

Compounds and lexical phrases in Dutch: complementarity and competition

to appear in Barbara Schlücker (ed.), *Complex lexical units: compounds and phrases*. Mannheim: Institut für Deutsche Sprache / Berlin: De Gruyter, 2018.

Geert Booij (Universiteit Leiden)

[version October 2017]

1. Introduction: morphological and phrasal lexical units

It is a generally accepted insight in linguistics that not only words, but also combinations of words (multi-word units, MWEs) may function as lexical units, and can be stored in the mental lexicon. MWEs may vary in size, from two words to a complete sentence (for instance, a proverb) (Hüning & Schlücker 2015). The existence of MWEs raises interesting questions about the organization of the grammar of natural languages, and their relationship to morphological word combinations. This is the topic of this article, with Dutch being the object language.¹

The number of MWEs in Dutch is enormous (cf. Schutz & Permentier 2016 for a recent survey). In this article I will discuss a specific subset of MWEs in Dutch, namely phrases that function as alternatives to compounds. Compound formation in Dutch serves to expand three major word classes, nouns, adjectives and verbs. They provide names for types of entities, properties, and events respectively. I will compare these types of compound with their phrasal counterparts with a similar naming function: noun phrases, adjectival phrases, and verbal phrases. As Koefoed (1993: 3) points out: “Naming is creating a link between an expression and a concept. The expression is often a word, but can also consist of more than one word”. The other function of phrases is that of description. Koefoed gives the phrase *vaderlandse geschiedenis* ‘national history’ as an example, it is the conventional name for a particular form of history, and may be contrasted with the phrase *de geschiedenis van het vaderland* ‘the history of the native country’ which is a description (Booij 2009a: 219).

In most cases, these two structural options for creating names complement each other, but there is also some competition. A comparison of these two options provides insight into the organization of grammar, the role of the lexicon, and the division of labour between morphological and syntactic devices.

The topic broached in this article may be qualified as a study of the relation between compounding and forms of ‘periphrastic word formation’. The latter term is used in Booij (2002c) as a characterization of the function of Dutch particle verbs. Traditionally, the term ‘periphrasis’ is applied to word combinations that fill cells of inflectional paradigms, for instance the cells for the perfect tense forms of Dutch verbs, combinations of an auxiliary (*hebben* ‘to have’ or *zijn* ‘to be’) and a past participle. As we will see below, phrasal word combinations can be used to fill in certain gaps in the

¹ The existence of such a wide range of MWEs also raises the psycholinguistic question which role they play in lexical processing. As far as Dutch is concerned, there are a number of psycholinguistic studies (Sprenger 2003, Sprenger, Levelt & Kempen 2006, Levelt & Meyer 2010, Nooteboom 2011) to which the reader is referred. However, this psycholinguistic dimension will be left out of consideration here.

word formation system and compete with synonymous complex words. This is the idea of complementarity between morphological and phrasal lexical units.

Investigating this relationship also makes sense from a diachronic perspective, since syntactic word combinations are the historical source of the various types of compounding that we find in Germanic languages like Dutch. Hence, it is important to understand the differences and similarities between phrasal and morphological constructions, and it may not always be easy to make this distinction due to this historical source of compounds. This demarcation problem has been pointed out by Herman Paul in chapter XIX of his *Prinzipien der Sprachgeschichte* (Paul 1880), where he argues that “[d]er Uebergang von syntaktischem Gefüge zum Kompositum ist ein so allmählicher, dass es gar keine scharfe Grenzlinie zwischen beiden gibt.” (p. 304). Paul’s observation on the blurred boundary between phrases and compounds implies that we need to investigate in more detail how we can distinguish compounds from phrases with a similar form and function. In this article, I will therefore first discuss the formal demarcation of compounds from phrases (section 2). In section 3, the naming functions of various types of compounds and their phrasal counterparts are discussed in detail. Section 4 shows how syntax plays a role besides compounding in the construction of complex numeral expressions. In section 5, it is briefly argued what these empirical findings imply for a proper theory of the organization of grammar, and why Construction Morphology (CxM) offers an insightful account of the relevant facts.

2. Demarcation of compounds and phrases

The demarcation of compounds and phrases in Dutch is based on a number of criteria: lexical integrity, orthography, phonological properties, and morphological properties. Before I discuss these criteria in detail, let me first give a number of relevant examples of compounds and their phrasal counterparts that consist of combinations of the same word classes:

(1)	<i>compound</i>	<i>phrase</i>
N+N	opoe+fiets ‘lit. grandma+bike, retro-bike’	opoe’s+fiets ‘grandma’s bike’
A+N	rood+baars ‘red bass’	rode+wijn ‘red wine’
A+A	donker+geel ‘dark-yellow’	rijk ² versierd ‘richly decorated’
N+V	raad+plegen ‘lit. advice+seek, to consult’	koffie+zetten ‘coffee make, to make coffee’
A+V	lief+kozen ‘lit. love+fondle, to caress’	schoon+maken ‘clean make, to clean’
P+V	over+komen ‘lit. over+come, to happen to’	over+komen ‘lit. over come, to come across’

The N+N and A+A phrases in (1) do not have a naming function, they are descriptive in nature. The A+N phrase *rode wijn* can be used as a name for a particular type of wine, or as a description. Yet, I discuss these phrases here because we are focusing on the formal differences between compounds and phrases, whether with a naming or a descriptive function.

Not all types of Dutch compounds have a counterpart in phrasal form; this applies to the following types:

(2)	V+N compounds	eet+kamer ‘dining room’
	N+A compounds	sneeuw+wit ‘snow-white’

² In this example, *rijk* functions as an adverb.

In these cases we cannot find phrasal counterparts because verbs cannot modify a nominal head, and nouns and verbs cannot modify adjectives in pre-adjectival position. Hence, for these types of word combinations there is no phrasal interpretation possible, and thus, the demarcation issue does not arise.

2.1. Lexical integrity

The first criterion that comes to mind for the demarcation of words and phrases is that of Lexical Integrity. The criterion of Lexical Integrity can be defined as follows: 'Syntactic rules cannot manipulate parts of words'. In other words, words are islands for syntactic operations. This narrow definition of Lexical Integrity as being restricted to syntactic operations does not exclude the possibility that the internal structure of words is accessible for other purposes, such as semantic interpretation, as should be the case (see Booij 2009b for detailed discussion of various definitions of Lexical Integrity).

The word combinations listed as phrases in (1) all allow for syntactic splits:

- (3)
- a. opoe's oude fiets 'grandma's old bike'
rode en witte wijn 'red and white wine'
 - b. rijk en kostbaar versierd 'richly and costly decorated'
 - c. Jan zet koffie 'John makes coffee'
Hij maakt de kamer schoon 'lit. He makes the room clean, He cleans the room'
Dit komt niet goed over 'lit. This comes not well over, This does not come across well'

In the cases (3a), the nominal head can be modified additionally, and hence we get a syntactic split between the first and the second word. The same applies to the adjectival head in (3b). The three verbal phrases in (3c) are all examples of so-called separable complex verbs (cf. section 3.3.). The non-verbal part is split off from the verb in main clauses (Booij 2010, Los et al. 2012). The word combinations in (1) that are classified as compounds, on the other hand, cannot be split. In the case of compound verbs this is clear from their not being split in main clauses:

- (4)
- *opoe-goede-fiets 'grandma-good-bike' / goede opoefiets 'good grandma's bike'
 - *rood-grote-baars 'red-big-bass' / grote roodbaars 'big red-bass'
 - *donker-diep-geel 'dark-deep-yellow' / diep donkergeel 'deeply dark-yellow'
 - *Jan pleegde zijn ouders raad / Jan raadpleegde zijn ouders 'John consulted his parents'
 - *Hij koost zijn vrouw vaak lief / Hij liefkoost zijn vrouw vaak 'He caresses his wife often'
 - *Dat komt mij niet weer over / Dat overkomt mij niet weer 'This will not happen to me again'

There are two cases where it seems as if parts of compounds can be split. First, Dutch features gapping of parts of words: a compound constituent can be omitted under identity with another constituent of the same prosodic form in a phrase, as in:

- (5)
- a. land- en tuinbouw 'agri- and horticulture'
 - b. voor- en achterkant 'front- and back-side'
 - c. ere- en eerste divisie 'lit. honour- and first division, premier and first league'
 - d. natuurbeheerders en -beschermers 'nature managers and -protectors'

However, as shown in Booij (1985), this kind of ellipsis is not syntactic in nature. Instead, it is a prosodic process in which one of two identical prosodic words is omitted. Both in compounds and phrases, the word constituents correspond to separate prosodic words (also referred to as ‘phonological words’). That is, this type of gapping is phonological in nature. This explains why a compound constituent like *divisie* in *eredivisie* can be omitted under identity with a separate word *divisie*, as in (5c): they are identical prosodic words, although their morpho-syntactic status is different.

The second type of split is found in phrases with coordinated relative compounds (cf. Hoeksema 2012) such as:

- (6) door- en doornat ‘through- and through-wet, very wet’
dood- en doodziek ‘dead- and dead-ill, very ill’

In relative compounds the first part functions as an intensifier. Again, this is not a case of syntactic gapping. We cannot assume underlying structures like *doornat en doornat* or *doodziek en doodziek* as the sources of the phrases in (6) since such phrases are ill-formed. Instead, what is at stake here is the repetition of an intensifier word in the left part of a compound, a case of word-internal coordination.

2.2. Orthography

A+A compounds and A+A phrases are not always that easy to distinguish. In A+A phrases, the first adjective functions as an adverb. However, Dutch adjectives can be used as adverbs without being morphologically marked as such. Hence, when we come across an A+A sequence such as *jong getrouwd* ‘lit. young married’ this word sequence can be interpreted either as a compound or as a phrase. The difference between compound and phrase is primarily a semantic one. When we spell *jonggetrouwd*, it is considered a compound with a naming, classifying function, and the meaning is ‘recently married’. When we use the phrase *jong getrouwd*, the phrase has a descriptive function ‘married at a young age’. In the latter case, we can modify the adjective *jong*:

- (7) Ze zijn nogal jong getrouwd
They are rather young married
‘They have married at a rather young age’

The orthography thus expresses a primarily semantic distinction here. Lexicalized word combinations may be felt as one word (the process of univerbation), have lost their syntactic flexibility, and are therefore spelled as one word. Thus, spelling may reflect lexicalization and univerbation.

However, orthography is not always revealing when we try to determine the status of Dutch word combinations. This is the case for separable complex words: the two parts of a separable complex verb are spelled as one word, without internal space, when they are adjacent:

- (8) Matthias was de kamer aan het schoonmaken ‘Matthias was cleaning the room’
Ik merkte dat de boodschap niet overkwam ‘I noticed that the message did not come across’

This spelling convention reflects that these word combinations are felt as lexical units, with often idiosyncratic meaning aspects. On the other hand, these separable complex verbs are not words in the morphological sense, as they cannot appear in second position in main clauses. In section 3.4. I will come back to this issue.

Dutch orthography requires compounds to be written without an internal space. However, many users of Dutch occasionally do insert a space between the two parts of a compound. This may be

partially due to the influence of English orthography in which many compounds are written with an internal space. Another factor might be that from a phonological point of view compounds are similar to phrases in that each constituent word forms a phonological word of its own. For instance, the N+N compound *tandextractie* ‘tooth-extraction’ consists of the phonological words /tand/ and /ɛkstraksi/. These two words form separate domains of syllabification. Hence, the first part *tand* is a syllable of its own. This implies that the underlying final /d/ of *tand* is in syllable-final position, and not in the onset of a syllable with the vowel /ɛ/ as its nucleus. It is therefore subject to the constraint of Dutch that obstruents are voiceless in coda position (Auslautverhärtung), and thus *tand* is pronounced as [tant], and the phonetic form of *tandextractie* is [tantɛkstraksi].

This phonological similarity between compound constituents and phrasal constituents, which both consist of more than one phonological word, may lead to uncertainty as to how spell compounds properly.

2.3. Phonological properties

Are there phonological properties that distinguish compounds from phrases? In the case of nominal compounds, main stress is in most cases on the first constituent, but there are exceptions, such as *boerenzoon* ‘farmer’s son’. In nominal phrases, on the other hand, main stress is on the head, except when contrastive stress is involved. That is, the location of stress is dependent on information structure. Thus, stress location may not always differentiate between nominal compounds and nominal phrases, but does so in pairs like *ópoefiets* (compound) versus *opoe’s fiets* (phrase). A+A compounds and A+A phrases also vary in stress location, again dependent on information structure, that is, on what counts as new and what as old information. For instance, the A+A compound *donker+geel* can be pronounced as *donker+géel* or, with emphatic or contrastive stress, as *dónker+geel*. Hence, stress location does not provide an unambiguous clue to the formal status of A+A sequences.

In verbal compounds of the type N+V and A+V, main stress is on the N and A respectively. The same applies to the corresponding separable complex verbs. Therefore, stress location cannot be used to distinguish between these compound verbs and the corresponding separable complex verbs. In verbal compounds with prepositions or adverbs as first constituents, however, main stress is on the second constituent, whereas in the corresponding separable complex verbs it is located on the non-verbal part. Thus we get a contrast between, for instance *over+kómen* ‘to happen to’ (compound) versus *óver+komen* ‘to come across’ (particle verb). Hence, stress can differentiate here between compounds and phrasal predicates. Because of this stress difference, the unstressed first constituents of these complex words may be considered prefixes, as Dutch native verbalizing prefixes such as *be-* do not carry the main stress of a complex verb either (cf. Section 3.4).

2.4. Morphological properties

Morphological properties can also be used to distinguish compounds from phrases. In present-day Dutch there is no regular case marking anymore. Hence, when morphemes such as *s*, *en*, or *e*, historically case or stem endings, appear in the middle of a word sequence, they are linking elements, as in:

- (9) *koning+s+zoon* ‘king’s son’
 her+en+huis ‘lit. gentleman’s house, mansion’

zonn+e+schijn³ ‘sun shine’

The presence of a linking element is a clear mark of compound status. The only apparent exceptions to this criterion are nouns used in the possessive construction (Booij 2010: 216-222). The N+N sequence *opoe-s fiets* ‘grandma’s bike’, for example, is a phrase: the *-s* is not a linking element here, but a marker of the possessive construction. This word sequence exhibits the normal flexibility of phrases, witness a phrase like *opoe’s zwarte fiets* ‘grandma’s black bike’. The stress pattern is also revealing, as in this word sequence the word *fiets* can carry main stress.

In the case of A+N sequences, the presence of the inflectional ending *-e* on the adjectives reveals the phrasal status of such sequences. In Dutch, prenominal adjectives have an ending *-e*, unless the noun phrase as a whole is singular indefinite, and the head noun is neuter. In the examples (10), the noun *boek* ‘book’ is neuter, and the word *vrouw* ‘woman’ has common gender:

- (10) een goed boek ‘a good book’
het goed-e boek ‘the good book’
(de) goed-e boeken ‘(the)good books’
- een mooi-e vrouw ‘a beautiful woman’
de mooi-e vrouw ‘the beautiful woman’
(de) mooi-e vrouwen ‘(the) beautiful women’

The inflection of prenominal adjectives indicates that these adjectives are words by themselves; within compounds an adjectival modifier cannot be inflected (compare the compound *snel+trein* ‘fast train, intercity train’ with *snelle trein* ‘fast train’). It is only the head of a compound that can carry inflectional markers.

There are two complications, however. The first one is that in some types of noun phrases the adjective does not carry an overt inflectional marker (Booij 2002a: 43ff, Tummers 2005). This applies to adjectives ending in *-en /ən/* (11a), where a sequence of two syllables with a schwa as vowel is avoided. It also holds for adjectives in A+N phrases that denote an individual (11b), the function of an individual (11c), or an institution (11d), where the presence of the inflectional marker *-e* is optional:

- (11) a. het open / *opene boek ‘the open book’
b. een wijs / wijze man ‘a wise man’
c. een toegepast / toegepaste taalkundige ‘an applied linguist’
d. het gemeentelijk / gemeentelijke museum ‘the municipal museum’

In these cases, the absence of the inflectional ending *-e* should not be taken as an indication of compound status. The stress pattern is that of noun phrases, with main stress on the nominal head.

The second complication is that some A+N phrases with inflected adjectives have undergone univerbation, and are now considered as one word, as reflected in the orthography:

- (12) a. jonge+mán ‘young man’
rode+kóol ‘red cabbage’
b. hóge+priester ‘high priest’

³ In *zonneshijn*, the final schwa of *zonne* ‘sun’ has disappeared in present-day Dutch, and *zon* is now the Dutch word for ‘sun’.

witte+brood ‘white bread’

The words in (12a) have final stress, like phrases, but the words in (12b) carry initial stress. The word status of these A+N sequences can be concluded from the way in which they form diminutives, in contrast to regular phrases:

- (13) een jongemannetje ‘a little boy’ *versus* een jong mannetje ‘a young little man’
een wittebroodje ‘a small white sandwich’ *versus* een wit broodje ‘a white small loaf of bread’

Diminutives are neuter nouns, and hence they require a prenominal adjective without *-e* in indefinite singular phrases of which they form the head. The examples in (13) show that both uses of the same A+N sequence are sometimes possible. In their use as words, they function as names, whereas in their phrasal use they have a descriptive interpretation.

A+N phrases frequently occur as left constituents of nominal compounds, as in

- (14) [oude+mannen]+huis ‘old men’s home’
[hete+lucht]+ballon ‘hot air balloon’
[zwarte+kousen]+kerk ‘lit. black stockings church, orthodox protestant church’

These sequences are words, and they are to be written without internal spaces: *oudemannenhuis*, *heteluchtballon*, *zwartekousenkerk*. The inflectional ending *-e* of the adjectives *oude*, *hete* and *zwarte* shows that here A+N phrases have been made parts of words. In the orthography, these compounds can be distinguished from phrases like *oude mannenhuis* ‘old house for men’ and *hete luchtballon* ‘air balloon that is hot’. The presence of a linking element *s* after the phrasal constituent confirms the compound status, as in *oude-dag-s-voorziening* ‘lit. old-day-s-provision, pension’.

In conclusion, there are a number of criteria for distinguishing between compounds and phrases. In a few cases two structural interpretations of two-word-sequences are possible, and in this case there is variation in the way language users deal with such word sequences.

3. Competition and complementarity in naming

In this section I discuss how compounds and phrases with a naming function complement each other, or are in competition. In section 3.1. I discuss the competition between A+N and N+N compounds on the one hand, and A+N phrases on the other. Section 3.2. deals with N+A compounds and phrases that express a comparison. In section 3.3 we have a look at the complementarity of N+V compounding and N+V phrases. Section 3.4. analyses the relation between particle verbs and compound verbs with a prepositional or adverbial first constituent, and section 3.5. deals with the nominalization of particle verbs by means of compounding.

3.1. Nominal compounds and A+N phrases

As pointed out by Schlücker (2014), the main, though not the only, function of A+N and N+N compounds is that of classification. These words create names for subclasses of entities. The same classifying function can be performed by A+N phrases (Booij 2002b, 2009, 2010: 183ff). Compare first N+N compounds with A+N phrases:

- (15) atoom+fysica ‘nuclear physics’ atom-aire fysica ‘nuclear physics’

structuur+analyse ‘structure analysis’	structur-ele analyse ‘structural analysis’
konings+huis ‘king-s house’	konink-lijk huis ‘royal house’
muziek+scholing ‘music(al) training’	muzik-ale scholing ‘music(al) training’
wetenschaps+beleid ‘science policy’	wetenschapp-elijk beleid ‘science policy’

In (15) we see that an N+N compound may correspond to an A+N phrase. Typically, in these phrases the adjective is a denominal adjective that belongs to the class of relational adjectives. This is a productive class of adjectives in Dutch, mainly, but not exclusively non-native in character. Both options are grammatical, and both types function as names. This may be expected for these A+N phrases since relational adjectives do not describe properties, but denote the relation between the head noun of the phrase and the base noun of the adjective. In principle both options are available, and which one is used is partially a matter of convention. For me as speaker of Dutch, *muzikale scholing* is the conventional name for this type of education, but *muziekscholing* is also found on the internet. The compound *konings-besluit* ‘king-s-decision’ is not used as an alternative for the A+N phrase *koninklijk besluit* ‘royal decision’, nor *koningsfamilie* ‘king-family’ besides *koninklijke familie* ‘royal family’, even though these N+N compounds are well-formed. The advantage of using the adjective *koninklijk* ‘royal’ instead of the compound constituent *konink* ‘king’ is that it may also be used for denoting queens.

This kind of competition between words and phrases is similar to the competition between words that is known as ‘blocking’. Blocking is the phenomenon that the formation of a complex word is blocked by the existence of another (simplex or complex) word with the same meaning. The formation of the deverbal noun *lieg-er* ‘liar’ in Dutch, for instance, is blocked by the existing complex word *leugenaar* ‘liar’. This does not mean that *lieger* is ill formed, but that it does not belong to the language convention (the norm) of the Dutch-speaking community. The fact that we find this type of competition between words and phrases as well confirms that both types of lexical units must be stored in the lexicon, and that the use of one of the relevant (morphological or syntactic) constructions for the formation of a new expression can be blocked by a stored instantiation of a competing construction. This implies that there cannot be a strict separation of morphology and syntax in the grammar of Dutch.

The second type of competition is that between A+N compounds and A+N phrases, a topic discussed in Hüning (2010), Hüning & Schlücker (2010), Schlücker (2014) and Schuster (2016). Here are some examples:

(16)	a.	<i>A+N compound</i>	<i>classifying or descriptive A+N phrase</i>
		rood+koraal ‘red coral’	rode koraal ‘red coral’
		rood+vos ‘red fox’	rode vos ‘red fox’
		*rood+wijn ‘red wine’	rode wijn ‘red wine’
	b.	<i>A+N compound</i>	<i>descriptive A+N phrase</i>
		hard+glas ‘safety glass’	hard glas ‘hard glass’
		hard+hout ‘hardwood’	hard hout ‘hard wood’
		rood+huid ‘redskin, Indian’	rode huid ‘red skin’

The compounds have initial stress on the first constituent, the phrases carry stress on the head noun, that is, final stress. The data in (16a) illustrate that both A+N compounds and A+N phrases are

possible as names, and do not necessarily block each other. A compound such as *roodwijn*, however, is odd. In some cases the compounds differ in semantic interpretation from the phrasal correlates, as shown in (16b): the compounds are names, but the corresponding phrases are used as descriptions.

A+N phrases that function as names have a restricted syntax compared to other A+N phrases (Booij 2010: 178): they cannot be modified, or split by another word. For instance, we cannot say **heel gele koorts* ‘very yellow fever’, and a phrase like *gele and hevige koorts* ‘yellow and high fever’ is also odd. When we coin the phrase *heel rode wijn* ‘very red wine’, we coerce *rode wijn* into a description, denoting wine with a very red color. This lack of syntactic flexibility of phrases with a naming function makes them more similar to compounds than other kinds of phrases.

Dutch more often opts for A+N phrases as names for entities in comparison to A+N compounds than German (Booij 2002b, Hüning 2010). There are two structural factors that play a role in this difference. First, given the rich adjectival inflection of German, A+N phrases in German have quite a number of different forms, whereas in Dutch there is only marginal variation in the shape of the adjective (usually ending in *-e*, occasionally in \emptyset). Hence, in the case of German the compound option has the advantage of reducing the form variation of the adjective, as only its stem is used (Hüning 2010). For instance, the Dutch phrase *rode wijn* ‘red wine’ and the German compound *Rotwein* both have a constant form for the adjective (*rode /rot*). This makes use of the phrasal alternative more feasible for Dutch. A second factor is that in Dutch A+N compounds the adjective has to be simplex (Schlücker 2014). This excludes the use of relational adjectives in A+N compounds. For instance, the compound *wetenscháppelijk+domein* ‘scientific domain’ is ill-formed, whereas this combination is fine as a phrase: *wetenschappelijk doméin*. This restriction also excludes the use of the various non-native relational adjectives in A+N compounds, a common pattern in German A+N compounds:

(17)	<i>Dutch phrase</i>	<i>German compound</i>	
	collectieve schuld	Kollektiv+schuld	‘collective guilt’
	nationale vlag	National+flagge	‘national flag’
	primaire literatuur	Primär+literatur	‘primary literature’
	sociale verzekering	Sozial+versicherung	‘social security’
	verbale aanval	Verbal+attaque	‘verbal attack’

This does not mean that A+N compounds with non-native adjectives are completely excluded in Dutch, but they are relatively rare, and often considered as loan translations from German (Schlücker 2014: 234). This applies to compounds such as *nationaal-socialist* ‘national-socialist’, *normaal+verdeling* ‘standard distribution’, *speciaal+zaak* ‘specialist shop’, and *spectraal+analyse* ‘spectral analysis’.

As to the choice between A+N compounds and A+N phrases, it has been argued for German that paradigmatic analogy plays an important role (Schlücker, this volume, Schlücker & Plag 2011, Rainer 2013). Schlücker & Plag (2011: 1546) argue that “the larger the compound family of an item, the more likely it is that participants choose the compound, and the larger the phrasal family of an item, the more likely it is that participants choose the phrase.” This role of paradigmatic analogy in the choice between compounds and phrases has been confirmed for Dutch by Schuster (2016) on the basis of an investigation of Dutch dictionaries and corpora.

The role of paradigmatic analogy can be observed in the use of colour adjectives. For example, Dutch colour adjectives such as *geel* ‘yellow’, *rood* ‘red’, and *zwart* ‘black’ are used in A+N compounds that function as names for animals and for human beings (in some cases with a possessive interpretation):

- (18) geel+bek ‘lit. yellow+mouth, fledgling’
 geel+gors ‘yellow hammer (type of bird)’
 geel+vink ‘serin finch’

rood+forel ‘red trout’
 rood+baard ‘lit. red+beard, person with red beard’
 rood+staart ‘lit. red+tail, redstart (bird with red tail)’

zwart+hemd ‘lit. black+shirt, fascist’
 zwart+kop ‘black-cap (type of bird)’
 zwart+rok ‘lit. black+coat, person wearing a blackcoat’

On the other hand, we find these colour adjectives in phrasal names such as *gele kaart* ‘yellow card’ and *rode kaart* ‘red card’, names for the cards used for indicating improper actions in a football match (a *kaart*-family). Likewise, there is a family of phrasal names with *zwart* ‘black’, as in *zwarte markt* ‘black market, *zwart geld* ‘black money’, *zwarte doos* ‘black box’, and *zwarte kunst* ‘black magic’, a *zwart*-family with *zwart* being used with the meaning ‘illegal, opaque’. These observations confirm that analogy to similar compounds or phrases, plays an important role in the choice between compound and phrase.

3.2. N+A compounds and adjectival phrases

Dutch N+A compounds can be used as an alternative to phrases that express a comparison (Hoeksema 2012: 7):

(19)	<i>compound</i>	<i>adjective phrase</i>	<i>gloss</i>
	dons+zacht	(zo) zacht als dons	‘soft as down’
	honds+trouw	(zo) trouw als een hond	‘faithful as a dog’
	ijs+koud	(zo) koud als ijs	‘cold as ice’
	kaars+recht	(zo) recht als een kaars	‘straight as a candle’
	sneeuw+wit	(zo) wit als sneeuw	‘white as snow’

According to Hoeksema (2012) the choice of the compound structure over the phrasal alternative is determined by two advantages of the compound option: compactness and expressiveness. There is always a phrasal alternative for the compound, but not vice versa. For instance, the comparison *sterk als een paard* ‘strong as a horse’ cannot be expressed by the compound *paardesterk*. The phrasal alternative might, however, not carry exactly the same meaning: *ijzersterk* ‘iron-strong’ can be used in contexts where the phrasal expression is odd. For instance, *een ijzersterk verhaal* ‘a very strong story’ cannot be properly paraphrased as *een verhaal sterk als ijzer* ‘a story strong as iron’ (Hoeksema 2012). Similar observations have been made for German (Schlücker, this volume), and Italian (Masini, this volume). The same applies to compounds with *reuze* (an allomorph of *reus* ‘giant’), as in *reuze-groot* ‘giant-big, very big’ where the phrase *zo groot als een reus* ‘as big as a giant’ may not be a proper paraphrase. In these compounds the nouns *ijzer* and *reuze* have acquired a more general meaning of intensification. These compounds are called elative compounds and express that the property denoted by the head is present to a high degree. This elative use is the source of the development of these nouns into intensifier affixoids. For instance, besides *bloed+rood* ‘red as blood’ we find compounds

like *bloed+saai* ‘lit. blood-boring, very boring’ and *bloed+mooi* ‘lit. blood-beautiful, very beautiful’, which cannot be paraphrased as *saai / mooi als bloed* ‘boring / beautiful as blood’.

This difference between compounds and phrases can also be observed for another class of N+A compounds of the type *dood+ziek* ‘lit. dead-ill, so ill that it may cause death’. Again, some of these nominal modifiers have acquired a more general meaning of intensification, and in such cases a phrasal paraphrase is not adequate:

- (20) *dood+gewoon* ‘very ordinary’
dood+simpel ‘very simple’

This development of nominal (and other) modifiers into affixoids, that is, words with a more abstract meaning of intensification when embedded in compounds, is discussed in detail in Booij & Hüning (2014) and Hüning & Booij (2014).

3.3. N+V compounds and phrases

Unlike nominal and adjectival compounding, the formation of verbal compounds is not a productive process in Dutch. This does not mean that there are no verbal compounds whatsoever. The main source of such compounds is backformation from nominal compounds with the form $[[N][V-er]_N]_N$ or $[[N][V-ing]_N]_N$. Examples are:

- (21) *beeld+houden* ‘lit. to image-hew, to sculpture’ < *beeld+houw+er* ‘sculptor’
honger+staken ‘lit. to hunger-strike, to go on hungerstrike’ < *honger+stak+ing* ‘hungerstrike’
vaat+wassen ‘to dish-wash’ < *vaat+wass+er* ‘dish washer’
tekst+verwerken ‘to text-process’ < *tekst+verwerk+ing* ‘text processing’

A second type of verbal compounds are verbs like *klapper+tanden* ‘lit. chatter-tooth.INF, to have chattering teeth’ and *kwispel+staarten* ‘lit. wag-tail.INF, to wag one’s tail’. They have the structure $[VN]_V$, and are exceptional in that they are left-headed. There are also a few VV compounds like *hoeste+proesten* ‘lit. to cough-sneeze, to cough and sneeze’, but again, this is not a productive process of word formation (Booij 2002a: 164-65).

The productive alternative for N+V compounds are phrasal word sequences that consist of a bare noun and a verb. An example is the N+V sequence *piano+spelen* ‘to play the piano’. This word sequence can be used as a verb phrase, but the noun can also be quasi-incorporated into the verb:

- (22) ... dat Julian {piano kan spelen / kan pianospelen}
... that Julian {piano can play / can play piano}
‘... that Julian can play the piano’

Verb phrases with a bare noun are often used as names for denoting a certain kind of activity. For instance, *piano spelen* is a specific type of musical activity. The word *piano* does not denote a specific referent here. This may be contrasted with a verb phrase like *de piano bespelen* ‘to play on the piano’, where, by using a definite noun phrase, the identifiability of a specific referent of *piano* is presupposed. When count nouns are used as bare nouns, without the normally expected determiner, this evokes an interpretation as name instead of description of the verbal phrase in which that bare noun is used. Note that in a compound like *pianospeler* ‘piano player’ the word *piano* likewise has no referential power.

In the second variant in (22), the noun and the verb form a syntactically closer unit than in the first variant, and are adjacent. This unit can be qualified as a case of quasi-noun incorporation. Noun incorporation is the process in which a noun is incorporated into a verb, and thus creates a verbal compound. However, in Dutch the incorporation process does not lead to compounds in the morphological sense. This is shown by the fact that the N+V sequence cannot appear in the position for finite verbs (the second position) in main clauses, unlike a real verbal compound like *beeldhouwen* ‘to sculpture’:

- (23) Julian { *pianospeelt graag / speelt graag piano } ‘Julian likes playing the piano’
 Amber beeldhouwt graag ‘Amber likes sculpturing’

This is why Dahl (2004) calls this process quasi-incorporation: there is incorporation and formation of lexical units, but these lexical units are not words. Quasi-noun incorporation in Dutch is discussed in detail in Booij (2010: Chapter 4), and the account below is mainly based on this chapter.

The strong bond between N and V in the incorporated variant can also be seen in two syntactic constructions, the verb raising construction and the progressive construction. In the verb raising construction the verb of the main clause forms a unit with the verb of the embedded clause. The incorporated noun can appear in between the two verbs (24a), whereas this is impossible for a full noun phrase (24b). The first option in (24a) is that with quasi-incorporation, and Dutch orthography requires the quasi-incorporated word combination to be spelled as one word, without an internal space:

- (24) a. ... dat Barbara { wil pianospelen / piano wil spelen }
 ... that Barbara { wants pianoplay / piano wants play }
 ‘...that Barbara wants to play the piano’
- b. ... dat Barbara { *wil de piano bespelen / de piano wil bespelen }
 ... that Barbara { wants the piano play / the piano wants play }
 ‘... that Barbara wants to play on the piano’

The second construction that functions as a litmus test for quasi-noun incorporation is the progressive construction of the form *aan het V-infinitive*:

- (25) Matthias is aan het lezen
 Matthias is at the read.INF
 ‘Matthias is reading’
- Matthias is { aan het pianospelen / piano aan het spelen }
 Matthias is at the piano-play.INF / piano at the play.INF
 ‘Matthias is playing the piano’
- Matthias is { de piano aan het bespelen / *aan het de piano bespelen }
 Matthias is { the piano at the PREF.play.INF / at the piano PREF.play.INF }
 ‘Matthias is playing on the piano’

Verbs with an incorporated noun can function as a unit in the progressive construction, and thus appear after *aan het*. This applies to the N+V sequence *piano+spelen*. On the other hand, the prefixed verb *bespelen* ‘to play on’ is an obligatorily transitive verb that does not allow for noun incorporation. Like verbal phrases with bare nouns, the quasi-incorporation structure is used to express

that the action referred to is a conventional action. In other words, it creates names for types of action. Whatever is considered as a conventional action by the language user can be expressed in this form. For instance, *auto+wassen* ‘to wash cars’ is a conventional action, whereas buying a car is not conceived as a conventional action, and therefore there is no verb phrase *auto kopen*, or quasi-compound *autokopen* (instead, the proper phrase for naming this action is *een auto kopen*, with an indefinite determiner). Hence the difference in syntactic behaviour between *auto+wassen* en *auto+kopen*:

- (26) ... dat Peter gaat {auto+wassen /*auto+kopen}
 ... that Peter goes {car+wash.NF / car+buy.INF}
 ‘... that Peter is going to {car+wash / *car+buy}’

Conventional actions can also be expressed with verbs + plural nouns. For instance, *aardappels schillen* ‘lit. potatoes-peel, to peel potatoes’ can be conceived as a conventional action, and hence we can say:

- (27) Geert is aan het aardappels schillen ‘Geert is peeling the potatoes’
 ... dat Geert wil aardappels schillen ‘that Geert wants to peel potatoes’

However, when the noun is plural, the N+V sequence is not spelled as one word.

The use of the term ‘quasi-incorporation’ may suggest that these quasi-compounds always derive from a regular phrase, but this is not the case. There are many N+V sequences where the bare noun cannot be interpreted as an object-NP. This applies to, for instance, the following cases (Booij 2010: 112):

- (28) buik+spreken ‘lit. to stomach speak, ventriloquizing’
 koord+dansen ‘lit. to rope dance, walking a tightrope’
 mast+klimmen ‘lit. to pole climb, climbing the greasy pole’
 steen+grillen ‘lit. to stone grill, stone-grilling’
 stijl+dansen ‘lit. to style dance, ballroom-dancing’
 vinger+verven ‘to finger paint’
 zak+lopen ‘lit. to bag walk, running a sack-race’
 zee+zeilen ‘lit. to see sail, ocean-sailing’

These quasi-compounds are referred to as immobile verbs in the linguistic literature (cf. Vikner 2005), because they cannot appear in second position, as illustrated here for *zee-zeilen* (29a). At the same time, they cannot be split (29b), but are fine if they are not split (29c,d):

- (29) a. *Mijn vader zee+zeilt vaak
 My father sea+sails often
 ‘My father often sails at sea’
 b. *Mijn vader zeilt vaak zee
 My father sails often see
 ‘My father often sails at sea’
 c. Mijn vader is vaak aan het zee+zeilen
 My father is often at the sea+sail.INF
 ‘My father often sails at sea’
 d. dat mijn vader vaak zee+zeilt

that my father often sea+sails
 ‘that my father often sails at sea’

The conclusion drawn from these facts in Booij (2010: Chapter 4) is that there are N+V combinations that are neither regular compounds nor regular syntactic phrases. Instead, they are quasi-compounds without a corresponding verbal phrase: a word sequence such as *zee zeilen* cannot be used as a well-formed phrase.

For a proper account of the distribution of quasi-compounds, their structure should be different from that of phrases and that of morphological compounds. They may be considered syntactic compounds. In a syntactic verbal compound a bare N⁰ is adjoined to a V⁰, and together they form a V⁰:

(30) [[zee]_{N0} [zeil]_{V0}]_{V0}

Their syntactic compound status prohibits them from being split in main clauses (29a). At the same time they cannot appear in second position in main clauses as this position allows only for a single verb (29b). When the bare noun functions as an object, as in *pianospelen*, the quasi-compound corresponds with a verbal phrase with a bare noun that can be split. Hence, the two possible word orders in sentences like (22). Thus, the grammar of Dutch provides three different structures for N+V combinations that function as names:

(31)	morphological compound:	[[honger] _N [staak] _V] _V	[[vaart] _N [was] _V] _V
	syntactic compound:	[[piano] _{N0} [speel] _{V0}] _{V0}	[[zee] _{N0} [zeil] _{V0}] _{V0}
	verb phrase:	[[[piano] _{N0}] _{NP} [speel] _{V0}] _{VP}	

Since quasi-compounds cannot be used as finite verbal forms in main clauses, the usual strategy is to use the progressive *aan het V-infinitive*-construction as an alternative, as illustrated by the sentences in (25).

This type of quasi-compound structure is also possible for A+V combinations:

(32) dood+vriezen ‘lit. dead+freeze, freeze to death’
 goed+keuren ‘lit. good+judge, to approve’
 schoon+maken ‘lit. clean+make, to clean’
 vreemd+gaan ‘lit. strange+go, to sleep around’
 vrij+geven ‘lit. free+give, to release’
 wit+wassen ‘lit. white+wash, money-laundering’
 zoek+maken ‘lit. missing+make, to mislay’

These A+V combinations are not words in the morphological sense, and are therefore split in main clauses, just like the N+V combinations. They exhibit the same word order variation as that shown in (22):

(33) ... dat de directeur het voorstel { wilde goedkeuren / goed wilde keuren }
 ... that the director the proposal { wanted good-judge / good wanted judge }
 ‘... that the director wanted to approve the proposal’
 ... dat Ton het boek { heeft zoekgemaakt / zoek heeft gemaakt }
 ... that Ton the book { has missing-made / missing has made }
 ‘... that Ton has mislaid the book’

In other words, what we see here are A+V combinations, often idiosyncratic in meaning, that are structurally interpreted either as verbal phrases with a bare adjective as complement, or as quasi-compounds.

Both types of compounds have past participles in which the participial prefix *ge-* appears before the verbal stem, which confirms their phrasal status:

- (34) Jan heeft piano+gespeeld ‘John has played the piano’
Wij hebben dit voorstel goed+gekeurd ‘We have approved this proposal’

The adjectives of the quasi-compounds cannot be modified, that is, they cannot head an adjectival phrase. When we add a modifier, this leads to an ungrammatical result, or another, more literal interpretation. For instance, the verb phrase *heel vreemd gaan* ‘lit. very strange go, to go very strange’, with the degree adverb *heel*, cannot be interpreted as ‘sleep around intensively’.

In conclusion, the lack of productivity of verbal compounding in Dutch is compensated by the availability of (i) verbal phrases with a bare noun or adjective as complement such as *piano spelen* and *goed keuren*, and (ii) by quasi-compounds with a verbal head and a nominal or adjectival adjunct (spelled without an internal space) such as *pianospelen*, *goedkeuren*, and *zeezeilen*. They function as names for conventional, nameworthy activities. The class of quasi-compounds is larger than that of the verbal phrases with bare complements, because in quasi-compounds the noun need not be licensed syntactically by the verb. For instance, in *zeezeilen*, the noun *zee* does not function as an object-NP, and hence its occurrence is not licensed by syntax. Nevertheless, it can combine with a verb into a syntactic compound.

3.4. Prefixed verbs and particle verbs

Dutch has a number of complex verbs which might be considered compounds because they consist of a preposition or an adverb followed by a verbal stem:

- (35) a. aan+bidden ‘lit. at+pray, worship’
 achter+halen ‘lit. behind+fetch, recover’
 voor+komen ‘lit. for+come, prevent’
- b. door+zoeken ‘lit. through+search, search through’
 om+geven ‘lit. around+give, surround’
 onder+schatten ‘under+estimate’
 over+spoelen ‘lit. over+wash, wash over’
- c. mis+lukken ‘lit. wrong+succeed, fail’
 weer+houden ‘lit. back+hold, restrain’
 vol+brenge ‘lit. full+bring, to finish’

The types of verb with *aan-*, *achter-* and *voor-* exemplified in (35a) are unproductive, just like those with the adverbs *mis-* and *weer-* and the adjective *vol-* shown in (35c). The types exemplified in (35b) with *door-*, *om-*, *onder-*, and *over-*, however, are productive. In reference grammars of Dutch they are usually considered prefixed words, because unlike what is normally the case for Dutch compounds, the main stress in these words is located on the second constituent (instead of the first constituent). Thus, from the point of view of stress location, they pattern with prefixed verbs such as *be-hálen* ‘to

acquire' and *ver-zóeken* 'to request'. Moreover, the meaning contribution of these morphemes in complex verbs may differ from that of the corresponding morphemes when used as words by themselves. In other words, these words have grammaticalized into prefix-like morphemes. Prefixes like *be-* and *ver-* also originate from words that are parts of compounds, but their phonological form has been reduced as well, with a reduced vowel /ə/. Hence, in present-day Dutch there are no identical lexical counterparts for these prefixes.

The number of productive processes of verbalizing prefixation in Dutch is quite restricted, and therefore, there is a huge range of meanings for the expression of which phrasal verbal predicates with a corresponding make up can be used. This is the class of particle verbs, with the particles corresponding to prepositions like *binnen* 'inside', postpositions like *mee* 'with, and adverbs like *neer* 'down'. The number of types is quite big, and I list here only a few for the purpose of illustration. Complete lists can be found in De Haas & Trommelen (1993), and on Taalportaal (www.taalportaal.org):

- (36) *binnen+komen* 'lit inside come, enter'
mee+vallen 'lit. with fall, turn out better than expected'
op+bellen 'lit. up phone, to phone up'
rond+lopen 'lit. around walk, walk around'
neer+vallen 'lit. down fall, to fall down'
weg+lopen 'lit. away walk, walk away'
voorop+lopen 'lit. in front walk, walk in front'

Particle verbs are lexical units, but phrasal in nature, just like verbal predicates such as *piano+spelen* and *schoon+maken* discussed in section 3.3. They are split in main clauses, and can function as verbal phrases. At the same time, they can also be used as quasi-compounds, that is, behave like a tight syntactic unit in verb raising constructions. In this latter use, they are spelled as one word. These two syntactic options are illustrated by the following sentences:

- (37) ... dat Hans zijn moeder {op wilde bellen / wilde opbellen}
 that Hans his mother {up wanted phone / wanted up-phone}
 '... that Hans wanted to call his mother'

Morphologically, particle verbs also behave as phrases, since the prefix *ge-* of the past participle appears in between the particle and the verb: *op-ge-beld*, not **ge-op-beld*. When we nominalize a particle verb by means of the prefix *ge-*, it also appears before the verbal stem, as in *op-ge-bel* 'calling'.

The proper grammatical analysis of Dutch particle verbs is discussed in detail in Booij (2010: Chapter 5), and in Los et al. (2012). The gist of this analysis is that each class of particle verbs has to be represented in the grammar of Dutch as a constructional idiom. Constructional schemas are schemas that specify the systematic correspondence between form and meaning of a construction. A constructional idiom is a constructional schema in which one or more slots are lexically fixed. Each type of particle verb will be represented by a constructional idiom with that particle specified. The meaning of the particle sometimes corresponds with that word used in isolation, but in other cases it has acquired a specific meaning. For instance, the particle *door* 'through' has acquired, among others, the aspectual meaning of 'to continue with', as in *door+fietsen* 'to continue cycling' and *door+eten* 'to continue eating', unlike the preposition *door* 'through'. Hence, I assume the following constructional idioms for *door+V*, one without, and one with quasi-incorporation. In the first case we have a phrasal verbal predicate, labeled as V', in the second case a syntactic compound:

parallelism, we should analyze *aankomst* as the compound $[[aan]_{\text{part}} [kom-st]_{\text{N}}]_{\text{N}}$. Because *komst* is listed as derived word, it can combine with a particle into a compound. This implies that we are confronted with an asymmetry between meaning and form, since the nominalizing suffix *-st* has semantic scope over the particle verb *aankom* (the stem of *aankomen* ‘to arrive’) as a whole. This systematic choice of an unproductive type of nominalization by particle verbs is shown in (41) (data from Booij 2015):

(41)	<i>verbal stem</i>	<i>nominalization</i>
	a. <i>no formal change (conversion)</i>	
	val ‘fall’	val ‘to fall’
	aan+val ‘attack’	aanval ‘attack’
	in+val ‘raid’	inval ‘raid’
	b. <i>with vowel change</i>	
	grijp ‘seize’	greep ‘grip’
	in+grijp ‘interfere’	ingreep ‘interference’
	mis+grijp ‘miss one’s hold’	misgreep ‘blunder’
	c. <i>stem change and/or suffixation</i>	
	gaan ‘go’	gang ‘going’
	af+gaan ‘fail’	afgang ‘failure’
	door+gaan ‘continue’	doorgang ‘taking place’
	neer+gaan ‘go down’	neergang ‘going down’
	op+gaan ‘rise’	opgang ‘rise’
	in+gaan ‘enter’	ingang ‘entrance’
	geef ‘give’	gave / gifte ‘gift’
	aan+geef ‘report’	aangifte ‘report’
	op+geef ‘state’	opgave ‘statement’
	uit+geef ‘spend’	uitgave ‘expense’
	kom ‘come’	kom-st ‘arrival’
	aan+kom ‘arrive’	aankom-st ‘arrival’
	op+kom ‘rise’	opkom-st ‘rise’

This observation concerning the selection of a particular unproductive type of nominalization for the particle verb is accounted for straightforwardly by an analysis in which nominalizations of particle verbs are compounds that consist of a particle plus the nominalized form of the simplex verb. Hence, the form part of the general construction schema for these particle verb nominalizations is:

(42) [Particle $[[x]_{\text{V}} z]_{\text{N}}]_{\text{N}}$

where $[[x]_{\text{V}} z]_{\text{N}}$ stands for the nominalized form of the simplex verb. The variable *x* stands for (an allomorph of) the verbal stem, and the variable *z* stands for a suffix or zero. All instantiations of unproductive types of nominalization have of course to be listed. Hence, listed nouns like *gang* and *komst* will be available for combining with a particle into a compound. Thus, it is predicted that the nominalized form of a particle verb corresponds to that of the nominalized form of the corresponding simplex verb.

The structure for compounds of the form (42) has to be available anyway in the grammar of Dutch, as there are a number of compounds of this form without a corresponding particle verb. This applies to, for instance, the following nouns:

(43)	<i>compound word</i>	<i>lacking particle verb</i>
	af+dronk ‘after-taste’	afdrinken
	bij+slag ‘bonus’	bijslaan
	toe+gang ‘access’	toegaan

The meaning of particle compounds has to be specified as being the nominalization of the corresponding particle verb, if available, which often has an idiosyncratic meaning. This is expressed by the following set of paradigmatically related constructional schemas:

(44)	form:	[Particle _i [[x] _{vj} z] _N] _N	≈	[Particle _i V _j] _{v_k}
	meaning:	Event of SEM _k		SEM _k

Recall that the symbol \approx denotes a paradigmatic relationship. The formal and semantic correspondences between the two schemas are specified by means of co-indexation. Such a schema of schemas is called a second order schema. For instance, given the particle verb *aankomen* with the meaning ‘to arrive’, second order schema (44) states that the compound noun *aankomst* is interpreted as the event of arriving.

This case shows that there might be an asymmetry between form and meaning in morphological constructions. The meaning of the particle compound is a compositional function of the meaning of the particle verb, even though the particle verb is not a formal subconstituent of the corresponding compound. Instead, there is a paradigmatic relationship between the particle compound schema and the schema for particle verbs. This kind of asymmetry can be accounted for by relating schemas paradigmatically in second order schemas (Booij & Masini 2015). Schema (44) is a second order schema, as it relates the constructional schema for particle compounds to the constructional schema for particle verbs.

This implies that the grammar of Dutch requires access to the meaning of phrasal lexical expressions in order to account for the meaning of particle compounds. This is another type of complementarity between compounds and phrasal lexical items, and shows again that we need a grammar in which morphological and phrasal lexical units can interact.

4. The construction of numeral expressions

Compounding and phrasal expressions are used in tandem in the construction of complex numeral expressions in Dutch (Booij 2010: Chapter 8). The use of compound structure can be observed in cardinal numbers like the following:

(45)	drie+honderd ‘three-hundred’
	vijf+duizend ‘five-thousand’

In the compounds *driehonderd* and *vijfduizend* there is a relation of multiplication between the first and the second constituent, the first constituent is the multiplier.⁴ These numerals are spelled as one word.

Phrasal structure is used in the form of coordination by means of the conjunction *en* ‘and’, as in:

⁴The word sequence *zes miljoen* ‘six million’ looks similar to these compounds, but is considered a phrase, as reflected by its spelling with an internal space. This means that *miljoen* is interpreted as a measure noun, similar to nouns like *gulden* ‘guilder’ and *uur* ‘hour’ which also appear in their singular form after a cardinal >1: *drie gulden*, *drie uur*. However, this interpretation is not chosen for words with *honderd* and *duizend*: *honderd*, *duizend*, en *miljoen* can all function as nouns, and may appear in plural form: *honderden*, *duizenden*, *miljoenen*.

- (46) drie+en+zestig ‘three and sixty, 63’ spelling: drieënzestig
 honderd+(en)+drie ‘hundred and three, 103’ spelling: honderd(en)drie

In (46) we see the use of syntactic coordination by means of *en*. This corresponds with the semantic effect of addition. This phrasal pattern is subject to a specific restriction, however, that does not apply to syntactic coordination in general: there is a fixed order in which the two numbers have to appear, the lower digit before the higher digit in numbers < 100, the higher digit before the lower one in numbers > 100. You cannot say *zestig-en drie* ‘63’ or *drie-en-honderd* ‘103’. Moreover, the conjunction *en* is optional in numbers > 100, an optionality that does not apply to regular coordination. In other words, phrasal coordination is used here for the expression of addition, but is subject to specific restrictions. Additional construction-specific properties for this use of coordination are that the conjunction *en* /en/ can be pronounced either as [ɛn] or as [əɛn] in numbers < 100, and can be optionally omitted in numbers > 100.

Compounding and phrasal coordination are used together in the formation of complex numerals: the numeral compounds are building blocks of the coordination construction, as in:

- (47) acht+honderd(en)drie+en+twintig ‘eight hundred three and twenty, 823’

with the structure:

- (48) [[[acht]_{Num} [honderd]_{Num}]_{Num} ([en]_{Conj}) [[drie]_{Num} [en]_{Conj} [twintig]_{Num}]_{NumP}]_{NumP}

where Num= Numeral, and NumP = Numeral Phrase.

The orthography of numerals reflects their hybrid nature. The compounds and the coordinated numerals are spelled as one word, except that there is a space after *duizend*. Moreover, the words *miljoen* and *miljard* are always spelled as separate words. Thus we get spellings like *achthonderd* (800), *drieëntwintig* (23), *achthonderdendrieëntwintig* (823), *tweeduizend drieënveertig* (2,043), and *vijf miljoen achthonderdduizend driehonderdentwintig* (5,800,320).

These numeral phrases seem to feed word formation in the construction of ordinals, as in:

- (49) acht+honderd(en)drie+en+twintig-ste ‘823th’

The spelling of this ordinal is *achthonderd(en)drieëntwintigste*. The ordinal suffix *-ste* is attached to the last word of this complex expression, but its semantic scope includes the part *achthonderd* as well. Hence, we see another asymmetry here between the formal structure and the semantic interpretation of complex expressions.

5. Construction Grammar and Construction Morphology

The data discussed in sections 3 and 4 provide strong evidence for a view of the organization of the grammar in which there is no strict separation between morphology and syntax. This is one of the core hypotheses of constructionist approaches to morphology and syntax. Here are the main points:

(i) Morphological and syntactic constructions may compete; both can be used for creating names, and hence, there are blocking effects between morphological and phrasal constructs.

(ii) Phrasal constructions may be subject to specific restrictions when used as names. For instance, in A+N phrasal names, the adjective cannot be separated from the head noun, nor be modified. In a constructionist

approach we can account for the properties of such phrasal names by phrasal constructional schemas which derive from general syntactic schemas, but with specific formal and semantic properties specified. The same applies to the description of specific forms of coordination in the construction of complex numeral expressions.

(iii) Morphological processes may be unproductive, or unavailable for the expression of certain types of names. In Dutch, phrasal structures fill those gaps, hence we may speak of periphrastic word formation. This is the case for separable complex verbs of various types: N+V, A+V and particle verbs. There is a clear complementarity between morphological and syntactic ways of creating names.

(iv) The interpretation of complex words may depend on the meaning of paradigmatically related phrasal lexical constructions. This is the case for nominalizations of particle verbs. Paradigmatic relationships between constructional schemas, morphological or phrasal, can be expressed by second order schemas.

These kinds of finding form underpinnings of the model of Construction Morphology proposed in Booij (2010), and further articulated in a number of publications on Dutch referred to in this article. In Construction Grammar (Hoffmann & Trousdale (eds., 2013)) and Construction Morphology, the grammar is seen as a multidimensional web of syntactic and morphological constructions of various degrees of abstractness. Constructional schemas form a hierarchy: more abstract schemas dominate more concrete ones, and constructions are instantiated by fully lexically specified constructions, which may be listed in the lexicon. For example, there are, in increasing order of concreteness, a general schema for Dutch right-headed compounds, a subschema for N+A compounds, a constructional idiom $[[dood]_N A]_A$ 'very A', and listed instantiations of this constructional idiom such as *doodziek* 'very ill' and *doodnormaal* 'very normal'. Syntactic constructions are also specified in terms of schemas. Phrasal names of the type A+N are specified by a subschema of the general schema for Noun Phrases, with certain restrictions imposed, such as linear adjacency of A and N and bareness of the adjective. Similarly, the grammar of Dutch contains a general syntactic schema for syntactic coordination, which dominates specific subschemas for numeral expressions in which the properties mentioned in section 5 are specified. Thus, the idea of periphrastic word formation finds its natural expression in Construction Grammar.

Since in Construction Grammar both morphological schemas and syntactic schemas, and their lexicalized instantiations are listed, there is potentially a competition between morphological and syntactic expression of the same meaning. This predicts the observed blocking effects.

Paradigmatic relations between schemas and between concrete constructions are expressed by means of co-indexation. They give expression to the existence of word families and phrase families. The presence of a network of paradigmatic relations in the grammar provides a natural interpretation for the observation that paradigmatic analogy co-determines the choice between compound and phrase when coning a name.

The claim that morphology and syntax cannot be separated in grammar does not mean that there is no formal distinction between morphological and phrasal constructions. This formal distinction is necessary for a proper account of the syntactic behaviour of the various types of names. At the same time, since compound schemas and phrasal schemas are not split in different components of the grammar, they can interact: phrasal constituents may form parts of compounds and vice versa, and compounds may function as nominalizations of particle verbs which themselves are phrasal expressions. These observations led to the conclusion that second order schemas (paradigmatic relations between constructional schemas) form part of the grammar.

Since morphology often derives historically from syntax, it should not come as a surprise that there are transitional cases such as quasi-compounds, verbs with incorporated particles, and cardinal numerals of the type *drieëntwintig* '23' where the conjunction *en* can also be interpreted as a linking

element. These phenomena underscore Hermann Paul's remarks on the blurred boundary between syntax and word formation quoted in the introduction of this article. As we saw above, a Construction Grammar approach can do justice to these transitional cases.

References

- Booij, Geert (1985): Coordination reduction in complex words: a case for prosodic phonology. In: Van der Hulst, Harry / Smith, Norval (eds.): *Advances in non-linear phonology*. Dordrecht: Foris. 143-160.
- Booij, Geert (2002a): *The morphology of Dutch*. Oxford: Oxford University Press.
- Booij, Geert (2002b): Constructional idioms, morphology, and the Dutch lexicon. In: *Journal of Germanic Linguistics* 14. 301-329.
- Booij, Geert (2002c): Separable complex verbs in Dutch: a case of periphrastic word formation. In: Dehé, Nicole / Jackendoff, Ray / Macintyre, Andrew / Urban, Silke (eds.): *Verb-particle explorations*. Berlin: Mouton de Gruyter. 21-42.
- Booij, Geert (2009a): Phrasal names: A constructionist analysis. *Word Structure* 2. 219-240.
- Booij, Geert (2009b): Lexical integrity as a morphological universal, a constructionist view. In: Scalise, Sergio / Magni, Elisabetha / Bisetto, Antonietta (eds.): *Universals of Language Today*. Berlin: Springer. 83-100.
- Booij, Geert (2010): *Construction Morphology*. Oxford: Oxford University Press.
- Booij, Geert (2015): The nominalization of Dutch particle verbs: schema unification and second order schemas. In: *Nederlandse Taalkunde* 20. 285-314.
- Booij, Geert / Hüning, Matthias (2014): Affixoids and constructional idioms. In: Boogaart, Ronny / Coleman, Timothy / Rutten, Gijsbert (eds.): *Extending the scope of Construction Grammar*. Berlin: Mouton De Gruyter. 77-105.
- Booij, Geert / Masini, Francesca (2015): The role of second order schemas in word formation. In: Bauer, Laurie / Körtvélyessy, Livia / Štekauer, Pavol (eds.): *Semantics of complex words*. Cham etc.: Springer. 47-66.
- Dahl, Östen (2004): *The growth and maintenance of linguistic complexity*. Amsterdam / Philadelphia: Benjamins.
- De Haas, Wim / Trommelen, Mieke (1993): *Morfologisch handboek van het Nederlands*. Den Haag: SDU Uitgeverij.
- Hoeksema, Jack (2012): Elative compounds in Dutch: properties and developments. In: Oebel, Guido (ed.), *Intensivierungskonzepte bei Adjektiven und Adverbien im Sprachenvergleich / Crosslinguistic comparison of intensified adjectives and adverbs*. Hamburg: Verlag dr. Kovač. 97-142.
- Hoffman, Thom / Trousdale, Graeme (eds., 2013): *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.
- Hüning, Matthias (2010): Adjective + noun constructions between syntax and word formation in Dutch and German. In: Onysko, Alexander / Michel, Sascha (eds.): *Cognitive perspectives on word formation*. Berlin / New York: Mouton de Gruyter. 195-215.
- Hüning, Matthias / Schlücker, Barbara (2010): Konvergenz und Divergenz in der Wortbildung. Komposition im Niederländischen und im Deutschen. In: Dammel, Antje / Kürschner, Sebastian / Nübling, Damaris (Hgg.): *Konvergenz und Divergenz in der Wortbildung*. (= Germanistische Linguistik 206–209). Hildesheim / Zürich / New York: Georg Olms. 783–825.

- Hüning, Matthias / Booij, Geert (2014): From compounding to derivation. The emergence of derivational affixes through 'constructionalization'. In: *Folia Linguistica* 48. 579-604.
- Hüning, Matthias / Schlücker, Barbara (2015): Multi-word expressions. In: Müller, Peter O. / Ohnheiser, Ingeborg / Olsen, Susan / Rainer, Franz (eds.): *Word formation. An international handbook of the languages of Europe*. Vol. 1. Berlin / New York: De Gruyter. 450-467.
- Koefoed, Geert (1993): *Benoemen. Een beschouwing over de faculté de langage*. Amsterdam: Meertens Instituut.
- Levelt, Willem J.M. / Meyer, Antje (2000): Word for word: Multiple lexical access in speech production. In: *European Journal of Cognitive Psychology* 12(4). 433-452.
- Los, Bettelou / Blom, Corrien / Booij, Geert / Elenbaas, Marion / van Kemenade, Ans (2012): *Morphosyntactic change. A comparative study of particles and prefixes*. Cambridge: Cambridge University Press.
- Masini, Francesca (this volume): Compounds and multiword expressions in Italian: demarcation and competition issues.
- Nooteboom, Sieb (2011): Self-monitoring for speech errors in novel phrases and phrasal lexical items. In: *Yearbook of Phraseology* 2.1-16.
- Paul, Hermann (1880): *Prinzipien der Sprachgeschichte*. Halle a S.: Max Niemeyer [1898³].
- Rainer, Franz (2013): Can relational adjectives really express any relation? An onomasiological perspective. In: *Skase Journal of Theoretical Linguistics* 10.12-40.
- Schlücker, Barbara (2014): *Grammatik im Lexicon. Adjektiv-Nomen-Verbindungen im Deutschen und Niederländischen*. Berlin: De Gruyter.
- Schlücker, Barbara (this volume). Compounds and multi-word expressions in German.
- Schlücker, Barbara / Plag, Ingo (2011): Compound or phrase? Analogy in naming. In: *Lingua* 121.1539-1551.
- Schuster, Saskia (2016): *Variation und Wandel. Zur Konkurrenz morphologischer und syntaktischer A+N-Verbindungen im Deutschen und Niederländischen seit 1700*. Berlin: Mouton De Gruyter. [<https://www.degruyter.com/view/product/456743>]
- Schutz, Rik / Permentier, Ludo (2016): *Met zoveel woorden. Gids voor trefzeker taalgebruik*. Amsterdam / Leuven: Amsterdam University Press / Davidsfonds Uitgeverij.
- Sprenger, Simone (2003): *Fixed expressions and the production of idioms*. Ph.D. dissertation. Nijmegen: Max Planck Instituut voor Psycholinguïstiek.
- Sprenger, Simone A. / Levelt, Willem J.M. / Kempen, Gerard (2006): Lexical access during the production of idiomatic phrases. In: *Journal of Memory and Language* 54. 161-184.
- Tummers, José (2005): *Het naakt(e) adjectief. Kwantitatief-empirisch onderzoek naar de adjectivische buigingsalternatie bij neutra*. Leuven: Katholieke Universiteit Leuven [Ph. d. dissertation].
- Vikner, Sten (2005): Immobile complex verbs in Germanic. In: *Journal of Comparative Germanic Linguistics* 8.83-105.