

# The relation of the head and the modifier in nominal compound names of medicinal herbs

Milada Walková, Technical University of Košice

*The paper presents research into nominal compound names of medicinal herbs in four languages – Dutch, English, French and Slovak. Only endocentric compounds are considered. From the morphological point of view, the head noun in these compounds is typically modified by a relational adjective, noun or noun phrase, genitive construction, or prepositional phrase. From the semantic point of view, the modifiers in herb names typically refer to various perceptual properties or environmental conditions, meaning that herbs are construed as natural kinds. However, the modifier can also refer to the purpose or effect of herbs; some other inherent properties or other plants and creatures.*

**Keywords:** *head-modifier relations, nominal compounds, herb names, natural kinds and artefacts, botanical terminology*

## 1. Introduction

Studies of English nominal compounds (e.g. Downing 1977; Wisniewski & Love 1998; Levin & Jurafsky 2013; Levin et al. 2014) show that the relation between the head and the modifier is systematic rather than idiosyncratic. In particular, the relation depends on whether the compound denotes a *natural kind* or an *artefact*. The distinction involves objects of nature, such as minerals, plants, and animals on the one hand and objects created by humans for a purpose, such as utensils and equipment, on the other hand. This philosophical distinction has linguistic consequences (Downing 1977; Wisniewski & Love 1998; Levin et al. 2014): natural kinds are typically named after an essential property of theirs, such as appearance or habitat, while artefacts tend to be named after an associated event of creation or use. In addition, borrowing predominates in natural kinds.

If there is a clear linguistic distinction between naming artefacts and natural kinds, it may be interesting to investigate the naming strategies of non-prototypical entities – ones that are on the borderline of the distinction. Therefore, the subject of the present study is the names of medicinal herbs, as they are an unusual case of natural kinds which are strongly associated with an event of use. The peculiarity of such use, or more generally effect on a living organism, of each herb has proved to be significant for treating maladies, and even a matter of life and death, especially in the case of poisonous plants. The aim of this paper is to explore how this peculiarity of medicinal herbs is reflected in the naming strategies and whether these strategies differ from naming strategies for other natural kinds. At the same time, I argue that herb names are compounds from the semantic point of view.

The paper studies nominal compound names of medicinal herbs in four languages, namely Dutch, English, French and Slovak (with examples marked hereafter as nl., en., fr. and sk., respectively) from both morphological and semantic point of view.<sup>1</sup> It is organized as follows. Section 2 presents the theoretical background and argues for treating herb names as compounds. The methodology is spelled out in Section 3. The morphology and semantics of

---

<sup>1</sup> Etymology is not considered here but see e.g. Machek (1954), van der Sijs (2010), and Harper (2001–2015).

herb names are discussed in Sections 4 and 5, respectively. The last section draws conclusions.

## 2. Theoretical background

### 2.1. Artefacts and natural kinds

Downing (1977) studies non-lexicalized yet attested noun-noun compounds. She reports on data and experiments involving a compound interpretation task and a compound naming task. Based on the results, Downing (1977: 828) proposes the following set of semantic relationships between the head and the modifier:

whole-part (*duck foot*), half-half (*giraffe-cow*), part-whole (*pendulum clock*), composition (*stone furniture*), comparison (*pumpkin bus*), time (*summer dust*), place (*Eastern Oregon meal*), source (*vulture shit*), product (*honey glands*), user (*flea wheelbarrow*), purpose (*hedge hatchet*), occupation (*coffee man*)

At the same time, however, she points out that the list is not exhaustive but merely reflects the most common naming strategies. In addition, Downing (1977: 831) notes that the particular relationship also depends on the denotee of the head noun – namely, synthetic objects (i.e. artefacts) are named after their purpose, e.g. *banana fork*, whereas natural entities (i.e. natural kinds) are named after an inherent property, such as identity for humans, e.g. *women officers*, appearance and habitat for flora and fauna, e.g. *trumpet plant*, *Salt Creek coyote*, and composition, origin and location for other natural objects, e.g. *granite outcroppings*, *cow hair*, *Montana beach*.

Wisniewski & Love (1998) investigated whether the modifiers in the compound names of natural kinds and artefacts express relations or properties. In their study, the relations included such categories as location (*mountain lion*, *wall clock*), human resource (*birthwort*), part (*bonefish*), named after (*executive chair*), made of (*latex gloves*), when active or used (*night snake*, *rain coat*), eats (*beargrass*), object or instrument nominalizations (*vacuum cleaner*), holds/contains (*pencil cup*), and depicts (*road atlas*), while the properties included the categories of object-part (*spiderflowers*), part-part (*porcupine fish*), colour/pattern (*leopard lizard*), overall shape (*peninsula table*), appearance (*bird's nest fungus*), and behaviour/function (*cangaroo rat*). Wisniewski & Love found that the modifiers in the names express properties in natural kinds more often (42.5%) than in artefacts (14.2%). The most typical property in natural kinds was that of location; the most common relation in artefacts was object nominalization. On recalculating their (1998: 195–196) data, we see that for the plants, location was the most frequent category (30%), followed by object-part (22.5%), human resource (7%), colour/pattern (6%), part (5%), and other minor categories.

More recently, Beth Levin and colleagues (Levin & Jurafsky 2013; Levin et al. 2014) have confirmed the tendency for natural kinds to be named after an essential property and for artefacts after an associated event, and shown that the distinction has important psycholinguistic grounding. They found that natural kinds are typically named after perceptual properties, such as colour (*lemon citrine*), dimension (*broad bean*), distinctive part (*cross necklace*), taste/smell (*sugar pea*), texture (*marrowfat pea*), and visual feature (*kidney bean*), after the environment they live in, including habitat (*garden pea*), season (*winter pea*),

and social/political (*African opal*), or they have names borrowed from other languages (*fava bean*). In contrast, artefacts are typically named after various aspects of events of creation and use, including categories made of (*chocolate cake*), method of creation (*skillet cake*), used by (*navy bean*), object-nom<sup>2</sup> (*banana slicer*), occasion (*birthday cake*), and purpose (*fish knife*). The category of ‘other’, including named after (*Queen Elizabeth cake*), other property (*hinged bangle*), part-whole relation (*beet greens*), and value (*wacky cake*), is not reserved for either natural kinds or artefacts.

## 2.2. Names of medicinal herbs as nominal compounds

I treat the names of medicinal herbs in my dataset as compounds, which may seem controversial. After all, the problem of proper definition of compounds and classification of various combinations as compounds or phrases is long-standing and notorious: a number of criteria have been proposed to delimit compounds from phrases, including spelling, stress, inflection of the whole compound, impossibility of modification by *very*, fixed sequence of constituents, impossibility of omission of constituents, inseparability of constituents, and lexicalized meaning – yet all of these admit many exceptions and/or are restricted in their application to only some types of compounds (e.g. Lieber 2005; Kavka & Štekauer 2006).

Let us now briefly probe the names of medicinal herbs according to these criteria. Starting with spelling, most herb names are, like phrases, spelled separately in all four languages, although (parts of) some names are spelled together, e.g. sk. *trojlistá* ‘three-leaved’, en. *bindweed*, nl. *bleekgele hennepnetel* lit. ‘pale-yellow hemp-nettle’, or via hyphens, such as fr. *mille-florins* ‘thousand-florins’.

Similarly, herb names are stressed like phrases rather than compounds, since they typically bear a level stress, e.g. *common yarrow*. However, some herb names are stressed on the left-hand constituent, e.g. *watercress*.

According to the inflection criterion, noun-noun combinations are clearly compounds. For instance, in nl. *bosaardbei* lit. ‘forest strawberry’ – *bosaardbeien*<sub>PL</sub> plural is marked on the head but not on the modifier. Left-headed names are also inflected on the head, e.g. *lilies*<sub>PL-of-the-valley</sub>, fr. *colchiques*<sub>PL d’automne</sub> ‘fall crocuses’, sk. *čaj z prasličky*<sub>GEN rol’nej</sub><sub>GEN</sub> ‘tea from common horsetail’. Some herb names contain internal inflection, such as names with genitive constructions, e.g. en. *cat’s valerian*, and a few adjectival modifiers in the comparative, e.g. sk. *lopúch väčší* ‘greater burdock’, en. *lesser celandine*. In general, however, adjective modifiers of herb names cannot be turned into the comparative, e.g. en. *sweet basil* – *\*sweeter basil*. Similarly, they cannot be modified by *very*, e.g. *\*very sweet basil*.

Another criterion is that the constituents in most compounds have a fixed sequence. This is true of herb names, e.g. en. *great masterwort* – *\*master great wort*/*\*masterwort great*/*\*wort great master*, etc., *European white birch* – *?white European birch*. In Slovak, this fixed sequence is different from the sequence of phrasal constituents: in phrases, adjectives precede head nouns, e.g. *pekňý*<sub>ADJ</sub> *chlapec*<sub>N</sub> ‘a handsome boy’, *dobrá*<sub>ADJ</sub> *kniha*<sub>N</sub> ‘a good book’, whereas in herb names the order is reverse, following the Latin model, cf. sk. *slněčnica*<sub>N</sub> *ročná*<sub>ADJ</sub> ‘common sunflower’ – Lat. *Helianthus*<sub>N</sub> *annus*<sub>ADJ</sub>.

<sup>2</sup> The label object-nom stand for cases in which the head contains a deverbal nominalisation.

As to the omission of constituents, in some (parts of) herb names the constituents cannot be omitted (especially when spelled together) without changing the meaning, e.g. sk. *hladkoplodá* lit. ‘smooth-fruited’ but the modifiers referring to species can be omitted (and in informal discourse they are indeed frequently omitted), although it means the loss of a specific reference to a particular species of a certain genus, e.g. sk. *lipa* ‘linden’ can refer to either *lipa malolistá* ‘little-leaf linden’ or *lipa veľkolistá* ‘large-leaved linden’.

What Lieber (2005: 377) considers as “perhaps the strongest test of compoundhood” is inseparability of compound constituents, meaning that an arbitrary modifier cannot be freely inserted into a compound. Herb names uniformly meet this criterion, e.g. nl. *witte* (\*maarts) *engbloem* lit. ‘white (\*March) vincetoxinum’, eng. *lily-of-the-(\*)green-valley*, fr. *herbe* (\*haute) *a Robert* lit. ‘herb (\*tall) of Robert’, sk. *šalvia* (\*lúčna) *lekárska* lit. ‘sage (\*meadow<sub>ADJ</sub>) officinal’. Names with the same head, however, can be coordinated, e.g. sk. *lipa malolistá a veľkolistá* ‘little-leaf and large-leaved linden’.

Lastly, let us look at the semantics. Herb names demonstrate a high degree of semantic unity with each name referring to a single generic concept, i.e. the respective species. The herb names in my dataset are largely compositional. This is mostly due to the fact that exocentric compounds, which are non-compositional, were not included in the study (see § 3).

To sum up our discussion so far, it appears that while some herb names are clearly compounds (especially noun-noun compounds), the status of many others – in particular of those whose modifiers are adjectives, prepositional phrases or genitive constructions – is questionable. Ten Hacken (2013) suggests that the problem of delimitation of compounds can be approached in two ways: we can either understand compounds as a purely formal category or we can take into account also semantic criteria, allowing for cross-linguistic correspondence. In the latter approach we can consider disputable combinations as compounds under certain conditions: first, if the combination refers to a single concept; second, if the adjective modifier in combinations with adjectives is relational; and lastly, if the preposition in prepositional phrases (PPs) or genitive constructions does not determine the relation between the constituents. Ten Hacken also gives some diagnostics for these criteria, which I will now discuss and apply to herb names.

According to one test, if a combination in French contains a definite article, it is probably not a compound, e.g. *donneur de sang* ‘donor of blood’ (compound) – *donneur du sang* ‘donor of the blood’ (phrase), although there are some exceptions, as Ten Hacken (2013: 104) points out. In my dataset, there are a number of French names containing definite articles, e.g. *fraisier des bois* lit. ‘strawberry of the forest’, which, however, pass another compoundhood test. This test is based on the already discussed inseparability of compounds. Namely, a combination is not a compound if a demonstrative can be inserted into the PP without change in meaning, e.g. *\*<sup>?</sup>usine de ces automobiles* ‘factory of these cars’ (*usine d’automobiles* is a compound). The French herb names containing definite articles do not allow demonstrative insertion, e.g. *\*fraisier de ces bois* lit. ‘strawberry of these woods’, cf. also en. *\*lily-of-this-valley*.

The last test distinguishes between relational and qualitative adjectives, and consequently between the respective compounds and phrases. The adjective is used predicatively to see if it is semantically felicitous. If the predicative use is not felicitous, the adjective is relational and the combination is a compound, e.g. *a nervous disease* (compound) – *the disease is nervous* (relational adjective); otherwise, the adjective is qualitative and the combination is a phrase, e.g. *a nervous applicant* (phrase) – *the applicant is nervous*

(qualitative adjective). The adjectives in herb names behave as relational rather than qualitative, e.g. sk. *lopúch väčší* ‘greater burdock’ – \**lopúch je väčší* ‘burdock is greater’ and *balota čierna* ‘black horehound’ – \**balota je čierna* ‘horehound is black’. In other contexts, though, these adjectives are typically used as qualitative, e.g. *a black dog* (phrase) – *the dog is black* (qualitative). I follow ten Hacken & Koliopoulou (2015) who argue that the semantic unity of a combination should take precedence over the formal criteria of compoundhood, and who consequently treat combinations of qualitative adjectives and head nouns with a specialized meaning as compounds, e.g. nl. *zwarte gat* ‘black hole’.<sup>3</sup> The meaning of *black hole* is not a simple sum of the meanings of *black* and *hole*. Rather, it has a very specific denotation. Similarly, just as *black hole* with a specialized meaning is not any hole which is black, *lopúch väčší* ‘greater burdock’ is not any burdock which is greater than another and *balota čierna* ‘black horehound’ is not a burdock which is black – in fact, the plant and its flowers are not black at all.

I conclude that names of medicinal herbs are compounds from the semantic point of view.<sup>4</sup> Their semantic unity is formally reflected in their inseparability, impossibility of modification by *very*, and fixed sequence of constituents. While many of these combinations are arguably non-prototypical compounds, taking semantics as a primary criterion of compoundhood allows for cross-linguistic correspondence (see ten Hacken 2013).

### 3. Data collection and analysis

To study nominal compound names of medicinal herbs in four languages which span several continents with varying flora, a common database of plants was created. A Slovak book containing a detailed description and illustrations of 256 medicinal herbs (Volák et al. 1987) served this purpose. To find the corresponding names in the other languages, I used Sedlářová (1994) and Wikipedia (<http://www.wikipedia.org>).

Although the database of herbs is the same for each language, the number of the head-modifier relations in each language differs for several reasons. First, some herbs have more than one name in a particular language. If these names contained two different modifiers, e.g. *common yarrow* and *nosebleed plant* (both naming *Achillea millefolium*), two relations were coded; if the plant names had the same modifier with two different heads, a single relation was coded, e.g. *common alkanet* and *common bugloss* (both naming *Anchusa officinalis*).<sup>5</sup> Second, for some names, there are several head-modifier relations, e.g. nl. *witte paardenkastanje* lit. ‘white horses-chestnut’, en. *European white birch*, sk. *rosička okrúhlostá* lit. ‘drosera round-leaved’. In such cases, each relation was coded separately. Third, exocentric and coordinative compounds, e.g. *holy ghost*, *lady’s mantle*, *church steeples*, and *lords and ladies*, were not taken into consideration due to the absence of an overt head-modifier relation. However, modifiers which modify an exocentric compound as a whole were included in the analysis, e.g. *field horsetail* or *common horsetail*. In addition, when the grammatical head referred to any plant, the compound was considered in the study,

<sup>3</sup> Notice an example from Slovak: *moja stará matka* can be a syntactic phrase meaning ‘my old mother’ or a compound meaning ‘my grandmother’. Only in the former case is the predicative use possible – *moja matka je stará* ‘my mother is old’.

<sup>4</sup> This will apply not only to botanical but also to zoological names, e.g. *red admiral butterfly*, *European robin*.

<sup>5</sup> It has to be pointed out that forms with the same base but distinct gender affixes were coded as a single type, e.g. *lekársky<sub>m</sub>* ‘medicinal’ – *lekárska<sub>f</sub>* ‘medicinal’.

as it was understood that a plant may stand for another on the basis of resemblance or relatedness, e.g. *deadnettle*, or hyperonymy/hyponymy, e.g. *dog grass*. Fourth, some names are not synchronically motivated, e.g. sk. *baza chabzdová* lit. ‘elder chabzda-ADJ (semantically non-transparent adjective)’ (*Sambucus ebulus*), the modifier serving to distinguish the particular species from other species of the same genus, cf. *baza čierna* lit. ‘elder black’ (*Sambucus nigra*) and *baza červená* lit. ‘elder red’ (*Sambucus racemosa*). Given the lack of synchronic motivation, the relation could not be assigned to any of the semantic categories and was therefore excluded from the study. Fifth, some names were excluded by the fact that they are not complex words, e.g. *fennel*. Lastly, some data are missing, as I was not able to find the names of a few herbs in a particular foreign language, possibly due to scarce geographical incidence. In total, I coded 231 semantic relations in Dutch, 456 in English, 421 in French, and 263 in Slovak.

Relying on the descriptions and illustrations of the plants in Volák et al. (1987) and Wikipedia (<http://www.wikipedia.org>), the plant names were coded for the semantic head-modifier relations according to Levin et al. (2014), see § 2. During the course of the research, the classification was slightly refined (see § 5). In Levin et al. (2014), a separate category (BORROWED) involved modifiers taken from another language, e.g. *genoise cake*, *fava bean*. In the present study, such a category was not created. Instead, all synchronically non-transparent modifiers were excluded from the study (see above) and modifiers which are calqued from Latin yet semantically transparent were coded into their respective category, such as PERCEPTUAL – COLOUR for lat. *Salix alba* lit. ‘willow white’ – en. *white willow* – sk. *vřba biela* lit. ‘willow white’ – fr. *saule blanc* lit. ‘willow white’.<sup>6</sup>

Some modifiers were ambiguous between different readings, e.g. *lemon* is ambiguous between referring to colour or taste/smell. In such instances, the relation was coded on the basis of the description of the particular plant, e.g. the plant named *lemon balm* has lemon/honey scent, so the relation was coded as PERCEPTUAL – TASTE/SMELL. Similarly, nl. *honing* ‘honey’ was coded as PERCEPTUAL – TASTE/SMELL in *honingklaver* lit. ‘honey clover’ but as OTHER – NAMED AFTER in *honingboom* lit. ‘honey tree’, as no connection to the perceptual properties of honey was found for this plant. On the other hand, some plants have names which may refer to both their visual properties and healing effect, a phenomenon known as the doctrine of signatures. Given the fact that the resemblance is attributed post-hoc (Bennett 2007), such cases were coded as EVENT – PURPOSE/EFFECT rather than for perceptual properties, e.g. nl. *longkruid* ‘lungwort’. However, the head-modifier relation was coded as PERCEPTUAL – VISUAL when the modifier describes the shape of the head’s referent, e.g. *liver-leaf hepatica*.

From the morphological point of view, the modifiers in the dataset were classified as one of the following: noun/noun phrase, genitive construction, adjective, numeral, verb, adverb, or prepositional phrase. Note that, following the tradition in Slovak linguistics, the category of numerals, i.e. parts of speech referring to specific quantities, has also been included, e.g. nl. *duizend* ‘thousand’, fr. *cent* ‘hundred’. Complex modifiers were broken down into individual components, e.g. fr. *trilobée* ‘three-lobed’ (number + adjective), sk. *starčeko-listá* lit. ‘old-man-leaved’ (noun + adjective), fr. *à grandes fleurs* ‘(preposition) big flowers’ (PP + adjective), unless their individual components do not add up to the meaning of

---

<sup>6</sup> Note that while one herb may have the same modifier in its names across languages, as in this example, it is typically not the case and a herb may be named after various properties in different languages, e.g. nl. *koningskruid* lit. ‘king-wort’ – en. *sweet basil* – fr. *basilic commun* lit. ‘basil common’ or *herbe royal* lit. ‘herb royal’ – sk. *bazalka pravá* lit. ‘basil true’ for the plant *Ocimum basilicum*.

the whole, e.g. *Saint John's* (genitive construction), nl. *Oost-Indische* 'East-Indian' (adjective), fr. *Notre Dame* 'Our Lady' (noun phrase). In total, I coded 227 modifiers in Dutch, 463 in English, 421 in French, and 262 in Slovak. It has to be pointed out that these numbers differ from the numbers of semantic relations, as there is not a neat one-to-one correspondence between morphology and semantics: for instance, what was coded as one semantic relation was sometimes coded as two different modifiers morphologically, e.g. fr. *d'or* lit. 'of gold' (genitive construction) – *doré* 'golden' (adjective) – both PERCEPTUAL – COLOR, *poison* (noun) – *poisonous* (adjective) – both OTHER – OTHER PROPERTY.

#### 4. Morphology

The results of morphologically classifying of modifiers are shown in Table 1 and Figure 1. As can be seen, adjective modifiers are the most common in all the studied languages. In fact, in Slovak almost all of the modifiers are adjectival. Dutch and English employ also many nominal modifiers and French utilizes a variety of other types of modifiers as well. I will now discuss the results in more detail.

Table 1: Morphological classification of modifiers in herb names.

Category/Language	Dutch	English	French	Slovak
Adjective	123	274	270	251
Noun/NP	95	174	19	3
Genitive construction	0	8	83	0
Prepositional phrase	1	2	40	0
Numeral	5	1	8	7
Verb	3	4	0	0
Adverb	0	0	1	1
<b>TOTAL</b>	<b>227</b>	<b>463</b>	<b>421</b>	<b>262</b>

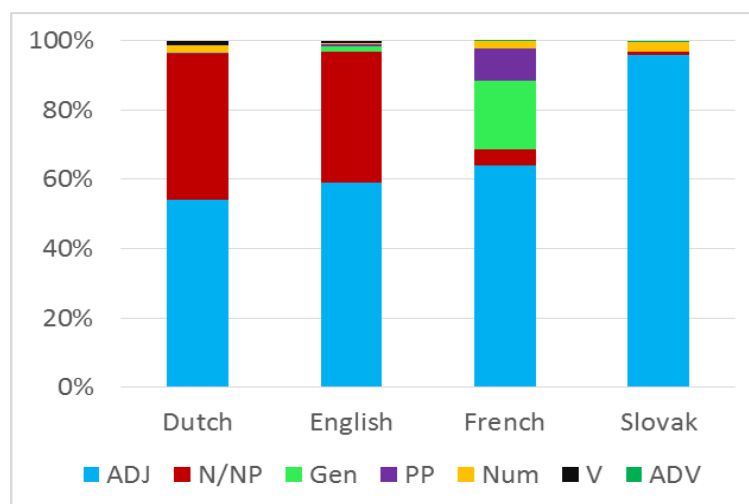


Figure 1: Morphological classification of modifiers in herb names.

In Dutch herb names, modifiers are typically adjectival (54%), e.g. *zure kers* 'sour cherry', *Canadese fijnstraal* 'Canadian horseweed', and nominal (42%), e.g. *kerstroos* 'Christmas

rose', *watermunt* 'water mint'. Adjectives agree with the head noun in gender, e.g. *geel walstro* 'yellow bedstraw' – *grote gele gentiaan* 'great yellow gentian'. A few nominal modifiers contain a linking element, e.g. *hondsroos* lit. 'dog-LINKING ELEMENT-rose', *koningskruid* lit. 'king-LINKING ELEMENT-wort'. Some modifiers include verbs (*ratel* 'rattle', *schiet* 'shoot'), numerals (*drie* 'three', *duizend* 'thousand'), or a prepositional phrase (*van-dalen* 'of valleys'). There are no adverbs or genitive phrases.

English names of herbs show the same pattern as those in Dutch: most modifiers are adjectives (59%), e.g. *English hawthorn*, *wooly burdock*, and nouns (38%), e.g. *field elm*, *milk thistle*. The other modifiers contain genitive constructions (*devil's*, *lady's*),<sup>7</sup> verbs (*bleed*, *sneeze*), prepositional phrases (*of the valley*, *of grace*), or a numeral (*single*). There are no adverbial modifiers.

In contrast, French employs modifiers of a wider variety. Adjectives are by far the most frequent (64%), e.g. *fumeterre officinal* lit. 'fumitory officinal', *lierre terrestre* lit. 'ivy ground-ADJ'. They agree with the head noun in gender, e.g. *marronnier blanc* lit. 'chestnut white' – *bryone blanche* lit. 'bryony white'. Adjectives are followed in frequency by genitive constructions (20%), e.g. *poivre d'eau* lit. 'pepper of water', *sorbrier des montagnes d'Europe* lit. 'rowan of the mountains of Europe', prepositional phrases (10%), e.g. *à balais* '(preposition) broom', *aux dents* '(preposition) teeth', and nouns (5%), e.g. *Marie*, *champêtre* 'country'. Numerals (*cent* 'hundred', *mille* 'thousand') and adverbs (*toujours* 'ever-') are marginal. There are no instances of verbs.

Finally, adjectival modifiers are almost exclusive in Slovak herb names (96%). The adjectives agree with the head noun in gender, e.g. *medovka lekárska* lit. 'lemon-balm officinal' – *yzop lekársky* lit. 'hyssop officinal'. Some of these modifiers are compounds themselves.<sup>8</sup> They are modified by numerals (*dvoj-* 'two-', *troj-* 'three-'), nominal (*starčecko-* 'old man', *hrabo-* 'hornbeam') or adverbial (*vždy-* 'ever-') modifiers. There are no verbal modifiers, prepositional or genitive phrases.

A final note has to be made on the headedness. Herb names in both Dutch and English are right-headed except for those with a PP (e.g. nl. *lelietje-van-dalen* 'lily-of-the-valley', en. *herb of grace*). In contrast, French and Slovak herb names are left-headed. Note that in Slovak this is not the canonical word order, as noted in § 2.2. The exceptions are Slovak complex modifiers, in which the modifier precedes the head, e.g. *okružlolistá* 'round-leaved', and the French modifiers *petit* 'small' and *grand* 'big' as in *petit boucage*, *grande capucine*, in line with grammar rules of French.

## 5. Semantics

The results of semantic coding are presented in Table 2 and Figure 2. As can be seen, medicinal herbs are construed as natural kinds in all the studied languages, as they are mostly named after environmental (Dutch 16%, English 29%, French 18%, Slovak 13%) and perceptual (Dutch 41%, English 29%, French 34%, Slovak 36%) properties. However, these percentages are lower than found by Levin et al. (2014), according to whose results

<sup>7</sup> It is interesting to see that modifiers which are genitive constructions typically fall into the semantic category of OTHER – NAMED AFTER in Dutch and English, unlike in French.

<sup>8</sup> Interestingly, in most of these complex adjectives the head belongs to the semantic category of distinctive part (see § 5), i.e. the head refers to a distinctive part of the plant and its modifier specifies it by referring to its shape, dimensions, etc., e.g. *velkokvetý* 'large-flowered', *úzkolistá* 'narrow-leaved'.



perceptual properties are the most common semantic coding for the names of greens/legumes (43%) and gemstones (61%) in English, followed by environmental properties (30% and 25%, respectively). Their percentages become even higher when borrowed names are disregarded (49%, 62%, 35%, and 25%, respectively), as in the present study. Since the proportion of perceptual and environmental properties in the names of medicinal herbs is lower than in greens/legumes and gemstones, one could expect that the proportion of associated events will be higher in medicinal herbs, but this is true only to some extent. Namely, the incidence of events in the names of medicinal herbs is 8% in Dutch, 7% in English, 12% in French, and 16% in Slovak, compared to 8% in the names of greens and legumes and 5% in gemstones in English (9% and 5%, respectively, when borrowed names are disregarded). A large number of the names of medicinal herbs fall into the category OTHER, whose incidence is the same in all the languages (Dutch 35%, English 35%, French 36%, Slovak 35%). In contrast, in Levin et al. (2014) this is a marginal category – 6% and 8% (7% and 8% when discarding borrowed names) in greens/legumes and gemstones, respectively. I will now discuss the subcategories of each of these larger categories in turn (see Table 2 and Figure 2).

Table 2: Semantic categories of modifiers in names of medicinal herbs.

Category/Language	Dutch	English	French	Slovak
habitat	19	51	38	17
season	9	12	6	8
social/political	10	70	31	10
<b>Total ENVIRONMENT</b>	<b>38</b>	<b>133</b>	<b>75</b>	<b>35</b>
colour	33	48	43	29
dimension	19	26	34	13
distinctive part	9	16	22	18
taste/smell	11	14	15	10
texture	7	9	9	6
visual	15	19	20	18
<b>Total PERCEPTUAL</b>	<b>94</b>	<b>132</b>	<b>143</b>	<b>94</b>
<b>Total natural kinds</b>	<b>132</b>	<b>265</b>	<b>218</b>	<b>129</b>
used by	0	3	34	36
purpose/effect	18	28	15	5
<b>Total EVENT</b>	<b>18</b>	<b>31</b>	<b>49</b>	<b>41</b>
other property	19	25	33	20
numerical	5	1	8	7
value	33	86	68	42
named after	24	48	45	24
<b>Total OTHER</b>	<b>81</b>	<b>160</b>	<b>154</b>	<b>93</b>
<b>TOTAL</b>	<b>231</b>	<b>456</b>	<b>421</b>	<b>263</b>

Starting with environmental properties, the most frequent subcategory is HABITAT, e.g. nl. *muur* ‘wall’, en. *marsh*, fr. *de mer* ‘of sea’, sk. *dedinský* ‘rural’, except in English, where SOCIAL/POLITICAL is the dominant subcategory. This is possibly due to the fact that English names span more land than the other languages, which is reflected in the names of herbs, e.g. *Irish/Cornish/Welsh oak* (all referring to the same plant *Quercus petraea*), *Scots pine*, *English yew*. The subcategory SEASON involves not only the four seasons of the year, e.g. sk.

*jarná* ‘spring’, but also other time frames, e.g. nl. *kerstroos* ‘Christmas rose’, en. *February*, fr. *toujours* ‘ever-’.

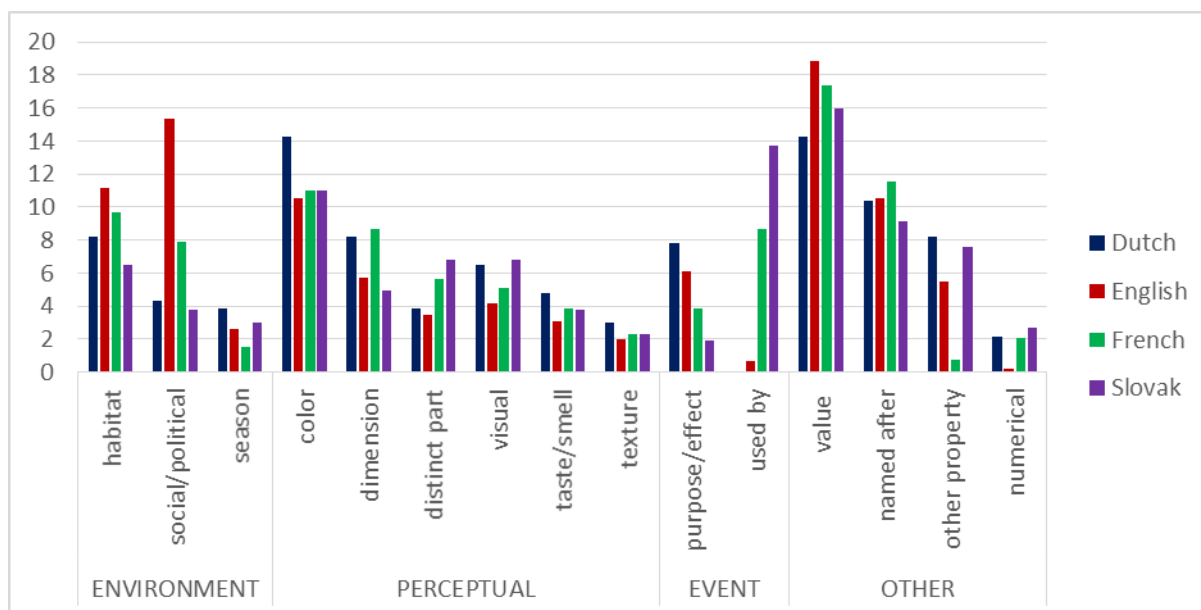


Figure 2: Percentages of semantic categories of modifiers in names of medicinal herbs.

From among perceptual properties, COLOUR (e.g. nl. *oranje* ‘orange’, en. *pale*, fr. *pourpré* ‘crimson’, sk. *pestrý* ‘colorful’) and DIMENSION (e.g. nl. *klein* ‘small’, en. *lesser*, fr. *élevé* ‘high’, sk. *úzko-* ‘narrow’) tend to be the most frequent, and TEXTURE (e.g. nl. *donzige* ‘fluffy’, en. *spiny*, fr. *laineuse* ‘wooly’, sk. *hladko-* ‘smooth’) and TASTE/SMELL (e.g. nl. *stinkende* ‘stinking’, en. *sweetscented*, fr. *acide* ‘acid’, sk. *sladkohorký* ‘bitter-sweet’) are the least frequent subcategories. The subcategory DISTINCTIVE PART includes names which refer to parts of herbs, e.g. en. *bulb*, fr. *tubéreux* ‘tuberous’, and which are frequently further modified, e.g. nl. *kleinbladige* ‘small-leaved’, sk. *jednosemenný* ‘single-seeded’. The data show that the subcategory VISUAL can be further refined, as the respective modifiers denote HABITUS/SHAPE, e.g. nl. *ronde* ‘round’, en. *kidney*, sk. *vzpriamený* ‘erect’, SPOTS, e.g. fr. *tacheté* and sk. *bodkovaný*, both ‘spotted’, or AESTHETICS, e.g. nl. *fraaie* and sk. *spanilý*, both ‘lovely’.

The category of EVENT is much less varied than in artefacts: my data do not include any instances of MADE OF, METHOD OF CREATION (naturally, since herbs are not created by humans), OBJECT-NOM (as in my dataset there are no endocentric compounds whose head is a verbal nominalization), or OCCASION. Instead, only the subcategories PURPOSE, adapted as PURPOSE/EFFECT (e.g. nl. *koorts* ‘fever’, en. *wound*, fr. *somnifère* ‘sleep-inducing’, sk. *prečisťujúci* ‘purging’) and USED BY (e.g. en. *dyer*’s, fr. *officinal* ‘medicinal’, sk. *kuchynská* ‘cuisine’, no instances attested in Dutch) are represented in the data. It is interesting to see that Dutch and English employ PURPOSE/EFFECT much more than USED BY, in stark contrast to French and Slovak, which have a higher incidence of USED BY. However, both French and Slovak rely on a few modifiers of this subcategory, which are employed heavily, especially fr. *officinal* and sk. *lekársky*, both ‘medicinal’.

In the category OTHER the most frequent subcategory is that of VALUE, such as nl. *echt* ‘true’, *wonder* ‘wonder’, en. *holy*, *of grace*, fr. *commun* and *vulgaire* both ‘common’, sk.

*královský* ‘royal’, *pravá* ‘true’. Apparently, this semantic relation is of high importance in naming plants because it reflects the respect people have held for the plants with strong healing effects. Other non-perceptual properties are subsumed under the subcategory OTHER PROPERTY. It includes properties inherent to the plant, e.g. nl. *wilde* ‘wild’, *vuurwerkplant* lit. ‘firework plant’ (plant which catches fire easily), en. *male*, *quaking*, fr. *femelle* ‘female’, *savauge* ‘savage’, sk. *siaty* ‘sowed’, *dvojdómá* ‘dioecious’. A number of these modifiers refer to the fact that a plant is (or is not) poisonous, e.g. nl. *gifsla* ‘poison lettuce’, en. *deadly*, *edible*, fr. *empoisonnée* ‘poisoned’, sk. *zlomocný* lit. ‘evil-powered’. The subcategory NAMED AFTER can be further refined: plants can be named after PERSONS, e.g. en. *Saint John’s*, sk. *Rostkovova* (named after the botanist Rostkovius), ANIMALS, e.g. nl. *paarden* ‘horses’, fr. *aux chats* ‘of the cats’, MYTHOLOGICAL CREATURES, e.g. en. *witch*, fr. *du diable* ‘of the devil’, other PLANTS, e.g. nl. *wilgen* ‘willow’, sk. *fazuľová* ‘bean’, or OBJECTS, e.g. en. *gas*, *shield* (in these cases, no connection to perceptual properties was detected). Lastly, I created an additional subcategory of OTHER that is absent in Levin et al. (2014), namely NUMERICAL, to cover modifiers such as en. *single*, nl. *twee* ‘two’, sk. *troj-* ‘three’, and fr. *mille-* ‘thousand’. While this is a marginal category in my data, it occurs in each of the studied languages, and it shows that the semantic classification of head-modifier relations may require adaptations for individual semantic fields. On the other hand, the subcategory PART WHOLE from Levin et al. (2014) is not represented in my data. This subcategory refers to parts of an entity, e.g. *beet greens* referring to the green parts of the beet plant. Naturally, my data do not include this category since names of plants do not refer to mere parts of herbs: a herb name always denotes the whole herb. (However, a herb may be named after its part – such modifiers are included in the DISTINCTIVE PART category, e.g. sk. *veľkokvetý* lit. ‘large-flowered’).

To sum up, essential properties prevail over event of use in herb names. These essential properties include not only environmental description and perceptual properties but also other properties inherent to herbs (subcategory OTHER PROPERTY). I conclude that medicinal herbs are construed as natural kinds.

## 6. Conclusion

The study analyses complex herb names in four languages from both morphological and semantic point of view. I argue that all the names in my dataset are compounds, with modifiers realized variously as adjectives, nouns or noun phrases, genitive constructions, and prepositional phrases, but also occasionally numerals, verbs, and adverbs. Herb names of both Dutch and English are typically right-headed, with the modifiers being most commonly adjectives and nouns. They differ from herb names in French and Slovak, which are typically left-headed and primarily adjectival, although French uses genitive constructions and prepositional phrases to a large extent as well. As these modifiers typically refer to essential properties, herbs are construed as natural kinds: the modifiers typically refer to perceptual, environmental and other essential properties. However, herb names also refer to medicinal uses of herbs.

## Acknowledgements

An earlier version of the paper was presented at the Typology and Universals in Word Formation III conference (Košice, June 2015). I thank the audience for questions and comments. I also thank two anonymous reviewers of this journal for their suggestions and comments which helped improve the paper considerably. The usual disclaimers apply.

## References

- Bennett, Bradley C. 2007. Doctrine of signatures: An explanation of medicinal plant discovery or dissemination of knowledge? *Economic Botany* 61(3). 246–255.
- Downing, Pamela. 1977. On the creation and use of English compound nouns. *Language* 53(4). 810–842.
- Harper, Douglas. 2001–2015. *Online Etymology Dictionary*. (<http://www.etymonline.com>) (Accessed January–June 2015.)
- Kavka, Stanislav J. & Štekauer, Pavol. 2006. *Compounds and compounding: An attempt at a complex view*. Ostrava: University of Ostrava.
- Levin, Beth & Jurafsky, Dan. 2013. Names for artifacts and natural kinds. Paper presented at the 9th Mediterranean Morphology Meeting, Dubrovnik, 15–18 September 2013.
- Levin, Beth, Glass, Lelia & Jurafsky, Dan. 2014. Corpus evidence for systematicity in English compounds. Paper presented at the 19th Sinn und Bedeutung conference, Göttingen, 15–17 September 2014.
- Lieber, Rochelle. 2005. English word-formation processes. In Štekauer, P. & Lieber, R. (eds.), *Handbook of word-formation*, 375–427. Dordrecht: Springer.
- Machek, Václav. 1954. *Česká a slovenská jména rostlin*. [Czech and Slovak plant names.] Praha: Československá akademie věd.
- Sedlářová, Lubica. 1994. *Názvy rastlín: Slovensko-latinsko-anglicko-nemecko-francúzsko-český slovník*. [Plant names: Slovak – Latin – English – German – French – Czech dictionary.] Piešťany: Výskumný ústav rastlinnej výroby.
- ten Hacken, Pius. 2013. Compounds in English, in French, in Polish, and in general. *SKASE Journal of Theoretical Linguistics* 10(1). 97–113.
- ten Hacken, Pius & Koliopoulou, Maria. 2015. Adjectival non-heads and the limits of compounding. Paper presented at the Typology and Universals in Word Formation III conference, Košice, 26–28 June 2015.
- van der Sijs, Noline (ed.). 2010. *Etymologiebank*. (<http://www.etymologiebank.nl>) (Accessed January–June 2015.)
- Volák, Jan, Stodola, Jiří & Severa, Frantšek. 1987. *Velká kniha léčivých rostlin*. [The big book of medicinal herbs.] Bratislava: Příroda.

*Wikipedia: The free encyclopedia*. FL: Wikimedia Foundation, Inc. (<http://www.wikipedia.org>)  
(Accessed January–June 2015.)

Wisniewski, Edward J. & Love, Bradley C. 1998. Relations versus properties in conceptual combination. *Journal of Memory and Language* 38(2). 177–202.

*Milada Walková*  
*Department of Languages*  
*Technical University of Košice*  
*Vysokoškolská 4*  
*Košice 042 00*  
*Slovakia*  
*milada.walkova@tuke.sk*

*In SKASE Journal of Theoretical Linguistics [online]. 2016, vol. 13, no.2 [cit. 2016-09-11]. Available on web page [http://www.skase.sk/Volumes/JTL32/pdf\\_doc/09.pdf](http://www.skase.sk/Volumes/JTL32/pdf_doc/09.pdf). ISSN 1336- 782X.*