

Compounding and construction morphology

Geert Booij

1. Introduction

Word formation patterns can be seen as abstract schemas that generalize over sets of existing complex words with a systematic correlation between form and meaning. These schemas also specify how new complex words can be created. For instance, the word formation process for endocentric compounds in English and Dutch can be represented as follows:

(1) $[[a]_X [b]_{Y_i}]_Y$ 'Y_i with relation R to X'

Schema (1) can be interpreted as the formal representation of a construction, that is, a particular structural configuration with a specific meaning correlate. The fact that the right constituent and the structure as a whole are dominated by the same syntactic category variable Y is the formal expression of the generalization that the syntactic category of the compound is determined by its right constituent. For instance, if Y_i has the value N, the compound as a whole is also an N. The relevant meaning correlate is that the right constituent functions as the semantic head of the compound, and that a semantic relation between the two constituents is invoked. The specific nature of that relation, however, is left unspecified in the schema, since it is not predictable on structural grounds.

The traditional notion of construction and its importance for theories of linguistic structure have recently received renewed attention within the theoretical framework of Construction Grammar (cf. Goldberg 2006 and the literature mentioned there). The basic idea of Construction Grammar can be summarized as follows:

“In Construction Grammar, the grammar represents an inventory of form-meaning-function complexes, in which words are distinguished from grammatical constructions only with regard to their internal complexity. The inventory of constructions is not unstructured; it is more like a map than a shopping list. Elements in this inventory are related through inheritance hierarchies, containing more or less general patterns.”

(Michaelis and Lambrecht 1996: 216)

As suggested by this quotation, both syntactic patterns and word formation patterns might be seen as constructions. This idea of morphological patterns as constructions has been developed in a number of publications (cf. Riehemann 1998, Koenig 1999, Jackendoff 2002: Chapter 6. Booij 2002a, b; 2005).

In this chapter I will present a constructional theory of compounding that makes use of some basic ideas of Construction Grammar, in particular constructional schemas, and the idea of a hierarchical lexicon (with multiple linking between words, and intermediate nodes between the most abstract schemas and the individual lexical items in order to express intermediate levels of generalization). Similar ideas have been developed in the framework of Cognitive Grammar (Langacker 1988, 1998a;b, Croft & Cruse 2003). More specifically, the purpose of this article is to provide morphological argumentation in support of such a view of morphology and the lexicon, a view that will be referred to as Construction Morphology. The empirical domain that I will focus on is, for obvious reasons, that of compounding. In section 2, some basic assumptions concerning the role of word formation schemas in the lexicon will be spelled out, and it will be shown how they can be applied in the analysis of endocentric compounds. In section 3, the issue of headedness will be discussed in relation to various categories of compounds. Section 4 focuses on synthetic compounds, and in section 5 I will show how Construction Grammar enables us to give an insightful account of compound-like phrases. Section 6 the presents a summary of our findings.

2. Word formation templates and the hierarchical lexicon

The idea of word formation templates in a hierarchical lexicon can be illustrated quite nicely by means of compounding. In many languages, compounding is a productive word formation process, and the notion ‘rule’ has therefore played an important role in accounting for this type of word formation. In Dutch, for instance, we find right-headed nominal, adjectival, and verbal compounds, and hence we might assume the following morphological rule for Dutch compounding:

$$(2) \quad X + Y \rightarrow [XY]_Y$$

In this rule, X and Y stand for the syntactic categories N, A, and V, and the label Y of the resulting compound expresses the generalization that the right constituent of a Dutch compound functions as its head, and thus determines the syntactic category of the word as a whole. In other

words, the generalization about the position of the head in endocentric compounds in Germanic languages such as English and Dutch expressed by the so-called Right-hand Head Rule (RHR; Williams 1981) can be expressed directly in the output form of this rule, without making use of an additional rule or principle.

Word formation rules such as (2) have two functions: they function as redundancy rules with respect to existing complex words, and specify how new complex words can be made (Jackendoff 1975). Note, however, that rule (2) in its present form does not say anything about the semantic regularities in the interpretation of such compounds.

Instead of speaking about word formation rules, we might also speak about word formation templates (or schemas). We then replace rule (2) with a word formation template for Dutch that generalizes about the structure of existing compounds, and that can be used for making new compounds as well. This is schema (1), repeated here for convenience:

(3) $[[a]_X [b]_{Y_i}]_Y$ ‘ Y_i with relation R to X’

The lower case variables *a* and *b* in this template stand for arbitrary sound sequences. The use of phonological variables indicates that phonological information does not play a restrictive role in this type of word formation. In (3) the general meaning contribution of the compound schema is also specified, since morphology is about form-meaning pairs. The nature of R is not specified, but is determined for each individual compound on the basis of the meaning of the compound constituents, and encyclopaedic and contextual knowledge (Downing 1977).

Template (3) does not yet express that it is not only the syntactic category of the head that is identical to that of the whole compound, but that the two Y-nodes are also identical with respect to properties such as gender and declension class for nouns, and conjugation class for verbs. Hence, we elaborate template (3) as template (3)’ in which $[\alpha F]$ stands for the set of relevant subclass features:

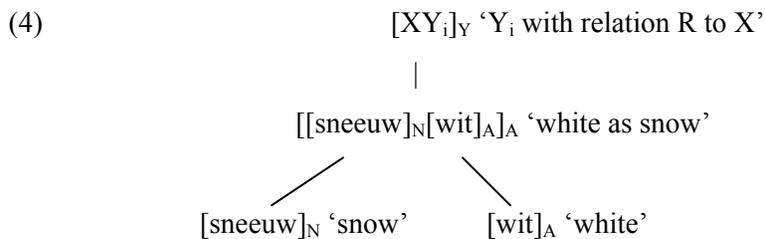
(3)’ $[[a]_X [b]_{Y_i}]_Y$ ‘ Y_i with relation R to X’
 $\begin{array}{c} | \quad | \\ [\alpha F] \quad [\alpha F] \end{array}$

Template (3)’ thus specifies the category of right-headed endocentric compounds of Dutch.

The format of (3)’ expresses that compounding is a construction at the morphological level with a systematic pairing of form and meaning. It specifies that the head Y is not only the

formal head, but also the semantic head: a Dutch compound denotes a certain Y, not a certain X. By specifying both the form and the meaning of this class of words, we are reminded of the fact that morphology is not a module of the grammar on a par with the syntactic or the phonological component. Instead, morphology is word grammar. Hence, it deals with three different and systematically related aspects of complex words: phonological form, formal structure, and meaning (cf. Jackendoff 2002). That is, we assume a tripartite parallel architecture of the grammar.

Template (3)' is to be seen as the node in the hierarchical lexicon of Dutch that dominates all existing compounds of Dutch. The individual compounds of Dutch inherit their non-unique (formal and semantic) properties from this dominating node, and from their constituent lexemes. For instance, the following substructure of the Dutch lexicon may be assumed for the adjectival compound *sneeuw wit* 'snow-white':



This tree is a 'multiple inheritance tree' with two types of relations: 'instantiation', and 'part of'. The word *sneeuw wit* is an instantiation of the general template at the top of the tree, and the lexemes *sneeuw* and *wit* form part of that adjectival compound (cf. Krieger and Nerbonne 1993). In this example, the relation R is interpreted as 'as' since the word means 'white as snow'.

A graph like (4) makes it clear that a complex word bears different kind of relations. It is not only an instance of a more abstract word formation schema, but it is also linked to other words in the lexicon. The compound *sneeuw wit*, for example, is linked to *wit* and to *sneeuw*. The lexemes *wit* and *sneeuw* will be linked to other complex words as well. Thus, we get the word families {*wit*, *sneeuw wit*, etc} and {*sneeuw*, *sneeuw wit*, etc}. The existence of such word families manifests itself in the family size effect: the larger the size of the family of a word, the faster that word can be retrieved in a lexical decision task (De Jong et al. 2000). The existence of the abstract word formation schema, on the other hand, manifests itself in the productive coining of new compounds by the language user.

This word formation schema itself is obviously based on a set of existing compounds. It is knowledge of existing compounds that is a precondition for the language user to develop the

abstract schema. That is, both the instantiation relation and the part-of-relation are based on paradigmatic relationships between words in the lexicon. The morphological structure assigned to the word *sneeuw* is a projection of such paradigmatic relationships on the syntagmatic axis of word structure.

One advantage of this approach to word formation is that generalizations about sub-patterns can be expressed in a straightforward way. The following generalization holds for the set of Dutch endocentric compounds (cf. Booij 2002a: Chapter 4 for the relevant data): only NN compounding (and under certain conditions to be discussed below, VN compounding as well) is recursive, both in the head position and the non-head position, that is, their constituents can be compounds themselves. Other types of endocentric compounds do not allow for recursivity. For instance, an adjectival compound such as *muislichtgrijs* ‘mouse light grey’ is felt as odd, although both *muisgrijs* ‘mouse grey’ and *lichtgrijs* ‘light grey’ are existing adjectival compounds of Dutch. Recursive NN compounding is illustrated in (5):

(5) *left constituent recursive:*

[[[[[ziekte]_N[verzuim]_N][bestrijdings]_N][programma]_N]_N

illness absence fight program ‘program for reducing absence due to illness’

[[[aardappel]_N [schil]_V][mesje]_N]_N

potato peel knife ‘knife for peeling potatoes’

right constituent recursive:

[[zomer]_N [[broed]_V[gebied]_N]_N]_N

summer breed area ‘breeding area for the summer’

both constituents recursive:

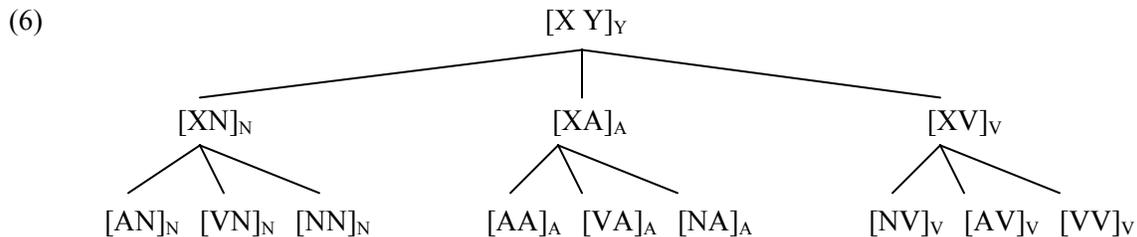
[[[grond]_N[water]_N][over]_P[last]_N]_N]_N

ground water over burden ‘groundwater problems’

Another example of a specific property of a subset of the Dutch endocentric compounds is that compounds with a nominal head allow certain types of phrases to function as the non-head constituent whereas other compounds do not. For instance, whereas *oudemannenhuis* ‘old men’s home’ is well-formed, an adjectival compound such as *oudemannenzwak* ‘lit. old men weak, weak as old men’ is odd. In short, we must be able to state generalizations about subclasses of compounds without obliterating the commonalities of all Dutch endocentric compounds.

Generalizations about subsets of words can be expressed in a hierarchical lexicon, by assuming intermediate levels of abstraction in between the most general template and the

individual existing compounds. That is, the following structure of the compounding part of the Dutch lexicon may be assumed:



Each of the lowest nodes in (6) will dominate a subset of the Dutch compounds with the relevant structural property. The second and the third of these nodes, $[NN]_N$ and $[VN]_N$ will be without restrictions on the internal complexity of their constituent lexemes, whereas to all other nodes, the condition ‘X and Y \neq compound’ will have to be added. This accounts for the first observation made above as to the larger range of structures that nominal compounding in Dutch allows for. Thus, we can make generalizations about subsets of compounds, while at the same time expressing the common properties of the whole set of Dutch compounds. The possibility of generalizations at different levels of abstractions is a clear advantage of this representation of word formation patterns, as we will also see below.

In the hierarchical lexicon approach, exceptions to generalizations are dealt with by assuming default inheritance: properties of higher nodes are percolated to lower nodes, unless the lower node bears a contradictory specification for the relevant property. As pointed out by Krieger and Nerbonne (1993: 91), “[t]he key advantage of default specifications is that they allow the description of sub-regularities, classes of items whose properties are largely, but not perfectly regular”. The use of inheritance trees is a prominent feature of Network Morphology (cf. Corbett and Fraser 1993, Hippisley 2001), Construction Grammar (cf. Taylor 2002, Croft and Cruse 2003, Goldberg 2006), and Typed and Construction-based Constituent Structure Morphology (TCCM, Koenig 1999). Examples of the use of inheritance trees in the domain of derivational morphology can be found in Riehemann (1998), and Hippisley (2001).

A basic feature of this approach to word formation is that abstract schemas and individual instances of those schemas co-exist. Once the abstract schemas have been discovered, the individual items on which they are based will not necessarily be lost from lexical memory (Langacker 1998a, Tomasello 2003). Thus, we avoid the rule/list fallacy: the fact that there is a productive rule for the formation of a certain linguistic construct does not imply that the outputs

of that rule should not be listed in the lexicon. It is obvious that we need this option for compounds. We have to specify the existing (or established) set of compounds of a language, while we also want to express that most of them have been formed according to a regular and productive schema that gives rise to novel compounds as well.

The idea of a hierarchical lexicon is also in line with the consideration mentioned above that people acquire the morphological system of a language, that is, the abstract morphological schemas, on the basis of their knowledge of a set of words that instantiate these patterns. Once they have come across a sufficient number of words of a certain type, they can infer an abstract schema, and will be able to expand the relevant class of words. The endpoint of language acquisition is therefore to be defined “in terms of linguistic constructions of varying degrees of complexity, abstraction, and systematicity” (Tomasello 2000: 238).

This view of word formation is in line with Vallés’ (2003: 141) position that “word formation patterns emerge from paradigmatic relations and the function of rule-learning might be to help organize the lexicon, to give it structure; its role is to express generalizations about what is part of the lexicon”. This means that the native speaker’s competence to create new compounds and derived words is based on abstractions over sets of existing complex words and the words that are paradigmatically related to them.

The need for intermediate generalizations for subclasses of compounds is particularly clear from the phenomenon of the blurred boundary between compounding and derivation (Booij 2005). The Dutch lexicon contains a number of compounds of the form $[[hoofd]_N[x]_N]_N$, with the meaning ‘main $[x]_N$ ’ (*hoofd* means ‘head’).

- (7) *hoofdingang* ‘main entrance’
 hoofdgebouw ‘main building’
 hoofdbezwaar ‘main objection’
 hoofdverdachte ‘main suspect’
 hoofdbureau ‘main office’

Therefore, the language user may assume a schema that generalizes over this set of complex words:

- (8) $[[hoofd]_N[x]_N]_N$ ‘main $[x]_N$ ’

In this schema, one of the positions is filled by a specific lexical item. Hence, it is a constructional idiom in the sense of Jackendoff (2002): a productive idiomatic pattern, with both variable and lexically fixed positions (cf. also Pitt & Katz and 2000).

This schema will be dominated by the general schema for NN compounding (that in its turn will be dominated by the general schema (3')). The semantic generalization for this schema is that the constituent *hoofd* 'head' in these words carries the specific meaning 'main': This specific meaning of *hoofd* 'head' when embedded in these compounds is quite clear from the fact that in the English glosses we cannot use the English translation of *hoofd* 'head', but have to use *main* instead. Since this meaning 'main' is strictly dependent on the occurrence of *hoofd* within compounds, this lexical morpheme becomes similar to affixes, whose meaning is also dependent on their occurrence in complex words: outside complex words affixes do not have meaning.

Another example of this situation is the productive use of the adjective *oud* 'old' in complex words, as in *oud-burgemeester* 'ex-mayor'. The meaning of *oud* in its bound use is 'former, ex-', and the literal meaning 'old, of high age' does not apply at all. This bound use of the morpheme *oud* is productive, and we can add this morpheme to all sorts of nouns.

These facts have raised the question whether we have to assume prefixes *hoofd-* and *oud-* respectively for Dutch. In De Haas and Trommelen (1993: 51), for instance, the morpheme *oud* with the meaning 'ex-, former' is indeed qualified as a prefix. This problem pops up repeatedly in the analysis of Dutch word formation patterns (cf. Booij 2005).

The terms 'affixoid' and 'semi-affix' have been introduced to denote morphemes which look like parts of compounds, and do occur as lexemes, but have a specific and more restricted meaning when used as part of a compound. The following Dutch words also illustrate this phenomenon (data taken from Booij 2005):

- (9) boer 'farmer' groente-boer 'lit. vegetables-farmer, greengrocer'
 kolen-boer 'lit. coal-farmer, coal trader'
 les-boer 'lit. lesson-farmer, teacher'
 melk-boer 'lit. milk-farmer, milkman'
 patat-boer 'lit. chips-farmer, chips seller'
 sigaren-boer 'lit. cigar-farmer, cigar seller'
 vis-boer 'fishmonger, fish dealer'
- man 'man' bladen-man 'lit. magazines-man, magazine seller'
 kranten-man 'lit. newspapers-man, newspaper seller'

ijsco-man 'lit. ice cream-man, ice cream seller'
melk-man 'milk man, milk seller'

The heads of the complex words have a specific and recurrent meaning when used in that context, just as was the case for the word *hoofd* above. The morpheme *boer* 'farmer' (etymologically related to the English morpheme *bour* in *neighbour*), when part of a complex word, has the meaning 'trader in', and no longer means 'farmer'. Crucially for a classification as semi-affix, the 'bound' use of these morphemes is productive (Becker 1994), as is illustrated here for *boer*. This morpheme is used in combination with nouns that do not denote agricultural products, and words with this morpheme form a series of words with a shared meaning component. A similar observation can be used for the lexeme *man* 'man' when used in compounds.

The observation of morphologically incorporated lexemes having specific meanings and being used productively with that specific meaning also applies to lexemes in the non-head position of compounds (data again from Booij 2005):

(10) *nouns used as pejorative prefixoids:*

kanker 'cancer'	kanker-school 'bloody school'
kut 'cunt'	kut-ding 'worthless thing'
kloot 'testicle'	klote-houding 'bad attitude'

nouns used as prefixoids of positive evaluation:

meester 'master'	meesterwerk 'very good piece of work'
wereld 'world'	wereld-vrouw 'fantastic woman'

nouns used as prefixoids with intensifying meaning:

steen 'stone'	steen-koud 'very cold', steen-goed 'very good', steen-rijk 'very rich'
beer 'bear'	bere-sterk 'very strong, bere-koud 'very cold', bere-leuk 'very nice'

The prefix *bere-*, for instance, derives from the noun *beer* 'bear', followed by the linking phoneme. The analytical problem that there is no sharp boundary between compounding and affixal derivation is not solved by postulating a category of semi-affixes or affixoids. That is just a convenient description of the fact that the boundary between compounding and derivation is blurred,

but does not in itself provide an explanation of why this is the case. What we need is a model of morphological knowledge that will enable us to explain these facts. It will be clear by now that the model of the hierarchical lexicon proposed above will be adequate for this purpose, when we combine it with the idea of constructional idioms, partially lexically specified productive patterns. For instance, the use of *man* with the specific meaning ‘trader’ exemplified in (10) will be accounted for by the following constructional idiom of Dutch:

(11) [[x]_N[*man*]_N]_N ‘trader in X’

This template will be dominated by the general schema for Noun-Noun compounding, and most of its properties derive from this general schema. It will also be linked to the noun *man*. The only specific property is that of the recurrent conventionalized interpretation of *man* as ‘trader’, which is a semantically richer specification than the meaning of *man*.

3. Headedness issues

Since Williams (1981) the importance of the notion ‘head’ for the analysis of morphological constructs has received new recognition and attention. It is clear that Williams’ Right Hand Rule cannot be a rule in the sense of a universal since many languages have left-headed compounds. Hence, one might consider the position of the head as a morphological parameter. For instance, Germanic languages may be qualified as right-headed, and Romance languages such as Italian as left-headed (Scalise 1984; 1992). The problem for such a parameter approach is that Italian and Spanish also have sets of right-headed compounds. Examples of such compounds in Spanish are the following:

(12) auto-escuela, ‘car school’, cine-club ‘cinema club’, eco-sonda ‘echo sounder’, tele-novela ‘television novel’, video-arte ‘video art’

Rainer & Varela (1992: 121) make the following remark with respect to these types of compound:

“Such right-headed n-n compounds would have to be limited to a fixed number of first elements [...] which is quite atypical since compound types can generally be defined at the categorical level. [...] Another property which they share with prefixes is their form: they are typically bisyllabic and end in a vowel”.

Therefore, they consider *auto-* etc. as prefixes. “The alternative would consist in setting up an independent sub-type of word formation characterized by right-headedness, a purely pragmatically driven rule of interpretation, and severe restrictions on the number and form of elements that can serve as left constituents”. (Rainer and Varela 1992: 121-22).

However, in the framework of construction morphology we can still analyse these words as compounds. We then assume a list of constructional idioms such as

(13) [auto [x]_N]_N ‘car with relation R to [x]_N’

instead of a general abstract template for right-headed compounds. By lexically specifying the left constituent of these compound schemas we express that the class of right-headed compounds in Spanish is restricted to compounds that begin with a word that is a member of a restricted and definable set of words. Moreover, these compounds will be specified as right-headed, and therefore, these constructional idioms will not be linked to the node for left-headed NN compounds in the lexicon of Spanish.

The same observation can be made for Italian. As pointed out by Schwarze (2005: 149), a restricted set of nouns can occur in right-headed compounds. The best example is that of compounds beginning with *auto* ‘car’, as in:

(14) autocentro ‘car center’, autoconvoglio ‘convoy’, autodromo ‘motor-racing circuit’, autolinea ‘bus route’, autoservizio ‘car service’, autostop ‘hitch-hiking’

They contrast with the regular left headed NN compounds such as *autoambulanza* ‘ambulance’, *autobotte* ‘tank truck’, *autocisterna* ‘tank truck’, and *autopompa* ‘fire truck’. Other types of right-headed compounds in Italian begin with words or roots such as *foto-*, *radio-*, and *tele-*.

Schwarze uses these data to argue that we must distinguish between rules and pattern-imitation:

“Constructed words arise from a generative system which may be described in terms of rules. These rules operate on morphological segments and their semantic representations in such a way that the forms and meanings of possible words are defined. [...]. Complex, non-constructed words, on the other hand, are brought about by “paragrammatical

procedures” [such as univerbation and pattern-based word formation] (Schwarze 2005: 137)

In the framework presented here these data receive a straightforward interpretation: these subsets of right-headed compounds can be accounted for by means of constructional idioms such as:

(15) $[[\text{auto}]_N[\text{x}]_N]_N$ ‘car with with relation R to $[\text{x}]_N$ ’

The same mechanism can be used to account for root compounds. In Italian, for instance, the root *tele-* can be used in compounds such as *telespettatore* ‘television watcher’. The morpheme *tele* is not a word by itself, but only occurs as part of complex words, hence it is to be qualified as a root. It also lacks a syntactic category. The fact that *tele* is a root can be expressed by assuming a constructional idiom $[\text{tele } [\text{x}]_V]_Y$ which expresses that *tele* can only occur as part of a complex word, and does not determine the syntactic category of the compound in which it occurs..

The next category to be discussed is that of exocentric compounds. A famous class is that of Italian VN compounds such as

(16) *porta-lettere* ‘lit. carry letters, post man’,
lava-piatti ‘lit. wash dishes, dish washer’

Neither of the constituents of these compounds is the head. A *lava-piatti* is neither a ‘wash’ nor a ‘dishes’. Hence, the formal structure of these words is as follows:

(17) $[\text{a}]_V [\text{b}]_{N_x}]_{N_y}$ ‘entity that performs action V on N_x ’

The different indices x and y indicate that there is no identity between the category of the right constituent and the category of the whole compound.

In sum, by analysing ‘wrong-headed’ and exocentric compounds in terms of specific morphological constructions, we can maintain the generalization that most Italian compounds are left-headed without obliterating the existence of productive patterns of compounding that do not conform to this generalization.

4. Synthetic compounds and embedded productivity

The productivity of a certain word formation pattern may be dependent on its being embedded in another morphological pattern. I refer to this phenomenon as embedded productivity. This is the case for verbal compounding in Dutch. That is, verbal compounds may be qualified as ‘bound compounds’. Let us first focus on NV compounding. This type of compounding appears to be non-productive in Dutch. What we do find for Dutch are the following classes of verbal compounds:

- (18) (i) conversions of nominal compounds such as [[[voet]_N[bal]_N]_N]_V, ‘to play football’;
(ii) isolated cases of back formation: [[beeld]_N[houw]_V]_V derived from *beeld-houw er* ‘lit. statue-cutter, sculptor’; [[woord]_N[speel]_V]_V ‘to play with words from *woordspeling* ‘word-play, pun’;
(iii) separable complex verbs such as *piano-spelen* ‘piano-play, to play the piano’. Such verbs are not verbal compounds, but phrasal verbs (cf. Booij 1990); the two parts are separated in main clauses, as in *Jan speelt heel goed piano* ‘John plays the piano well’
(iv) defective NV verbs that mainly occur in the infinitival form, such as [[worst]_N[happ-en]_V]_V; ‘to sausage-eat’ some of these verbs have finite forms but only in embedded clauses where the two parts are linearly adjacent (... *dat Jan worsthapte* ‘that John sausage ate’)

Embedded in nominal compounds, however, NV compounds appear to be quite productive as observed in Booij (2002a: 150); the following examples illustrate this pattern:

- (19) [[[aardappel]_N[schil]_V]_V[mesje]_N]_N ‘lit. potato peel knife, potato peeler’
[[[brand]_N[blus]_V]_V[installatie]_N]_N ‘lit. fire extinguish installation, fire extinguisher’
[[[koffie]_N[zet]_V]_V[apparaat]_N]_N ‘lit. coffee make machine, coffee maker’

In these nominal compounds the left constituent is a NV compound in which the N functions as the Patient of the verb. Remember that, as mentioned in section 2, only compounding with nominal heads is recursive. Note, however, that Dutch does not have the corresponding compound verbs *aardappelschil* ‘to potato-peel’, *brandblus* ‘to fire-extinguish’, and *koffiezet* ‘to coffee-make’ as NV compounds (*koffiezet* does occur, however, but as a phrasal, separable verb). That is, this NV pattern is only productive when morphologically embedded.

[NV]_V compounding in Dutch is not only triggered by VN compounding (the NV compound feeds the non-head V position of these VN compounds) , but also by suffixation with the deverbal suffixes *-er*, *-ster*, *-ing* and *-erij*:

- (20) *aandacht-trekk-er* ‘attention drawer’
 brand-bluss-er ‘fire extinguisher’
 gif-meng-er ‘poison mixer, poisoner’
 grappen-mak-er ‘lit. jokes maker, comedian’

kinder-verzorg-ster ‘children’s care worker (fem.)’
kranten-bezorg-ster ‘newspaper deliverer (fem.)’
rokkenn-naai-ster ‘skirts sewer (fem.)’
vee-hoed-ster ‘cattle herd (fem.)’

evangelie-verkondig-ing ‘gospel preaching’
hand-oplegg-ing ‘hands imposition’
kinder-verzorg-ing ‘child care’
tempel-reinig-ing ‘temple cleansing’

bijen-houd-erij ‘bee keeping’
bloem-kwek-erij ‘flower nursery’
vlaggen-mak-erij ‘flag makery’
wijn-zuip-erij ‘excessive wine drinking’

These types of complex words are usually referred to as synthetic compounds since both compounding and derivation seem to be involved in these word formation processes.

The proper account of complex words of the type [NV-*er*]_N has evoked a lot of discussion in the morphological literature on English and Dutch (cf. Lieber 1983, Booij 1988). The main objection raised against assuming NV compounds as bases for these kinds of derivation (the hypothesis put forward by Lieber 1983 for English) is that, both in English and Dutch, NV compounding is not productive (Booij 1988, Lieber 2004: 48). On the other hand, from a semantic point of view the assumption of NV bases is attractive because the N in these examples functions as the Patient of the V. Therefore, another analytical options has been proposed: in Booij (1988), a word such as *brandblusser* ‘fire extinguisher’ is analysed as a case of NN

compounding in which the head noun is a deverbal N. That is, the semantic unit corresponding with NV is not reflected by a structural unit NV. Instead, the notion ‘inheritance of argument structure’ is invoked: the deverbal noun inherits the Patient argument of the verb, and the left constituent receives this Patient role.

Once we accept the idea that productivity of a certain word formation pattern may be linked to its occurrence in certain morphological constructions, another attractive analytical option is offered: these words have been derived by means of the conflation of NV compounding and the suffixation of *-er* to these NV compounds. The resulting structure gives direct expression to the generalization that the noun and the verb belong together from the semantic point of view. The Dutch word *grappenmaker* ‘comedian’, for instance, refers to someone who makes jokes, and an *aandachttrekker* ‘lit. attention drawer’ is someone who draws attention. This analytical option provides a way of overcoming the problem that NV compounding is at first sight unproductive. Thus, the structural analysis is the same as that proposed by Lieber (1983), but combines this with template conflation. Through unification of the templates $[NV]_V$ and $[V\ er]_N$ we get the template $[[NV]_V\ er]_N$. This latter, unified template can be qualified as productive, unlike the NV template in isolation. More generally, the observed productivity boost of NV compounding in de-verbal word formation can be expressed by unification of the relevant templates. Hence, the following productive unified templates can be postulated for Dutch:

- (21) $[[[N][V]]_V\ er]_N$
 $[[[N][V]]_V\ ster]_N$
 $[[[N][V]]_V\ ing]_N$
 $[[[N][V]]_V\ erij]_N$

These templates will be dominated by two different word formation templates, one for NV compounding, and one for deverbal nouns in *-er*, *-ster*, *-ing* and *-erij*. The templates will be instantiated by complex words such as those listed in (20).

The words in (20) are referred to as synthetic compounds since they seem to be cases of compounding and derivation at the same time. The account outlined above makes this notion of simultaneity more precise: structurally there is a hierarchy in that the compound is the base of a derived word, and the systematic co-occurrence of the two word formation processes is expressed by template unification. In order to use the notion of conflation, we need templates or schemas for the specification of these recurrent combinations of word formation patterns. Such templates, in

which both variable and lexically specified positions occur, are in fact constructional idioms at the word level, and thus provide additional evidence for a constructional approach to morphology.

5. Compound-like phrases

In many languages certain types of phrases perform similar functions as words (Jackendoff 1997; 2002, Booij 2002a,b). For instance, the following Dutch AN phrases function as terms to refer to (classes of) objects:

- (22) dikk-e darm ‘large intestine’
 hog-e hoed ‘lit. high hat, top hat’
 vrij-e trap ‘free kick’
 zwart-e doos ‘black box’

We know for sure that the AN combinations are phrases, since the adjectives are inflected: the final *-e* is an inflectional ending (a schwa). The specific property of these AN phrases is that the A position cannot project into a full AP with modifiers. For instance, in the phrase *hele zwarte doos* ‘very black box’ the specific meaning of ‘black box’ as used for the registration device in air planes gets lost. Similarly, Giegerich (2005) presents ample evidence from English why certain classes of AN combinations such as *polar bear* and *mental institution* must be considered lexical even though they are phrasal in nature.

Ralli and Stavrou (1998) discuss similar data from Modern Greek. They argue for a distinction between A-N compounds and A-N constructs for this language. For instance, the AN combination *psixros polemos* ‘cold war’ is qualified as a construct rather than a compound since there is inflection of the adjective, and agreement of the adjective with the noun. On the other hand, the AN combination *aghri-o-ghatos* ‘wild cat’ is a real compound consisting of the non-inflected adjectival root *aghri* followed by a linking element *-o-* and the noun *ghatos* ‘cat’. What we therefore need to account for these AN-constructs is a constructional schema $[AN]_N'$ that is dominated by the more general constructional schema for Greek NPs, and generalizes over the set of ‘non-projecting’ AN combinations that functions as conventionalized terms rather than as descriptions.

In Hoekstra (2002), a study of Frisian NN combinations, a distinction is proposed between two classes of NN compounds: the regular, garden variety ones, and a specific class of genitive compounds. In the latter type of NN compounds, the first noun is inflected and bears

genitive case. That is, such compounds have phrasal characteristics because case marking of a word-internal constituent by another constituent of the same complex word should be ruled out. Hoekstra therefore concludes that such genitive compounds must be analyzed as lexical phrases, and he points out that we should use the idea of the constructional idiom for this purpose. That is, these NN combinations have phrasal status, which explains why the head noun can impose genitive case marking on the other noun. Yet, they are lexical, since they form a set of established expressions.

Similar syntactic, yet lexical constructions can be found for modern Hebrew (Borer 1988). The difference between Frisian and Hebrew is that in Frisian it is the modifier that is case-marked (with the genitive ending *-s*), whereas in modern Hebrew it is the head noun that is marked as having a relation with another noun (*beyt* is the construct state form of *bayit* ‘house’). Hence, the head noun appears in the construct state form:

- (23) *Frisian* kening-s-dichter ‘king’s daughter’
Hebrew beyt xolim ‘lit. house sick, hospital’

Another class of compound-like complex expressions is that of the particle verbs in Germanic languages such as Dutch, Frisian, and German. As has been shown in a number of publications (cf. Booij 1990, Blom 2005 and the references mentioned there), these particle verbs cannot be considered verbal compounds of the type [Particle V]_v. They are phrasal in nature because the particle and the verb are split, with other words intervening, in root clauses. Yet they require lexical specification. Construction Grammar provides a natural account for these ‘lexical phrases’: they can be accounted for in terms of constructional idioms. For instance, the Dutch particle *door*, which corresponds with the polysemous preposition *door* ‘through, by’, evokes a specific meaning of continuation when combined with verbs. Examples are *dooreten* ‘to continue eating’ and *doorwerken* ‘to continue working’. Hence, the following phrasal constructional idiom has to be assumed in the lexicon of Dutch:

- (24) [[door [x]_v]_v] ‘to continue V-ing’

The framework of Construction Grammar thus provides the means to do justice to both the phrasal and the lexical nature of such word combinations. More generally, by making use of the notion ‘construction’ in morphology and in syntax alike, we are able to give an adequate treatment of both morphological and syntactic word combinations that functions as terms and

account for both their similarities and their differences (complex word versus phrase) in a model of the grammar in which there is no separation of syntax and the lexicon (cf. Jackendoff 2002: Chapter 6). We thus get rid of hybrid concepts such as ‘syntactic compound’.

6. Conclusions

In this chapter I showed how some basic insights of Construction Grammar can be applied to the domain of compounding, and how compounding provides empirical support for this grammatical model. First of all, we saw how (semantic or formal) generalizations about subsets of compounds can be made while at the same time the common or default properties can still be expressed.

The regularities concerning the position of the head in compounds are more complicated than can be handled with a simple left-right parameter. The notion ‘constructional idiom’ appeared to be the key to a proper account of the relevant complications, as well as for dealing with prefix-like constituents of compounds..

The possibility of conflation of word formation schemas provides a formal account of the co-occurrence of word formation patterns, in particular in the domain of synthetic compounding.

Finally, the difference between compounds that are words in the morphological sense and lexical phrases receives a straightforward representation in such a model of the grammar.

References

- Becker, Thomas. 1994. ‘Back-formation, cross-formation, and ‘Bracketing Paradoxes’ in Paradigmatic Morphology’, in Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology 1993*. Dordrecht: Kluwer, 1-26.
- Blom, Corrien 2005. *Complex predicates in Dutch. Synchrony and diachrony*. Utrecht: LOT.
- Booij, Geert 1988. ‘The relation between inheritance and argument linking: deverbal nouns in Dutch’, in Martin Everaert, Arnold Evers, Riny Huybregts, and Mieke Trommelen (eds), *Morphology and modularity*. Dordrecht: Foris, 57-74.
- Booij, Geert 1990. ‘The boundary between morphology and syntax: separable complex verbs in Dutch’, in Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1990*. Dordrecht: Foris Publications, 45-63.
- Booij, Geert 2002a. *The morphology of Dutch*. Oxford: Oxford University Press

- Booij, Geert 2002b. 'Constructional idioms, morphology, and the Dutch lexicon', *Journal of Germanic Linguistics* 14, 301-327.
- Booij, Geert 2005. 'Compounding and derivation: evidence for Construction Morphology.' In Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer, and Franz Rainer (eds.), *Morphology and its demarcations*. Amsterdam / Philadelphia: John Benjamins, 109-32.
- Borer, Hagit 1988. 'On the morphological parallelism between compounds and constructs', in Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1988*. Dordrecht: Foris Publications, 45-65.
- Corbett, Geville and Fraser, Norman M. 1993. 'Network Morphology: a DATR Account of Russian Nominal Inflection'. *Journal of Linguistics* 29, 113-42.
- Croft, William and Alan Cruse 2003. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- De Haas, Wim and Mieke Trommelen 1993, *Morfologisch Handboek van het Nederlands*. Den Haag: SDU Uitgeverij.
- De Jong, Nivja H., Rob Schreuder, and R. Harald Baayen 2000. 'The morphological family size effect and morphology.' *Language and Cognitive Processes* 15, 329-365.
- Downing, Pamela 1977. On the creation and use of English compound nouns. *Language* 53, 810-842.
- Giegerich, Heinz 2005. 'Associative adjectives in English and the lexicon-syntax interface'. *Journal of Linguistics* 41, 571-591.
- Goldberg, Adele 2006. *Constructions at Work*. Oxford: Oxford University Press.
- Hippisley, Andrew 2001. Word formation rules in a default inheritance framework: a Network Morphology account of Russian personal names', in Geert Booij and Jaap van Marle (eds.), *Yearbook of Morphology 1999*. Dordrecht: Kluwer, 221-262.
- Hoekstra, Jarich 2002. 'Genitive compounds in Frisian as lexical phrases', *The Journal of Comparative Germanic Linguistics* 6, 227-259.
- Jackendoff, Ray S. 1975. 'Semantic and morphological regularities in the lexicon'. *Language* 51, 639-71.
- Jackendoff, Ray S. 1997. *The architecture of the human language faculty*. Cambridge Mass.: MIT Press.
- Jackendoff, Ray S. 2002. *Foundations of language*. Oxford: Oxford University Press.
- Koenig, Jean-Pierre 1999. *Lexical relations*. Stanford: CSLI.

- Krieger, Hans-Ulrich and John Nerbonne 1993. 'Feature-based inheritance networks for computational lexicons', in Ted Briscoe, Ann Copestake, and Valeria de Paiva (eds.), *Inheritance, defaults and the lexicon*. Cambridge: Cambridge University Press, 90-136.
- Langacker, Ronald W. 1988. 'A usage-based model'. In Brygida Rudzka-Ostyn (ed.), *Topics in cognitive linguistics*. Amsterdam / Philadelphia: John Benjamins, 127-161.
- Langacker, Ronald W. 1998a. 'Conceptualization, symbolization, and grammar. In Michael Tomasello (ed.) *The new psychology of language. Cognitive and functional approaches to language structure*. Mahwah NJ. / London: Lawrence Erlbaum Ass. Publ, 1-39.
- Langacker, Ronald 1998b. 'Linguistic evidence and mental representations', *Cognitive Linguistics* 9, 151-173.
- Lieber, Rochelle 1983. 'Argument linking and compounds in English'. *Linguistic Inquiry* 14, 251-286.
- Lieber, Rochelle 2004. *Morphology and lexical semantics*. Cambridge: Cambridge University Press
- Marchand, Hans 1969. *The categories and types of present-day English word formation*, 2nd ed. München: Beck.
- Michaelis, Laura A., and Knud Lambrecht 1996. Toward a construction-based theory of language function: The case of nominal extraposition. *Language* 72, 215-247.
- Pitt, David and Jerrold J. Katz. 2000. 'Compositional idioms', *Language* 76, 409-432.
- Ralli, Angela and Melita Stavrou 1998. 'Morphology-syntax interface: A-N compounds vs. A-N constructs in modern Greek', in Geert Booij and Jaap van Marle (eds.) *Yearbook of Morphology 1997*. Dordrecht: Kluwer, 243-264.
- Riehemann, Suzanne 1998. 'Type-based derivational morphology', *The Journal of Comparative Germanic Linguistics* 2, 49-77.
- Rainer, Franz and Soledad Varela 1992. Compounding in Spanish. *Rivista di Linguistica* 4, 117-142.
- Scalise, Sergio 1984. *Generative morphology*. Dordrecht: Foris.
- Scalise, Sergio 1992. Compounding in Italian. *Rivista di Linguistica* 4, 175-200.
- Schwarze, Christoph 2005. Grammatical and para-grammatical word formation. *Lingue e Linguaggio* 2, 137-162
- Taylor, John 2002. *Cognitive Grammar*. Oxford: Oxford University Press.
- Tomasello, Michael 2000. 'Do young children have adult syntactic competence? *Cognition* 74 (2000), 209-253.

- Tomasello, Michael 2003. *Constructing a language. A usage-based theory of language acquisition*. Cambridge Mass. / London Engl.: Harvard University Press. :
- Vallès, Teresa 2003. Lexical creativity and the organization of the lexicon. *Annual Review of Cognitive Linguistics* 1, 137-160.
- Williams, Edwin 1981. 'On the notions 'Lexically related' and 'Head of a word''. *Linguistic Inquiry* 12, 245-74.