Constructional Idioms, Morphology, and the Dutch Lexicon

Geert Booij

*Vrije Universiteit Amsterdam*

Syntactic constructions may form an alternative to, or compete with the morphological expression of semantic and grammatical content. This applies to the passive forms of verbs, the progressive form, analytic causatives, adjective-noun sequences, and particle verbs in Dutch. In this article I develop a view of the Dutch lexicon in which this interaction between syntax and morphology can be understood. The central notion used is that of the constructional idiom, a construction with a (partially) non-compositional meaning, of which not all terminal elements are fixed. These constructional idioms, like morphological word formation, serve to extend the fund of expressions that are available for concatenation in the syntax.*

1. Introduction.

It is well known that syntactic structures sometimes perform the same function as morphological structures in the same or another language. Periphrasis is the standard term for this morphological function of syntactic units within the inflectional system of a language. In the domain of word formation, linguists often contrast analytic constructions with synthetic constructions, and distinguish, for instance, between analytic causatives (multiword units) and synthetic, that is, morphological causatives.

In this paper I argue that the notion “constructional idiom” should be used in order to get a better insight into the kind of syntactic expressions that function as alternatives to morphological expressions. The basic claim is that it is syntactic expressions that qualify as constructional idioms that play a role in the division of labor between syntax and morphology.

---

* I would like to thank the anonymous referees of this journal, and Farrell Ackerman, Arie Verhagen, and the audience at the MMM3 meeting in Barcelona, September 2001, for their comments on previous drafts of this paper.

© Society for Germanic Linguistics
Constructional idioms are syntactic constructions with a (partially or fully) noncompositional meaning contributed by the construction, in which—unlike idioms in the traditional sense—only a subset (possibly empty) of the terminal elements is fixed. The idea of constructional idioms can be found in the work of Langacker (1987), in the framework of Construction Grammar (cf. Goldberg 1995; Fillmore, Kay, and O’Connor 1988; Kay and Fillmore 1999; Pitt and Katz 2000), and in recent work by Jackendoff (1995, 1997, 2001, 2002). Other terms used are “lexical phrases with a generalized frame” (Nattinger and DeCarrico 1992:36), and “idiomatic pattern” (Everaert 1993:9).

A telling example of a Dutch constructional idiom is the *een schat van een kind* construction, well known among Dutch linguists since it plays a prominent role in the work of the Dutch grammarian Paardekooper. The examples in 1 illustrate this constructional idiom (cf. Everaert 1992:48).

(1)  

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>een schat van een kind</em></td>
<td><em>(lit.) a sweetheart of a child, a sweet child</em></td>
</tr>
<tr>
<td><em>een kast van een huis</em></td>
<td><em>(lit.) a cupboard of a house, a big house</em></td>
</tr>
<tr>
<td><em>een boom van een kerel</em></td>
<td><em>(lit.) a tree of a chap, a big chap</em></td>
</tr>
</tbody>
</table>

The formal syntactic structure of such phrases is that of an NP with a PP complement. However, semantically the noun of the PP complement functions as the head, and it also determines the gender of the relative pronoun for which it is the antecedent as shown in 2.

(2)  

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>een kast van een huis</em></td>
<td>a big house that needs to be painted</td>
</tr>
</tbody>
</table>

Note that the noun *kast* is non-neuter, whereas *huis* is neuter; the relative pronoun *dat* is the pronoun for antecedents with neuter gender. This clearly shows that it is not the formal syntactic head that determines the gender of the relative pronoun, but the noun of the complement. Another specific property of this construction is that the two nouns have to agree in number. For instance, the plural of *een schat van een kind* is *schatten van kinderen*, with both nouns in their plural form: both *schatten van een kind* and *een schat van kinderen* are ill-formed in the interpretation under discussion here (the literal interpretations, however, are well formed).
This class of constructional idioms can be extended, and hence they do not form a fixed list of expressions. The first noun has to be a noun that expresses an evaluation of properties of the noun in the PP complement. For instance, it is possible to coin the phrase *een godin van een vrouw* ‘(lit.) a goddess of a woman, a ravishing woman’ as a new instantiation of this constructional idiom. Nevertheless, this construction does not lend itself to unlimited extension, and the example *een godin van een vrouw* is perceived as a case of creative language use. That is, the notion “restricted productivity” applies, a notion that is standardly used for describing morphological patterns.¹

The implication of the existence of such constructional idioms is that the lexicon, the list of fixed linguistic expressions, has to be extended with partially underspecified idioms, in this case the NP-type *een N₁ van een N₂* with the meaning ‘N₂ who/which is an N₁’.

The *een schat van een kind* construction is mentioned here only as an illustration of the notion “constructional idiom,” and is not to be seen as an alternative to morphological expressions. In this article I focus on those constructional idioms that do function as alternatives to morphological expressions, and I argue that it is typically constructional idioms that may perform that function. In section 2 I discuss periphrastic expressions, both in the inflectional and the derivational domain, and in section 3 I discuss constructional idioms that function as alternatives to morphological word formation. In section 4 I summarize and discuss my findings.

2. Periphrasis.

2.1. Inflectional Periphrasis.

In the inflectional domain, it is quite clear that we need the theoretical concept of periphrasis, the expression of inflectional information by means of a combination of words. Periphrastic constructions are the prototypical cases of analytic lexical expressions.

A well-known case of periphrasis is the expression of the perfective passive form in Latin by means of a combination of the past participle plus an appropriate form of the verb *esse* ‘to be’, as in *laudatus est* ‘he

¹ Similar constructional idioms are found in English (*a brute of a man*), German (*ein Teufel von einem Mann* ‘a devil of a man, a brute man’), Spanish (*esa mierda de libro* ‘that shit of a book, that shitty book’) and French (*une drôle d’histoire* ‘a strange story’).
Booij has been praised’ (Börjars et al. 1997; Sadler and Spencer 2001). These periphrastic combinations are only used for the perfective passive, whereas synthetic forms are used for expressing the imperfective passive, as illustrated in 3 (from Sadler and Spencer 2001: 74).

(3) Paradigm of 3sg. forms of *laudare* ‘to praise’

<table>
<thead>
<tr>
<th></th>
<th>IMPERFECTIVE</th>
<th>PERFECTIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>Present</td>
<td>laudat</td>
</tr>
<tr>
<td></td>
<td>Past</td>
<td>laudabat</td>
</tr>
<tr>
<td></td>
<td>Future</td>
<td>laudabit</td>
</tr>
<tr>
<td>Passive</td>
<td>laudatur</td>
<td>laudatus/a/um est</td>
</tr>
<tr>
<td></td>
<td>laudabantur</td>
<td></td>
</tr>
<tr>
<td></td>
<td>laudabitur</td>
<td></td>
</tr>
</tbody>
</table>

The fact that this periphrastic form is the only possible form for expressing the perfect passive shows that the form fills a cell in the inflectional paradigm. Moreover, as pointed out by Börjars et al. (1997), in the case of deponentia (verbs with a passive form and an active meaning) such as *loquor* ‘to speak’, the periphrastic form has an active meaning, just like the other, synthetic, forms: *locutus est*, for instance, means ‘he has spoken’. Börjars et al. (1997) propose to account for the functional equivalence of such word combinations to synthetic morphological forms in the inflectional paradigm of Latin verbs in terms of unification of the functional structures of the two words into one functional structure at the level of f(unctional)-structure. However, as Sadler and Spencer (2001:78) argue, there is a problem with this compositional approach: the forms of *esse* ‘to be’ that are used in this construction are imperfective forms, and yet the whole construction bears perfective aspect. Hence, it is the periphrastic construction as a whole that has to be assigned the perfective aspect.

The notion “periphrasis” can also be used in a looser sense, namely for the analytic expression of information in a certain language that is expressed morphologically in other languages (cf. Haspelmath 2000). This applies to the expression of information with respect to voice, aspect, Aktionsart, and similar categories. This kind of analytic expression is a widespread property of natural languages, as is also clear from the grammaticalization studies in Bybee and Dahl 1989, and Bybee et al. 1994. It is the very phenomenon of grammaticalization that makes
us expect to find such patterns of analytic expression of grammatical information: lexical words can develop into grammatical words (and these in turn may subsequently develop into bound grammatical morphemes).

A well-known case of this morphology-like use of syntax is the passive construction in Indo-European languages, which developed from regular syntactic patterns. However, it is not always the case that there are also synthetic passive forms (as is the case in Latin as discussed above), and hence there is no strict argument for considering the periphrastic passive constructions of such languages as filling cells of the inflectional verbal paradigm.

The Dutch passive construction is of particular interest because it shares a property with the Latin periphrastic passive: the perfective forms are expressed with imperfective forms of the verb *zijn* ‘to be’. Consider the data in 4.

(4) 3sg. forms of the Dutch verb *doden* ‘to kill’

<table>
<thead>
<tr>
<th></th>
<th>Active</th>
<th>Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Present</strong></td>
<td>doodt</td>
<td>wordt gedood</td>
</tr>
<tr>
<td><strong>Past</strong></td>
<td>doodde</td>
<td>werd gedood</td>
</tr>
</tbody>
</table>

**PERFECTIVE**

<table>
<thead>
<tr>
<th></th>
<th>heeft gedood</th>
<th>is gedood</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Present</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Past</strong></td>
<td>had gedood</td>
<td>was gedood</td>
</tr>
</tbody>
</table>

As is clear from these data, the passive participle *gedood* does not carry inherent perfective aspect since in combination with forms of *worden* it occurs with imperfective aspect. The passive participle is also used with imperfective interpretation in combination with certain verbs, such as *krijgen* ‘to get’, as illustrated by sentence *De burgemeester krijgt een fiets aangeboden* ‘The mayor is offered a bicycle’. Yet in combination with the imperfective forms of the verb *zijn* ‘to be’, it expresses perfective aspect. If we expressed perfective aspect compositionally, we would expect forms such as *is gedood geworden* since the verb *zijn* functions as auxiliary for the perfective aspect forms of the verb *worden*. Actually, equivalent aspectual forms do occur in English (*has been killed*) and in German (*ist getötet worden*), but such forms are normally
not used in standard Dutch. In earlier stages of Dutch the verb *zijn* was also used in combination with the passive participle to express imperfective aspect (as in present-day English), but gradually the verb *worden* has taken over this role (Van der Wal 1986), and *zijn* now triggers a perfective interpretation.

There is no doubt that the English and the German passive periphrastic forms are also constructional idioms because of the specific passive interpretation imposed on the combination of the verbs *to be* / *worden* in combination with a past participle. The Dutch passive construction, however, is, as shown above, even more idiomatic because of the mismatch between the aspectual properties of the finite verb and that of the passive construction as a whole.

### 2.2. Progressives.

An example of a constructional idiom with a progressive meaning in Dutch is a form of the verb *zijn* ‘to be’ followed by a PP of the form *aan het* + verbal infinitive; the verb must be durational, and hence the stative verb *wonen* ‘to live’ and the punctual verb *sterven* ‘to die’ cannot be used in such a construction; compare the examples in 5.3.

(5)  

a. Jan is aan het fiets-en  
    John is at the cycle-INF  
    ‘John is cycling.’

---

2 Detailed argumentation for this analysis of the passive construction in Dutch can be found in De Haan 2000. He argues against previous analyses in which the passive interpretation of this construction was derived as much as possible (though not completely) from the meaning of its constituents, as in Cornelis and Verhagen 1995. Given the fact that the passive participle is always identical to the perfect participle, one might even go a step further, and argue that the passive meaning of the participle is not an inherent property of the passive participle, but a property of the construction. However, the passive meaning is also present if the participle is used as an attributive adjective, without a passive auxiliary, as in *het gedode paard* ‘the killed horse’. This suggests that we should assume a conversion rule that converts perfect participles into passive participles, with the effect that the Agent argument of the verb is suppressed.

3 See Booij forthcoming for a detailed analysis of this construction.
b. Mijn zus was aan het studeren
   My sister was at the study-INF
   ‘My sister was studying.’

c. Jan is de aardappels aan het schillen
   John is the potatoes at the peel-INF
   John is peeling the potatoes.’

d. *De hond was aan het sterven
   ‘The dog was dying.’

e. *Jan is aan het wonen in Amsterdam
   ‘John is living in Amsterdam.’

As noted by Bybee and Dahl (1989:78–82), the use of to be + PP for expressing action in progress is widespread crosslinguistically. The formal structure of the PP complement of zĳn is that of a PP with the preposition aan and the neuter determiner het followed by the verbal infinitive which can function as a neuter noun. The specific property of the zĳn [aan het + verb]_{PP} construction is that it functions as a predicate with the meaning ‘continuous V-ing’. Therefore, we have to assume a construction in which the first three words are fixed, and with an open slot for the verbal infinitive:

\[(6) \text{zĳn } [\text{aan}]_{P} [\text{het}]_{Det} [\text{V-INF}]_{N}]_{NP} \text{PP}\]

‘be V-ing continuously’

The constructional idioms discussed above have an inflectional function. The important point to bear in mind is that they are the only grammatical expressions available for expressing this information: there is no synthetic morphological passive or progressive form. The situation is similar to English where the combination of to be and the -ing-form of a verb function together as the progressive form.

The NP-part (het + infinitive) of the progressive construction does not exhibit the normal properties of an infinitival nominalization. Normally a verbal infinitive preceded by a determiner allows for being projected, and takes a prepositionless NP-object, or a PP-complement as in 7.
(7) het een appel kopen
   ‘(lit.) the an apple buying, the buying of an apple’
   het kopen van een appel
   ‘the buying of an apple’

However, this is impossible in the *aan het*-construction, as the (un)grammaticality of the following sentences attests. As shown in 8a, the direct object *een appel* ‘an apple’ can appear before the sequence *aan het V-INF*. This shows that the whole construction functions as a verbal unit.

(8) a. Jan is een appel aan het kopen
    ‘John is buying an apple.’

   b. *Jan is aan het een appel kopen
      ‘John is buying an apple.’

   c. *Jan is aan het kopen van een appel
      ‘John is buying an apple.’

On the other hand, it is possible to use multiword expressions of the type N V (with the N receiving a generic interpretation) after *aan het*.

(9) a. Jan is aan het koffie zetten
    ‘John is making coffee.’

   b. Jan is aan het aardappels schillen
    ‘John is peeling potatoes.’

In section 3 we see that *koffie zetten* and *aardappels schillen* are themselves constructional idioms that function as lexical units, and this explains why they can occur in the verbal position after *aan het*.

We thus observe that in this constructional idiom the normal syntactic projection possibilities of the verb are blocked. In section 3 we come across other cases of nonprojecting categories within constructional idioms.

In Dutch there is a division of labor between present participles and the *aan het INF*-construction: present participles (with the morphological
form stem + -end) also receive a progressive interpretation, but these present participles can only be used in attributive position; compare 10.4

(10) a. De fiets-end-e man
    ‘The cycling man’

b. De man is fietsend / aan het fietsen
    ‘The man is cycling.’

This example shows that the aan het INF-construction functions as a periphrastic form of verbs in order to express progressive aspect in predicate position. Present participles in Dutch still have a full verbal potential even when used as adjectives in attributive position. In predicate position, however, it is the aan het-INF-construction that has this full verbal potential:

(11) a. De zijn vader beledigende jongen
    ‘the boy who is insulting his father’

b. *De jongen is zijn vader beledigend
    ‘The boy is his father insulting’

c. De jongen is zijn vader aan het beledigen
    ‘The boy is insulting his father.’

On the other hand, the periphrastic progressives cannot be used in attributive position. This follows from the fact that their formal status is that of a PP, since PPs cannot be used as attributive modifiers in Dutch:

(12) a. *De in het bos jongen
    ‘the in the wood boy’

---

4 Note that there are adjectives with the form of present participles that can be used in predicative position such as schokk-end ‘shocking’, woed-end ‘(lit.) raging, angry’, and lop-end ‘(lit.) walking, on foot’. However, these are all lexicalized cases of present participles with an unpredictable meaning that function as adjectives. They cannot be used in predicative position in their literal interpretation.
b. *de aan het fietsen jongen
   ‘the cycling boy’

Thus, we observe here a specific division of labor between morphology and syntax: progressive aspect is expressed by a synthetic form (the present participle) in attributive position, and by an analytic form (the aan het INF-construction) in predicative position. This functional equivalence and complementarity between a participle and a syntactic construction can be understood since the syntactic construction is a constructional idiom: constructional idioms are lexical units, and hence they can be expected to interact with morphology, another source of the lexical units of a language.

2.3. Causatives.
The Dutch causative verb laten ‘to let’ combines with verbs to form a complex causative predicate. Consider the following sentence (with the SOV word order of embedded clauses, the underlying word order of Dutch):

(13) dat Jan zijn vader een boek laat zien
    that John his father a book lets see
    ‘that John shows his father a book’

In this sentence, the verbal complex laat zien, consisting of the finite verbal form laat ‘lets’ and the infinitive zien ‘to see’, has the same meaning and syntactic valency as its gloss shows in English. Yet, it is a multiword unit, and the two words can be separated from each other in root clauses such as Jan laat zijn vader een boek zien ‘John shows his father a book’. In the early days of generative grammar this construction was analyzed in terms of a syntactic transformation of verb raising: the verb laten and a number of other verbs trigger the raising of the verb of the embedded clause to the higher clause (Evers 1975). That is, the following surface structure was assumed:

(14) … Jan [zijn vader]ₜ [PROₜ het boek tₜ]ₜ /[laat]ₜ/[ zien]ₜ \v

A classic example of complementarity between syntactic and morphological expressions is the comparative form of the adjective in English; a synthetic form for monosyllabic adjectives and for disyllabic adjectives ending in a light syllable, an analytic form more A in all other cases.
This analysis accounts for the fact that the verbal complex *laat zien* consisting of two verbs has three NP arguments: two from the verb *laten*, and two (of which one is coindexed with an argument of the main verb) from the verb *zien*. Thus, a syntactic rule of Verb Raising is assumed to create the complex predicate *laat zien*.

An alternative syntactic analysis of such complex predicates, without making use of a transformation, is one in which the verb *laten* is subcategorized in the lexicon as taking a complement V. In addition, we then need a mechanism of argument inheritance (functional composition) that expresses that the verbal complex has three argument positions, those of the main verb and those of the complement verb. Moreover, there may be coindexing between an argument of the main verb and one of the complement verb. In this example, the NP *zijn vader* may be considered as an argument of both *laten* and *zien*. This is the kind of analysis of complex predicates that has been proposed in non-derivative syntax of the HPSG type (see Hinrichs et al. 1998). The advantage of this analysis is that it avoids the need to assume transformational rules that are triggered by the presence of specific lexical items. The possibility of forming complex predicates is expressed through the subcategorization properties of lexical items, in combination with the mechanism of functional composition, in which the argument structures of a head and its complement are combined.

A third option, and the one that I defend here for the verb sequence *laten* + V, is that the existence and creation of such verbal complexes are accounted for by means of a syntactic template of the form [*laten* V], a template with two terminal positions of which one is fixed and the other one is a variable into which all kinds of verbs can be inserted. This constructional idiom analysis has the same advantage as the HPSG analysis in that it properly accounts for the lexical governedness of the construction. It has the added advantage that it allows for a natural account of the irregularities in this constructions as well. For instance, as observed in detail in Coopmans and Everaert 1988, many combinations of *laten* + V have idiosyncratic meanings, and often the complement V has a meaning that it only has in combination with *laten*, as illustrated by the examples in 15 from Coopmans and Everaert 1988.

(15) a. dat Jan deze kans liet lopen
    that John this chance let slip away
    ‘that John missed this opportunity’
b. Hij heeft haar laten barsten
   He has her let burst
   ‘He has left her in the lurch.’

The verb *lopen* ‘to walk’ in the first example can only be used with the special meaning ‘slip away’ in combination with *laten*: a sentence such as *De kans liep* cannot receive the interpretation ‘The chance slipped away’. Similarly, the sentence *Zij barst* ‘(lit.) She bursts’ cannot have the meaning ‘She is in the lurch’. In the same vein, there are verbs that only occur in combination with *laten*, for instance the verb *geworden*, as in 16.

(16) iemand laten geworden
    ‘to let somebody alone’

Such verbs are therefore bound elements, but not in the morphological sense of bound morphemes (they are still separable from finite forms of *laten* in root clauses, and hence independent words), but in the sense that they only occur as part of a lexically listed *laten* + V complex. Therefore, I assume a productive lexical template [**laten** V] in the lexicon, and a list of instantiations of this template, with the V-position lexically specified, for those cases that have idiosyncratic properties.

This lexical template analysis of the *laten*-construction is also defended in Kemmer and Verhagen 1994 and Verhagen 1997. They observe that the causal predicate *laten* does not introduce its own specific semantic roles, as is expected for fully elaborated predicates that project their own clause. Instead, the verb *laten* can be said to create ditransitive predicates when combining with transitive verbs, as illustrated in 17.

(17) a. Ik liet mijn boek aan mijn collega zien
    I let my book to my colleague see
    ‘I showed my book to my colleague.’

b. Zij lieten dit aan niemand lezen
    they let this to nobody read
    ‘Nobody was allowed to read this.’

In these sentences the NPs *mijn collega* and *niemand* bear the recipient role, which is marked by the preposition *aan*. This recipient role is not available for *laten* as such, nor to its verbal complements *zien* or *lezen*: it is only available for the combination of *laten* + V.
The existence of a productive analytic causative construction has apparently preempted the rise of a productive morphological process for the formation of causative verbs in Dutch. The only causative suffix in Dutch is the nonnative suffix -iseer ‘-ize’, which can only be used for the formation of causative verbs from nonnative adjectival bases. For the creation of deverbal causatives, there is no affixation process available, and thus we see again that, as in the case of the progressive construction, it is constructional idioms that appear to influence the range of possibilities and use of the morphological system of a language.

3. Phrases and Compounds.

In this section I show that certain kinds of constructional idioms are functionally similar to compounds. They behave like the constructional idioms discussed above in the sense that they form alternatives to morphological expression of lexical and/or grammatical content. However, the phrases discussed in this section have an additional property in that they also feed word formation. For instance, the Dutch AN phrase rode kool ‘red cabbage’ can feed compounding, as in rode-koolgerecht ‘red cabbage dish’. Similarly, the AN phrase harde kern ‘hard core’ can be used as a base for the formation by means of the suffix -er of the derived noun harde kerner ‘hard core member’. In this respect they differ from the periphrastic/analytic constructions discussed in the preceding section, which do not feed word formation, as illustrated by the examples in 18.6

(18) passive: *[[gedood word]v [huis]sN], compare [[sterf]v [huis]sN ‘house of mourning’
causative: *[[laat zien]v[er]]sN, compare [[toon]v[der]]sN ‘person who shows something’

The kind of multiword units that will be discussed here are AN phrases, number names, and particle verbs.

6 The only exception in the case of complex predicates consisting of two verbs that I know of is the noun zitten-blijv-er ‘(lit.) sit-stay-er, repeater’.
3.1. Adjective Noun Phrases.

As has been pointed out time and again, more recently for English by Jackendoff (1997, 2001), noun phrases of the type A+N may have the same function as compounds. They provide names for a relevant class of entities for which the language user needs an expression with which these entities can be named. It is obvious that many of these phrases have to be listed in the lexicon because of their being the conventional, established names for these entities, and because of the unpredictable meaning aspects. This applies to English NPs such as *hard disk, big toe, yellow pages*, and *red tape*, and equally to Dutch phrases such as the following:

(19) dikke darm
   ‘large intestine’
dood spoor
   ‘(lit.) dead trail, deadlock’
hoge hoed
   ‘(lit.) high hat, top hat’
vrede trap
   ‘free kick’
open haard
   ‘(lit.) open hearth, fireplace’
vaste benoeming
   ‘(lit.) fixed appointment, tenure’
zure regen
   ‘acid rain’
zwarte doos
   ‘black box’

In the literature on English compounds and phrases of the type A+N there is often confusion or uncertainty about the formal status of individual cases of such word sequences: is it a compound or a phrase? In Dutch, however, the grammatical status of A+N sequences can be determined unambiguously due to the fact that prenominal adjectives in NPs are inflected. For instance, in Dutch we find the pair *zuurkool* ‘sauerkraut’ vs. *rode kóol* ‘red cabbage’. The first one is a compound, the second one a phrase with an inflected adjective ending in schwa. Note also the difference in stress: in the compound main stress is on the first constituent, in the phrase it is located on the second constituent, the head
of the phrase. The phrasal nature of *rode kool* can also be deduced from its diminutive form *een rood kooltje* ‘a small red cabbage’, in which the adjective is schwaless. This inflectional schwa is absent in singular indefinite NPs with a neuter noun as their head (the diminutive noun is neuter) (Booij 2002a: ch. 2). For some lexicalized ANs, their being conventional units is reflected by the orthography because they are written as one word, without internal spacing, as in *blindedarm* ‘blind gut’ and *jongeman* ‘young man’. The phrasal origin of these lexical units is still visible because the adjective ends in an inflectional schwa. Synchronically they are not phrases anymore, but words, as can be deduced from the form of their diminutives: *blindedarmpje*, *jongemannetje*, instead of the forms *blind darmpje* and *jong mannetje* which would be the correct forms if these expressions were phrases.

When AN phrases are used as names, that is, as terms for categories of entities, they have specific formal characteristics: the adjective is always a bare A, without modifiers. In other words, in this case the A does not project a full AP. As soon as we coin a phrase with a modified adjective like *een zeer rode kool* ‘a very red cabbage’, the phrase loses its classificatory function, its status as a name for a specific kind of cabbage. It becomes a descriptive expression that describes the (color) properties of a single cabbage. That is, we observe again that within constructional idioms, the lexical categories lose their normal projection possibilities.

The class of AN phrases with naming function can be readily extended, and is not a closed set of lexicalized phrases. For instance, as soon as there is a kind of tea to be named with its distinguishing color *yellow*, the name *yellow tea* can be coined. Therefore, it is appropriate to assume a constructional idiom in the lexicon with the form [A N]_{NP}, with two open positions, and no terminal element fixed. This template is a specific instantiation of the general syntactic template for NPs. The relation between the two can be expressed by making use of inheritance trees (Goldberg 1995). In such a tree most properties of the AN phrasal names are inherited from NPs in general as defined by the syntactic module. It is only the name function that needs to be specified for the AN phrases. By establishing a formal relation of inheritance between constructional idioms and the canonical syntactic templates as defined by the syntax, we express the generalization that most constructional idioms
are not completely arbitrary units of words, but form a subset of the types of word combinations that are defined as well formed by the syntax.7

Dutch AN phrases of the type discussed, with a naming or classificatory function, exhibit a morphological peculiarity in that in some of these phrases, there is no overt inflectional ending at the end of the adjective. This is illustrated by the examples in 20 (see Booij 2002a:47–48 for the complicated details).

(20) een medisch-Ø student
    ‘a medical student’/ ?een medische student
    de wetenschappelijk-Ø directeur
    ‘the scientific director’/ ?de wetenschappelijke directeur

Speakers of Dutch vary as to when they omit the inflectional schwa, but it is at least clear that it can be omitted. Since the main stress is on the noun, we can be certain that these units are phrases because AN compounds have main stress on the first constituent. This lack of inflectional schwa suggests that such AN phrases are becoming more and more similar to AN compounds in that they have no internal inflection. This can be seen as a symptom of AN phrases having the status of classificatory lexical expressions.

In languages that have both AN phrases and AN compounds, there is a competition between the two kinds of names. In German AN compounding is productive, whereas in English this kind of compounding is unproductive, and AN phrases are used instead. As is so often the case for Dutch, it is in between German and English with respect to the use of phrases or compounds, and it uses both options. Compare the following cases, most of them taken from De Caluwe 1990. In German we systematically have AN compounds, Dutch varies, but is rather similar to English, and English systematically has phrases (AN compounds do occur in English but this pattern is no longer productive; cf. Marchand 1969:63):

7 However, there are exceptions to this generalization. For instance, the English syntactic pattern the more-A/A-er, the more A/A-er (as in the more expensive, the better) is not a canonical syntactic pattern of present-day English. Note, by the way, that this type of expression presupposes the functional identity of the morphological and the periphrastic comparative.
In Dutch the difference between compounds and phrases is reflected systematically (with some exceptions, as mentioned above) by the orthography: there is no spacing within compounds, and always between the constituents of phrases. The orthography thus unambiguously indicates whether a Dutch AN sequence is a phrase or a compound.

The complementarity and competition of these two ways of coining expressions has often been observed (cf. De Caluwe 1990, Heynderickx 2001). If one form is chosen, the other is sometimes blocked, and this may be taken as an indication of the lexical status of this phrasal pattern. Normally, syntax does not block morphology. For instance, the fact that we can coin the PP *zonder gewicht* ‘without weight’ does not block the use of the morphological process of *-loos*-suffixation: the coining of *gewicht-loos* ‘weightless’ is not impeded by the existence of a syntactic alternative. The following examples from Dutch illustrate the blocking effects. The question mark indicates that, although the expression is well formed, native speakers feel unhappy about it, and will not use it.

<table>
<thead>
<tr>
<th>(21)</th>
<th>German</th>
<th>Dutch</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dunkelkammer</td>
<td>donkere kamer</td>
<td>dark room</td>
<td></td>
</tr>
<tr>
<td>Festplatte</td>
<td>harde schijf</td>
<td>hard disk</td>
<td></td>
</tr>
<tr>
<td>Kleinkind</td>
<td>klein kind</td>
<td>small child</td>
<td></td>
</tr>
<tr>
<td>Grosskind</td>
<td>kleinkind</td>
<td>grandchild</td>
<td></td>
</tr>
<tr>
<td>Kleinbus</td>
<td>kleine bus</td>
<td>small bus</td>
<td></td>
</tr>
<tr>
<td>Rotwein</td>
<td>rode wijn</td>
<td>red wine</td>
<td></td>
</tr>
<tr>
<td>Roteiche</td>
<td>rode eik</td>
<td>red oak</td>
<td></td>
</tr>
<tr>
<td>Tiefdruck</td>
<td>lage druk</td>
<td>low pressure</td>
<td></td>
</tr>
<tr>
<td>Hochspannung</td>
<td>hoogspanning</td>
<td>high tension</td>
<td></td>
</tr>
<tr>
<td>Hochsaison</td>
<td>hoogseizoen</td>
<td>high season</td>
<td></td>
</tr>
<tr>
<td>Schnellzug</td>
<td>sneltrein</td>
<td>fast train</td>
<td></td>
</tr>
</tbody>
</table>

(22) AN phrase    NN compound

| academisch jaar | ?academiejaar | ‘academic year’ |
| ?academisch lid | academielid   | ‘academy member’ |
| koninklijk besluit | ?koningsbesluit | ‘royal decision’ |
| koninklijk huis | koningshuis   | ‘royal family’   |
| ?koninklijk blauw | koningsblauw | ‘royal blue’ |
| muzikale scholing | ?muziekscholing | ‘musical training’ |
| muzikaal talent | muziektalent  | ‘musical talent’ |
| ?muzikale school | muziekschool | ‘school of music’ |
As can be concluded from these data, there is not always a blocking effect. The absence of absolute, systematic blocking is the normal situation for productive patterns of coining lexical expressions (Rainer 1988). But since blocking is typically an effect of the competition of lexical expressions, these blocking phenomena support the interpretation of AN phrases as lexical expressions.\(^8\)

The adjectives in the AN phrases above are all relational adjectives. Typically, these adjectives cannot be modified when used relationally (that is, they are nonprojecting), and, with some exceptions, they are only used in attributive position. Most European languages make use of such relational adjectives for the construction of NPs with a naming function (Heynderickx 2001).

It has been observed time and again that certain types of phrases can feed word formation, in particular nominal compounding, and not only lexicalized (that is, established) syntactic expressions (cf. Booij 2002a for an overview of the possibilities of Dutch). However, it is not the case that all kinds of phrases can be used in this position (Hoeksema 1988). It is precisely AN phrases (quantifiers are considered here a subset of adjectives) with a classificatory function that form the most frequently used type of phrase in the nonhead position of Dutch nominal compounds:

\[(23) \quad \text{[} \text{oude mannen} \text{]} \text{NP} \text{[} \text{huis} \text{]} \text{N} \]

‘old men’s home’

\[(23) \quad \text{[} \text{blote vrouwen} \text{]} \text{NP} \text{[} \text{blad} \text{]} \text{N} \]

‘nude women magazine’

\[(23) \quad \text{[} \text{hete lucht} \text{]} \text{NP} \text{[} \text{ballon} \text{]} \text{N} \]

‘hot air balloon’

\[(23) \quad \text{[} \text{zevende dag} \text{]} \text{NPs} \text{[} \text{adventist} \text{]} \text{N} \]

‘seventh day adventist’

Suppose now that the following generalization holds: lexical expressions with a classificatory function can feed word formation. If we assume that the lexicon contains a constructional idiom \([A N]_{NP}\) for the

\(^8\) The nature of ANs as lexical chunks also manifests itself in patterns of codeswitching, as observed by Backus (2000:99). For instance, a Turkish student in the Netherlands used the NP *verkorte opleiding* ‘shortened study program’ as a chunk in a Turkish sentence, presumably because this NP expresses a specific concept in the Dutch educational system.
construction of classificatory expressions, it follows that the left constituent of a nominal compound can be an NP of the AN type, even in cases where this AN combination has no idiosyncratic properties, as in the case of *blote-vrouwenblad* with the non-lexicalized constituent *blote vrouwen* ‘nude women’.

Classificatory AN-phrases of Dutch may also feed productive derivation, such as the creation of nouns by means of suffixation with *-er*. The examples in 24 illustrate this (the ordinals *vierde* and *achtste* in the b and c examples behave as adjectives):

(24) a. [harde kern]er
   ‘(lit.) hard corer, who belongs to the hard core’

b. [vierde klass]er
   ‘(lit.) fourth classer, who belongs to the fourth class’

c. [achtste groep]er
   ‘(lit.) eighth grouper, who belongs to the eighth group’

In sum, by enriching the lexicon with constructional idioms, templates with one or more open positions, the behavior of phrases that are functionally similar to compounds can be accounted for adequately.9

3.2. Number Names.

Dutch number names exhibit interesting properties that also shed light on the division of labor between syntax and morphology. The Dutch

---

9 The similarity in behavior between NPs and nominal compounds, in combination with the lexicalization and classificatory use of such NPs, also plays an important role in the debate on the proper analysis of certain kinds of nominal lexical expressions in Romance languages (cf. Scalise 1992) and Greek. For instance, Ralli (1992) showed that Greek word combinations such as *atomika vomva* ‘atomic bomb’ have the status of loose or phrasal compounds (AN constructs with a relational adjective followed by a bare noun). They behave as syntactic atoms (cf. also Ralli and Stavrou 1998 for a detailed analysis) except that there is internal inflection: the adjective agrees with the noun. Ralli and Stavrou (1998:258) therefore conclude that these AN constructs have to be generated by the syntax. However, if we extend the notion of the lexicon, and enrich it with constructional idioms, it is clear that such AN constructs will be accounted for by a template of the form [A N]_{NP} in the lexicon of Greek.
Booij

patterns for cardinal number names above 20 are as follows (Booij 2002a):

(25) 21–99     een-en-twintig, twee-en-twintig, ...
       101– …    honderd-(en)-een, honderd-(en)-twee, ...

That is, between 21 and 99 the ones come before the tens (like in German, but unlike English). The coordinative conjunction *en* ‘and’ is used, and is optional above 100. Such complex expressions are not endocentric since it is impossible to designate one of the constituents as the head. So they look like cases of regular syntactic coordination (and semantically they are, because the quantity referred to by the number name is the sum of that of its constituents!), except that the two patterns are lexicalized: it is only above 100 that *en* is optional. Moreover, the conjunction *en* is pronounced as [en] when used in normal syntax, and in numbers above 100, but obligatorily as [en] in numbers between 21–99. In such numbers the conjunction cliticizes prosodically to the first number word. The order in which the ones and the tens appear is also a matter of convention, as we have seen above. This implies that these two coordinative patterns for number names are constructional idioms that have to be stored in the lexicon.10

Just like classificatory AN-phrases cardinals of this form feed word formation, in particular the formation of ordinals, but also compounding, as in 26.

(26) een-en-twintig-ste ‘twenty-first’
       honderd-en-vier-de ‘hundred-and-fourth’
       vier-en-twintig-uurs-bijeenkomst ‘24 hour-conference’

Thus, our claim that constructional idioms may feed word formation is supported by the behavior of number names.

3.3. Particle Verbs.

Particle verbs or separable complex verbs (SCVs) are combinations of a preverb and a verb that function as complex verbs. Preverbs in Modern Dutch and German are quite similar in their behavior. Most of them derive from adpositions and adverbs. In addition, there are some nouns

---

10 More examples of copulative compound-like expressions are discussed in Olsen 2001.
and adjectives that pattern in the same way as preverbs. Preverb-verb sequences in these languages differ from prefixed verbs and verbal compounds in that the preverb is separable from the verb. Dutch and German have two different word orders, XvSOV in main clauses (where v stands for the finite verb), and SOV in embedded clauses. This difference in word order has the effect that preverbs can be stranded at the end of the main clause, as a result of finite verb movement to second constituent position of the separable verb complex. The separability of the preverbs is illustrated in 27 by means of examples from modern Dutch (taken from Booij 2002a: ch. 6).

(27) … Hans zijn moeder opbelde / Hans belde zijn moeder op
   ‘Hans phoned his mother.’
   … de fietser neerstortte / De fietser stortte neer
   ‘The cyclist fell down.’
   … Jan het huis schoonmaakte / Jan maakte het huis schoon
   ‘John cleaned the house.’
   … Rebecca pianospeelde / Rebecca speelde piano
   ‘Rebecca played the piano.’
   … dit resultaat ons teleurstelde / Dit resultaat stelde ons teleur
   ‘This result disappointed us.’

In the first example the word op ‘up’ that combines with the verb is also used as an adposition. In that case, the nonverbal element is also referred to as a particle, and the combination is referred to as a particle verb. Particle verbs form a productive class of separable complex verbs (SCVs). In the second example the word neer ‘down’ is also used as an adverb. The next two examples show that adjectives (like schoon) and nouns (like huis) can also occur in SCVs. In the last example the word teleur ‘sad’ does not occur as an independent word. The fact that SCVs are perceived as word-like units is reflected by Dutch orthography, which requires SCVs to be written as one word, without internal spacing, if the two constituents are adjacent.

The separability of SCVs also manifests itself in the position of the infinitival particle te that occurs between the two constituents of SCVs, as in op te bel len, and in the form of the perfect/passive participle, with the prefix ge- in between the particle and the verbal stem: op-ge-beld. In derivational morphology, SCVs behave similarly; for instance, the ge-
nominalization of *opbellen* is *opgebel*, with the prefix in between the particle and the verbal stem.

A number of these particles correspond to bound morphemes with an identical phonological form; these are real prefixes that cannot be separated from the verbal stem. These prefixed verbs carry main stress on the verbal stem, not on the prefix, whereas the SCVs carry main stress on the nonverbal constituent. Thus we get minimal pairs like the following:

(28) SCV prefixed verb
    dóorboor doorbóor
    ‘to go on drilling’ ‘to perforate’
    ómblaas ombláas
    ‘to blow down’ ‘to blow around’
    ónderga ondergá
    ‘to go down’ ‘to undergo’
    óverkom overkóm
    ‘to come over’ ‘to happen to’
    vôorkom voorkóm
    ‘to occur’ ‘to prevent’

As is the case for phrasal verbs in English (cf. Brinton 1988), the meaning of the preverb-verb combination (PV-V) in Dutch is often not fully predictable, and this implies that at least these combinations are lexical units of some sort. Typically, the preverbs contribute, through their specific lexical meaning, to the aspectual properties of the PV-V, in particular inherent aspect (Aktionsart) such as telicity or durationality. Thus they may also influence the syntactic valency of the verb because verbal valency is codetermined by the Aktionsart of a verb (cf. Tenny 1994). For instance, the Dutch verb *lopen* ‘to walk’ is intransitive, whereas the SCV *aflopen* can be used as a transitive verb, as in the VP *de straten aflopen* ‘to tramp the streets’. In this respect preverbs are quite similar to verbal prefixes, which also influence the aspectual and valency properties of a verb.

A second domain in which the unitary character of the PV-V combination manifests itself is that of word formation: PV-Vs can feed word formation, both compounding and derivation, as illustrated by the following examples from Dutch (with SCVs in the left column):
(29) _deverbal suffixation_

- _aanbied ‘to offer’_ → _aanbied-er ‘offerer’_,
- _aanbied-ing ‘offer’_
- _aankom ‘to arrive’_ → _aankom-st ‘arrival’_
- _aantoon ‘to prove’_ → _aantoon-baar ‘provable’_
- _aantrek ‘to attract’_ → _aantrekkelijk ‘attractive’_

_deverbal prefixation:

- _invoer ‘to introduce’_ → _her-invoer ‘to reintroduce’_
- _uitgeef ‘to publish’_ → _her-uitgeef ‘to republish’_
- _uitzend ‘to transmit’_ → _her-uitzend ‘to retransmit’_

*compounding with verbal left constituent:*

- _doorkies ‘to dial through’_ → _doorkies-nummer ‘direct number’_
- _doorkijk ‘to see through’_ → _doorkijk-bloes ‘(lit.) see-through blouse, transparent blouse’_
- _opberg ‘to store’_ → _opberg-doos ‘store box’_

As argued in Booij 2002a, b, the notion “constructional idiom” can be used to do justice to both the phrasal and the word-like properties of SCVs. The basic claim is that SCVs all have the following syntactic structure:

(30) \[ Y \ [x] V \]

By assigning a V'-node to SCVs, we represent their phrasal nature, and hence their syntactic separability. The node V' indicates a first level of projection above the V-node. It cannot be equated with the VP-node in the classic sense, since we must be able to distinguish between SCVs and VPs that contain NPs. In standard Dutch, VPs of embedded clauses cannot be raised to their matrix clauses, unlike SCVs. Note furthermore that the left constituent is a single lexical category and does not form a phrase. This correctly implies that they cannot be modified. Again, we observe here the nonprojecting behavior of lexical categories within constructional idioms.

In structure 30, the verbal position is open, and can in principle be filled by any verb. The nonverbal constituent, however, is specified. That is, there are as many different constructional idioms of this kind as there are words that can fill the left position. For instance, we will have the following constructional idioms:
They give rise to particle verbs that begin with *af*, *door*, and *op* respectively, with a fixed terminal node for the particle constituent. This has two advantages. First, the notion “particle” has no role outside the construction under discussion here, and therefore such words need not be specified independently as particles in the lexicon. Secondly, if a specific particle verb combination is no longer productive, we will not have the corresponding constructional idiom in the lexicon, but only a list of the individual existing cases of that type. Note that there are also cases where the verb only occurs in the SCV-construction, cases like *nabootsen* ‘to imitate’ and *omkukelen* ‘to fall down’. In these cases, we are dealing with lexicalized instantiations of a constructional idiom, with all terminal nodes fixed.

For each constructional idiom of this kind, its meaning will also be specified. For instance, the meaning of the constructional idiom *af-V* will be specified as ‘to finish V-ing’, and *door-V* will be specified as ‘to go on V-ing’.

In sum, the constructional idiom approach to complex predicates with preverbs can do justice to the fact that they function as periphrastic word formation. This account is also fully in line with the lexical analysis of such complex predicates in German, argued for in Ackerman and Webelhuth 1997, 1998.

This paper focuses on the effect of constructional idioms on morphology. Therefore, we want to know if the existence of particle verbs restricts the system of derivational morphology? Indeed, in Dutch the use of affixes to form derived verbs is very restricted. Dutch has only one productive verbalizing suffix, -*iseer*, which, as mentioned above, is almost exclusively attached to nonnative adjectives. The only verbalizing prefixes of Dutch are *be-*, *ver-*, and *ont- ‘de-’, and they are not very productive, unlike most of the particles. In other words, it appears that preverbs are functionally similar to verbalizing aspectual affixes, and this probably impedes the rise of a full morphological system for the expression of Aktionsart and aspectual distinctions.

This claim of functional similarity is supported by the observation that Dutch preverbs are also used in combination with nouns and adjectives that are then converted to verbs:

(31) $[[af]_P[x]_V]. \quad [[door]_P[x]_V]. \quad [[op]_P[x]_V]_V.$
4. Discussion and Conclusions.

In this article I have argued that syntactic constructions may serve as alternatives to (and hence compete with) the morphological expression of information if they have the status of constructional idioms. This was applied to the periphrastic expression of aspect, the progressive meaning, and causatives, as discussed in section 2. In section 3 I discussed three kinds of constructional idioms that also exhibit this competition with morphology, but have the additional property that they are even more like words in that they feed word formation. This applies to adjective-noun phrases, number names, and particle verbs.

The lexicon of a language can be defined as the provider of the expressions that are, at least potentially, available for concatenation in the syntax, and for word formation processes. It consists of a list of established expressions (words and idioms), and of the list of means for extending the fund of expressions through morphological operations and the use of constructional idiom templates. By locating constructional idioms in the lexicon, we provide an explanation for the restricted competition between syntax and morphology that we saw above, since blocking phenomena typically belong to the realm of the lexicon.

Word formation may then be assumed to take expressions in the lexicon as its input, not only words, but also larger units. We then have to ask ourselves why all types of constructional idioms do not feed word
formation. This may have to do with the formal structure of the constructional idiom in question. In the case of AN phrases and particle verbs, there is an unambiguous head, N and V respectively. This makes it easy for these constructions to participate in denominal and deverbal word formation respectively. Moreover, the head appears on the right, which implies that a denominal suffix is indeed preceded by a noun, and a deverbal suffix by a verb, as the suffixes require. Indeed, suffixation of a particle verb in Dutch is easier and commoner than in English, where the particle appears after the verb, and hence the suffix -er does not appear right after the verb. A noun such as put-outer (from the Harry Potter novels) is possible, but clearly more marked than similar Dutch nouns such as aambied-er ‘provider’.

As to number names, the whole expression has the lexical status of quantifier, and thus they appear in complex words in those positions where quantifiers can appear.

In the case of the analytic constructions discussed in section 2, it is hard to decide which word of the construction is the head, and hence the basis for word formation. The “verbal auxiliary” seems to be the head from the syntactic point of view, but the other constituent is the head from the point of view of lexical content. This may explain why it is impossible to use such constructions as input for word formation.

Finally, a recurrent property of constructional idioms that distinguishes them from normal syntax is the nonprojecting nature of the lexical categories involved. This underscores their special status as syntactic chunks with a lexical status.

REFERENCES


Jackendoff, Andrew McIntyre, and Silke Urban. 21–42. Berlin and New York: Mouton de Gruyter.


Vrije Universiteit Amsterdam
Faculteit der Letteren
De Boelelaan 1105
NL-1081 HV Amsterdam
[ge.booij@let.vu.nl]