Adjectival non-heads and the limits of compounding

Pius ten Hacken & Maria Koliopoulou, Innsbruck[1]

Nominal compounds are compounds with a noun as their head. In the most prototypical cases, the non-head is also a noun. However, we also find A+N expressions that are compounds. As adjectives are more commonly found as syntactic modifiers of nouns, this raises the question how to distinguish [A N] compounds from [A N] syntactic constructions. In this context, we address the issue of the borderline between word formation and syntax. Section 1 delimits the scope of the discussion, characterizing the types of A+N expressions we want to consider here. Section 2 discusses English, German and Dutch data and section 3 the corresponding data in Greek. In section 4, we turn to a special type of construction, which seems to be intermediate between compounds and phrases. Section 5 proposes an analysis in which not all A+N expressions that belong to word formation are compounds.

Keywords: *[A N] compounds, word-formation and syntax, English, German, Dutch and Greek*

1. Scope

The delimitation of compounding is a matter of debate. We will assume here that what falls into the category of compounding is determined by a combination of formal and semantic criteria and that the precise contribution of each type of criterion is in part an empirical matter. In this respect, the question of the delimitation of compounding is similar to the question of the boundary between inflection and derivation. As argued in ten Hacken (2014), the interplay of empirical and theoretical considerations is crucial. We should not let theoretical considerations override the data, but there is no theory-neutral approach, because any classification depends on analysis and the data do not include their own analysis. As such, questions such as the delimitation of compounding are of a type that occurs more generally for what ten Hacken (2015a) calls *terms in the narrow sense*. The issues involved are similar to those arising for concepts such as *planet*.

In this contribution, we focus on nominal compounds with adjectival non-heads. The boundary we consider is the one between the structures in (1).

(1)   a. $[A\ N]_N$
        b. $[A\ N]_\alpha$

We will call the structure in (1a) a compound. In (1b), the question is how to label α. If the structure is phrasal, α should not be N. However, NP is not an adequate label either. If noun phrases are the maximal projections of N, (1b) is lacking a determiner. In a version of the X-bar theory as assumed by Chomsky (1981), α should rather be labelled as N'. If we adopt Abney's (1987) DP analysis, the problem with the missing determiner disappears, but in modern versions, e.g. Alexiadou et al. (2007), the structure of a phrasal A+N combination is rather $[_{FP}\ AP\ [F\ NP]]$. In this structure, A and N are each embedded in a maximal projection, the NP is the complement of a functional category F and the AP is the specifier of F. Here we

will use α in (1b) as indicating a phrasal analysis, excluding α = N, without committing ourselves to a particular alternative value.

There are many different types of adjective. Some of these are illustrated in (2).

(2)   a. cheap car
       b. solar energy
       c. alleged thief

The adjectives in (2) belong to semantically different types. In (2a), we have a qualitative adjective, which expresses a property. This will be the type we concentrate on in this contribution. (2b) illustrates another type which is highly relevant in the context of compounding. Here, *solar* is a relational adjective, which expresses a relation to a corresponding noun, in this case *sun*. (2b) is parallel to *wind energy*, a typical [N N]$_N$ compound. Therefore, it is often argued, e.g. by Levi (1978) and ten Hacken (2013a), that (2b) should be considered a compound. We will not focus on relational adjectives here.

Apart from qualitative and relational adjectives, there are other types, of which (2c) is an example. Here, *alleged* is not a property assigned to *thief*, but a qualification of the label *thief*. Such adjectives can never be part of a compound with a structure (1a).
A final restriction on the domain of discussion is headedness. We will not be concerned here with exocentric compounds of the type *loudmouth*. The delimitation problem does not pose itself in the same way for such items. In ten Hacken (2010), it is argued that they should be analysed as the result of conversion of structures of the type in (1b), which makes them instances of derivation rather than compounding.

2. [A N] structures in Germanic languages

In English, Marchand (1969: 63-65) distinguishes six types of endocentric [A N]$_N$ compounds, illustrated in (3).

(3)   a. blackbird
       b. grandfather
       c. Englishman
       d. freedman
       e. New England
       f. blacksmith

For all of these types, synchronic productivity is dubious. Marchand (1969: 64) claims in particular that (3a) has not been productive for 100 years and (3c) for 300 years. (3b) is limited to a very small set of words, where *grand* could be analysed rather as an affix than as an adjective, because it has a specialized meaning that does not correspond to the meaning as an adjective. (3d) only appears in Old English poetry. (3e) is only used in proper names. The difference between (3a) and (3f) is the relation between the two components of the compound. Whereas in (3a), the A gives a property of the N, in (3f) the A designates a property of something related to the N. Bauer et al. (2013: 434-435) briefly mention [A N] compounds and give some examples corresponding to Marchand's types (3a) and (3f).

In German, the situation is rather different. We also find compounds of the types (3a) and (3f), but in contrast to English the rule responsible for them is without any doubt productive. We can therefore find many contrasts of the type in (4).

(4) a. Altstadt   ('old_town/city')
     b. alte Stadt   ('old town/city')

The semantic difference between (4a) and (4b) is significant. Whereas the phrasal (4b) designates a *Stadt* ('town/city') and predicates the property *alt* ('old') over it, the compound (4a) designates a part of the *Stadt* for which *alt* holds. This can be seen in the contrast in (5).

(5) a. Innsbruck hat eine Altstadt.
         'Innsbruck has an old_town/city'
         i.e. Innsbruck has a historical centre
     b. Innsbruck ist eine alte Stadt.
         'Innsbruck is an old town/city'

Whereas the semantic contrast in (5) is obvious, in many cases it is rather more subtle. A good example is (6).

(6) a. Rotwein   ('red_wine')
     b. roter Wein   ('red wine')

Because red wine has a red colour, the two expressions in (6) end up having almost the same meaning. The difference in meaning is still significant, however. Whereas the compound (6a) designates a type of wine, the phrase (6b) designates a wine with a red colour. This interpretation is supported by the corresponding contrast in (7).

(7) a. Weisswein   ('white_wine')
     b. ²weisser Wein   ('white wine')

As white wine does not have a white colour, the phrase (7b) designates wine of a rather unnatural type. This judgement is supported by data from the *Deutsche Referenzkorpus* (DeReKo).² Whereas the compound (6a) has 25,192 occurrences, the phrase (6b), including all inflected forms of the adjective, has only 518. For (7), the overall figures are lower, but the proportion is even more striking. We found 2,888 occurrences for (7a) against 36 for (7b). The compound is 49 times more frequent in (6) and 80 times in (7).³

In English, the contrast between *blackbird* and *black bird* is of the same type as (4). For *red wine* and *white wine*, there is only one form, which is used in both senses illustrated in (6) and (7). The question is whether we should then consider *red wine* and *white wine* compounds when they designate a type of wine, i.e. when they correspond to (6a) and (7a). In

---

² The *Deutsche Referenzkorpus* ('German reference corpus') is a large, dynamic corpus of German maintained by the Institut der Deutschen Sprache. It is available under https://cosmas2.ids-mannheim.de/cosmas2-web/ . Figures for query data are correct as of 7 December 2015.
³ We do not know why (6) is so much more frequent than (7). Interestingly, when the data are divided by country, the figures for Switzerland are 2,686 for (6a) and 2,777 for (7a). Switzerland accounts for 10.7% of the occurrences of (6a), but 96.2% of the occurrences of (7a). All occurrences of (7b) are from Switzerland.

terms of stress, they tend to behave like phrases. However, stress is a notoriously difficult criterion to apply (cf. ten Hacken 1994: 32-34). It is not only language-specific, but also subject to variation due to syntactic and pragmatic factors. A much more robust criterion is the lack of inflection on the adjective. However, English adjectives do not have inflectional endings. Hüning (2010: 206) proposes the hypothesis that the English use of phrases as in (1b) is linked to the lack of inflection on the adjective. In any case, inflection cannot be used to determine whether *red wine* has a structure as in (1a) or (1b).

In Dutch, adjectives have inflection, but it is restricted to the contrast between the presence or absence of the ending *-e*. This is illustrated in (8).

(8)   a. een hoog huis   ('a high house')
      b. een hoge boom   ('a high tree')
      c. hoge huizen   ('high houses')
      d. het hoge huis   ('the high house')

The ending *-e* appears when the adjective modifies a non-neuter noun, e.g. *boom* in (8b), when it modifies a plural noun as in (8c), and when the noun phrase it appears in is definite as in (8d). In (9) we give a Dutch example of a contrast similar to the German contrast in (4), in which the same A+N combination has the two structures in (1).

(9)   a. hoogspanning   ('high_tension')
      b. hoge spanning   ('high tension')

The form in (9a) is clearly the preferred form when *spanning* refers to electric tension (i.e. *voltage*). Here *hoog* is not a scalar modifier, but classifies the tension as belonging to the type that is used for the long-distance transmission and distribution of electricity. When describing, for instance, the tension of tyres, only (9b) can be used. The compound (9a) is also used by extension for the type of tension that can arise, for instance, between parties in negotiations. In many other contexts, *spanning* is rather modified by *groot* ('great') than by *hoog*. This explains that in the *Corpus Hedendaags Nederlands*, the frequency of (9b) is much lower than that of (9a).[4]

The productivity of A+N compounding in Dutch is a matter of debate. Hüning (2010: 204-205) describes that A+N compounds marked by the absence of inflection on the A have a long history in Dutch. In the 20th century, they were often associated with German, so that they were avoided in favour of phrasal alternatives. At the time, A+N compounds were sometimes claimed to be German borrowings in Dutch, but the attestation of fairly large numbers of A+N compounds in 15th and 16th century Dutch makes such claims dubious. Following Corbin (1987), we take the central question of productivity to be whether the word formation process in question is synchronically available for the naming of new concepts. Especially for processes with a long historical pedigree, the perspective may easily be blurred by the existence of words formed at earlier times.[5]

---

[4] The *Corpus Hedendaags Nederlands* is a large corpus of Dutch developed by the *Instituut voor Nederlandse Lexicologie*, https://portal.clarin.inl.nl/search/page/help . For *hoogspanning* it gives 634 hits, for *hoge spanning* only 31 (searches 13 Dec. 2015).
[5] Schuster (2016) presents a historical contrastive study of Dutch and German A+N compounds. As her analysis does not take as its starting point the type of productivity we consider as central and also departs from our

Corresponding to (6), Dutch does not have *roodwijn*, but only *rode wijn*. This implies that, like in English, the semantic distinction between the two German expressions in (6) cannot be expressed in Dutch. In many cases, lexicalized A+N compounds in German correspond to expressions using different naming patterns in Dutch. An example is (10).

(10)	a.	Hochhaus	('high-rise building, sky scraper')	DE
	b.	torenflat	('tower block')	NL
	c.	wolkenkrabber	('sky scraper')	NL

Whereas German (10a) is an A+N compound corresponding to the phrase in (8a), Dutch has different naming strategies. In (10b) we find an N+N compound using a different image and in (10c) a synthetic compound using the same image as in English.

Evidence of another type comes from the naming of new, non-lexicalized concepts. Ten Hacken (1994: 99-100) discusses the German example in (11).

| (11) | a. | Rotbus | ('red_bus') |
| | b. | roter Bus | ('red bus') |

Even if it may not be straightforward to identify the concept in the actual world, it is clear that (11a) refers to a particular type of bus. This contrasts with (11b) referring to a bus with a particular colour. (11a) can be chosen as the name for a bus in a particular programme to improve public transport and increase its use. If in an upgrade of the style the buses are no longer red, but for instance white with a red logo for the programme, they can still bear the name *Rotbus*. In Dutch, it would not be conceivable to use **roodbus* parallel to (11a). If a naming pattern with an adjective is chosen, only phrases such as *rode bus* corresponding to (11b) are possible.

On the basis of the data discussed here, we propose that A+N compounding of the type exemplified in (9a) is no longer productive in Dutch. The corresponding naming function has at least in part been taken over by A+N combinations of the type exemplified by *rode wijn* and *rode bus*, in the same way as we saw for English. In other cases, as illustrated in (10), other naming patterns are used. This raises the question of whether A+N expressions with a naming function should be analysed as compounds despite the inflection of the adjective, i.e. whether the ambiguity between naming expression and descriptive expression corresponds to an ambiguity between word formation and syntax for these expressions.

3. [A N] Structures in Greek

In order to broaden the basis of the discussion beyond the Germanic languages, we consider Modern Greek (hereafter Greek). Greek has many compounds and an extensive system of inflection. Pairs illustrating the contrast between compounds and phrases, corresponding to German *Altstadt* and *alte Stadt* in (4) are given in (12) and (13).

---

theoretical assumptions in other ways, a detailed comparison of the results would be beyond the scope of this article.

(12) a. ελαφρόπετρα[6]   ⇐   ελαφρ(ιά)   πέτρ(α)
          pumice stone   light   stone
      b. τυφλοπόντικας   ⇐   τυφλ(ός)   ποντικ(ός)
          mole   blind   mouse
      c. κοκκινόχωμα   ⇐   κόκκιν(ο)   χώμ(α)
          clay soil   red   soil

(13) a. ελαφριά πέτρα   ('light stone')
      b. τυφλός ποντικός   ('blind mouse')
      c. κόκκινο χώμα   ('red soil')

A+N expressions in Greek are mainly one-word compound formations, e.g. *ελαφρόπετρα*
(12a), or phrases like *ελαφριά πέτρα* (13a). Both types of structure are very productive and
right-headed. However, they can easily be distinguished because of their different morpho-
syntactic, phonological and semantic characteristics. In many cases there are comparable
structures, as shown in (12) and (13), but not all syntactic phrases have an equivalent
compound structure. There is no corresponding compound, for instance, for the phrase
μεγάλος κήπος ('big garden', *μεγαλόκηπος*). Moreover, as shown by the glosses in (12) and
(13), compounds have in most cases a different meaning compared to the meaning of the
corresponding syntactic phrase. Therefore, the distinction between compounds and phrases is
clear in these cases.

Compounds in Greek are one-word formations mostly formed out of stems. Due to
the preference for stem constituents, Greek compounds are written as one word. Their
graphemic representation correlates with the strength of the morphological unity as well as
with the phonological properties of these structures (Ralli 2013: 22). They display a high
degree of lexical integrity, not permitting syntactic operations access to their internal
structure. A key difference between Greek compounds and their Germanic counterparts in
section 2 is that the former consist of stems and the latter of words.

In Greek, a compound is a morphological and phonological word with a single main
stress. Stress is orthographically marked in Greek. The position of the stress varies according
to the morphological category of the right-headed component. Specifically, in case the right-
headed component is already an inflected word, it keeps its properties and bears the stress of
the whole structure. If the right-headed constituent is a stem, a compound-specific rule is
applied according to which stress is assigned to the antepenultimate syllable of the compound
(Nespor & Ralli 1994: 201, 1996: 357, Revithiadou 1997, 1999: 183). A+N compounds are
formed on the basis of two patterns, [stem-stem] and [stem-word]. In stem-stem formations,
like those given in (12), stress is assigned to the antepenultimate syllable, whereas the stress
position of the stem-word structures depends on the stress of the head noun, as shown in (14).

(14) a. χαζοπαρέα   ⇐   χαζ(ή)   παρέα
          pals, gang, regular company   stupid   company
      b. μεγαλοκοπέλα   ⇐   μεγάλ(η)   κοπέλα
          unmarried, old woman   old   girl

---

[6] Greek examples are given in the Greek script with an English translation. All compounds are decomposed into
their constituents. Suffixes are given in parentheses only when a compound constituent appears as a stem in the
compound structure.

c. μεγαλομανία   ⇐   μεγάλ(η)   μανία
   mania for grandiose performance   big   mania

In both types of structure there is a linking element. The appearance of a linking element is a common characteristic in the formation of compounds in several languages, including also German and Dutch. In Greek compounding, there is a single form of the linking element, *-o-*. Its appearance between the compound constituents is obligatory, since it is responsible for the linking of these components, which are mostly stems. Its exceptional absence is determined by specific word formation rules (cf. Ralli 2013: 47-53, 56-71). In the case of A+N compounds, this element links the adjectival stem with the second compound constituent. The linking element cannot be considered as an inflectional ending since in most of the cases presented above (12a-b, 14a-c) the inflectional suffix of the adjective is different from the form of the linking element.

      Contrary to compounds, the syntactic phrases in (13) consist of two independent words. In these [A N] structures, the adjective bears an inflectional ending, which obligatorily agrees with the inflectional features of the noun, namely of the head of this structure. In (13a), the adjectival stem *ελαφρ-* is followed by the feminine inflectional suffix *-ια* which agrees in number, case and gender with the noun *πέτρα*. Stress is assigned in both constituents; therefore the structure has a phrasal stress. With regard to the semantic properties, *ελαφριά πέτρα* refers to a stone in general which has the property of being light. The meaning of the structure is fully compositional, as in most syntactic phrases.

      As opposed to syntactic phrases, which usually display a transparent meaning, the meaning of compound structures is not always predictable on the basis of the meaning of the two components. Taking as an example *ελαφρόπετρα* (12a), we observe that the meaning of the compound is specialized. *Ελαφρόπετρα* denotes a particular type of stone. Because of its porous nature and its low density, even a rather big pumice stone is particularly light. This means that *ελαφρόπετρα* is a rather light stone, as the meaning of the compound constituents declare, not because of its size but because of its nature. Specialized meaning is a typical characteristic of Greek compounds, as also shown in the comparable structures between (12) and (13). Therefore, compositionality of meaning can be used as indicative evidence to distinguish compounds as morphological structures from syntactic phrases.

      Semantic compositionality in compound formations is a scalar characteristic since the degree of compositionality varies among compounds (Ralli 2013: 18-19). Highly opaque compounds have a meaning which does not directly derive from the components, as for instance in examples given under (12). Semantically transparent compounds can be interpreted easily since the whole formation designates a concept that can be predicted fairly well on the basis of the sum of the meanings of the two components, as for instance in *ασχημόπαπο* ('ugly duckling'), whose meaning is not significantly different from the combination of the meanings of the compound constituents (*άσχημ(ο)* 'ugly', *παπ(ί)* 'duckling').

      As far as A+N compounds are concerned, most structures display an intermediate degree of semantic compositionality, since the meaning of the entire formation is related to the meaning of the two components. However, comparing the compound structures in (12) with the corresponding phrases in (13), it can be argued that compound formations have an additional or differentiated meaning, as, for instance, in *ελαφρόπετρα* (12a) compared to *ελαφριά πέτρα* (13a). Another example is the compound *μεγαλομανία* (14c) which denotes a certain type of mania and not only a 'big mania' as the corresponding phrase *μεγάλη μανία*

would designate. Here we have a relation between the two components that is similar to *blacksmith* in (3f), where the A does not modify the N, but something related to the N.

In some cases, we find a further semantic difference in the sense that the meaning of the compound has a positive connotation, whereas the meaning of the corresponding phrase has a negative one, or vice versa. For instance, the structure χαζοπαρέα (14a) has a rather positive connotation, denoting a group of friends. The phrase consisting of the same components, namely χαζή παρέα ('stupid company'), has a negative meaning. Conversely, the compound μεγαλοκοπέλα (14b) clearly has a negative meaning – at least in the way the formation is usually used in a Greek context – while the corresponding phrase μεγάλη κοπέλα ('big girl') has rather a positive meaning.

Comparing the two A+N structures in Greek presented here, we argue that there is a clear distinction between compounds and syntactic phrases. The distinguishing criteria are based on their different morpho-syntactic, phonological and semantic properties, summarized under (15).

(15)   a.   Compounds are one-word formations mainly formed out of stems, whereas phrases consist of two independent words;
       b.   Compounds have a single main stress, as opposed to the phrasal stress of syntactic phrases;
       c.   Compound constituents are linked with the element *-o-*, whereas phrases display inflectional agreement;
       d.   Compounds tend to have a specialized meaning, contrasting with the compositional meaning of phrases.

4. Intermediate [A N] structures in Greek and Dutch

In addition to the contrast between compounds and phrases discussed so far, we have an intermediate type of A+N construction, for which the diagnostics are particularly clear in Greek. They are included in what Booij (2010: 169-192) calls *phrasal names*, but Booij's concept is much broader and also encompasses combinations with relational adjectives. Some Greek examples are given in (16).

(16)   a.   ψυχρός πόλεμος   ('cold war')
       b.   τρίτος κόσμος   ('third world')
       c.   μαύρη τρύπα   ('black hole')
       d.   μαύρη αγορά   ('black market')

Anastassiadis-Symeonidis (1986, 1994) shows that these structures have been introduced in the Greek vocabulary relatively recently and originate from corresponding English and French names. They often cover terminological needs, e.g. (16c). The classification of expressions such as (16) has been a matter of debate among Greek linguists. In order to assign them a label that is distinct from *compound* and *phrase*, various names have been proposed, including *lexical phrase* (e.g. Anastassiadis-Symeonidis 1986), *multi-word compound* (e.g. Ralli & Stavrou 1998), *loose multi-word compounds* (e.g. Ralli 2007, Koliopoulou 2008, 2009) and *phrasal compounds* (e.g. Ralli 2013, Bağriaçik & Ralli 2015).

Structurally, the expressions in (16) look like syntactic phrases, because they are composed of an inflected adjective and a noun. Both components are phonological words and the adjective agrees with the noun in number, case and gender. Semantically, however, they are quite different from regular syntactic [A N] phrases, because they display a high degree of semantic opacity, which is more typical of compounds. It is especially the adjectival component that is interpreted in a non-literal sense which contributes to the specialized meaning.

In Dutch, we find similar expressions. The Dutch translations of (16) are given in (17).

(17)   a. koude oorlog   ('cold war')
        b. derde wereld   ('third world')
        c. zwart gat   ('black hole')
        d. zwarte markt   ('black market')

The expressions in (17) are very similar to the ones discussed in section 2, e.g. *rode wijn* ('red wine'). In Greek, the semantic specialization of the expressions in (16) correlates with a number of characteristic structural properties that distinguish them from regular [A N] phrases. These properties have been studied, for instance, by Anastassiadis-Symeonidis (1986), Ralli & Stavrou (1998: 244-149), Koliopoulou (2009: 62-63), Ralli (2013: 244-248) and Bağriaçik & Ralli (2015: 348-349). We will consider the tests in (18).

(18)   a. independent modification of the adjectival non-head
        b. insertion of an element between the components
        c. doubling of the definite article of the structure
        d. reversing of the order of the adjective and noun

Let us first consider (18a-b) and their application to (16a). In (19), we find the contrast between the intermediate structure *ψυχρός πόλεμος* and the syntactic phrase *ψυχρό κλίμα* ('cold climate').

(19)   a. πολύ ψυχρό κλίμα   ('very cold climate')
        b. *πολύ ψυχρός πόλεμος   ('very cold war')

When the intensifier *πολύ* ('very') is added, *ψυχρός* ('cold') can no longer be interpreted in the specialized sense it has in (16a). This effect can also be observed in the English translation and in Dutch. The only case in which (19b) and its translations may occur is when a word play on the literal and specialized senses of the adjective is intended. However, the use of intensifiers is not possible for all adjectives. Thus, for *τρίτος* ('third') in (16b), an intensifier is also excluded with regular syntactic phrases. Apart from ordinal numerals this can also be observed for inherent superlatives (e.g. *huge*), degree adjectives (e.g. *absolute*) and modal adjectives (e.g. *alleged*). As the test is semantically based, it gives similar results in many languages.

The test in (18b) can be carried out by inserting another adjective in between the two components. However, in many cases, this triggers the insertion of coordinating *και* ('and'). Therefore, the contrast in (20) is not a pure reflection of this test.

| (20) | a. | ψυχρό και υγρό κλίμα | ('cold and humid climate') |
| | b. | *ψυχρός και αιματηρός πόλεμος | ('cold and bloody war') |

In (20a) it would be very unnatural to leave out *και*. What (20b) shows is rather that *ψυχρός* ('cold') and *αιματηρός* ('bloody') have a different type of semantic relation to *πόλεμος* ('war') and that *ψυχρός* and *πόλεμος* cannot be separated. The semantic nature of the constraint is shown by (21).

| (21) | τρίτος και τέταρτος κόσμος | ('third and fourth world') |

If *τέταρτος κόσμος* ('fourth world') is interpreted in a way parallel to *τρίτος κόσμος* ('third world'), i.e. as an extension of the three-world model, (21) is perfectly grammatical. Coordination is not an adequate test for compoundhood, as it is generally available to compounds in languages such as German, Dutch and English, e.g. *travel and history books*.

  Examples of noun phrases with two modifying adjectives that are not coordinated occur in Greek as well, but they are in other ways restricted. Whereas *και* can hardly be omitted in (20a), (22) is unproblematic.

| (22) | ένα καινούργιο μπλε αυτοκίνητο | ('a new blue car') |

Whereas in (20a), *ψυχρό* and *υγρό* are first combined into one constituent, which then modifies *κλίμα*, in (22), the last adjective is combined with the noun in a first step and *καινούργιο* ('new') then modifies *μπλε αυτοκίνητο* ('blue car'). Obviously, we expect that no adjective can intervene in the expressions in (16) in the way *μπλε* ('blue') does in *καινούργιο αυτοκίνητο* ('new car'). However, this criterion is of limited use. First, as we saw in the context of (20a), it is not always possible to insert an adjective without adding *και*, and this does not make *ψυχρό κλίμα* ('cold climate') a compound. Secondly, reversing the two adjectives in (22) results in a rather unnatural expression, but this does not make *μπλε αυτοκίνητο* ('blue car') a compound. As in the case of (18a), also the observations concerning (18b) apply in a similar way to a wide range of languages. The Dutch translations of (21) and (22) have broadly the same properties.

  The test in (18c) refers to a peculiar construction in Greek, illustrated in (23).

| (23) | a. | η μαύρη τσάντα | ('the black bag') |
| | b. | η μαύρη η τσάντα | ('the black the bag') |

The meaning of (23b) is not different from that of (23a). The difference is one of emphasis. (23b) highlights the property designated by the adjective. It can be used for contrast, but this is not necessary. It is then not entirely surprising that the construction in (23b) is not available for the expressions in (16). Thus, for (16c) we find the contrast in (24).

| (24) | a. | η μαύρη τρύπα | ('the black hole') |
| | b. | *η μαύρη η τρύπα | ('the black the hole') |

Another way of highlighting the property expressed by the adjective is to reverse the order of the noun and the adjective. For the same pair of A+N expressions as (23) and (24), (25) gives an example of this contrast.

(25) a. Μια τσάντα μαύρη αγόρασε η Μαρία.
          'A bag black bought the Maria'
          i.e. Maria bought a black bag
      b. *Μια τρύπα μαύρη εντοπίστηκε πρόσφατα.
          'A hole black was found recently'
          intended: A black hole was discovered recently

The construction in (25a) highlights the adjective even more than (23b). It is stylistically marked and often used in contrast. It can be compared to the French contrast in (26).

(26) a. une soirée agréable   ('an evening pleasant')
      b. une agréable soirée   ('a pleasant evening')

It is difficult to describe the difference in meaning between the two expressions in (26). In French, adjectives normally follow the noun, but the reverse order is possible if the adjective has a strong affective load (Grevisse 1980: 433).[7] Clearly, this is not possible for the type of intermediate constructions illustrated in (16).

Word order variation in the A+N combination is crosslinguistically less marked than the article-doubling in (23), but in Germanic languages it does not generally occur. A construction that can be used to highlight the property designated by the adjective in a somewhat comparable way is the use of a relative clause as in (27) for Dutch.

(27) a. een zwarte tas   ('a black bag')
      b. een tas die zwart is   ('a bag that black is', i.e. …that is black)

A relative clause highlights the property in a way that is not compatible with the naming function of the A+N combination in cases such as (28).

(28) a. een gat dat zwart is   ('a hole that black is')
      b. wijn die rood is   ('wine that red is')

The description in (28a) cannot be interpreted as referring to a black hole in the astronomical sense, but only to a hole that has a black colour. Similarly, whereas *rode wijn* ('red wine') is ambiguous between the type of wine and wine with a particular property, a contrast we found expressed in (6) for German, (28b) can only refer to wine with a particular property, *roter Wein* in German.

A property that cannot be used as a test, but suggests that A+N combinations of the type in (16) are similar to compounds is that these expressions can be input to derivation in the same way. We can compare the derivations from A+N compounds in (29) to the derivations from (16) in (30).

---

[7] In the original formulation: "il [the adjective] peut le [the noun] précéder s'il a beaucoup de force affective." In a later edition, Grevisse & Goosse (2008 : 408-409) replace this remark by an account of the attempts to express "force affective" by more theoretically developed notions.

(29) a. χαζοπαρέα   ⇒   χαζοπαρεάκι
          'pals…'   'pals …DIM'
      b. μεγαλοεπιχειρηματίας ⇒ μεγαλοεπιχειρηματικός
          'big entrepreneur'   'big entrepreneurADJ'

In (29a) we find the compound discussed as (14a) above. In the derivation it gets a diminutive suffix -άκ(ι). In (29b), a relational adjective is derived from an A+N compound by means of the suffix -ικ(ός).

(30) a. τρίτος κόσμος   ⇒   τριτοκοσμικός
          'third world'   'third worldADJ'
      b. μαύρη αγορά   ⇒   μαυραγορίτης
          'black market'   'black marketeer'

In (30a), the same suffix as in (29b) is applied to (16b). In (30b), the suffix -ίτ(ης) is applied to (16d).[8] Whether this behaviour distinguishes the expressions in (16) from other A+N combinations is a further question. In Dutch, we find examples such as (31).

(31) a. Franse taal   ⇒   Franstalig
          'French language'   'French-speaking, in French'
      b. twee talen   ⇒   tweetalig
          'two languages   'bilingual'

The suffix -ig is a general adjective-forming suffix in Dutch, comparable to -ed in English and -ικ(ός) in Greek. In (31a), it is applied to an A+N combination for which we might wonder whether it belongs to the same class as (17). The resulting adjective can apply to people as well as to texts, illustrating its general meaning. In (31b), the same suffix is applied to a phrase consisting of a numeral and a noun, which definitely does not belong to the same class. In ten Hacken (2003), it is argued that the inputs to the formation process in (31) are Morphological Phrases, a combination of two words that does not have a syntactic distribution because it lacks a syntactic category. It should be noted that the phrases given as input in (31) contain inflectional material that is not realized in the output. As syntactic phrases, *Frans taal and *twee taal are ungrammatical. Therefore, these phrases have to undergo a word formation process that assigns them a syntactic category. In Greek, the absence of inflection on the adjectives in τριτο- and μαυρ- in (30) can be interpreted in the same way.[9]

Greek A+N expressions of the type in (16) are in many ways similar to A+N compounds and unlike regular A+N phrases. A central property that aligns them with compounds is that they are used for naming rather than for description (Anastassiadis-

---

[8] The suffix -ίτ(ης) selects a nominal base to indicate a person. According to the Triantafyllidis (1998) Dictionary of Common Modern Greek, it signifies the origin or the residence of a person (e.g. Ρόδος 'Rhodes' ⇒ Ροδίτης 'male person from Rhodes') also used as a last name, or a property related to the meaning of the nominal base (e.g. ισόβια 'life sentence' ⇒ ισοβίτης 'lifer'). The nominal base selected by the suffix -ίτ(ης) can also be a A+N expression, as in (30b), or an A+N phrase, as in ξενομερίτης ('outlander, foreigner') deriving from the syntactic phrase ξένο μέρος ('foreign place').

[9] In (30a), the -o- in τριτο- is not an inflectional ending, but a linking element. In (30b), the linking element is elided in front of the initial vowel of the head.

Symeonidis 1986: 142-143). This semantic property is then reflected in a range of syntactic properties listed in (18). Although not all of these can be used as tests and in several cases there are independent explanations for the observations, they provide evidence that the way the meaning is acquired influences syntactic possibilities. In phonological and morphological properties, however, the expressions in (16) behave like regular phrases. They have phrasal stress and inflectional agreement. The discussion of the Dutch counterparts in (17) shows that the syntactic properties are not in all cases the same, but that the division between semantic and syntactic similarities with compounds and phonological and morphological similarities with regular phrases is broadly the same.

5. The boundary between word formation and syntax

For linguists who want to distinguish word formation and syntax, the ambiguity of A+N expressions such as English *red wine* is somewhat of an embarrassment. As we saw in section 2, there are two German equivalents, repeated here as (32).

(32)   a. Rotwein   ('red_wine')
        b. roter Wein   ('red wine')

If *red wine* is taken as a naming unit, it is equivalent to (32a). If it is a descriptive phrase, it is equivalent to (32b). In our view, naming units have a different status to descriptive phrases. The formation of a naming unit is a change in competence, whereas the formation of a descriptive phrase is a performance process. In ten Hacken (2013b), it is argued that this difference justifies a separate word formation component. How, then, do we propose to account for expressions that are ambiguous as to the place of their formation?

There are several possible 'simple' solutions we would like to exclude. One is giving up the distinction between word formation and syntax, as proposed by Booij (2010). Another is to reduce word formation to a phenomenon based on phonological and inflectional properties. On this basis, Matthews (1974: 35) argues that Latin *tribunus militaris* (lit. 'tribune$_N$ military$_{ADJ}$', a function in the Ancient Roman army) is an idiom rather than the result of word formation. As noted before, combinations of relational adjectives and nouns, e.g. *solar energy*, get their meaning in the same way as N+N compounds, e.g. *wind energy*. There is no problem in considering both as the result of word formation, i.e. the rule-based formation of naming units.

Our objective was first of all to explore the borderline area between word formation and syntax for A+N combinations in a number of languages. In German, A+N compounding is productive in the sense that it is available for new formations. For English and Dutch, there are good reasons to assume that the corresponding process is no longer productive. In these languages, A+N expressions are used for providing naming units, thus creating the ambiguity of phrases such as *red wine*. Greek turned out to be particularly interesting. On one hand, A+N compounding is productive in Greek. On the other, A+N phrases are also used as naming units. Moreover, there are a number of syntactic properties these phrasal naming units share with compounds rather than with descriptive phrases. Not all of these are crosslinguistically applicable as tests, but this does not mean that the distinctions they demonstrate does not exist in other languages.

A crucial issue in this context is the question of lexicalization. As shown by Thomas (2013a), there are a number of conflicting views of what lexicalization is and how it takes place. Here we will interpret the term as referring to the process of storing an expression in the lexicon. In a competence-based view of language, this means that a particular speaker changes their lexicon to include an additional entry. Two speakers may have a different mental lexicon without noticing this in communication, because one of them has a complex expression stored in competence, whereas the other constructs this expression online. The performance is the same in both cases. Jackendoff (2002) and Booij (2010) assume that the only reason why lexicalization of a complex expression takes place is because it is prominent enough. It is certainly plausible that speakers make unconscious decisions to lexicalize frequent or otherwise prominent expressions. This is obviously the case for what Wray (2002) calls *formulaic language*. Examples are the expressions in (33).

(33) a. ladies and gentlemen
      b. it is a great pleasure for me to introduce

Most speakers of English will have (33a) stored in their lexicon, as can be shown quite easily by the uncommonness of the expression with the two nouns in reverse order. Also the expression in (33b) will be stored by many speakers. These expressions are not naming units, they are not words, and (33b) is not even a phrase. They do not have a specialized meaning. However, they are frequent enough to make it efficient to store them.

However, formulaic language is not the only source of extending the lexicon. Ten Hacken (2013b) gives the example of *image converter* as an instance of ostensible naming. Even if we do not know what this word means, it presents itself as a name of something. The word formation process used in it gives clues about the range of meanings, but a crucial step in determining its meaning is what ten Hacken (2015b) calls *onomasiological coercion*. In naming, we first have a concept. Whatever name is adopted will have as its meaning this concept. If word formation rules are used to produce a name, they will determine the appropriateness of the name by constraining the expected range of meanings, but such constraints can be overridden. We can consider *spaceship* a headed compound, even though a spaceship is not really a ship. It is onomasiological coercion that distinguishes naming from description.

Adopting Jackendoff's (2002) Parallel Architecture (PA) as a background framework, ten Hacken (2015b) distinguishes seven possible types of rule that would change existing lexicon entries. Each of the phonological, syntactic and conceptual structures can be modified or left unchanged independently of the others, but if none of the three is changed, we do not have a rule. If word formation is motivated by naming, only the types that change conceptual structure can be included in the word formation component. There are four such types, corresponding to whether or not they change the phonological and the syntactic components. Three of them are exemplified in (34).

(34) a. Type 7 (111): *drive* $\Rightarrow$ *driver*
      b. Type 5 (101): *able* $\Rightarrow$ *unable*
      c. Type 3 (011): *clean*$_A$ $\Rightarrow$ *clean*$_V$

135

The three types in (34) correspond to category-changing affixation, non-category-changing affixation, and conversion.[10] The numbers of the types indicate in their binary forms which of the three structures are affected by the rule. They assume the common representation of PA with phonological structure on the left, syntactic structure in the middle and conceptual structure on the right. The absence of phonological change in (34c) makes it easy to identify it in contrast to (34a-b). The contrast between (34a) and (34b) can be further illustrated by the diminutive in Dutch and in Spanish. In Dutch, all diminutives are neuter, so that the rule belongs to Type 7. In Spanish, the gender depends on the gender of the noun that undergoes the rule, so that the rule is in Type 5.

There is one more type in which conceptual structure is changed, Type 1 (001). Here, the conceptual structure is the only representation that is changed by the rule. As the study of word formation has always been preoccupied mainly with form, this type has not usually been recognized as a type of word formation. Ten Hacken (2015b) proposes to use it, for instance, for the regular process/result alternation. The idea is illustrated in (35).

(35) a. In many ways, the crusades helped to civilize Europe.
      b. The civilization of Europe made significant progress in the period of the crusades.
      c. Western civilization is nowadays upheld as a standard.

In (35b) we have a pure transposition. *Civilization* refers to an action that is referred to by the verb *civilize* in (35a). In (35c), *civilization* designates the result of this action. The process that turns the verb *civilize* in (35a) into the noun *civilization* in (35b) is a pure transposition. It does not change the meaning. Both refer to the same event, but one is a noun and the other a verb. In ten Hacken's (2015b) system, it is a Type 6 (110) rule, which does not belong to word formation. The process that changes *civilization* as an event in (35b) to *civilization* as the result of this event in (35c) is a Type 1 rule. It only changes the meaning. The extensive analysis of English and French nouns in *-ation* by Thomas (2013b) demonstrates that the formation of event nouns as in (35b) is more basic than the formation of result nouns as in (35c), which supports the idea that the Type 1 rule builds on the result of the Type 6 rule.

When we now consider the implications for A+N naming units such as *red wine*, we can analyse them as the result of a Type 1 rule operating on a syntactic construction. We propose that the classification of A+N expressions as belonging to word formation or to syntax should be based on their function. The domain of word formation encompasses rule-based expressions intended as names. For A+N expressions this means that they belong to word formation if they are used for naming, and to syntax if they are used for description. In their descriptive use, they are analogous to the transposition in (35b), but in their naming use, they belong to word formation. It should be noted that in this approach, A+N naming expressions are not necessarily compounds. For an expression such as Dutch *zwart gat* (17c), we propose an analysis in terms of a Type 1 (001) rule rather than as a compound.

---

[10] There are three types in which the conceptual structure is not affected. Type 6 (110) is transposition as in *ugly* and *ugliness*. Both designate a property, in fact the same property, but one is an adjective and the other a noun. In type 2 (010), we find transposition by conversion, as in *check*$_V$ and *check*$_N$. Again, both designate an action, it is the same action, but they have different syntactic categories. Type 4 (100) is stem change without syntactic or conceptual implications. Linking categories in German compounds and thematic extensions in French are good candidates for this rule type.

References

Abney, Steven Paul. 1987. *The English noun phrase in its sentential aspect*. Ph.D. Dissertation, Massachusetts Institute of Technology.

Alexiadou, Artemis, Haegeman, Liliane & Stavrou, Melita. 2007. *Noun phrase in the generative perspective*. Berlin: Mouton De Gruyter.

Anastassiadis-Symeonidis, Anna. 1986. *Η νεολογία στην κοινή νεοελληνική* [*Neology in Standard Modern Greek*]. Thessaloniki: Aristotle University of Thessaloniki.

Anastassiadis-Symeonidis, Anna. 1994. *Νεολογικός δανεισμός της νεοελληνικής* [Neological borrowing in Modern Greek]. Thessaloniki.

Bağriaçik, Metin & Ralli, Angela. 2015. Phrasal vs. morphological compounds: Insights from Modern Greek and Turkish. *Language Typology and Universals: STUF – Sprachtypologie und Universalienforschung* 68. 323-357.

Bauer, Laurie, Lieber, Rochelle & Plag, Ingo. 2013. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Corbin, Danielle. 1987. *Morphologie dérivationnelle et structuration du lexique*. Tübingen: Niemeyer (2 vol.).

Grevisse, Maurice. 1980. *Le bon usage: Grammaire française avec des remarques sur la langue française d'aujourd'hui, onzième édition revue*. Paris/Gembloux: Duculot.

Grevisse, Maurice & Goosse, André. 2008. *Le bon usage: Grammaire française, 14ᵉ édition*, Bruxelles: De Boeck Duculot.

ten Hacken, Pius. 1994. *Defining morphology: A principled approach to determining the boundaries of compounding, derivation, and inflection*. Hildesheim: Olms.

ten Hacken, Pius. 2003. Phrasal elements as parts of words. In Hajičová, Eva, Kotěšovcová, Anna & Mírovský, Jiří (eds.), *Proceedings of CIL17*, CD-ROM, Praha: Matfyzpress, MFF UK (18 pp.).

ten Hacken, Pius. 2010. Synthetic and exocentric compounds in a Parallel Architecture. *Linguistische Berichte Sonderheft* 17. 233-251.

ten Hacken, Pius. 2013a. Compounds in English, in French, in Polish, and in general. *SKASE Journal of Theoretical Linguistics* 10. 97-113.

ten Hacken, Pius. 2013b. Semiproductivity and the place of word formation in grammar. In ten Hacken, Pius & Thomas, Claire (eds.), *The semantics of word formation and lexicalization*, 28-44. Edinburgh: Edinburgh University Press.

ten Hacken, Pius. 2014. Delineating derivation and inflection. In Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford handbook of derivational morphology*, 10-25. Oxford: Oxford University Press.

ten Hacken, Pius. 2015a. Terms and specialized vocabulary: Taming the prototypes. In Kockaert, Hendrik J. & Steurs, Frieda (eds.), *Handbook of terminology*, vol. 1, 3-13. Amsterdam & Philadelphia, PI: John Benjamins.

ten Hacken, Pius. 2015b. Transposition and the limits of word formation. In Bauer, Laurie, Körtvélyessy, Livia & Štekauer, Pavol (eds.), *Semantics of complex words*, 187-216. Cham: Springer.

Hüning, Matthias. 2010. Adjective + noun construction between syntax and word formation in Dutch and German. In Onysko, Alexander & Sascha, Michel (eds.), *Cognitive perspectives on word formation*, 195-215. Berlin & New York, NY: Mouton de Gruyter Mouton.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Koliopoulou, Maria. 2008. The loose multi-word compounds of Modern Greek under the prism of Construction Morphology. In Lavidas, Nikolaos, Nouchoutidou, Elissavet & Sionti, Marietta (eds.), *New perspectives in Greek linguistics*, 213-224. Cambridge: Cambridge Scholars Publishing.

Koliopoulou, Maria. 2009. Loose multi-word compounds and noun constructs. *Patras Working Papers in Linguistics* 1. 59-71.

Levi, Judith N. 1978. *The syntax and semantics of complex nominals*. New York, NY: Academic Press.

Marchand, Hans. 1969. *The categories and types of Present-day English word-formation: A synchronic-diachronic approach*, 2nd edition. München: Beck.

Matthews, Peter H. 1974. *Morphology: An introduction to the theory of word structure*. Cambridge: Cambridge University Press.

Nespor, Marina & Ralli, Angela. 1994. Stress domains in Greek compounds: A case of morphology-phonology interaction. In Phillipaki-Warburton, Irene, Nicolaidis, Katerina & Sifianou, Maria (eds.), *Themes in Greek linguistics*, 201-208. Amsterdam & Philadelphia, PI: John Benjamins.

Nespor, Marina & Ralli, Angela. 1996. Morphology-phonology interface: Phonological domains in Greek compounds. *The Linguistic Review* 13. 357-382.

Ralli, Angela. 2007. *Η σύνθεση λέξεων: διαγλωσσική μορφολογική προσέγγιση* [Compounding: A cross-linguistic morphological approach]. Athena: Patakis.

Ralli, Angela. 2013. *Compounding in Modern Greek*. Dordrecht: Springer.

Ralli, Angela & Stavrou, Melita. 1998. Morphology - syntax interface: A-N compounds vs. A-N constructs in Modern Greek. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of morphology 1997*, 243-264. Dordrecht: Kluwer.

Revithiadou, Anthi. 1997. Prosodic domains in Greek compounding. In Drachman, Gaberell, Malikouti-Drachman, Angeliki, Klidi, Sila & Fykias Jiannis (eds.), *Greek linguistics 95: Proceedings of the 2nd International Conference on Greek Linguistics*, 107-116. Graz: Neugebauer.

Revithiadou, Anthi. 1999. *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems*. Den Haag: Holland Academic Graphics.

Schuster, Saskia. 2016. *Variation und Wandel: Zur Konkurrenz morphologischer und syntaktischer A+N-Verbindungen im Deutschen und Niederländischen seit 1700*. Berlin & New York, NY: Mouton de Gruyter.

Thomas, Claire. 2013a. Lexicalization in generative morphology and conceptual structure. In ten Hacken, Pius & Thomas, Claire (eds.), *The semantics of word formation and lexicalization*, 45-65. Edinburgh: Edinburgh University Press.

Thomas, Claire. 2013b. *Characterizing the polysemy of French and English deverbal nominalization suffixes*. Ph.D. Dissertation, Swansea University.

Triantafyllidis. 1998. *Λεξικό της Κοινής Νεοελληνικής* [Dictionary of Common Modern Greek]. Thessaloniki: Institute of Modern Greek Studies (Manolis Triantafyllidis Foundation), Aristotle University of Thessaloniki.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Pius ten Hacken, Maria Koliopoulou
Institut für Translationswissenschaft
Leopold-Franzens-Universität Innsbruck
Herzog-Siegmund-Ufer 15
6020 Innsbruck, Austria
{pius.ten-hacken, maria.koliopoulou}@uibk.ac.at