

## **The edges of definition**

### **Abstract**

There is a strange discrepancy between the vague and flexible concepts that words evoke in the mind and that are described by lexical semanticists and lexicographers as the meaning of the words and on the other hand the unequivocal referential use that speakers and hearers make of those words. In this paper it is suggested that concept and referential valency should be treated as mutually independent things.

### **1. Something on sense and reference**

The most significant distinction between definitions of word meaning in formal logic and in lexical semantics is that only in the former is quantification used. So in logical terms a noun like *bachelor* is defined as

$$\text{Ax}[\text{bachelor}(x) \quad \text{human}(x)\&\text{male}(x)\&\text{adult}(x)\&\text{unmarried}(x)]$$

while the word's linguistic definition can be something like

*bachelor*, unmarried man (where 'man' can be further analysed as 'adult male human being').

The meaning of this difference is that the logical definition aims directly at the referential value of the word, while in lexical semantics reference is only of secondary importance. What the logical definition does is try to determine for each individual the necessary and sufficient condition to be called truly a bachelor. The lexicographical definition is rather concerned with an idea, a concept of an individual having the right properties to be called a bachelor; reference is determined by correspondence of concrete individuals to that concept.

In fact lexical semantics has never shown excessive interest in the phenomenon: usually it has been wiped under the rug of something considered peripheral, like *parole*, performance or pragmatics. And this is stranger than it seems at first glance, if one takes into account the semantic properties of words in general. In fact the idea that word meaning is – in the first place – a concept implies that the word is an independent meaningful unit. But this implication is not exactly what should be expected from the very facts of language. Words are simply never used independently to express a message; they are used as parts of sentences (or as elliptic sentences), and as such they are only components of other truly independent

meaningful expressions. Words are in the first place – at least in affirmative sentences – elements in the calculation of a truth value or a truth condition. It should be evident that in such a calculation the word's referential value is involved, not the concept that it evokes. Or, to say it somewhat more tentatively: that the concept can only be involved insofar as it determines reference. This means that a sentence like *John is a bachelor* can be given an interpretation based on the concept evoked by *bachelor* if and only if this concept determines entirely the referential valency of the word. In other words: only an Aristotelian concept, in terms of necessary and sufficient conditions for reference, can be useful in sentence interpretation. Recent literature, however, has made clear that the above definitions for *bachelor* do not offer a sufficient condition.

This matter will be taken up again in Section 3, but first something should be said about the proper nature of reference in natural language.

## 2. Reference in natural language

The interest in and the treatment of referential matters are not innate to lexical semantics. They have been adopted from analytic philosophy and they still show their origin. Discussion of reference is restricted to clear cases of words having concrete and verifiable referents in the world, such as – only seemingly, as has been said – *bachelor*, or to undisputably analytic sentences like *All lizards are reptiles*, or even to the referential identity of proper names like *Tully* and *Cicero*. The unspoken idea behind this is that there are some clearly verifiable relations between natural language expressions and things in reality, or in some model of it which can be studied and that for less clear cases it holds that *wovon man nicht sprechen kann, darüber muss man schweigen*. Sentences of which the truth value cannot be proved should not be taken into consideration by semanticists. It should be clear, however, that this is only a valid restriction in cases where the treatment of philosophical problems in terms of natural language is concerned. But also where this is not the case, a judgment in the form of a natural language sentence on the world or on what is thought of as the world is meant by both speaker and hearer to have a truth value.

Consider such sentences as *John is an imposter* or *John made a good bargain*, in a normal context of use. In spite of the presence of words denoting vague or subjective concepts (*imposter* being an invective, *good* a subjective quality judgment) clearing the way for arbitrary judgments, these sentences are not noncommittal. They are clear and unequivocal judgments on situations in the world, and as such their meaning (i.e. their "truth value") is the same as their correspondence to those situations. A hearer can agree or disagree with these judgments, and this is not just an arbitrary act, but a real acceptance or rejection, based on the situation in the world.

Now let us assume that this meaning is derived compositionally from the sentence's parts. In that case a derivation into a clear and unequivocal

statement would be impossible if it were true that *imposter* and *good* only denote vague concepts with unverifiable reference. Compositional derivation is only possible on the assumption that *imposter* defines a clear dividing line between persons who are imposters and persons who are not, and that *good* divides clearly the things that are bargains into good bargains and others. From this it follows that *imposter* and *good*, apart from their conceptual content, also must have a referential function. In order to make the above sentences interpretable, *imposter* should be able to define a subset of persons, *good* to define a subset of bargains. This definition requires, apart from the conceptual content of the word, a set of application conditions, like f.i. clear answers to questions like “What must one have done to be an imposter?” or “When is a bargain good?”. One can guess that only part of these conditions will be quite general, some will rely on the beliefs of only part of the linguistic community (when is one an imposter?), some on specific contexts (when is a bargain good?), some even on subjective judgments. Of course it is impossible to incorporate this set of conditions in the conceptual analysis of the word, but that is not the point. The point is that words are not free expressions and that they should not be treated as if they were. Words only occur as parts of a larger meaningful unit, the sentence. For that reason a linguistic description of word meaning should focus on the part the word plays in sentence meaning. The meaning of a sentence is a truth value or a truth condition. And so the most central aspect of word meaning is the way it brings about reference, not the concept that it is the name of in some kind of philosophical or folk ontology. In a linguistic context making the concept the central part of word meaning is only defensible if that concept completely determines the referential valency of the word.

But, as already has been said, there are some problems with that.

### 3. The paradox of conceptual definition

Recent literature, from Fillmore (1982) on, has made clear that both the above definitions of the word *bachelor* are inadequate. There happen to be individuals who satisfy the definition but who cannot appropriately be called bachelors. This is mainly due to two causes: (a) some individuals may possibly have other properties than those mentioned in the word definition, which block the applicability of the word; so f.i. a Roman Catholic priest would rather be called a celibate than a bachelor, and (b) also to the type of language in which the word has to be used plays a part; at the Registrar's even someone being in an irreversible coma for ten years may be a bachelor, while of course the man's relatives would be horrified by the use of that term.

Obviously both definitions make too strong a generalisation to predict the reference of the word correctly, and so, inevitably, a referential paradox arises: someone can at the same time be a bachelor (according to the conceptual definition) and not be a bachelor (according to the empirical facts). Wierzbicka (1990) tries to solve the problem by extending the

definition of *bachelor* to “an unmarried man thought of as someone who could marry” (349). This, however, is only an elegant way of explaining the difficulties away. Concerning the reference of the word the first interesting question is when someone can be thought of as someone who could marry. Wierzbicka only manages in turning an analytic concept into a fuzzy one – as is indeed her intention.

It does not require intense thinking to see that the dependence of reference on fuzzy or prototypical categories is even more problematic, there being no more unequivocal conceptual criteria to decide upon reference. It is out of the question that there could be a paradox here, since the conditions to call something a paradox have been taken away: there are no more verifiable judging criteria. A notorious example of a word that is undefinable in terms of a single unambiguous, sufficient and verifiable set of features is *fruit*. The category denoted by that word and the concept it evokes are discussed elaborately in f.i. Rosch (1975) and recently in Geeraerts (1993); a brief and superficial impression of the problems will do for our purpose. So one characteristic of fruits is that they are the seed containing parts of plants, but that is not the case for all fruits: strawberries are false fruits, and according to Webster’s Third even rhubarb, the petiole of a leaf, belongs to the category. Another typical property is that fruits should be soft, juicy and sweet, but bananas are not juicy, a grapefruit is not exactly what one calls sweet, compared to a prune an apple is not exactly soft, and a tomato, more or less having all the above properties is not considered as a fruit. None of the above characteristics is valid for all the members of the category and consequently it is quite bewildering how a language user can derive the referents of the word *fruit* from this kind of concept. The astonishing thing is that he does and that he seems to have little problem doing it.

The conclusion from all this is that in at least many cases deriving the reference of a word from the concept it denotes either may lead to a paradox (in the case of analytic definition) or may be comparable to simple gambling (in the case of prototypical concepts).

In standard lexicographic practice there are two specific techniques to bridge the gap between concept and reference: the use of encyclopaedic definitions and the application of usage notes. A beautiful example of the first technique is the definition of *bird* in Webster’s Third: “a member of the class Aves all differing from the ancestral reptiles in possession of a covering of feathers instead of scales, a completely four-chambered heart served by a single (the right) aortic arch, fully separate systemic and pulmonary circulations, a warm-blooded metabolism...” and so on. But such a definition can neither be considered as a natural language concept, nor can it be used as a reference criterium for the language user. It can only be understood by specialists, and for that reason it can only function in a “division of labour”-model for natural language, as has been advocated in Putnam (1975). It is not a linguistic explanation of linguistic facts.

Usage notes add supplementary information to the conceptual definition in order to correct its referential predictions. In some dictionaries, like the *Oxford Advanced Learner's Dictionary of Current English* usage notes are separate articles in which the distribution of quasi-synonyms is described (for an elaborate discussion of this type see DiMarco and Hirst 1993). But it is another type of usage note that is relevant here: the type that can be found in the Webster's Third entry *fruit*, where the conceptual definition "the reproductive body of a seed plant consisting of one or more seeds and usu. various protective and supporting structures" is modified by the restrictive usage note "used esp. of edible bodies". The problem with this treatment is that referentially absolutely normal and regular phenomena are dealt with as if they were exceptions. And for theoretical lexical semantics this is not a satisfactory solution. In fact it creates an even worse paradox than the previous one. On the one hand one states that a thing can have a certain name although it does not correspond to the concept evoked by that name. But on the other hand by giving the name to the thing one asserts that it does correspond to that concept. Given, of course, that only one sense of the word is involved.

#### 4. A possible solution

With respect to reference there are two types of sentences in natural languages like English and Dutch. The difference is illustrated by the pair *That beaver is building a dam* and *Beavers build dams*. The first one is a referential sentence: it is about some entity in the world to which some properties are attributed: being a beaver and building a dam. The interpretation of the sentence depends on whether both these properties are true of that entity. The second sentence is a so-called categorial sentence. It does not refer to any beaver in the world, but to beavers as they are thought of: beavers as they are imagined to be in their typical form.

But not only are the sentences referentially different, also the word *beaver* must have different senses in both sentences, since it entails different sets of properties. In the first sentence an animal of the species 'beaver', belonging to the class 'mammals', is meant, and for the rest it does not matter whether the animal has three or four legs, one or two eyes and so on. For the use in the second sentence these properties are crucial. The categorial beaver has four legs, two eyes, a flattened tail and so on.

The valency to be used categorially is a (nearly?) universal property of words. At least nouns, verbs, adjectives and prepositions can systematically be used in a categorial sense. It follows from this that all words have at least two senses: a referential one and a categorial one. Now what can be said "truly" in categorial sentences is exactly what belongs to the concept evoked by the word. *Beavers build dams* will be accepted, but not *Beavers eat lions*. Like in a (prototypical) concept there is a greater tolerance for slight inconsistencies in categorial sentences than in referential ones. In other

words: conceptual definitions explain the facts of categorial use, not those of referential use. And because of these facts it seems a legitimate idea that there may be no direct relation between concept and reference. Given that every word has basically two senses, the concept may be thought of as the categorial sense. For the referential use of the word one must then assume some referential distinguishing criterium as a sense, determining the use conditions of the word in any imaginable context.

It should be admitted that it is hard to prove the mutual independence of concept and reference, but it is even harder to prove the dependence of reference on the concept, given the paradoxical implications of that stance.

### 5. The referential criterium

One of the (many) questions remaining to be answered is what should be understood by a referential selection criterium. This question should be answered by an analysis of a sufficient sample of use instances of one or more words, and, complementarily, by something alas much less easily available: a survey of the instances where the word cannot be used, which is something that of course cannot be extracted from any corpus. One can, however, rectify this inconvenience by taking a group of near-synonyms as the subject of the inquiry: it offers the possibility of mutual substitution tests. The group that has been selected is a set of near-synonyms of Dutch *goed* 'good'. This choice is of course not so self-evident: these words all denote vague, subjective, unquantifiable quality judgments and it is hard to imagine any referential function for them. But, on the other hand, if for these words the existence of an independent referential "life" can be shown, then this should also be possible for all less abstract cases.

The basis for the selection has been the collection of synonyms of *goed* in Van Sterkenburg a.o. (1991). From this collection those adjectives have been taken that express merely or mainly a **degree** of 'goodness' and that have hardly any additional conceptual structure. Conceptual information in Van Sterkenburg a.o. (1991) has been supplemented by the conceptual definitions of two other Dutch dictionaries: Van Sterkenburg and Pijnenburg (1984) – henceforth NN – and *Van Dale Groot Woordenboek der Nederlandse Taal* – henceforth VD.

An additional selection criterion has been the relation of the words to the different senses of *goed*. The adjective *goed*, just like its English equivalent *good*, exhibits two basic types of relations to its subject, as can be illustrated by the English examples *a good bargain* and *the good mrs Simpson*. The first phrase is about something that is good as a bargain, the second phrase refers to someone who is good in some respect – derivable from the context of utterance – and is called mrs Simpson.

To avoid complications also adjectives having a very common homonym or having 'good to some degree' only as a less prominent sense have been kept out of the selection.

This resulted in the following set. Additional conceptual information obtained from the dictionaries is given between brackets. (a) Meaning 'good to an absolute degree': *optimaal*, *perfect*, *vlekkeloos* (VD: without any shortcomings or inaccuracies); (b) meaning 'good to a very high degree': *excellent*, *steengoed*, *superbe*, *uitmuntend*, *uitnemend*, *uitstekend* (VD: surpassing comparable persons or things), *voortreffelijk*; (c) meaning 'good to a satisfactory degree': *degelijk* (VD: well fit to serve its purpose and stripped of all its frills), *deugdelijk*; (d) meaning 'good to an acceptable degree': *behoorlijk* (VD: presentable, usable), *billijk* (VD: as is commonly felt to be decent under the given circumstances), *fatsoenlijk*, *redelijk*, *schappelijk* (NN: not too demanding), *verdienstelijk* (VD: worthy of praise).

These conceptual descriptions have been compared to a corpus of some 10,000 sentences, each containing one of these adjectives,<sup>1</sup> in order to investigate whether the conceptual description is sufficient to account for the use phenomena, and, if this is not the case, what kind of supplementary information is necessary to account for them. For convenience's sake only attributive use of the adjectives has been selected, just because it is an easy means to determine the reference of synonyms of *good* in their *as a* sense.

A first thing that should be considered is the frequency criterium. In three cases (b, c and d) one word is found much more frequently in the corpus than the other ones. For (b) this is *uitstekend* (2032), compared to *voortreffelijk* (528), and the other words less than 100. In group (c) *degelijk* (1656) wins from *deugdelijk* (114). In group (d) *behoorlijk* (1938) and *redelijk* (1903) seem to have comparable scores, but these words show the same polysemy as their English equivalents *decent* and *reasonable*. In the sense 'good to an acceptable degree' *behoorlijk* is twice as frequent as *redelijk*. One should expect from this that the most frequent word will be the general term for the degree of "goodness" denoted and that the other words have only specialised applications. But this is only more or less so in (b). In the case of (d) both most frequent words often show some special connotations: for *behoorlijk* this is 'according to some – relatively low – standard', for *redelijk* 'being a good bargain between what one was hoping for and what one was to expect'. In the group (c) it is the more frequent word *degelijk* that has the specialised meaning: very often it is used with the connotation 'old-fashioned' and even 'dull'. *Deugdelijk* generally is not used with these connotations. Even in group (b) there are some problems. All words in this group show referential specialisation compared to *uitstekend*, except *voortreffelijk*. Both have a very broad range of application. And yet the corpus suggests a subtle difference. *Uitstekend* seems to involve to a preference for the thing denoted by the subject, *voortreffelijk* mostly just affirms that this thing is very good. Anyway, relative frequency in this set of synonyms does not seem to correlate fully with the difference between specificity and generality.

Of course the kind of connotations that have been found here can be considered as part of the concept evoked by these words. But it should be

remarked that some of them are not observed when one considers the word meaning in isolation but are only noticed and activated in use.

The most important observation, however, is that contexts often show a – sometimes exclusive – preference for one out of a set of synonyms that is independent of the concept. This can be illustrated by the words in group (a).

So the use of *vlekkeloos* in the corpus sentences is restricted to cases in which the speaker judges perfection not exceptional, even easily attainable or even normal. Actions, achievements or the persons who perform or realise them have a preference for *perfect*. The kind of subject generally combined with *optimaal* can be described as something involved in some deliberately planned or otherwise regular and/or controllable process, either as an instrument or some other component of that process or as its result. In all instances *optimaal* can be replaced by *perfect* but the reverse is not the case: so, f.i., changing *perfect geintje* ‘perfect joke’ into *optimaal geintje* results into something odd; *perfect ritme* ‘perfect rhythm’ and *optimaal ritme* would have different meanings. It appears from these cases that referents (i.e. the subjects of these adjectives) have some properties that create conditions on the use of a word, independently of that word’s conceptual content.

It is these conditions, together with those elements of the concept that are valid for all referents (or something else containing the same information) that in the case of this set of adjectives probably constitute the referential distinguishing criterium.

The substitution tests indicate that the borderlines of the referential application of these words appear to consist of some prohibiting norms. But that does not mean that the referential valency of a word is clearly delimited. In general, the referential behaviour of words appears to show some formal resemblance to prototypical concepts. There are clearly some cores of referents for which the word is optimally usable, and around them there is a fuzzy mishmash of peripheral applications, and around these there are just fragments of a borderline. But this image of fuzziness is of course the result of looking at a whole corpus, a collection of different utterances by different speakers in different circumstances. It might well be the case – and for communication’s sake it should be the case – that in a given context of utterance a speaker using *good* or one of its synonyms knows very well what he is talking about and is able to put forward clear criteria for his use of the word.

#### Note

- 1 The corpus was obtained from the INL Language Data Base. I would also like to thank the INL colleagues for discussions on the topic discussed here and Nigel Barclay for revising my English. All remaining mistakes are of course mine.

#### References

- DiMarco, Chrysanne and Graeme Hirst 1993. ‘Usage notes as the basis for a representation of near-synonymy for lexical choice’ in *Ninth Annual Conference of the UW Centre for the New*

- OED and Text Research Making Sense of Words. Proceedings of the Conference.* Waterloo:UW Centre for the New OED and Text Research.
- Fillmore, Charles 1982. 'Towards a descriptive framework for spatial deixis' in R. Jarvella and W. Klein (eds.), *Speech, Place and Action: Studies in Deixis and Related Topics*. Chichester: Wiley.
- Geeraerts, Dirk 1993. 'Vagueness's puzzles, polysemy's vagaries' in *Cognitive Linguistics 4-3*. *Oxford advanced learner's dictionary of current English*, fourth edition 1989. Oxford:University Press.
- Putnam, Hilary 1975. 'The meaning of Meaning' in H. Putnam, *Mind, Language and Reality*. Cambridge:University Press.
- Rosch, Eleanor 1975. 'Cognitive representations of semantic categories' in *Journal of Experimental Psychology: General* 104.
- Sterkenburg, P.G.J. van, and W.J.J. Pijnenburg 1984. *Van Dale Groot woordenboek van hedendaags Nederlands*. Utrecht:Van Dale Lexicografie.
- Sterkenburg, P.G.J. van, a.o. 1991. *Groot woordenboek van Synoniemen en andere betekenisverwante woorden*. Utrecht:Van Dale Lexicografie.
- Van Dale Groot Woordenboek der Nederlandse Taal*. Twaalfde, herziene druk door prof. dr. G. Geerts en dr. H. Heestermans 1992. Utrecht: Van Dale Lexicografie.
- Webster's Third = Gove, Philip Babcock (ed.) 1976. *Webster's Third New International Dictionary of the English Language, Unabridged*. Springfield:Merriam.
- Wierzbicka, Anna 1990. 'Prototypes save: On the uses and abuses of the notion 'prototype' in linguistics and related fields' in S.L. Tsohatzidis (ed.), *Meanings and Prototypes*. London: Routledge.