated'. This also illustrates the Cinderella status of labels in lexicographic circles.

Finally, even if dictionary users understand the label, they are often insecure about what the relevance of this information is for the kind of decision they have to make about the text they are working on. We have reasons to believe that this last problem will grow rather than diminish in the future. That is because of the tendency to replace the explicit normative labels by descriptive counterparts, like in the case of the Dutch word *karaktermoord* as defined in Van Dale. We did not do a test but we suspect that our students can do nothing with the label information: (loan translation from Eng. *Character assassination*, character [=reputation]). At least we cannot.

4. Final remarks and conclusion

Empirical evidence shows that labels are hardly used, for a couple of reasons: they are often given in not directly clear abbreviated form, they are typographically hidden and the information given is mostly already known to the user. This situation might change with the coming of lexical databases. Over the last two decades, there has been a rapid progress in the development of digital versions of dictionaries, called *lexical databases*. Lexical Databases are commonly adapted versions of paper dictionaries, and hence share many of their properties. Still, the shift to a digital medium has many fundamental consequences, including some for the use of labels.

The most direct consequence is that the abbreviatory form of labels can disappear: given the increase of space, there is no longer a need to give labels in their abbreviated form. The unwieldy abbreviations can be evaded because their meaning can be made a (metalinguistic) part of the definition of the word in question. By supplying every dictionary entry with a list of sentences in which the word in question is actually used in a context that can be recognized as belonging to a certain register or as used by a certain group, lexicographers can *show* the usage of a word: if a word is only used in 'vulgar' contexts, it is clearly a vulgar word. By showing this, there is less of a need to also indicate it by a label. For educational purposes, showing is much better than telling. Still, as a quick reference, the presence of a label could still be desirable.

The objection might be raised that our treatment of scalar labels is at odds with our remarks about the questionable usability of label terms just mentioned. As observed earlier, dictionary users refuse to read the preface and the instructions carefully, so it follows that they will not understand the scales. We suggest that scalar labels of the -2 -1 0 1 2 kind have an advantage for the lexicographer. After all, even if all labels are abandoned, some of them return as a part of the definitions. This means that scales will be used structurally in the definitions themselves and this is to be preferred over the unstructured set of label terms which are currently in use. If one prefers to use labels in their abbreviatory form it is very easy to explain scales like '(very informal), (informal), (common), (formal), (very formal)'.

What we have tried to show in this article is that the information labels convey should be omnipresent in dictionaries, but in their present form they appear not as useful as they are often taken to be. Labels should be given the right place in making definitions. As soon as one realizes that they are basically metalinguistic devices, their proper treatment dictates itself dependent on the room a lexicographer has.

We have pointed out that a coarse distinction can be made between two kinds of labels: group labels (indicating a word as group-specific), and register labels (guiding the user to choose between alternatives with a different pragmatic load). The group labels can be further analysed in terms of the kind of group the word belongs to: a social group, a geographical area or a time period. A subdivision within the register labels is less easily established,

since words can be coloured by all kinds of interrelated things like level of formality, kind of text, figurative use, offensive use, attitude, etc.

Register labels indicate deviation from a norm, relative to more neutral alternatives. Hence, register labels only appear for words that have (near)-synonymous alternatives. Also, where there is more than one alternative, the various words can deviate in degrees from the norm. To capture this, it is best to have scaled labels. We have argued that the use of the (fig.)-label can best be understood against the idea of scalarity.

References

- Atkins, Sue & Varantola, K. 1997. Monitoring Dictionary Use. *International Journal of Lexicography*, 10:1, 1-45.
- Crenn, Tiphaine. 1996. Register and Register Labeling in Dictionaries. PhD thesis, School of Translation and Interpretation, University of Ottawa
- Hausmann, Frans Josef. 1989. Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht. In: Hausmann, Reichmann, Wiegand & Zgusta (eds.), *Wörterbücher, Dictionaries, Dictionnaires*. Berlin: de Gruyter.
- Höhne, S. 1991. Die Rolle des Wörterbuchs in der Sprachberatung. *Zeitschrift für Germanische Linguistik*. 19: 293-321.
- Norri, Juhani. 2000. Labelling of derogatory words in some British and American dictionaries. *International Journal of Lexicography*. 13:71-106.
- Rundell, M. 1999. Dictionary use in production. *International Journal of Lexicography* 12.1: 35-53.
- Scholfield, P. J. 1999. Dictionary use in reception, *International Journal of Lexicography* 12:1, 13-34
- Wolf, Brigitta. 1992. Wörterbuch und Benutzer, Versuch einer empirischen Untersuchung. Berlin: Brause & Viehweger.
- Zgusta, Ladislav. 1971. *Manual of Lexicography*. Den Haag: Mouton.