Phrasal names: A constructionist analysis

Geert Booij

Abstract

Some types of phrases share the naming function with complex words. Hence both phrases and words can be lexical units stored in the lexicon. This article discusses how the functional equivalence between words and phrases can be accounted for without ignoring their formal differences. Such types of phrases can be characterized in terms of phrasal schemas with specific properties, that is, as constructions. The article focuses on the formal properties of adjective+noun sequences with a naming function, in particular in Modern Greek and Dutch. The constructionist approach is able to do justice to the lexical unit properties of phrasal names, and highlights the parallelism between phrasal and morphological constructions.

1 Naming and description

Words are the linguistic expressions par excellence for the function of naming. Words function as names for concepts. This holds for both underived and derived words. However, the naming function is not restricted to words: certain types of phrases can also function as names. This functional equivalence between words and phrases is a challenge for linguistic theory, because we should maintain the formal distinction between words and phrases, and yet do justice to their common properties.

The following definition of naming is the starting point of this article: ‘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.’ (Koefoed 1993) [my translation]. As an example of a phrase with a naming function, Koefoed mentions the Dutch noun phrase vaderlandse geschiedenis ‘national history’, which is the conventional name for a particular form of history, namely that from the perspective of one’s native country. This phrase can be opposed to the phrase geschiedenis van het vaderland ‘history of the native country’, a descriptive phrase that refers to the history of one’s native country.
These two different functions of phrases are important from the perspective of a theory of the architecture of the grammar. Phrases used as names are often conventional expressions, and hence lexical units. However, when a linguistic expression is a lexical unit, this does not imply that it is a word. Words and phrases need to be distinguished carefully in an adequate linguistic theory. On the other hand, the functional similarity or even equivalence between words and certain types of phrases as names has the effect that such phrases may have specific formal properties, as will be discussed below. Hence, a proper theory of the relation between morphological and syntactic naming constructions is called for.

The role of word formation as an alternative to phrasal expression has been discussed in detail in Kastovsky (1988), who calls this recategorization. It applies to both compounding and derivation, and sometimes it has to do with the need for stylistic variation in texts. In the following examples, the use of ringer-up is a means of avoiding the clumsy phrase, the person who rang up. In the second example, repetition of the word civil is avoided in the root clause by not using the phrase make civil but the complex word civilize:

(1) Miss Pride is convinced that the ringer-up was Miss Cost
    If that’s not civil, civilize it and tell me (Kastovsky 1988: 595)

Nominal compounds are often used as short alternatives to syntactic descriptions. This is found in particular in the headlines of newspapers, to such an extent that one might speak of ‘headline-morphology’. Here are some examples of such descriptive compounds from Dutch newspaper headlines:

(2) zwemles-regeling ‘swimming lessons arrangement’
    het Cruyf-interview ‘the Cruyf-interview’
    Schelde-verdieping lit. Schelde-deepening, ‘deepening of the river Schelde’.4

In sum, both complex words and phrases can be used for two functions: naming and description.

The formation of new names for concepts can take place through a number of mechanisms (De Caluwe 1990): word formation, the construction of phrases, borrowing of names from other languages, the creation of new simplex words (brand names), the formation of acronyms which turns descriptions into names, as in gsm, sms, semantic extension (metaphor, metonym), and clippings from phrases, for instance Dutch mobiel (often used in its diminutive form mobiel-tje) clipped from mobiele telefoon ‘mobile phone’.5

Phrasal names, once conventionalized, belong to the class of fixed expressions (Wray 2002). Fixed expressions may be qualified as follows:

(3) Fixed expressions (FEs) refer to specific combinations of two or more words that are typically used to express a specific concept. Typical examples of FEs that are referred to in the literature often have an opaque meaning or a deficient syntactic structure, for example, by and large or kick the bucket. However, these properties are not essential. The defining feature of a FE is that it is a word combination,
stored in the Mental Lexicon of native speakers, that as a whole refers to a (linguistic) concept. This makes FEs “non-compositional” in the sense that the combination and structure of their elements need not be computed afresh, but can be retrieved from the Mental Lexicon. However, the degree of lexical and syntactic fixedness can vary. (Sprenger 2003: 4)

In this quotation, it is rightly stressed that the notion ‘fixed expression’ encompasses more than idioms. Idiomatic expressions have one or more non-compositional properties, whereas a fixed expression may be completely compositional, but nevertheless stored because it is a conventional name for a particular concept. Thus, this quotation reminds us of the fact that knowledge of a language encompasses both knowledge of the grammatical system of a language, and knowledge of the conventions involved in using that language, as has been stressed in the work of Eugenio Coseriu. The lexicon is the place where the conventional use of linguistic expressions is encoded, and hence, the lexicon is the meeting point of language as system, and language as convention.

As mentioned above, the notion ‘lexical unit’ is not to be identified with the notion ‘word’ (in the sense of ‘lexeme’ or in the sense of ‘syntactic atom’). Lexical units can be constructed by means of syntactic rules. In European languages we find the following patterns for phrasal names (not an exhaustive list):

(4) (a) NPs of the form A+N or N+A: Spanish media luna ‘half moon’, luna nueva ‘new moon’
(b) NPs of the form N+PP: French moulin à vent ‘windmill’
(c) NPs of the form N-GEN+N or N+N-GEN: English woman’s magazine, Greek zona asfalies ‘belt safety’, GEN, safety belt’
(d) NPs of the form N+NP (apposition): Dutch directeur artistieke zaken ‘director of artistic affairs’

Such constructions are sometimes referred to as loose compounds, but this is a misleading term since these expressions are not compounds in the morphological sense. Words are subject to the principle of Lexical Integrity which states that ‘the syntax neither manipulates nor has access to the internal structure of words’ (Anderson 1992: 84). In an expression like Spanish media luna in (4a), for instance, we observe that the syntactic rule of gender agreement has applied: the form of the adjective is medi-a because it has to agree with the feminine gender of luna. Therefore, media luna cannot be interpreted as a morphological compound.

NPs of the form (4a) will be discussed in more detail in section 2. NPs of the type (4b) are used a lot in Romance languages. The following French phrases (and many more) that can all be used as names can be found in Fradin (2003: 199):

(5) (a) N de N: fil de fer ‘iron wire’
(b) N à N: moulin à vent ‘windmill’
(c) N à Det N: sauce à l’ail ‘garlic sauce’
(d) A N: moyen âge ‘Middle Ages’
(e) N A: poids lourds ‘heavyweight’
The patterns in (5) have a certain degree of productivity. In particular the construction \( N \text{ à } N \) is extremely productive in French for coining names, as illustrated in (6):

(6)  

- *moulin à poivre* ‘pepper mill’
- *verre à vin* ‘wine glass’
- *bois à feu* ‘firewood’
- *fruit à confiture* ‘jam fruit’
- *moteur à essence* ‘petrol engine’

Note also the difference between *verre à vin* ‘wine glass’ and *verre de vin* ‘glass of wine’. The construction with \( à \) has typically a classifying role.

The use of NPs with a PP complement consisting of a P and a bare noun as names is also found in Spanish and Italian:

(7) Spanish (Rainer & Varela 1992):

- *telón de acero* ‘iron curtain’ (between Eastern and Western Europe)
- *piano de cola* ‘grand piano’
- *gafas de sol* ‘sun glasses’

(8) Italian: (Bisetto & Scalise 1999):

- *permesso di pesca* ‘fishing licence’
- *mulino a vento* ‘windmill’
- *occhiali da sole* ‘sun glasses’

The Spanish examples are traditionally classified as *compuestos improprios* (= ‘improper compounds’). However, Rainer and Varela (1992) rightly stress the point that they are phrases, albeit with specific formal properties:

> The naming function of such phrases can be attributed to the pragmatic component in a modular conception of the grammar. Global pluralization, accentual unity and opacity to syntactic rules like adjectival / adverbial modification, properties which some of these formations – among them *media luna* – share with simple words, should be viewed as what they are, namely consequences of the naming function or lexicalization which one also finds in idioms. (Rainer & Varela 1992: 120)

These formal effects of the naming function of phrasal lexical units are the focus of this article, in particular as far as *A+N* phrases are concerned.

An example of the genitive construction mentioned in (4c) can be found in English. The genitive noun is in pre-nominal position, and has a classifying role: *a boy’s hat, a woman’s/women’s magazine*. These genitive constructions function like compounds but the modifying noun is marked by a morpheme that is historically a genitive case ending (Rosenbach 2007). Note that the genitive can also be used in descriptive phrases, as in *the boy’s hat* ‘the hat of the boy’.
Modern Greek features N+N constructions with a genitive marker as well. The second noun is marked with genitive case, and the whole expression functions as a name (cf. Ralli (2007) for details):

(9)  
*ikos anoxis* 'house tolerance.GEN, brothel'
*oni asfalias* 'belt safety.GEN, safety belt'
*ikos eveygrias* 'house old.PL.GEN, old people’s home'

The fact that both words and phrases can function as names creates a problem of demarcation for languages with left-headed compounds: left-headed N+N sequences (the class of expressions (4d)) could be interpreted either as left-headed compounds or as phrases in which the head noun is followed by an appositive noun. For instance, French and Italian feature N+N/NP sequences that may look like left-headed compounds, but may also be seen as phrases with appositional Ns:

(10) French

=cigarette filtre 'filter cigarette'
=sortie piétons 'pedestrians’ exit'
=impression laser 'laser printing'
=cuisinière quatre feux ‘four-burner stove’
=assurance tous risques ‘all risks insurance’

Italian

effetto serra ‘greenhouse effect’
=smaltimento rifiuti ‘garbage disposal’

A possible argument for interpreting these N+N/NP sequences as phrases is that the second constituent can be modified, as in the last two examples. However, phrases in the non-head position of compounds cannot be excluded by a universal constraint (Booij 2007: 188), and hence this observation does not exclude the word combinations in (10) to be interpreted as left-headed compounds. When we pluralize these expressions, it is the left noun that is pluralized: *cigarettes filtre ‘filter cigarettes’, effetti serra ‘greenhouse effects*. If these expressions were compounds, we would expect plural marking at the right edge; hence, pluralization indicates that a phrasal interpretation of these expressions is called for.

Additional examples of Italian phrasal expressions with a compound-like function are the following (Bissetto & Scalise 1999: 39):

(11) (a) *produzione scarpe* ‘lit. production shoes, shoe production’
(b) *caduta massi* ‘lit. falling stone, falling of stone’

Baroni *et al.* (2008) also argue that these N+N sequences are phrases, subject to specific principles of ‘headline syntax’, a specific style or register in which determiners and prepositions are omitted. The complement of such N-headed phrases can also be a phrase (12a, b), and the complement is sometimes even used in a non-generic manner (12b):

(12) (a) *approvazione nuovi parametri* ‘checking of new parameters’
(b) *accertamento posti disponibili* ‘checking for some/the available seats’
The use of appositional constructions without determiners and prepositions for names can also be observed in Dutch personnel advertisements. Dutch compounds are right-headed, and hence the following left-headed expressions must be seen as phrases with a special syntax used for coining names for particular jobs (source NRC-Handelsblad 28 Sept. 2008):

(13) senior-adviseur installatietechniek ‘senior adviser equipment technics’
    manager kredietrisicomanagement ‘manager credit risk management’
    lid raad van commissarissen ‘member board of supervisory directors’
    sectormanager infrastructuur ‘sector manager infrastructure’
    hoofd communicatie ‘head communication’
    adjunct-directeur artistieke zaken ‘deputy director artistic affairs’

In conclusion, phrases can function as names, and there is no simple one-to-one correlation between the form of a linguistic expression and its function as name or description. Yet, there appear to be specific correlations between the form and the naming function of phrases. In section 2, these formal properties are discussed for A+N phrases, in particular those of Greek and Dutch. In section 3 I argue in favour of a constructionist account of these phrasal names. Thus, justice is done to their being lexical units without obliterating the boundary between syntax and morphology.

2 A+N phrases as names for concepts

The use of A+N/N+A phrases as names for concepts or categories is widespread in European languages. The following examples of A+N phrases from Booij (2002a) illustrate this use for Dutch:

(14) dikke darm ‘lit. thick intestine, large intestine’, dood spoor ‘lit. dead trail, deadlock’,
    hoge hoed ‘lit. high hat, top hat’, magere yoghurt ‘lit. lean yoghurt, fat-free yoghurt’,
    open haard ‘lit. open hearth, fireplace’, rode kaart ‘red card’,
    fixed appointment, tenure’, zure regen ‘acid rain’ vrije trap ‘free kick’, witte was
    ‘white laundry’, zwarte doos ‘black box’

We are certain that these A+N sequences are phrases because the adjectives are inflected, and agree with the head noun with respect to the features gender, number and (in)definiteness. The pre-nominal adjective ends in the suffix -e, unless the NP is indefinite and the head noun is singular and neuter (in the latter case the ending is zero). According to the principle of Lexical Integrity, the syntactic rule of agreement cannot affect parts of words, and hence these A+N sequences must be phrasal in nature.

The functional similarity between such phrases and compounds is nicely shown by the comparison between German compounds and their Dutch glosses (source Die Welt 31/05/2008):

(15) Dutch AN phrase German A-N compound gloss
    bijzonder-e zitting Sonder-sitzung ‘special session’
    gebruikt-e batterijen Alt-batterien used batteries’
    (het) geheim-e nummer (die) Geheim-nummer ‘(the) secret number’
    mobiel-e telefoon Mobil-telefon ‘mobile phone’
    nieuw-e auto Neu-auto ‘new car’
This comparison between Dutch and German should not be taken to imply that Dutch does not have A-N compounds. Examples of such compounds are *fijn-stof* 'lit. fine dust, fine-grained dust’ and *vroeg-geboorte* ‘lit. early-birth, premature birth’ in which the A is not inflected. However, they are far less frequent than A-N compounds in German. The A in the Dutch A-N compounds has to be simplex; German also imposes that restriction, but with some exceptions such as adjectives in *-al* (*National-staat* ‘nation state’), *-iv* (*Suggestiv-frage* ‘suggestive question’, and *-ig* (*Niedrig-wasser* ‘low tide’) (Hünig, pers. comm.). This implies a certain division of labour between the phrasal and the compound coining of such names: in Dutch (and partially also in German) these two different A+N constructions only compete when A is simplex.8

The A in these Dutch A+N phrases cannot be modified. That is, it does not project a phrase of its own, and can therefore be qualified as a non-projecting category (cf. Toivonen 2003). Hence, it is an A0, and not an AP. Moreover, A and N cannot be split by some other word, they must be adjacent. A possible structural interpretation for these A+N sequences is that of a syntactic compound, or that of a noun phrase.9

(16) (a) *[donker-e]A0 [kamer]N0 ]N0 ‘dark room’
    (b) *[donker-e]A0 [kamer]N0 ]_N ‘dark room’

The structure in (16a) is that of a syntactic compound (as opposed to the morphological compound structure [A N]N which can be inserted as a whole in a syntactic N0 position, and in which agreement between A and N does not apply). Hence, this structure predicts that no other word can intervene between the A and the N. This syntactic compound structure has also been proposed for similar English A+N sequences (Sadler & Arnold 1994). Alternative (16b) will be discussed below.

Qualifying adjectives such as *rood* ‘red’ and *geel* ‘yellow’ cannot be modified when used in A+N sequences that function as names, and hence they are non-projecting adjectives:

(17) (a) *erg rode kool* ‘very red cabbage’
    (b) *heel gele koorts* ‘very yellow fever’

The phrasal name *rode kool* does not simply denote the intersection of the class of cabbages, and that of red things, so it has a non-intersective reading. The same applies *mutatis mutandis* to *gele koorts*. The non-projectivity of the adjective is a reflection of this semantic property. The phrase (17a) is a grammatical expression, but not when *rode kool* is meant to refer to a specific conventional subtype of cabbage, red cabbage (note that the actual colour of red cabbage is more purple than red). In (17b), the use of the degree adverb *heel* forces the adjective *gele* to be interpreted as referring to the set of yellow things, resulting in a uninterpretable reading.

Relational adjectives, that is, adjectives (often denominal) that express a relation between the head noun and a nominal entity such as *Nederland* ‘Netherlands’ in *Nederland-se regering* ‘Dutch government’, and *burger* ‘civilian’ in *burger-lijke vrijheid* ‘civilian liberty’, always function as non-projecting adjectives (Heynderickx 2001), and only occur in NPs with a naming function.
The use of A+N phrasal constructs as names is widespread in European languages;
examples are given in (18):

(18) (a) Dutch: Booij (2002a; Hünig forthcoming); cf. (14);
(b) English: Jackendoff (1997, 2002; Sadler & Arnold 1994): Arabian horse, blue
cheese, electrical outlet, modern art, natural childbirth;
(c) German (Hünig forthcoming): saure Sahne ‘sour cream’; saurer Regen ‘acid
rain’, grüne Welle ‘phased traffic lights’;
(d) French (Fradin 2003), cf. the examples in (5);
(e) Italian (Semenza & Mondini 2006): NA: febbre gialla ‘yellow fever’, natura
morta ‘still life’; AN: alta moda ‘haute couture’, mezza luna ‘half moon’;
(f) Spanish (Rainer & Varela 1992): NA: luna nueva ‘new moon’; AN: media luna
‘half moon’;
(g) Catalan: NA llibre blanc ‘white book’, peix blau ‘bluefish’, NA belles arts ‘fine
arts’;
(h) Greek (Ralli & Stavrou 1998): psixros polemos ‘cold war’, tritos kosmos ‘Third
World’.

A short survey of such ‘tight’ AN phrases in Celtic and Romance languages is given
in Dahl (2004: 228–30). In some cases, the semantic tightness of these phrases manifests
itself in the pre-nominal position of the adjective in languages where the default position
of the adjective is post-nominal, as in Spanish un gran hombre ‘a great man’ versus
un hombre grande ‘a big man’ (Dahl 2004: 229).10

In section 2.1 I discuss Greek AN combinations in more detail, and the properties of
Dutch AN phrases are discussed in section 2.2. Section 3 draws some conclusions from
these facts and their analysis for the architecture of the grammar.

2.1 Greek A+N combinations

In Ralli & Stavrou (1998) two types of A+N combinations are distinguished: A+N
compounds and A+N constructs. Note that the term ‘compound’ is used by these
authors in the sense of ‘lexical unit’, in contrast with our use of the term compound
as referring to a morphological class of words.11

(19) (a) A+N ‘compounds’
    psixr-os polemos ‘cold war’
    trit-os kozmos ‘third world’
    mavr-i lista ‘black list’
    mikr-i othoni ‘small screen’
(b) A+N ‘constructs’
    atomik-i vomva ‘atomic bomb’
    musik-i kritiki ‘musical review, music review’
    odhik-o dhiktio ‘road network’
    pirinik-i dhokimi ‘nuclear testing’
The difference between (19a) and (19b) is that in (19b) relational adjectives are used, whereas (19a) contains qualifying adjectives. In both constructions, the As are non-projecting, and they cannot be separated from the head N by the definite determiner, nor follow the head noun, unlike what is the case for descriptive A+N phrases in Greek, where double marking of definiteness is possible:

\[
\begin{align*}
(20) \quad & (a) \quad \text{o polemos o psixros} \\
& \quad \text{the war the cold ‘the cold war’} \\
& (b) \quad \text{o kafes o zestos} \\
& \quad \text{the coffee the hot ‘the hot coffee’}
\end{align*}
\]

According to Ralli & Stavrou (1998), the difference between the two types of A+N combinations is that (i) A+N ‘compounds’ are idiosyncratic in meaning whereas A+N ‘constructs’ are semantically regular, and that (ii) in the A+N constructs (21b) but not in the A+N compounds (21a) the order of A and N may be reversed in indefinite phrases:

\[
\begin{align*}
(21) \quad & (a) \quad \text{mia pedh-iki xara} \quad \Rightarrow \quad \text{*mia xhara pedh-iki} \\
& \quad \text{a child-ADJ delight} \quad \text{a delight child-ADJ} \\
& \quad ‘\text{a playground’} \\
& (b) \quad \text{atom-iki vomva} \quad \Rightarrow \quad \text{vomva atom-iki} \\
& \quad ‘\text{atom-ic bomb’}
\end{align*}
\]

Moreover, the Greek A+N ‘constructs’ can be split by parenthetical expressions, unlike A+N ‘compounds’.

\[
\begin{align*}
(22) \quad & i \quad \text{vionixaniki, opos oli borite na dhite, zoni} \\
& \quad \text{the industrial, as all you can see, area}
\end{align*}
\]

For these reasons, Ralli & Stavrou consider the first class of A+N combinations as created by a morphological rule of compounding, whereas A+N constructs are considered to be phrasal, and hence syntactic in nature.

Ralli & Stavrou also note that not only the A+N ‘compounds’, but also the A+N combinations classified as ‘constructs’ behave as units with respect to qualifying adjectives, since they can follow or precede the qualifying adjective as a whole as shown in (23):12

\[
\begin{align*}
(23) \quad & (a) \quad \text{sinexis psixros polemos ‘continuous cold war’} \\
& \quad \text{*psixros sinexis polemos ‘cold continuous war’} \\
& (b) \quad \text{mia theatriki kritiki kali} \\
& \quad \text{a drama review good ‘a good drama review’} \\
& \quad \text{*mia theatriki kali kritiki}
\end{align*}
\]

The main problem for a morphological analysis of the Greek A+N ‘compounds’, as proposed by Ralli and Stavrou, is that there are formal indications that they are phrasal: the adjectives are inflected, and they agree with the head noun in gender, number, and case. Hence, interpreting these A+N sequences as morphological compounds is in conflict with the principle of Lexical Integrity which forbids syntactic rules to apply within words. Therefore, I opt for a phrasal interpretation of both subsets of A+N
combinations. A possible way of differentiating between the two classes is to consider the A+N compounds as syntactic N₀ compounds, and the A+N constructs as phrasal. These structures may be contrasted with that of morphological compounds, which would have the form in (24c):

(24) (a) constructs: \([A^0 N^0]_N\)  e.g. \([[\text{atomiki}]_0 [\text{vomva}]_0]_N\)
(b) syntactic compounds: \([A^0 N^0]_N^0\)  e.g. \([[\text{psixros}]_0 [\text{polemos}]_0]_0\)
(c) morphological compounds: \([N N]_N\)  e.g. \([[\text{nixt}]_0 [\text{puli}]_0]_N\)  ‘night bird’

Morphological AN compounds in Greek are also different from the other two types of construction in that the first constituent is a stem, followed by a linking element –ο, and that the compound as a whole forms one phonological word (Ralli 2007). A morphological compound headed by the category label N can occupy a N₀ position in a syntactic structure. This structural differentiation can be used to account for the differences in syntactic behaviour of the different types of A+N sequences.

The structure proposed here for these A+N ‘compounds’ is the same as that proposed for similar A+N constructions in English by Sadler and Arnold (1994). A similar structure has been proposed for verbs with quasi-noun incorporation in Japanese (Iida & Sells 2008) and Dutch (Booij 2009a). These verbal units are N+V sequences with certain phrasal properties that behave as very tight units. Hence, their structure can be assumed to be \([N^0 V^0]_0\).

Ralli & Stavrou (1998) point out that a morphological compound analysis of the Greek A+N ‘constructs’ is also out of the question because they cannot be used as the nominal bases of relational adjectives, as we would expect under a morphological compound interpretation of these A+N sequences. This is one of the reasons why Ralli and Stavrou do not qualify them as compounds, but give them phrasal status (constructs). It is only the corresponding stems without the inflectional endings that can form the basis of such relational adjectives, as illustrated in (25):

(25) \(\text{psixr-οs polemos-οs} \text{ ‘cold war’} \quad \text{psixr-ο-polem-ιkos} \text{ ‘cold-war-like’}\)

This is to be expected if derivation is restricted to taking words (in their stem form) as bases, as is usually the case. Word-based derivation excludes the phrasal name \(\text{psixros polemos}\) from functioning as a base for complex adjectives. The base \(\text{psixr-ο-polem-}\) in \(\text{psixropolemikos}\) has the form of a compound: its first constituent is the adjectival stem \(\text{psixr-}\) followed by the generally used linking element –ο and the nominal stem \(\text{polem-}\). What we observe here is a paradigmatic relationship between an A+N phrase and a nominal relational adjective with a compound base. The meaning of the complex adjective is a compositional function of that of the paradigmatically related phrasal name \(\text{psixros polemos}\) and that of the suffix –ικοσ. The morphological schema for such Greek relational adjectives with a compound base is:

(26) \([[\text{A-o-N}]_N \text{ ik-οs}]_A\)

This schema is a unification of the schema for denominal relational adjectives with the schema for Greek AN compounds.
In section 3, I discuss how these construction types, with both phrasal and lexical properties, can be accounted for in the grammar of Greek.

2.2 Formal properties of Dutch A+N phrases with a naming function

Let us now have a more detailed look at how the naming function of A+N phrases affects the form and behaviour of such phrases in Dutch.

a. A first relevant observation is that of a blocking effect: the coining of N-N compounds is often blocked by the existence of a competing A+N phrase, and vice versa. The following examples are from Booij (2002b):

\[(27)\] A+N phrase N-N compound
- *academisch jaar* ‘academic year’ *academie-jaar* ‘academy year’
- *koninklijk lid* ‘royal decision’ *konings-lid* ‘king decision’
- *muzikale scholing* ‘musical training’ *muziek-scholing* ‘music training’
- *muzikaal talent* ‘music talent’ *muziek-talent* ‘music talent’
- *muzikale school* ‘musical school’ *muziek-school* ‘music school’

The adjectives in these A+N phrases are denominal relational adjectives. That is why they constitute an alternative to modifying a noun by means of another noun in N-N compound structures. Some of the phrases and compounds are marked with a question mark. This does not mean that these word combinations are ungrammatical, but that they are infelicitous because of the existence of a conventionalized synonymous (and hence competing) expression. Blocking is the effect of competition between synonymous lexical units. It is not a formal principle that qualifies constructs as ungrammatical. We also find individual variation, which is to be expected since languages users differ in which lexical units they have stored. The blocking effects observed in (27) show that coining such A+N phrases is coining names which therefore, as lexical units, compete with synonymous compounds.

b. A+N phrases with a classifying role can appear in the non-head-position of compounds and as bases of certain types of derived words. In this position, they function as names for categories, and hence receive a generic interpretation:

\[(28)\] Dutch ANs that form lexical collocations the adjective may appear without the expected inflectional ending -e (Booij 2002a: 47–48; Tummers 2005). In the following
examples, the nouns have common gender, and normally an adjective preceding such nouns ends in a schwa:

(30)  
\[
\text{een/de geheim-ø agent } \quad \text{‘a/the secret agent’}
\]
\[
\text{een/de taalkundig-ø onderzoeker } \quad \text{‘a/the linguistic investigator’}
\]

The absence of the expected inflectional schwa strengthens the naming function of these NPs: ‘the common meaning of the construction resides in the fact that the uninflected adjective in these constructions focuses less on the individual properties of the person or thing referred to, and more on general or categorial properties or stereotypes deducible from certain specialized usages of the substantives’ (Blom 1994: 81). The effect of the absence of the inflectional schwa is that they look more like compounds. Yet, they are still phrasal since they have main stress on the noun (in A-N compounds, main stress is on the adjective).

There are also A+N phrases with a neuter noun as head that function as names. In that case, the inflectional ending -e tend to be omitted in the singular form only, but we might find uninflected adjectives before plural neuter nouns as well:

(31)  
\[
\text{het oudheidkundig(e) museum } \quad \text{‘the archeological museum’}
\]
\[
\text{de oudheidkundig musea } \quad \text{‘the archelogical museums’}
\]
\[
\text{het koninklijk(e) paleis } \quad \text{‘the royal palace’}
\]
\[
\text{de koninklijk paleizen } \quad \text{‘the royal palaces’}
\]

The phrasal nature of these A+N names is clear from their stress pattern (stress on the last constituent), and the optionality of the absence of the inflectional ending.

d. A+N phrasal names can occasionally be modified as if they are the head constituent of a nominal compound, with a modifying word or prefix in the non-head position. The head of a compound or a prefixed word in Dutch cannot be a phrase: Dutch compounds are right-headed, and hence, a phrasal head would imply phrasal status for the construction as a whole. Yet, we can find A+N sequences with a naming function, preceded by a noun with a modifying function, or preceded by a prefix that normally attaches to nouns. Thus, these A+N sequences behave like heads of a compound or derived word. The following examples (some of them from Ackema & Neeleman 2004: 125) illustrate this phenomenon:

(32) Compounds
\[
\text{namaak mobiele telefoon } \quad \text{‘imitation mobile phone’}
\]
\[
\text{wereld rode wijn } \quad \text{‘world red wine, superb red wine’}
\]
\[
\text{deeltijd pastoraal medewerker } \quad \text{‘part-time pastoral assistant’}
\]

Prefixed words
\[
\text{ex-aanstormend talent } \quad \text{‘ex up-and-coming talent’}
\]
\[
\text{pseudo-taalkundig onderzoeker } \quad \text{‘pseudo linguistic researcher’}
\]
\[
\text{pseudo-epileptische aanval } \quad \text{‘pseudo epileptic fit’}
\]

These cases suggest that such A+N phrasal names may be reinterpreted as being N0, and hence usable as heads of the constructs in (32). The A+N heads of these constructs
still exhibit inflection of the adjective. This suggests considering these A+N constructs in this context as cases of syntactic compounds dominated by N°.

e. Classifying A+N phrases (i.e. phrasal names) can be coordinated with compounds but not with descriptive phrases (Heynderickx 2001); the common constituent of the coordinated constituents can be gapped (deletion of a prosodic word constituent under identity with another one (Booij 1985)):

(33) Classifying phrase+compound

Amerikaanse (talen) en Papoeatalen
'American (languages) and Papua-languages'

Descriptive+classifying phrase

*het grote (paleis) en koninklijke paleis
'the large and royal palace'

Compound+classifying phrase

ijs(beren) en bruine beren
'ice bears and brown bears'

Classifying phrase+descriptive phrase

*Amerikaanse (talen) en moeilijke talen
'American languages and difficult languages'

These observations show that classifying phrases behave like compounds as to their coordination possibilities, and differ in this respect from descriptive phrases. That is, we need a specific semantic category ‘name’ that generalizes across compounds and phrasal names, in order to express the constraint on coordination that the two non-heads must exhibit semantic parallelism.

f. The order of adjectives in Dutch is such that descriptive adjectives precede classifying adjectives:

(34) vieze rode kool ‘filthy red cabbage’ / *rode vieze kool
goedkope witte wijn ‘cheap white wine’ / *witte goedkope wijn
dure mobiele telefoon ‘expensive mobile phone’ / *mobiele dure telefoon

Thus, descriptive adjectives are peripheral to the AN unit that functions as a name. The same observation can be made for adjective sequences in other languages, as mentioned in section 2.1 for Greek. It also holds for the other type of phrases with a naming function listed in (4), as illustrated here for the Italian \[N_a N\] name sedia a rotelle ‘wheel chair’ (Semenza & Mondini 2006: 92):

(35) sedia a rotelle rotta / *sedia rota a rotelle ‘broken wheel chair’

These distributional facts follow from the status of Dutch A+N phrasal names as either syntactic compounds with the structure \([A^0 N^0]_{N_0}\), or as phrases of the type \([A^0 N^0]_{N'}\). Both structures require the adjective to be adjacent to the noun.

The two different possible dominating nodes (N° or N') may be used to differentiate between phrasal names with qualifying adjectives (type rode kool ‘red cabbage’, gele koorts ‘yellow fever’) and A+N sequences with relational adjectives. The latter type of
adjectives exhibit more syntactic freedom: they can, for instance, be coordinated, as in een juridisch en fiscaal advieskantoor ‘a law and tax advice office’. This phrase cannot be considered as a case of coordination reduction from een juridisch kantoor en een fiscaal kantoor since only one kantoor ‘office’ is meant. Hence, this is a case of coordination of adjectives. Coordination of modifiers in phrases is a regular syntactic option, and hence I assign such A+N phrasal names the status of N’. Coordination is odd with qualifying adjectives: *de tropische en gele koorts ‘the tropical and yellow fever’, hence the interpretation of gele koorts ‘yellow fever’ as a syntactic compound with N0 status.

g. A+N phrasal names behave like compounds in allowing for a particular type of semantic reanalysis, that of semantic concentration, in which the meaning of a whole compound is projected onto its first constituent (Meesters 2004). The notion of semantic concentration can be illustrated by the following example. In Dutch, the verbal stem scharrel ‘to scratch, to potter around’ acquired the meaning ‘free range, eco-’ starting from the compound scharrel-kip:

(37) scharrel-kip ‘scratch chicken, free range chicken > scharrel-ei ‘free range egg’ scharrel-vlees ‘free range meat, eco-meat’

Whereas chickens can potter around, this is not the case for eggs or meat. This shows that a process of semantic reinterpretation has been in effect here. Such semantic concentration effects may lead to the rise of new affixes from compound constituents (Booij 2005). The semantic concentration effect can also be observed in A+N phrasal names, as illustrated here for the adjective onbespoten ‘unsprayed’ that acquired the meaning ‘eco-’:

(38) onbespoten groente ‘unsprayed vegetables’ = ‘eco-vegetables’
onbespoten restaurants ‘lit. unsprayed restaurants’ = ‘eco-restaurants’
onbespoten mensen ‘lit. unsprayed people, eco-minded people’ (Dutch daily Trouw 14 June 2008)
onbespoten idealen ‘lit. unsprayed ideals, ecological ideals’

This use of onbespoten in A+N names has led to a meaning ‘environment-friendly. This phenomenon shows that we have to distinguish a semantic class of A+N phrasal names, in order to explain why, like compounds, these phrases may exhibit these semantic concentration effects.

In sum, we have seen a number of phenomena in which A+N phrasal names in Dutch behave as a unit. Their formal and corresponding semantic structure should therefore be represented in such a way that their parallelism with morphological A-N and N-N compounds as names for categories is made clear. A first, admittedly rather informal approach to the semantics of these construction types is the following:

(39) (a) syntactic compounds with qualifying adjectives:

\[ A^0, N^0 \] \( k \leftrightarrow \) [NAME for \( SEM \), with some relation \( R \) to \( SEM \) \( k \)]

(b) phrases with denominal relational adjective:

\[ A^0, N^0 \] \( k \leftrightarrow \) [NAME for \( SEM \), with some relation \( R \) to entity \( E \) of \( SEM \) \( k \)\]
The name interpretation of these constructional schemas is specified to the right of the double arrow, which stands for the correspondence between form and meaning. SEM stands for the meaning of the constituents, and co-indexation specifies the relation between formal constituents and corresponding meanings. In (39b), the meaning of the relational adjective involves a relation with some entity which may have a corresponding base noun in this relational adjective. In a relational adjective like academisch ‘academic’, the evoked entity academie ‘academy’ corresponds with a base noun, academie. In the case of the adjective juridisch ‘legal’ there is no identifiable base noun that corresponds with the entity ‘law’ but the meaning of juridisch implies the existence of such an entity.

Just as in morphological compounds, the semantic relation R between the head and the modifier is not predictable from the linguistic structure as such, and has to be filled in for each individual compound on the basis of contextual and world knowledge (Downing 1977). The same holds for these A+N phrases. For instance, in gele koorts ‘yellow fever’ the fever has the effect of a yellow skin, whereas in het Rode leger ‘the Red army’, the army adheres to principles symbolized by the color red.

The representations (39) are constructional schemas, in which a particular structural configuration is linked to a particular semantic interpretation. The distinguishing formal property of these constructional phrasal schemas for names is the non-projecting A, which does not project a full phrase with modifiers.

2.3 Phrasal lexical units and word formation

Let us now look at the interaction of phrasal A+N names with word formation. A+N phrases can function as non-heads of compounds, as mentioned above. There are a few Dutch denominal suffixes that accept phrases as bases: the denominal suffixes -er ‘id.’ and -achtig ‘-like’. They occur with A+N phrases:

\[(40)\]  
\(\text{derde-klass-er ‘third class-er, member of the third class’}\)  
\(\text{zeventiende-eeuw-er ‘17th century-er, someone living in the 17th century’}\)  
\(\text{zwarte-band-er ‘black belt-er, judoka with a black belt’}\)  
\(\text{jonge-meisjes-achtig ‘young girls-like, like young girls’}\)  
\(\text{oude-mannen-achtig ‘old men-like, like old men’}\)

The Dutch diminutive suffix, on the other hand, does not accept phrasal bases. Yet, we can diminutivize phrasal A+N names. The diminutive suffix has semantic scope over the whole phrase, but formally attaches to its head N:

\[(41)\]  
\(\text{hoge hoed ‘high hat, tall hat’}\)  
\(\text{volle maan ‘full moon’}\)  
\(\text{rode wijn ‘red wine’}\)  
\(\text{witte was ‘white laundry’}\)

In the example rode wijn, the diminutive is used either as a suffix of endearment or individuation (in order to express the meaning ‘a glass of’). The scope of the diminutive
suffix in all examples is clearly the whole A+N. For instance, *een hoog hoedje* is a small top hat, not a small hat that is high, and *een rood wijntje* is a nice red wine or a glass of red wine, not a nice wine or a glass of wine that is red.15

This kind of mismatch is a much more general phenomenon, however. Classic examples are phrases like *transformational grammarian* related to *transformational grammar*, and *physical scientist* related to *physical science*. A transformational grammarian is not a grammarian who is transformational, but a linguist who adheres to the theoretical framework of transformational grammar. Hence, the semantic scope of the suffix *-ian* is *transformational grammar*, whereas its formal scope is the noun *grammar* only. This phenomenon is usually referred to as a ‘bracketing paradox’ since two different bracketing structures seem to be necessary (Spencer 1988):

(42) Form: [transformational]A[grammar-ian]N

The analysis proposed by Spencer is an essentially paradigmatic analysis of these facts since he invokes the notion ‘proportional analogy’ to account for this type of word formation (Spencer 1988: 675):

(43) \[
\begin{align*}
\text{grammar} & \leftrightarrow \text{grammian} \\
\updownarrow & \updownarrow \\
\text{transformational grammar} &  \\
? &
\end{align*}
\]

In this pattern, the question mark will be filled in as *transformational grammarian*.

This kind of interaction of lexical phrasal expressions with word formation can be observed in many other cases. Particle verbs in Germanic languages are phrasal in nature, even though they are lexical units (Booij 1990). The same applies to Italian phrasal verbs (Iacobini & Masini 2006). Some examples of this kind of mismatch between form and meaning are:

(44) (a) English particle verbs may carry the suffix *-er* on their verbal head: *look-er-on, runn-er-up, digg-ing up, switch-ing off the lights* (Elenbaas 2007: 9);
(b) The past participle of Dutch particle verbs is formed by prefixing *ge-* and suffixing *t/d/en* to the stem form of the verbal head: *aan-vallen ‘to attack’* → *aan-ge-val-len*, *op-bellen ‘to phone up’* → *op-ge-bel-d*; *ge*-nominalization also applies to the head: *rondspringen ‘jump around’* → *rondgespring ‘jumping around’* (Müller 2003, 2006);
(c) When German particle verbs undergo nominalization with the affix combination *ge-e*, this affix combination is attached to the verbal head of the particle verb: *herum-hopsen ‘to jump around’* → *Herum-ge-hops-e ‘jumping around’* (Müller 2003, 2006);
(d) In Italian nominalizing suffixes are attached to the verbal head (Masini, pers. comm.): *venire giù ‘to come down’* → *la venuta giù ‘the coming down’, mangiare fuori ‘to eat out’* → *la mangiata fuori ‘the meal at a restaurant’*;
(e) French *mettre en scène* ‘to direct a play’ – *metteur en scène* ‘director’ (example provided by an anonymous referee).
We can also observe this difference between the semantic scope and the formal scope of affixes in the following Dutch noun phrases:

(45) (a) muzikal-e vaardig-heid ‘musical ability’
     (b) digital-e competent-ie ‘digital competence’ (Booij 2009b)

In example (45a), the suffix -heid has semantic scope over muzikaal vaardig ‘lit. musically able, being musical’. In (45b) the suffix -ie has scope over digitaal competent ‘digitally competent’. Yet, the adjectives are not part of the base of the suffixes, as we can see from their being inflected.

These cases form evidence for the existence of paradigmatic relationships between constructional schemas. For instance, the schema for syntactic compounds of the type (39a), is paradigmatically related to the schema for diminutive nouns. This relation, symbolized by the sign $\approx$, may be expressed as follows (the brackets $<$ and $>$ demarcate the boundaries of a constructional schema):

(46) $< [A^0_i N^0_0 N^0_k] N^0_0 k > \approx < [A^0_i [N^0_j-tje] N^0_0] N^0_n m > \leftrightarrow \text{SEM}_k$ $\leftrightarrow \text{SMALL} \text{SEM}_m$

This paradigmatic correlation between two constructional schemas expresses that the diminutive suffixes attached to the head noun has semantic scope over the whole A+N combination: the semantic operator SMALL of the diminutive suffix has as its scope SEM$_k$, which stands for the meaning of the A+N phrasal name construction.

The paradigmatic analysis must be extended to the other types of phrasal expressions mentioned in (4). For example, if we pluralize the Italian $N N$ expression coda di cavallo ‘pony tail’, this will have the result code di cavallo, not coda di cavalli. That is, it is the head noun that is formally pluralized. Yet, the pluralization operation has semantic scope over the whole expression.

### 3 Theoretical implications: lexical phrasal constructions

The Greek and Dutch A+N phrasal names discussed above need to be listed in the lexicon, even though they are formed according to the rules or schemas of phrasal syntax, if they are the conventional names for certain concepts. Hence, these rules function as redundancy rules with respect to conventionalized A+N sequences. In this respect syntactic rules are not different from morphological rules or schemas. Recall that morphological rules function as redundancy rules with respect to existing, listed complex words (Jackendoff 1975), and can also be used to coin new words. The same holds for syntactic schemas. This implies that the lexicon may contain phrasal schemas.

Schemas are abstract patterns that express generalizations about sets of listed linguistic expressions. The morphological schema $[XN]_S$, for instance, generalizes over all (right-headed) Dutch nominal compounds. Such schemas form part of a hierarchical lexicon, in which schemas of different degrees of abstraction dominate their concrete instantiations as words (Booij 2009c). The individual words inherit by default the properties of the schemas to which they linked in the hierarchical lexicon (Sag et al. 2003: Chapter 16).
In the case of A+N phrases with naming function, the schemas 39 are instantiations of the more general schema for Dutch NPs. These constructional schemas are pairings of form and meaning. The specific meaning contribution of these constructions is that they are used for making names for concepts. This correlates with the specific formal property of the adjective being non-projecting.

The implication of this analysis is that there is no boundary between syntax and the lexicon: syntax permeates the lexicon because phrases can be lexical. The grammar is a network of syntactic and morphological constructions, with conventionalized instantiations of both types of schemas listed in the lexicon (Booij 2002b; Jackendoff 2002, 2007). A similar conclusion was reached by Sadler & Arnold in their analysis of AN combinations in English as ‘syntactic small constructions’:

If the analysis is right, it poses a serious challenge to the view that there is a strict separation of lexical and syntactic aspects of grammar: grammatical theory should recognize a kind of construction which is neither fully syntactic nor fully lexical, but has properties of both. (Sadler & Arnold 1994: 225).

Similar analyses may apply to the other types of phrasal names mentioned in section 2 under (4).

The link between a specific syntactic pattern and the naming function can also be observed in the case of apposition: an N followed by an N or NP without intermediate grammatical words such as prepositions or determiners. It is only when such patterns are used as names that the omission of grammatical words is permitted (and, we should add, in the special register of headline syntax, as mentioned above). As observed in Dutch, apposition is used for coining job names; it can also be used for names of committees, and all kinds of notions in the domain of government administration, etc.:

(47) [directeur]N0 [artistieke]A0 [zaken]N0N0 ‘director (of) artistic affairs’
    [Commissie]N0 [Kok]N0 ‘committee Kok, committee chaired by mr Kok’
    [Wet]N0 [Arbeids-Ongeschiktheid]N0 (WAO) ‘law labour-disabledness, law for people being unable to work’

These names are phrasal constructs, dominated by N0 (that is, I propose an interpretation of these constructs as syntactic compounds). The first example shows that the expression following the head noun can also be complex; it can be a phrasal name itself (in this example artistieke zaken ‘artistic affairs’). Such names are not just lexicalized phrases since these patterns are very productive for coining new (job) names.

Nice examples of the possible complexity of the post-nominal modifier in appositional names are found in Italian: naming expressions with coordinated nouns in the post-nominal position:

(48) (a) un marito pipa e pantofole (Bisetto & Scalise 1999: 32)
    a husband pipe and slippers
    ‘the type of husband characterized by smoking a pipe and wearing slippers’
(b) un marito tutto pipa e pantofole
    a husband entirely pipe and slippers
    ‘a type husband very much characterized by smoking a pipe and wearing slippers’
(c) una ragazza acqua e sapone (Masini, pers. comm.)
    a girl water and soap
    ‘a girl who does not use make-up’

These expressions are obviously phrasal names for specific concepts.16

In sum, (i) the notion ‘construction’ as used in various grammatical models (cf. Goldberg 2006; Jackendoff 2008; Sag et al. 2003) for both morphological and syntactic constructions, (ii) modeling the grammar of natural languages as a network of constructions (pairs of forms and meanings) of various degrees of abstractness, and (iii) assuming paradigmatic relationships between constructions, are essential ingredients for coping with the relation between the form and function of naming expressions.

Notes

1. I would like to thank Matthias Hüning, Barbara Schlücker, Francesca Masini, Angela Ralli, the anonymous referees, and the audience at the Freie Universität Berlin workshop on naming strategies for their constructive comments on earlier presentations of this text.
2. The term ‘word’ is to be understood here in the sense of ‘lexeme’, the abstract lexical unit which is realized by one or more concrete word (form)s.
3. The original Dutch text reads as follows: ‘Benoemen is het tot stand brengen van een verbinding tussen een talige uitdrukking en een begrip (bewustzijnsinhoud). De uitdrukking is vaak een woord, maar kan ook uit meerdere woorden bestaan.’ (Koefoed 1993: 3).
5. Exactly the same process is found in Russian mobil’nyj telefon > mobil’niki (Francesca Masini, pers. comm.).
6. In Booij (2009b) I argue that this principle is too strong in that syntactic rules may have access to word-internal structure but should not be allowed to manipulate parts of words, which is the essential point here.
7. As argued in Östman (2005), the form of constructions may be sensitive to specific styles or registers. Östman therefore argues that the principles of Construction Grammar must be extended to the discourse level.
8. In German A-N compounding is far more productive than in Dutch. An attempt to explain this difference can be found in Hüning (forthcoming).
9. The phenomenon of non-projecting words is also discussed in detail under the heading of légèreté (lite-ness) in recent work by Abeillé & Godard (Abeillé & Godard 2004, 2006), and this notion is also very relevant for the analysis of syntactic forms of noun incorporation (Abeillé & Godard 2004; Booij 2009a), and the analysis of particle verbs (Blom 2005; Booij 2002; Los et al. forthcoming).
10. Similar pre-nominal adjectives of French are analysed as ‘lite’, non-projecting adjectives by Abeillé & Godard (Abeillé & Godard 2004).
11. The spelling of the Greek words is that given in my source.
12. This observation concerning the order in which adjectives appear applies to such A+N combinations in many European languages (Heynderickx 2001).
13. For a similar analysis, see Abeillé & Godard (2004).
14. Thanks to Matthias Hüning for bringing to my attention that some language users accept plural neuter nouns with uninflected adjectives.
15. AN phrases may lexicalize into words. This is the case for expressions like Dutch wittebrood in which the inflectional ending -e is still there. Its diminutive is wittebroodje which shows that synchronically it is one word, otherwise we would get _wit broodje since broodje is a neuter noun. For some speakers of Dutch the same applies to rode kool ‘red cabbage’ which has to be written as one word according to the Dutch rules of orthography. Yet, rood kooltje ‘small red cabbage’ is the usual diminutive form of rode kool for Dutch language users, as a Google search will show. Other examples of the development of such A+N phrases into words are jongeman ‘young man’ and blindedarm ‘blind gut’ (cf. Booij 2002a: 314).
16. The construction of names by means of phrasal expressions is not restricted to the domain of nominal entities. Several languages have a similar mechanism for creating names for events, so-called quasi incorporation (Dahl 2004: Chapter 10). An example from Dutch is the NV combination piano spelen ‘to play the piano’, with the bare noun piano. These phrases with a bare noun are used for denoting nameworthy activities. That is, they have the function of names for events (for a more detailed analysis, cf. Booij 2009a). Again, we observe a correlation between the function of linguistic expressions as names and their formal properties.

References


Elenbaas, Marion 2007. The synchronic and diachronic syntax of the English verb-particle combination. Utrecht: LOT.


Author’s address: (Geert Booij) Leiden University Faculty of Letters P.O. Box 9515 2300 RA Leiden The Netherlands E-mail: g.e.booij@hum.leidenuniv.nl

DOI: 10.3366/E1750124509000427