

Compounds in dictionaries: A semantic perspective

Rufus H. Gouws

The entries in a monolingual alphabetic dictionary are selected from the standard language of native speakers, and the lexicographer has to include an extensive selection of simplex words. The mere occurrence of a simplex word as item in the standard language compels the lexicographer to accommodate it as a lemma. However, compounds are not treated in the same simplistic way, and their selection confronts the lexicographer with a wide range of linguistic problems. Although compounding is a very productive word-formation process in Afrikaans, not all compounds can be accounted for in a dictionary. To qualify for inclusion a compound has to satisfy, among others, some semantic criteria not applied to simplexes. In contrast with simplexes the mere inclusion of a compound in a dictionary has definite semantic implications with which the lexicographer has to be well acquainted. The semantic criteria that apply to the selection of compounds are by no means well defined and the lexicographer needs a sound semantic judgement to ensure successful treatment of this category of words. Even the most common criterion, that of word transparency, still offers many problems (cf. Stein 1985: 39). It is therefore understandable why Cowie (1983: 103) says: "There are few features of dictionary organization on which editorial policy differs as strikingly from one dictionary – or lexicographical tradition – to another as the treatment of compounds."

This is yet another field where the collaboration of linguist and lexicographer is needed. This paper will focus on some of the semantic problems confronting the lexicographer. The main emphasis will be on the Afrikaans particle verb as one example of the category of compounds.

1. The selection and definition of compounds

In Afrikaans dictionaries the distinction between transparent and intransparent compounds is still used as an important criterion for the inclusion of compounds. In spite of all the problems involved in defining this criterion, it still offers a valuable guideline for the lexicographer. Intransparency can be the result of a polysemous compound. Polysemy, however, does not only apply to independent words. It can also apply to the components of a complex word. Quite often the different components of Afrikaans compounds have polysemous characters. The morphological context in which this component occurs, determines the applicable sense. If one component of a compound is polysemous that compound will be intransparent and has to be included in the dictionary.

It has been shown by Van Jaarsveld (1976: 84) that in the majority of Afrikaans compounds the morphological context activates the most dominant sense of that component's occurrence as simplex. The results of this study by Van Jaarsveld could be of great value to lexicographers. The way in which a lexicographer treats compounds should be clearly explained in the preface of the dictionary and there it can be stated that compounds are not included if both their components have the meaning listed as the dominant sense of their occurrence as simplexes. This will also imply that before finalizing the arrangement of the different senses of a polysemous simplex the lexicographer should study the semantic value of that word in its occurrence as part of a compound. Although many compounds activate non-dominant senses of their components, the occurrence of a word as part of a compound could give some indication of the possible dominant sense of the simplex.

The lexicographer's omission of these first sense compounds can only be justified when the compound itself is not polysemous. Some first sense compounds have further senses which are either a combination of non-first senses or a first and a non-first sense. It can also be the result of metaphorical usage. If a first sense compound is included because of either semantic or other reasons, e.g. spelling or pronunciation difficulties or morphological peculiarity, it is not necessary to define the first sense. A code, explained in the preface, can be used to identify the entry as a first sense compound. The definition of compound meaning can be reserved for those with a semantic plus value, i.e. a value additional to the first sense of the different components' occurrences as simplexes.

According to Stein (1985: 43), more emphasis is given in present-day English dictionaries to items that function as modifiers in compounds than to those that function as heads. The alphabetical principle is mainly responsible for this position. According to Selkirk (1982: 22), "the notion 'head' is crucial in characterizing the semantics of compounds. In general, in endocentric compounds, the class of elements denoted by the compound is a subset of the class of elements that would be denoted by the head noun on its own. The nonhead constituent of the compound in some way further defines the head."

A compound entry should not contain a redefining of the simplexes that form the components of that compound. Both components can have a metaphorical use. If the head is used metaphorically the definition has to be more extensive because it should include more than just an indication of the way in which the first component modifies the head.

Selkirk (1982: 48) argues convincingly that a compound and its components are of the same category level and that it is categories of the level 'word' that are involved. This implies that the components of a compound can also qualify for inclusion as separate dictionary entries — not functioning as simplexes but functioning as parts of a compound. A lexical item functioning as part of a compound can have a sense restricted to that item's occurrence as component of a compound. A dictionary has to account for these lexicalized senses. Because

Afrikaans compounds are written as one word, this semantic plus value cannot be indicated in the main entry of the simplex. An additional hyphenated lemma has to be entered with this new sense as one of its senses. A cross-reference should also be made to those senses of the simplex that can also apply in the item's occurrence as component of a compound. For example: the Afrikaans simplex *hoof* can have several senses including "chief" and "head". As the first component of a compound it can also mean "main" as in *hoofskakelaar* (main switch). This sense is restricted to its occurrence as part of a compound and compels the lexicographer to include *hoof-* as a hyphenated lemma with a definition of this and other polysemous senses. Senses defined for the simplex lemma need not be repeated but can merely be indicated by way of cross-reference.

Hyphenated lemmas in Afrikaans dictionaries represent more occurrences of first components than of second components. There are more post- than pre-hyphenated lemmas. This is also determined semantically. Because the majority of Afrikaans compounds are endocentric and right-headed the second component's meaning can usually be related to its meaning as a simplex. Where the meaning of a head differs from its meaning as simplex this difference normally leads to intransparency and requires a separate compound entry. This semantic plus value is usually restricted to the item's occurrence in a single compound, and because of this unproductive occurrence there is no need for the lexicographer to enter such a head as a hyphenated lemma. In many examples this type of compound is metaphorical, with neither the first nor the second component's literal sense applicable, e.g. *varkhok* (pig sty) meaning "a dirty place".

If, however, a lexical item has a semantic value restricted to its use in compounds and that item can combine with more than one first component yielding the same sense, the lexicographer will do well to include that item as a pre-hyphenated lemma. Afrikaans has a noun *lui* (folk, people). In present-day Afrikaans this word is very rarely used on its own, although it still occurs regularly in compounds, e.g. *seelui*, *vaklui*, *werkslui* (the plural of *seaman*, *craftsman*, *workman*). There should be a lemma *-lui mv. mense . . . (. . . pl. people)*. Unfortunately Afrikaans dictionaries wrongly include this element as an independent simplex.

2. Particle verbs

The Afrikaans lexicon includes a large number of particle verbs. According to Selkirk (1982: 19) it could be argued that for English there is one endocentric but left-headed compound-type: the verb-particle collocation where the verb on the left is clearly the head. The treatment of particle verbs in English dictionaries can emphasize the semantic head to a larger extent than in the treatment of other compounds because here the alphabetical arrangement is in accordance with the compound head.

In Afrikaans the particle verb (hereafter *p.v.*) has the verbal element on the right and the *p.v.* is written as a one-word compound. Some syntactic environments can compel a different distribution of particle and verb. Afrikaans is a SVO language. When a *p.v.* occurs in the V2 position the verbal element precedes the particle and it can be separated from the particle by an object or an adjunct. The *p.v.* *opkom* (to rise) can be used in the following ways:

Die son *kom op* (The sun rises)

Die son *kom soggens op* (The sun rises in the morning)

Die son sal môre weer *opkom* (The sun will rise again tomorrow).

Although in actual speech the particle is often preceded by the verb, *p.v.*'s are entered in dictionaries alphabetically according to the particle. This inclusion of *p.v.*'s as single and unified lexical entries is the correct way of treating them. However, it is not sufficient as the only way of treating *p.v.*'s. In spite of the syntactic patterns separating particle and verbs in some environments, the *p.v.* remains a lexical unit, but on semantic grounds some particles also have to be included as separate lemmas, and as a result of the Afrikaans spelling system they have to be entered as post-hyphenated lemmas. This possible hyphenated inclusion, on which the lexicographer has to decide for each individual *p.v.*, will be determined by the meaning of the *p.v.*, the particle and the verb.

2.1. Particle verbs and semantic relations

Particle verbs have a high usage frequency in Afrikaans and this contributes to the fact that they often are polysemous words. Many Afrikaans *p.v.*'s have synonyms for one or more than one of their polysemous senses. The synonyms are either other *p.v.*'s or ordinary verbs. The *p.v.* *uitdelg* (to destroy) can have the *p.v.* *uitroei* or the verbs *vernietig*, *verdelg* as synonyms of the same sense, and in the dictionary the lexicographer should give an indication of all the applicable synonyms. In the treatment of *p.v.*'s the lexicographer has to pay attention to the relation between the *p.v.* and its components. These components are also elements of the lexicon and therefore part of a network of semantic and syntactic relations, and the lexicographer may not be ignorant of these relationships. When dealing with *p.v.*'s, the application of the criterion of transparency causes additional problems for the lexicographer. If two verbs have a semantic relationship and both combine with the same particle to form *p.v.*'s, the lexicographer has to study this combination in an attempt to see if the same semantic relationship still holds. The results of this probing should be indicated in the dictionary. As a linguist the lexicographer will soon find out that the relationship between two verbs does not necessarily still hold when they combine with the same particle. The result of such a combination cannot be predicted, and because of this semantic instability the dictionary should include these *p.v.*'s in spite of an apparent lack of intransparency.

The Afrikaans p.v.'s *uitpraat* (finish what one has to say) and *uitspreek* (pronounce) are polysemous but don't offer major semantic problems. The dominant senses of these two words are not synonymous. The verbs *praat* and *spreek* differ when it comes to the degree of formality, but the dominant senses of these two verbs are synonymous: "to talk". In combination with the same particle *uit-* the synonymous relationship is no longer predominant. This is due to the polysemous character of the particle *uit-*. In *uitpraat* the particle means "to finish" and in *uitspreek* it means "to utter". As with other compounds the lexicographer cannot expect his dictionary users to know this meaning and these p.v.'s should be accounted for as separate entries.

A second reason for the inconsistent relation between two simplex verbs and their occurrence in combination with the same particle is the semantic unity of the p.v. As with other compounds, p.v.'s become near semantic units. The Afrikaans p.v.'s *indien* and *inlewer* share a synonymous sense "submit". It can be argued that the same sense of the particle *in-* applies in both these p.v.'s. The simplex verbs *dien* and *lewer* are never synonyms. However, in combination with the same particle two new lexical items are formed and a set of semantic relations and possibilities not applying to the components are enforced. Once again, the lexicographer has to give account of these p.v.'s because their meaning cannot be constructed from the meaning of their components.

2.2 Particle verb components included as hyphenated lemmas

A transparent p.v. does not have to be included as a lemma, and where the components of a p.v. have no semantic plus value restricted to their occurrence as p.v. components they do not have to be included as hyphenated lemmas. The lexicographer has to negotiate with the same set of possibilities that applies to other compounds.

Intransparency of p.v.'s can often be traced to the fact that a new sense emerges for either of the components. In some p.v.'s the head gets a definite new sense of which the lexicographer has to give account. Afrikaans has a verb *faseer* and two p.v.'s *infaseer* (to phase in) and *uitfaseer* (to phase out). The verb *faseer* has a very limited use, which is generally restricted to physics with the meaning "to synchronise the phase of alternating currents." The p.v.'s *uit-* and *infaseer* have a much more general use and meaning: "to discontinue/introduce gradually". In these two p.v.'s the verb *-faseer* has a definite new sense that does not apply to its occurrence as simplex and therefore should be accounted for in a hyphenated lemma.

Other p.v.'s have a verbal element that can be traced to a simplex but without any semantic resemblance. The p.v. *uitbrei* (to extend) has the particle *uit-* with the meaning "outwards". The element *-brei* has no clearly definable meaning that can be related to any of the senses of the Afrikaans verbs *brei* (to knit) or

brei (to coach). But there is another verb *verbrei* (distribute) and this *-brei* seems to be related. The lexicographer has to decide whether a new sense could be attributed to *-brei*, and this decision will motivate the possible inclusion of a new hyphenated lemma.

In some p.v.'s the intransparency is of an even higher degree. The verbal element *-roei* in the p.v. *uitroei* cannot be related semantically to the simplex verb *roei* (to row). No direct sense can be allocated to *-roei* in *uitroei*. No further evidence exists for this sense of *-roei*, and therefore there is no need to include this occurrence of *-roei* as a hyphenated lemma. In *uitroei* a meaning can be allocated to the particle *uit-*. *Uitroei* has *uitdelg* as one of its synonyms and it is the same sense of *uit-* ("to finish") that applies in both these p.v.'s. This meaning of *uit-* has to be accounted for in its occurrence as hyphenated lemma. The p.v. *uitroei* is in defiance of Selkirk's claim that the verbal element is the head of a p.v. The particle does not satisfy the conditions for a compound head either, and this p.v. could be regarded as a non-headed compound. It has a unified meaning allocated to the p.v. as a separate lexical entry, and because this meaning cannot be reconstructed from the meanings of the different components, the lexicographer has to define the p.v.'s meaning without using any of the components in his definition.

The dictionary has to give account of the semantic relations holding between lemmas. If an indication is given of an independent word's synonym paradigm, the lexicographer should treat the hyphenated lemmas in the same way. Afrikaans has a set of synonyms with the meaning "to blurt out": *uitblaker*, *uitflap* and *uitlap*. Not one of the simplexes *blaker*, *flap* or *lap* has a sense remotely related to these p.v.'s. In all these p.v.'s the particle *uit-* has the same meaning "outwards". Each of these p.v.'s should have a cross-reference to the other members of this synonym paradigm. The verbal elements of these p.v.'s have an identifiable sense for which a hyphenated lemma should be entered. In the articles of these hyphenated lemmas there should be a definite cross-reference to the other hyphenated lemmas with the same sense. This will enable the dictionary user to get a complete semantic picture and to know how this specific meaning is lexicalized in Afrikaans.

The treatment of compounds confronts the lexicographer with a wide range of problems. In order to give a sound semantic analysis that will enable the user to have a clear understanding of all the lexical items involved, the lexicographer has to compare the meanings of the compound and its components and evaluate them as potential dictionary entries. Each sense has to be accounted for: either as part of the compound's entry, the entries of the components as simplexes or as a separate hyphenated lemma. The applicable semantic relations must also be indicated. The lexicographer hereby adheres to the principle that semantics de-

termines not only the treatment but also the selection of compounds and hyphenated lemmas as dictionary entries.

References

- Cowie, Anthony P. (1983), "On specifying grammar", in: Hartmann, R.R.K. (ed.), *Lexicography: principles and practice*, London: Academic Press, 99–107.
- Selkirk, Elizabeth O. (1982), *The syntax of words*, Cambridge, Massachusetts: MIT Press.
- Stein, Gabriele (1985), "Word-formation in modern English dictionaries", in: Ilson, R. (ed.), *Dictionaries, lexicography and language learning*, Oxford: Pergamon Press, 35–44.
- Van Jaarsveld, Gert J. (1986), "Semantiese aspekte by nominale komposita", in: *Tydskrif vir Geesteswetenskappe*, 16,2: 77–85.