Productivity of form and productivity of meaning in N+N compounds

Elizaveta Tarasova, IPU New Zealand

This paper reports on the results of a corpus-based study of English N+N compounds that are hyponyms of their rightmost elements (Bloomfield 1933: 235), e.g. library book*,* hen house*,* coffee cup*, with the aim of investigating the relation between productivity, on the one hand, and, on the other hand, the structural pattern of N+N compound formation (with a particular reference to the relationship between the elements of the compound). The discussion of the results addresses the question of whether productivity of a noun in forming a paradigm of compounds (constituent word family) could be connected with the productivity of the relation realised in the compound paradigm containing this noun. This question has its roots in the ongoing discussion of the role of analogy in the formation of complex lexical items, and its investigation may shed light on the issue of analogy in both formal and semantic patterning.*

Keywords: *N+N compounds*, *productivity*, *constituent word family*, *analogy*, *semantic relations*

1. Introduction

Research into the formation of meaningful sequences consisting of two adjacent nouns (henceforth N+N), e.g. *door handle*, *pencil sharpener*, *biology research*, *team management*, *singer-poet*, etc. in English (as well as in other languages) has taken a new turn in the last few years with the focus of the study shifting from the argument on what formal characteristics make an N+N sequence a compound to the investigations on how the meaning of a compound is construed in the mind of an individual, and how the knowledge that native speakers possess is applied to form new sequences that follow the same pattern. This brings us to the question of morphological productivity, the question addressed in a large body of research in word-formation (e.g. Schultnik 1961; Aronoff 1976; Bauer 1983, 2001; Baayen 1992, 1993, 2003; Plag 1999; Rainer 2004, to name just a few).

Despite a considerable amount of work in the area of morphological productivity and the development of various tools (quantitative and qualitative) for measuring productivity (Baayen 1992, 1993; Baayen & Lieber 1991; Baayen & Renouf 1996; Bauer 2001; Hay 2003; Hay & Baayen 2005 Hay 2003; Fernández-Domínguez 2009, etc.), productivity in the formation of N+N compounds has not been analysed in much detail. Only a limited number of researchers focus on productivity in compound formation in different languages (e.g. Zwitserlood 1994; Krott et al. 1999; Hein & Engelberg 2017).

In this paper, I look only at one group of compounds, i.e. English subordinative N+Ns; however, the suggestions proposed here can probably be applied to other kinds of N+N sequences. What follows is the discussion of the issues of productivity in compound formation with the special focus on individual nouns when used as constituents of N+N compounds, with the aim of connecting the productivity of a noun as a compound constituent with the productivity of a relational meaning that a given constituent realises. Based on this, we can formulate two hypotheses that are tested in this study:

i. A compound constituent is used more productively in one position (modifier or head) but not the other.
ii. There is a connection between the family size of the compound constituent and this constituent's concentration on one semantic relation.

2. Limiting the scope of structures

One of the most heated ongoing debates in compounding research is devoted to the question of whether a sequence of two nouns is a unit of morphology (a lexeme) or a unit of syntax (phrase), since compounding borders on affixation and on syntax (Bauer 2017; ten Hacken 2017). In order to be able to distinguish between compounds and other phenomena, a number of different criteria have been offered and quite a few of them have been shown to have limited applicability in Bauer (1998, 2017) and Giegerich (2015). Another question in this regard concerns the importance of the ability to apply such criteria crosslinguistically, since, although it is not universal, compounding is widely spread in world languages (Štekauer et al. 2008). Ten Hacken (2017: n.p.g.), in his overview of different approaches to delimiting compounds from other phenomena, outlines the importance of the theoretical framework for the selection and formulation of criteria, because "[…] if, for instance, word-formation and syntax are strictly separated and compounding is in word formation, it is crucial to draw this borderline precisely". In the case of the present investigation, the borderline between syntax and morphology is not important, as the question of whether N+N compounds are words or phrases is not relevant for this discussion, which looks at the process of combining two noun concepts into a single meaningful unit (be it a complex word or phrase). In this research, I accept the opinion shared by some linguists (e.g. Warren 1978; Bauer 1978, 1983, 1998; Olsen 2000; Bell 2005, 2011; Bell & Plag 2012) that the evidence for a class of syntactic N+N constructions is limited. Here, the term *N+N compound* is applied (quite loosely) to any N+N construction that consists of two adjacent nouns, in which one modifies the meaning of the other, or where together they have a single meaning different from the meaning of either constituent individually. Proper names (e.g. *James Smith*), constructions with an appositive modifier (e.g. *lady Jane*), and sequences like *letter A*, *number nine* are excluded.

Thus, this investigation is concerned with English endocentric N+N compounds, in which a concept denoted by the lexical complex N1+N2 is an instance of the concept denoted by N2. In my research I follow the commonly accepted definition offered by Bloomfield (1933: 235), who identifies such compounds semantically, by means of hyponymy, i.e. the referents of such compounded structures are hyponyms of their head elements. For example, *apple tree* is a kind of tree, *cookie jar* is a kind of jar, and *sun glasses* are a kind of glasses. I focus on analytical structures, also called *root compounds* or *primary compounds* (Harley 2009: 139), in which the head element is not derived from a verb, and the whole sequence does not yield a Verb-Object interpretation, i.e. *synthetic compounds* in Bloomfield's (1933) terms, or *verbal-nexus compounds* in Marchand's (1969) terms. This means that expressions like *taxi-driver* and *language teacher* are not included into this analysis.

It is important to note a certain fuzziness in terms of categorizing compounds into different types. Fabb (1998: 67), in the discussion of the matter gives the example of *greenhouse* whose being endocentric or exocentric depends on whether you think of it as a kind of house or not. This makes it obvious how the interpretation (and also the opinion of the analyst) may completely change our understanding of whether one and the same compound should be analysed as an endocentric or an exocentric one. Due to the limitations of space, I will not focus on the matters of exocentricity, as well as other issues regarding categorising nominal compounds into different types here and will refer to Bauer (2017) for a most comprehensive overview and discussion on this issue. In this study, I follow the assumption expressed in Bauer et al. (2013: 465) that an endocentric compound is one "[…] in which the compound as a whole denotes a subset of the head element of the compound".

3. Issues of productivity

Bauer et al. (2013: 451) state that English endocentric N+N compounds are the most productive kind of nominal compounds in English. By this they seem to consider that the pattern of forming such sequences is characterised by its *availability*, i.e. the scope of a word-formation rule, and *profitability*, i.e. the actual use of this rule in the process of forming new coinages (Kastovsky 1986: 586). The dichotomy between these two notions is discussed in Fernández-Domínguez (2015: 220), who demonstrates the inter-relation between availability and profitability of a word-formation rule and the ability of developing paradigmatic and syntagmatic relations in the language respectively. Consideration of both of these two notions is important for the synchronic analysis of a given word-formation pattern (in our case N+N compounding), since it allows the combination of the qualitative (availability) and quantitative (profitability) aspects of productivity.

The issue that seems to be overlooked in the research on productivity of compounds is productivity of nouns as constituents of compounds, and there are very few publications that look at the productivity of individual members of compounded structures in English. This question, however, is important for the discussion of how compounds are produced and what mechanisms are involved. A recent quantitative study by Hein & Engelberg (2017) looks at productivity values for nouns that are commonly used as constituents of German compounds, where the authors analyse such factors as *semantic proximity* and *frequency of the head noun* in order to account for the degree of productivity with which certain nouns are used as constituents of compound words.

The question of productivity of nouns as constituents of compounds is relevant for English too. One of the opinions presented in the literature is that the distribution of nouns between modifier and head roles is relatively equal in the BNC (e.g. Maguire et al. 2010). However, a closer look at the proportion of noun occurrences in combination by modifier and head, as shown in Maguire et al. (2010: 61), reveals that nouns belonging to such categories as substance, plant, location and etc. demonstrate the preference to be used as modifiers. Thus, we can see that on the one hand, nouns in general seem to be used equally frequently in both N1 and N2 positions. On the other hand, the semantic content of certain nouns may promote their use in only one of the roles.

A quick search in the corpus for one of the frequent nouns denoting substance, i.e. *water*, supports this observation. According to the data extracted from the BNC, *water*, when used as a modifier, occurs three times more often and in a wider variety of combinations than as a head noun. To be more specific, the modifier family for the constituent *water* contains 48 types (e.g. *water authority*, *water industry*, *water services*, *water board*, *water charges*, *water temperature*, *water condition*, *water garden*, *water shares*, *water park*, etc.), whereas the head family contains only 15 types (*sea water*, *surface water*, *bath water*, *salt water*, *waste water*, *rain water*, etc.).

Baayen's (2010) study reports that out of 2200 compound constituents, 710 were used only as heads, 902 only as modifiers, and 588 both as heads and modifiers. Despite the considerable number of nouns that occur both as heads and modifiers, the regression models presented in the study demonstrate significant correlations of productivity ranks of the constituents with degrees of constituent productivity. This also means that the prevalence for one position cannot be limited to modifiers. The corpus data for the word *problem*, for example, shows that it can be found in the head position in 116 compounds (*back problem*, *crime problem*, *money problem*, *business problem*, *quality problem*, *family problem*, *disease problem*, *city problem*, *attitude problem*, *acquisition problem*, *housing problem*, etc.), but its occurrence as a modifier is very limited, with only 10 compounds where it is used in this role (e.g. *problem behaviour*, *problem children*, *problem situation*, *problem area*, *problem families*, etc.). This means that its head family size is much larger than that of the modifier.

This puts Maguire et al.'s suggestion about the equal distribution of nouns between the modifier and head roles into question and calls for further investigation of the issue with a more detailed analysis of individual nouns that are used as elements of compounds. However, consideration of the semantics of N+N compounds and their connection to the elements' preferential use in either of the positions should also be considered. Looking at the issue of productivity from the position of the constituents comprising a compound as well as the relations that these constituents realise, may help shed light on the question of how compounds are formed.

4. Issues of semantics

It is a distinguishing feature of English N+N compounds that as linguistic forms, they combine two or more discrete lexemes into a semantic whole, despite the fact that there are no grammatical markers to indicate what the relation between these two (or more) parts is, or how this relation has been obtained. It has been noted that endocentric compounds that have a common constituent realise different relations. Bauer (1979: 45) provides the example of the word *pill*, which, when used in combination with different modifiers, realises different meanings, e.g. a *sleeping pill* CAUSES somebody to sleep, a *seasickness pill* is used FOR seasickness, and an *antihistamine pill* CONTAINS antihistamines. The same happens if we make changes to the first constituent, e.g. *citizen army* vs. *liberation army*; *riot police* vs. *peace police*, *jewellery box* vs. *bargain box*. However, on the level of structural representation, we are still dealing with the combination of two nouns and there are no formal markers to indicate the changes in meaning.

This ability of nouns to realise different relations when they are put together into a syntagm has been the central topic of discussion in the research on the semantics of compounds for over half a century. The question of the semantics of N+Ns has been considered in a number of theoretical frameworks bringing up a number of important issues, the analysis of which helps shed light on the nature of this phenomenon (Lieber & Štekauer 2009). However, there is still no unanimous opinion on how many of such relations there are in language or how to classify them (for a brief discussion of the existing problems see Bauer & Tarasova 2013: 1–3). The number of semantic relations in different classifications ranges from four (Granville Hatcher 1960) to over a hundred (Brekle 1976), with the most influential being Levi's (1978) classification. Though not without drawbacks, Levi's set of nine fundamental relations, aka recoverably deletable predicates (RDPs), provides a manageable number of relations that can be applied to N+N compounds, as well as other nominal structures (Bauer & Tarasova 2013). Despite a number of problems with Levi's (1978) classification (see ten Hacken 1994: 44–49 and ten Hacken 2016: 4–5 for discussion), the practicality of Levi's set makes it the choice of convenience for this research as well, since it allows for demonstrating the trends important for the present discussion on the connection between productivity on the level of structural representation with the productivity on the semantic level. The list of relations with examples are given in Table 1.

Table 1: Levi's (1978) classification of RDPs

Meaning	RDP	Examples
N1 CAUSE N2	CAUSE1	*sex scandal*, *withdrawal symptom*
N2 CAUSE N1	CAUSE2	*tear gas*, *shock news*
N1 HAVE N2	POSSESSION1	*lemon peel*, *school gate*
N2 HAVE N1	POSSESSION2	*camera phone*, *picture book*
N1 MAKE N2	COMPOSITION1	*snowball*
N2 MAKE N1	COMPOSITION2	*computer industry*, *silkworm*
N2 USE N1	INSTRUMENT2	*steam iron*, *wind farm*
N2 BE N1	ESSIVE2	*island state*, *soldier ant*
N2 IS IN N1	LOCATION2	*field mouse*, *letter bomb*
N2 IS FOR N1	PURPOSE2	*arms budget*, *steak knife*
N2 IS FROM N1	SOURCE1	*business profit*, *olive oil*
N2 IS ABOUT N1	TOPIC2	*tax law*, *love letter*

As can be seen from Table 1, Levi's (1978) set of RDPs includes such central, highly abstract predicates as CAUSE, HAVE, MAKE, BE, USE, FOR, FROM, IN, ABOUT. The semantic relations in the compounds have a number marker, which signifies the directionality of the relation. Notably, at least three of the semantic relations (CAUSE, HAVE, MAKE) demonstrates bidirectionality, e.g. *tear gas* – 'gas CAUSE tears' vs. *drug deaths* – 'drug(s) CAUSE deaths' (Levi 1978: 76). The necessity for distinguishing between the directions of the relations is dictated by the difference in the readings of such compounds as *heat birth* and *problem children*, since in the first case N1 CAUSES/INDUCES N2, and in the second one N2 CAUSES N1. In the course of this study it was noticed that such oppositions are not limited to the three relations pointed out by Levi (1978: 76), and examples of compounds realizing contrasting directions of the same relation may require reconsideration of this issue, as shown in examples in (1)–(4).

(1) SOURCE: *photon energy* (N1 IS FROM N2) vs. *heart sounds* (N2 IS FROM N1)
(2) ESSIVE: *mansion house* (N1 IS N2) vs. *tower house* (N2 IS N1)[1]

[1] The distinction in the direction of the ESSIVE relation in compounds seems to be quite subtle and generally follows the principle: ESSIVE1 – every N1 is N2; ESSIVE2 – N2 which is N1. Looking at the above examples, we can see that every mansion is a house, which does not seem to be the case for every tower. However, this principle does not always seem to work. For example, for the compound *building business* the relation is

(3) TIME: *work day* (N1 IN/DURING N2) vs. *crisis decisions* (N2 IN/DURING N1)
(4) INSTRUMENT: *farm machinery* (N1 USE N2) vs. *wind farm* (N2 USE N1)

At the same time, it is necessary to point out that usually one of the directions is actively utilised, whereas the cases of the opposite direction can be marginal. For example, in the corpus collected for this research, compounds realising CAUSE2 relation occur only 22 times, whereas CAUSE1 is much more frequent, with 244 cases.

Another issue with Levi's (1978) classification concerns the necessity to distinguish between the SPACE and TIME relations (Levi views TIME as a subtype of SPACE). In the current research the relations of TIME and SPACE are separated and viewed as two distinct categories. The parallelism of the concepts of SPACE and TIME has been long discussed and the overall historical tendency of spatial expressions to develop temporal meanings is recorded across languages. In the cognitive approach, this parallelism is explained by Mapping Theory (Lakoff 1993; Lakoff & Johnson 1999; Boroditsky 2000; Radden 2003), according to which the abstract domain of time gets its structure from a more concrete domain of space. Studies on locative prepositions in English (Sandra & Rice 1995; Kemmerer 2005) suggest that language users clearly distinguish temporal meanings from spatial ones. The results of Kemmerer's (2005) research suggest that although the spatial and temporal meanings of prepositions are historically linked by virtue of the TIME IS SPACE metaphor, they can be represented and processed independently of each other in the brains of modern adults. It seems logical to suggest that the concepts of TIME and SPACE probably have separate entries in the mental lexicon and the same can be said about the corresponding semantic relations. Matlock et al. (2005) point out that although the domains of space and time share conceptual structure, with frequent use, mappings between space and time come to be stored in the domain of time and so thinking about time does not necessarily require access to spatial schemas.

5. Number of relations per noun

The next question addressed here is whether a given noun can realise any possible relation when combined with another noun to form an N+N compound, or whether the number of relations for a single noun is limited. According to data from psycholinguistic research, semantic relations by which the overall meaning of an N+N compound is characterised are conceptual rather than lexical (Gagné et al. 2009) and should be viewed as "independent representational units that can be utilised by various and dissimilar concepts" (Estes & Jones 2006: 89). This makes it possible to assume that in theory, any relation can be used with any noun, e.g. *bear paw* (POSSESSION), *bear cave* (LOCATION), *bear medicine* (PURPOSE), *bear scare* (CAUSE), etc. A number of researchers support this opinion, and argue that different interpretations are possible for a single compound (e.g. Lieber 1992; Coulson 2001). At the same time, the fact that that different relations differ in their degrees of availability for individual nouns has been demonstrated in some empirical studies. For example, in the experiment described in Štekauer (2005), the participants were presented with novel compounds and were requested to propose and rate all possible interpretations for these

---

defined as ESSIVE1 because it is possible to say that building (as industry) is a kind of business, but business is building is somewhat clumsy. The decision on the directionality of the ESSIVE relation was taken based on the more natural interpretation of a compound.

compounds. Interestingly, almost all compounds received only one reading that was dominant.

This tendency goes in line with predictions of Competition Among Relations In Nominals (CARIN) model (Gagné 2000; Gagné & Shoben 1997), which is based on the assumption that availability of different relations varies for different concepts. For example, when the word *mountain* is used in the modifier role, we are more likely to expect that the entity denoted by the head will be LOCATED near/around/in a mountain (e.g. *mountain cloud*), rather than telling us ABOUT the mountain (e.g. *mountain magazine*) (Gagné 2000). It has been often suggested that the modifier element has more weight than the head in determining the relational reading a compound receives (Gagné 2001; Gagné & Shoben 1997, 2002; Gagné & Spalding 2004; van Jaarsveld et al.1994). Speakers apply the knowledge of how the modifier is used in various combinations thus building a database of all possible relations that are associated with the modifier. The psycholinguistic evidence obtained by Gagné & Shoben (1997, 2002) suggests that the selection of a relation for a novel compound is affected by the speakers' linguistic and statistical knowledge about how the modifier is generally used in the language, which is directly connected to the modifier constituent word family (henceforth CWF) size. The same was noticed for compounds in German and Dutch. Krott (2009) claims that the modifier family appears to be more influential than the head family, and its influence is clearly seen in the choice of linking elements, which is shown to be determined by the semantics of the modifier rather than the semantics of the head. Based on this, Krott (2009) assumes that the semantic properties of the modifier element are more valuable for compound processing. Moreover, this importance is attributed not to the first element in the compound but to the modifier. Storms & Wisniewski (2005) performed a psycholinguistic study of compounds in Indonesian, a language in which the modifier constituent follows the head. Their results suggest that the position of the modifier does not influence compound processing. This means that the subcategorisation process and meaning construal takes place on the conceptual level and may not be connected with the structural peculiarities of the language.

While the arguments regarding the role of the modifier in the process of interpretation are more than persuasive, the same is not necessarily true for the process of creation of compounds. The number of works that focus on the process of compound formation is limited and the evidence collected in empirical studies is inconclusive at this stage. The present investigation follows the assumptions of the slot-filling models (Murphy 1988; Wisniewski 1997), which suggest that the process of compound formation is head-oriented. This means that the semantics of the head noun serves as the schema whose properties are responsible for attracting certain modifiers based on the slots that are more available for filling. The availability of slots is determined primarily by the semantic properties of the head noun and only then by other factors, including our previous linguistic and non-linguistic experience (Tarasova 2013).

6. Data collection

In order to investigate the issue of productivity of a noun as a constituent of a compound and the issue of productivity of the semantic relation realised by this constituent, a database of CWFs was collected. This section provides information on the stages of the data collection process and limitations of the search, and presents the details of the statistical analysis that was performed to check the general trends in relation to the issues above.

500 examples of non-lexicalised nominal N+N compounds were picked from New Zealand printed media. The 500 examples were analysed and classified in accordance with Levi's (1978) set of RDPs; therefore, it was important to choose items that are semantically transparent and do not include constituents that are extensions of the meaning of existing lexemes by metaphor or metonymy. Established compounds were included as long as they satisfied these two criteria. Such caution was dictated by the necessity for consistency in the analysis of the semantic relations, as Levi's set of RDPs is aimed at non-lexicalised compounds (1978: 8); however, some of her examples are clearly institutionalised (in Bauer's (1983) understanding). The purpose of the analysis was not to define the relations between the constituents as precisely as possible (since the interpretation of the meaning of a compound is relatively subjective), but to show that similar patterns recur and to suggest the reasons why this might be the case. The constituents of compounds from the corpus were then checked for their frequency of occurrence in the BNC, and the ones whose frequency of occurring in the corpus (irrespective of whether they are used as single words or elements of compounds) was 1000 or more occurrences per 100,000,000 were considered candidates that could form the basis for collecting a corpus of CWFs. The frequency limitation was set in order to be able to collect a representative corpus of compounds in which a given word may occur either in the head or modifier position, and a higher frequency of occurrence of a noun in the corpus was viewed as a predictor of the noun's ability to form extended CWFs.

Overall, the results of the semantic analysis of the relations connecting the elements of compounds in the collected corpus showed that the distribution of the relations in compounds is not even, and some of the categories occur more often than others. Almost two thirds of all the compounds in the collected data set realised only three semantic relations: PURPOSE (34%), POSSESSION (16.2%) and LOCATION (14.4%). Such uneven distribution of the semantic categories is assumed to be natural for English endocentric N+N compounds in general, and some semantic categories seem to occur more frequently in certain nominal constructions (Bauer & Tarasova 2013).

Such an uneven distribution of semantic relations required stratification of the sample in order to include compounds from different categories. Proportionate allocation strategy was used; i.e. a sampling fraction (the ratio of the sample size to the size of the stratum) was picked in each of the strata that was proportional to that of the total collected corpus. 50 compounds were chosen from the set of 500 items for the corpus experiment.

100 different constituents comprising the selected 50 items served as the basis for extracting the paradigms of compounds, i.e. CWFs, in which each of the constituents occurs. The semantic relations realised in compounds comprising CWFs were determined based on Levi's (1978) classification. In the process of picking the sample, the distinction into the different directions of a relation was not made, since for some categories the number of nouns was not large enough to be included in the stratified sample. However, some recent works in compounding (e.g. Plag et al. 2007; Kunter 2011) demonstrate the necessity to extend Levi's list in order to make a distinction in the direction of the relation.

The search engine of *BYU-BNC: The British National Corpus* (Davies 2004–) was used to collect N+N combinations in which every single noun of the 100 nouns that constitute the selected compounds may occur. Individual nouns were typed into the *search* field and [nn*] was typed into the *collocates* field. The collocations were limited by one noun before and one noun after the searched item. Then, where there was one, the plural form of the same noun was put into the search engine in order not to miss compounds in which the searched constituent in the head position does not occur in the BNC in the singular form.

The search of the BNC database was targeted and the examples were hand-picked in order to avoid the combinations that would not satisfy the criteria, as in examples like *century house* and *umbrella back* in (5) and (6), which either have a different bracketing structure (as in the first case) or are not compounds at all.

(5) so we found a minister who was very excited by the idea of marrying us in a seventeen [sic] **century house**, and in fact, on the morning of the wedding he was more worried about what he was wearing than what I was!

(6) Give me my **umbrella back**.

Since the current research is concerned with the constituent family size, rather than the constituent family frequency, there were no limitations concerning the frequency of the collected items. As this study is concerned with non-lexicalised compounds, a high frequency of occurrence of all individual compounds within a CWF cannot be expected; moreover, it is generally believed that lexicalisation is strongly associated with high frequency. At the same time, lower frequency compounds, especially compound hapaxes, i.e. the types that have only one token in a given corpus, provide important information about the development of the morphological pattern at the current time (Baayen & Lieber 1991; Baayen 1992, 1993; Ricca 2010), and, hapaxes in particular, are viewed as indicators of productivity.

Manual picking of the examples was also necessary to exclude synthetic compounds. It does not seem possible to set the limitations for these using the corpus tools, as the morphological criterion offered by Lieber (1992) does not always work. For instance, if we set the limitations on nouns with the agentive endings *-er* and *-or* in the head position, then compounds with head words like *professor* and *minister* will be excluded from the analysis. In the current research I do not follow Lieber's (1992) suggestion that all compounds in which the head element is a deverbal noun should be considered synthetic. In this research compounds are considered synthetic only if the modifier functions as a direct object of the derivational base of the head noun in the interpretation, e.g. *literature teacher* – teach literature, *company management* – manage a company. Combinations like *art competition*, *tourist information*, *state prosecutor*, *weight division*, etc. are not considered to be synthetic since they require relational interpretations. The so-called non-affixal deverbal compounds, i.e. NDVCs (Lieber 2010) like *history research*, *business talk* are included in the corpus if their interpretation involves one of Levi's (1978) semantic relations.

As a result of the data collection, 7332 compounds that make up 197 head and modifier constituent word families were collected for further analysis. Although the expected number of constituent word families is 200, three of the nouns (*matter*, *future*, *figure*) appear only in one position in N+N combinations in the collected corpus. These were still included in the statistical analysis. The semantic relations realised in the collected N+N sequences were analysed and the number of occurrences of each of the compounds in a given constituent word family was checked.

7. Data analysis

The final stage of the data collection involves statistical analysis in order to obtain a picture of the general trends.

7.1 *Productivity of form*

The first hypothesis tested through statistical analysis was aimed at answering the question of whether a compound constituent is used more productively in one position (modifier or head) but not the other.

In order to test this the collected corpus data were coded on the basis of a table that listed for each compound constituent the number corresponding to the family size of this constituent as it is used in the head position (famSizeN2) and in the modifier position (famSizeN1). Table 2 gives an overview of the distribution of family sizes. Family sizes of 0 occurred only once in N2 position, and only twice in N1 position (with each position having 100 different nouns).

Table 2: Distribution of family sizes

Constituent	Minimum	Maximum	Mean	Standard Deviation
N1	0	83	31.79	20.47
N2	0	116	45.95	30.26

To reduce the skewing of the distribution of the family sizes, and in order to alleviate possibly harmful effects of extreme values on our statistical models, the family sizes were log transformed (see Baayen 2008: 38). In doing this, we also follow the standard practice with compound family size measures in other studies (e.g. Bell & Plag 2012). Since some of the frequencies were 0, we added 1 to all family sizes before log transformation. The statistical analysis at this stage was carried out using the statistical package R (Pinheiro & Bates 2000; Bates et al. 2007; R Development Core Team 2011).

A linear regression model with the log family size of N1 as the independent variable and the log family size of N2 as the dependent variable was devised. The model with the best fit included a quadratic term for the independent variable and showed highly significant correlations between the terms for family size of N1 and family size of N2. Table 3 gives the coefficients of the model.

Table 3: Model coefficients for linear model showing the relationship between N1 and N2 family sizes. N1 = 98, Adjusted R2 = 0.3796

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	3.43574	0.42932	8.003	2.65e-12 ***
logFamSizeN1	1.07457	0.32638	3.292	0.00139 **
$(\text{logFamSizeN1})^2$	-0.30143	0.06094	-4.946	3.17e-06 ***

The graph in Figure 1 below shows for each noun the size of its family in N1 position and in N2 position. The vertical and the horizontal axes represent the logs of family sizes for constituents used as heads (logFamsizeN2) and modifiers (logFamsizeN1) respectively. The dots represent the observed values for nouns that are attested both in N1 and N2 position. The solid line is the regression line, with the 95 percent confidence interval given by the two dotted lines. The negative coefficient for the quadratic term gives us the downward slope of the regression line (see, for example, Baayen 2008: Ch 4). As can be seen from the graph, the dots are not very concentrated. However, the distribution and the shape of the regression line clearly demonstrate that the larger the head family is for a given noun, the smaller the modifier family is for the same noun. This means that the more often a given noun occurs as a head of an N+N sequence, the less often it occurs as a modifier. As can be seen from the graph, the effect is not very pronounced for smaller N1 family size and becomes more pronounced with rising N1 family sizes.

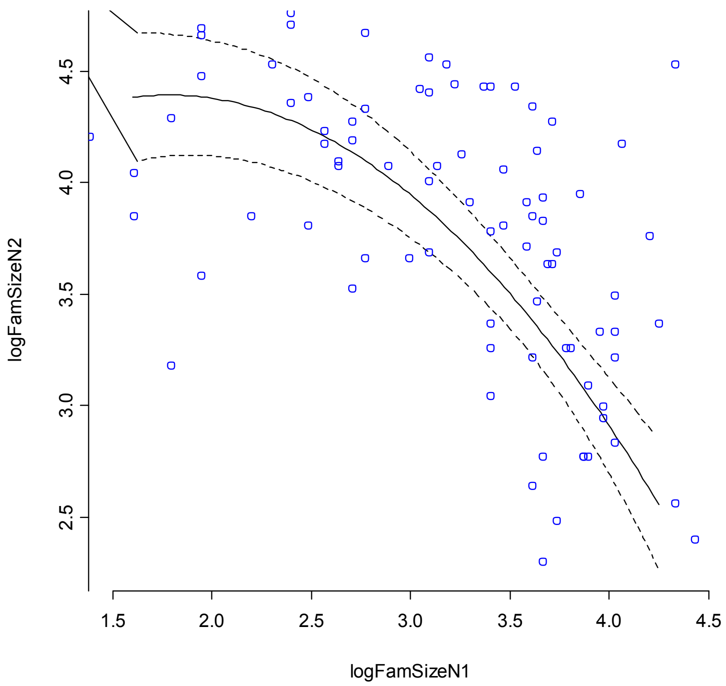

Figure 1: Relationship between family sizes

This means that it is possible to say that a noun when used as a constituent of a compound tends to be used more productively either as a modifier or as a head (but not both).

7.2 *Productivity of meaning*

Now I will turn to the next hypothesis, according to which we expect to see a connection between the family size and the constituent's concentration on one semantic relation. In order to test this, it is necessary to look at the two suggestions that it implies: (a) the constituent should demonstrate a preference for certain relations, or even one particular relation, over other relations; (b) this tendency should be stronger the larger the constituent family is.

To show that there is a preference for one semantic relation over the others within a given constituent word family, a measure called *instantiation index* is used. As a means to determine the instantiation index, the number of instantiations of different relations (DiffInstN1 and DiffInstN2) and the most frequent relation (HighestInstN1 and HighestInstN2) in each constituent family was calculated and added to the coded data table. Then the proportion of the highest instantiation among all family members was computed.

The diversity of meanings present in the family was factored in by multiplying the proportion of the highest instantiation in a CWF with the number of different relations attested in the family. For example, if we have a family size of 20 with eight different relations, and 10 compounds have the most frequent relation A, the index is (10/20)*8 = 4. If this family had only two different relations, the fact that 10 out of 20 have relation A is not so striking. This is exactly what we can see from the instantiations index, which will be much lower in this case: (10/20)* 2 = 1. So, we can assume that the higher the instantiations index is, the more concentrated the constituent family is on one relation.

In reference to suggestion (b), we should also expect the effect of the family size, i.e. the index should be higher for larger constituent families.

In order to test the above suggestions, the instantiation indices for constituents N1 and N2 were calculated and used for the analysis. The obtained measures were used for testing the family size effect. Two linear regression models (one for each constituent) were fitted in order to see if there is a connection between the family size and the constituent's concentration on one semantic relation. Log family size was used as an independent variable and log instantiation index as a dependent variable. Tables 4 and 5 below give the coefficients of the two models respectively.

Table 4: Model coefficients for linear model showing the relationship between family size N1 and instantiations of relations index for N1. N1 = 98, Adjusted R2 = 0.6503

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.34443	0.08427	4.087	9.06e-05***
logFamSizeN1	0.33816	0.02511	13.468	< 2e-16***

Table 5: Model coefficients for linear model showing the relationship between family size N2 and instantiations of relations index for N2. N2 = 99, Adjusted R2 = 0.4333

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0.40984	0.11722	3.496	0.000713***
logFamSizeN2	0.27787	0.03189	8.715	7.95e-14***

As can be seen from the tables, the family size is a highly significant predictor in both models, which, together with the positive coefficients, supports the suggestion that the constituent's concentration on one semantic relation is stronger in larger constituent families. Note also that these models have a very good fit and can explain a large proportion of the variance ($R^2 = 0.65$ for N1, $R^2 = 0.43$ for N2).

Figures 2 and 3 below represent the findings in graphical form. The vertical axes in both graphs show the measures for instantiation indices of semantic relations in a constituent family, and the horizontal axes show the log measures for the respective family size. The lines are the regression lines of the respective models.

The lines show that with increasing family sizes, the instantiation indices become higher. This also means that compounds with larger family sizes tend to be more concentrated on one relation than compounds with smaller constituent family sizes.

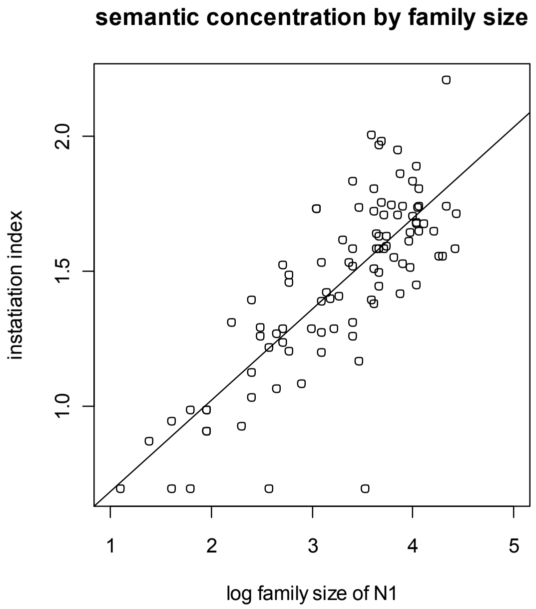

Figure 2: Relationship between N1 family size and the ratio of instantiations of different semantic relations

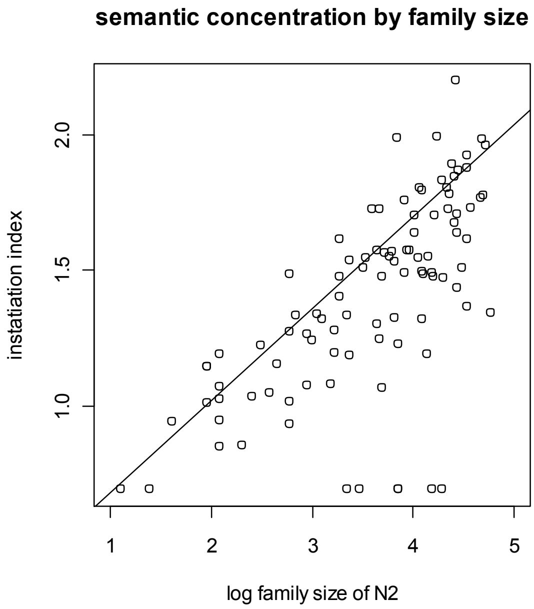

Figure 3: Relationship between N2 family size and the ratio of instantiations of different semantic relations

Thus, the obtained results demonstrate that there is a possible connection between the productive use of a compound constituent in one position and productive use of the semantic relation, which can be considered evidence of interaction between the levels of structural and semantic representation in compounds.

8. Discussion and conclusions

The first hypothesis that the analysis tested is whether, as a constituent of a compound, a lexeme is productive in one position and not the other. The obtained results clearly indicate a negative correlation between the sizes of the modifier and head families for one constituent. This means that with the increase of the head family size for a given noun, the size of the modifier family for the same noun decreases. This suggests that any given noun is used more productively either as a modifier or as the head of the compounded structure. This result is in line with Baayen's (2010) claim that the constituents of lexicalised compounds are position-bound. The fact that compounds used in the current study are not lexicalised, makes it possible to suggest that this claim holds true for non-lexicalised compounds too. This allows for a speculation that in the process of forming a new compound on the level of structural representation, for each noun or a group of semantically similar nouns, we follow a certain pattern that predetermines the use of a noun as a head or modifier. The development and solidification of this pattern in language can be connected with speakers' experience of how a noun is used in an N+N compound, and this experience may also be considered responsible for the extension of the constituent family paradigm, which becomes a dominant for this noun. Inevitably, the speakers' experience will depend on the amount of exposure to how a noun is used. As discussed in Bauer et al. (to appear), the analysis needs to consider a combination of factors, rather than one single factor. The possible factors include (but are not limited to) high frequency of use of a limited number of established compounds within a paradigm, the number of such compounds in a paradigm, the use of compounds using a specific pattern (e.g. *water + N*, *N + problem*, etc.) in a wide range of contexts and genres, the importance of the concept denoted by a compound for the life of the language community, etc. Another preliminary speculation that can be put forward here concerns the morphological properties of the noun that demonstrates the tendency to be used in one position. For example, items that are adjective-like, e.g. *future*, *animal*, *chocolate*, *lemon*, have a higher probability of occurring in the roles typical of an adjective (attributive role in our case) and, therefore, are more often used in the modifier position than in the head position. The analysis of a larger data set might help shed light on this issue.

Another issue that needs to be pointed out here is that the degree of preference for being used in one position differs for different nouns and ranges from absolute (100%) to low (≥51%), which allows for the suggestion that morphological productivity of an element in one position should be viewed as a scalar phenomenon that may or may not change under the influence of the factors outlined above. The differences in the degree of productivity may be dependent on the number of factors that are involved in each individual case, with the productivity of a noun as a constituent of N+N compound increasing when more than one factor is involved.

This research has also looked into the semantic relations realised by noun concepts as constituents of compounds in order to check whether there is an overall preference for a limited number of semantic relations to be realised in a constituent family. The statistical analysis of the collected data demonstrates this trend. It has also been found that the size of the constituent family can be considered a strong predictor of the degree to which a constituent concentrates on one semantic relation. The results suggest that the possibility of instantiation of one particular relation over the others increases as the family size increases, notwithstanding a logical supposition that large CWFs should have a wider distribution of semantic relations. This tendency is slightly stronger for head CWFs than for modifier CWFs (See Tarasova (2013) for the discussion on possible reasons for this). The consistency in realising one relation varies for different concepts; but, as suggested by the statistical analysis, the preference for one relation is not random and is consistent in the collected corpus. This implies that there should be a connection between the productivity of a noun as a constituent of a compound and the consistent realisation of one semantic relation by this noun within a constituent family.

This finding has important implications for further research in this area. One of the areas for further exploration concerns the use of analogy in the process of coining a compound. In the cognitive framework analogy is defined by Blevins & Blevins (2009: 2) as "[…] a general cognitive process that transfers specific information or knowledge from one instance or domain (the analogue, base, or source) to another (the target). Sets of percepts, whether visual images, auditory signals, experiences, or dreams, are compared, and higher-order generalisations are extracted and carried over to new sets". It is obvious from this definition that analogy is not limited to the structural patterns but also involves other levels, e.g. level of conceptual representation.

Analogy is widely discussed in derivation since it can be seen in affixes that are used for deriving morphologically complex words. Affixes usually alter the meaning of a derivational base in a specifically prescribed way. Extending this to compounds, Booij (2008, 2010) claims that analogical compounding is based on an individual compound (model word) with an idiosyncratic meaning. This meaning must be known for new compounds formed by analogy from the model compound to be understood. Analogy can be traced using semantic evidence when a particular idiosyncratic interpretation recurs in newly coined words (Booij 2008: 37). However, following the opinion expressed in Krott et al. (2007: 27) and Schlücker & Plag (2011: 1542), it is assumed here that analogy does not have to be driven by individual model words, but also involves paradigms that function as the basis for analogy. In this case not a single compound but a set, i.e. a paradigm (in our terms, constituent word family and its characteristics), predetermines the formation of new items. Krott et al. (2007: 27–28) use the term *paradigmatic analogy* to refer to the type of analogy in which "[…] the selection is based on the similarity of the target compound to a set (i.e., paradigm) of compounds, opposed to its similarity to a single exemplar, i.e., a single compound".

The studies to date, including the study on the use of linking elements in Dutch and German (Krott et al. 2007), the studies on stress assignment in compounds (Plag et al. 2007; Plag 2010), and the studies on the interpretation of the semantic relations (Gagné & Shoben 1997; Gagné 2001; Estes & Jones 2006, 2008), provide sufficient evidence that a paradigmatic approach to analogy provides a good account for the patterns noted in the formation of compounds. The analogical approach claims that the formation of new complex lexemes is based on the paradigms of similar existing complex lexemes and their formal properties rather than on abstract rules (Schlücker & Plag 2011: 1540). Therefore, the form of new compound coinages relies on the formal and semantic properties that the constituent words in the new combinations share with other compounds these constituents occur in.

The results obtained in the course of the current study indicate that the use of analogy in the formation of N+N compounds possibly takes place on at least two levels of representation: structural and semantic. It may also be the case that there should be interaction between the levels of structural and semantic representation in the process of forming compounds. Further research in this direction would helpful for a better understanding of the processes involved in the formation of compounds.

Acknowledgements

I would like to express my thanks and gratitude to Laurie Bauer, Natalia Beliaeva, Ingo Plag, and Gillian Claridge for their comments, advice and recommendations.

References

Aronoff, Mark. 1976. *Word formation in Generative Grammar*. Cambridge: Cambridge University Press.

Baayen, Harald. 1992. Quantitative aspects of morphological productivity. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of morphology 1991*, 109–149. Dordrecht: Kluwer.

Baayen, Harald. 1993. On frequency, transparency and productivity. In Booij, Geert & van Marle, Jaap (eds.), *Yearbook of morphology 1992*, 181–208. Dordrecht: Kluwer.

Baayen, Harald. 2003. Probabilistic approaches to morphology. In Bod, Rens & Hay, Jennifer & Jannedy, Stefanie (eds.), *Probabilistic linguistics*, 229–287. Massachusets: The MIT press.

Baayen, Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, Harald. 2010. The directed compound graph of English: An exploration of lexical connectivity and its processing consequences. In Olsen, Susan (ed.), *New impulses in word-formation* (Linguistische Berichte Sonderheft 17), 383–402. Hamburg: Buske.

Baayen, Harald & Lieber, Rochelle. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29(5). 801–844.

Baayen, Harald & Renouf, Antoinette. 1996. Chronicling the Times: Productive lexical innovations in an English newspaper. *Languages* 72(1). 69–96.

Bates, Douglas & Maechler, Martin & Dai, Bin. 2007. lme4: Linear mixed-effects models using S4 classes. (http://lme4.r-forge.r-project.org/) (Accessed 2018-08-27.)

Bauer, Laurie. 1978. *The Grammar of nominal compounding with special reference to Danish, English and French*. Odense: Odense University Press.

Bauer, Laurie. 1979. On the need for pragmatics in the study of nominal compounding. *Journal of Pragmatics* 31(1). 45–50.

Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press.

Bauer, Laurie. 1998. When is a sequence of two nouns a compound in English? *English Language & Linguistics* 21(1). 65–86.

Bauer, Laurie. 2001. *Morphological productivity*. Cambridge: Cambridge University Press.

Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge: Cambridge University Press.

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

Bauer, Laurie, & Tarasova, Elizaveta. 2013. Meaning link in nominal compounds. *SKASE Journal of Theoretical Linguistics* 10(3). 12–18. (http://www.skase.sk/Volumes/JTL24/pdf_doc/01.pdf) (Accessed 2018-08-27.)

Bauer, Laurie & Beliaeva, Natalia & Tarasova, Elizaveta. (To appear). Productivity revisited. *Zeitschrift für Sprachwisenschaft.*

Bell, Melanie. 2005. Against nouns as syntactic premodifiers in English noun phrases. *Working Papers in English and Applied Linguistics* 11(1). 1–48.

Bell, Melanie. 2011. At the boundary of morphology and syntax: Noun noun constructions in English. In Galani, Alexandra & Hicks, Glyn & Tsoulas, George (eds.), *Morphology and its interfaces* (*Linguistics Today* 178), 137–167. Amsterdam: John Benjamins.

Bell, Melanie & Plag, Ingo. 2012. Informativeness is a determinant of compound stress in English. *Journal of Linguistics* 48(3). 485–520.

Blevins, James & Blevins, Juliette (eds.). 2009. *Analogy in grammar: Form and acquisition*. Oxford: Oxford University Press.

Bloomfield, Leonard. 1933. *Language*. New York: Holt.

Booij, Geert. 2008. Construction morphology and the lexicon. In Montermini, Fabio & Boyé, Gilles & Hathout, Nabil (eds.), *Selected proceedings of the 5th Décembrettes: Morphology in Toulouse*, 34–44. Somerville, MA: Cascadilla Proceedings Project.

Booij, Geert. 2010. Compound construction: Schemas or analogy? A Construction Morphology perspective. In Scalise, Sergio & Vogel, Irene (eds.), *Cross-disciplinary issues in compounding*, 93–108. Amsterdam: Benjamins.

Boroditsky, Lera. 2000. Metaphoric Structuring: Understanding time through spatial metaphors. *Cognition* 75(1). 1–28.

Brekle, Herbert. 1976. *Generative Satzsemantik im System der englischen Nominalkomposition*. München: Fink.

Coulson, Seana. 2001. *Semantic leaps: Frame-shifting and conceptual blending in meaning construction.* Cambridge: Cambridge University Press.

Davies, Mark. *2004–*, *BYU-BNC* (Based on the British National Corpus from Oxford University Press). (http://corpus.byu.edu/bnc/) (Accessed 2018-08-27.)

Estes, Zachary & Jones, Lara. 2006. Priming via relational similarity: A COPPER HORSE is faster when seen through a GLASS EYE. *Journal of Memory & Language* 55(1). 89–101.

Estes, Zachary & Jones, Lara. 2008. Relational processing in conceptual combination and analogy. *Behavioral & Brain Sciences* 31(4). 385–386.

Fabb, Nigel. 1998. Compounding. In Spencer, Andrew & Zwicky, Arnold (eds.), *The handbook of morphology*, 66–83. Oxford: Blackwell.

Fernández-Domínguez, Jesús. 2009. *Productivity in English word-formation: An approach to N+N compounding*. Bern: Peter Lang.

Fernández-Domínguez, Jesús. 2015. The syntagmatic and paradigmatic axes of productivity: An application to the onomasiological model of word-formation. *SKASE Journal of Theoretical Linguistics* 12(3). 218–237. (http://www.skase.sk/Volumes/JTL29/pdf_doc/10.pdf) (Accessed 2018-08-27.)

Gagné, Christina. 2000. Relation-based combinations versus property based combinations: A test of the CARIN theory and dual process theory of conceptual combination. *Journal of Memory & Language* 42(3). 365–389.

Gagné, Christina. 2001. Relation and lexical priming during the interpretation of noun-noun combinations. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 27(1). 236–254.

Gagné, Christina, & Shoben, Edward. 1997. The influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 23(1). 71–87.

Gagné, Christina, & Shoben, Edward. 2002. Priming relations in ambiguous noun-noun combinations. *Memory & Cognition* 30(4). 637–646.

Gagné, Christina, & Spalding, Thomas. 2004. Effect of relation availability on the interpretation and access of familiar noun-noun compounds. *Brain & Language* 90(1–3). 478–486.

Gagné, Cristina & Spalding, Thomas & Figueredo, Lauren & Mullaly, Allison. 2009. Does snowman prime plastic snow? The effect of constituent position in using relational information during the interpretation of modifier-noun phrases. *Mental Lexicon* 4(1). 41–76.

Giegerich, Heinz. 2015. *Lexical structures: Compounding and the modules of grammar*. Edinburgh: Edinburgh University Press.

Granville Hatcher, Anna. 1960. An introduction to the analysis of English noun compounds. *Word* 16(3). 356–373. (https://doi.org/10.1080/00437956.1960.11659738) (Accessed 2018-01-22.)

Harley, Heidi. 2009. Compounding in Distributed Morphology. In Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford handbook of compounding*, 129–144. Oxford: Oxford University Press.

Hay, Jennifer. 2003. *Causes and consequences of word-structure*. London: Routledge.

Hay, Jennifer & Baayen, Harald. 2005. Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences* 9(7). 342–348.

Hein, Katrin & Engelberg, Stefan. 2017. Morphological variation: The case of productivity in German compound formation. In Koutsoukos, Nikos & Audring, Jenny & Masini, Francesca (eds.), *Morphological variation: Synchrony and diachrony*, 11th Mediterranean Morphology Meeting, Online proceedings 2018, 36–50. (http://mmm.lis.upatras.gr/index.php/mmm/article/view/2871) (Accessed 2018-08-27.)

Kastovsky, Dieter. 1986. The problem of productivity in word formation. *Linguistics* 24(3). 585–600.

Kemmerer, David. 2005. The spatial and temporal meanings of English prepositions can be independently impaired. *Neuropsychologia* 43(5). 797–806.

Krott, Andrea. 2009. The role of analogy for compound words. In Blevins, James & Blevins, Juliette (eds.), *Analogy in grammar: Form and acquisition*, 118–136. Oxford: Oxford University Press.

Krott, Andrea & Schreuder, Robert & Baayen, Harald. 1999. Complex words in complex words. *Linguistics* 37(5). 905–926.

Krott, Andrea & Schreuder, Robert & Baayen, Harald & Dressler, Wolfgang. 2007. Analogical effects on linking elements in German compound words. *Language & Cognitive Processes* 22(1). 25–57.

Kunter, Gero. 2011. *Compound stress in English: The phonetics and phonology of prosodic prominence.* Berlin: de Gruyter.

Lakoff, George. 1993. The contemporary theory of metaphor. In Ortony, Andrew (ed.), *Metaphor and thought* (2nd edition), 202–251. Cambridge: Cambridge University Press.

Lakoff, George, & Johnson, Mark. 1999. *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.

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

Lieber, Rochelle. 1992. *Deconstructing morphology: Word formation in syntactic theory*. Chicago: University of Chicago Press.

Lieber, Rochelle. 2010. On the lexical semantics of compounds: Non-affixal (de)verbal compounds. In Scalise, Sergio & Vogel, Irene (eds.), *Cross-disciplinary issues in compounding*, 127–144. Amsterdam: Benjamins.

Lieber, Rochelle & Štekauer, Pavol. 2009. Introduction: Status and definition of compounding. In Lieber, Rochelle & Štekauer, Pavol (eds.), *The Oxford handbook of compounding*, 3–18. Oxford: Oxford University Press.

Maguire, Phil & Wisniewski, Edward & Storms, Gert. 2010. A corpus study of semantic patterns in compounding. *Corpus Linguistics & Linguistic Theory* 6(1). 49–73.

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

Matlock, Teenie & Ramscar, Michael & Boroditsky, Lera. 2005. On the experiential link between spatial and temporal language. *Cognitive Science* 29(4). 655–664.

Murphy, Gregory. 1988. Comprehending complex concepts. *Cognitive Science* 12(4). 529–562.

Olsen, Susan. 2000. Compounding and stress in English: A closer look at the boundary between morphology and syntax. *Linguistische Berichte* 181(1). 55–70.

Pinheiro, José & Bates, Douglas. 2000. *Mixed Effects Models in S and S-Plus*. New York: Springer.

Plag, Ingo. 1999. *Morphological Productivity*. Berlin: Mouton de Gruyter.

Plag, Ingo. 2010. Compound stress assignment by analogy: The constituent family bias. *Zeitschrift für Sprachwisenschaft* 29(2). 243–282.

Plag, Ingo & Kunter, Gero & Lappe, Sabina. 2007. Testing hypotheses about compound stress assignment in English: A corpus-based investigation. *Corpus Linguistics & Linguistic Theory* 3(2). 199–233.

R Development Core Team 2011. Vienna: R Foundation for Statistical Computing.

Radden, Günter. 2003. The metaphor TIME AS SPACE across languages. In Baumgarten, Nicole & Böttger, Claudia & Motz, Markus & Probst, Julia (eds.), *Übersetzen, Interkulturelle Kommunikation, Spracherwerb und Sprachvermittlung – das Leben mit mehreren Sprachen. Festschrift für Juliane House zum 60. Geburtstag. Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 8(2/3). 226–239. (http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.91.7412&rep=rep1&type=pdf) (Accessed 2018-08-27.)

Rainer, Franz. 2004. Studying restrictions on patterns of word-formation by means of the Internet. *Italian Journal of Linguistics* 15(1). 131–139.

Ricca, Davide. 2010. Corpus data and theoretical implications: With special reference to Italian V-N compounds. In Scalise, Segio & Vogel, Irene (eds.), *Cross-Disciplinary Issues in Compounding*, 237–254. Amsterdam: Benjamins.

Sandra, Dominiek & Rice, Sally. 1995. Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language users? *Cognitive Linguistics* 6(1). 89–130.

Schlücker, Barbara & Plag, Ingo. 2011. Compound or phrase? Analogy in naming. *Lingua* 121(9). 1539–1551.

Schultnik, Henk. 1961. Produktiviteit als morfologish fenomeen. *Forum der Letteren* 2(1). 110–125.

Štekauer, Pavol. 2005. *Meaning Predictability in Word Formation: Novel, context-free naming units*. Amsterdam: Benjamins.

Štekauer, Pavol & Valera, Salvador & Körtvelyessy, Livia. 2008. *Universals, tendencies and typology in word-formation: A cross-linguistic study*. Košice: Josef Šafárik University.

Storms, Gert & Wisniewski, Edward. 2005. Does the order of head noun and modifier explain response times in conceptual combination? *Memory & Cognition* 33(5). 852–861.

Tarasova, Elizaveta. 2013. *Some new insights into the semantics of English N+N compounds*. Wellington: Victoria University of Wellington. (Doctoral dissertation.)

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

ten Hacken, Pius. 2016. Introduction: Compounds and their meaning. In ten Hacken, Pius (ed.), *The semantics of compounding*, 1–11. Cambridge: Cambridge University Press.

ten Hacken, Pius. 2017. Compounding in Morphology. In Aronoff, Mark (ed.), *The Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. (http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-251) (Accessed 2018-08-27.)

van Jaarsveld, Henk J. & Coolen, Riet & Schreuder, Robert. 1994. The role of analogy in the interpretation of novel compounds. *Journal of Psycholinguistic Research* 23(1). 111–137.

Warren, Barbara. 1978. *Semantic Patterns of Noun-Noun Compounds*. Goethenburg: Acta Universitatis Gothoburgensis.

Wisniewski, Edward. 1997. When concepts combine. *Psychonomic Bulletin & Review* 4(2). 167–183.

Zwitserlood, Pienie. 1994. The role of semantic transparency in the processing and representation of Dutch compounds. *Language & Cognitive Processes* 9(2). 341–368.

Elizaveta Tarasova
Department of BCIS
IPU New Zealand
Palmerston North
etarasova@ipu.ac.nz

In SKASE Journal of Theoretical Linguistics [online]. 2019, vol. 16, no. 1 [cit. 2018-16-01]. Available on web page http://www.skase.sk/Volumes/JTL39/pdf_doc/04.pdf. ISSN 1336-782X.