

How to translate compounds into Russian?

Anonymous author

Abstract

The present study contributes to contrastive Germanic-Slavic linguistics through an empirical investigation of Norwegian compounds and corresponding constructions in Russian. First, it is demonstrated that five Russian constructions function as frequent equivalents of Norwegian compounds: simplex word, adjective + noun, noun + noun in the genitive, noun + prepositional phrase, and compound. Second, it is shown that the five Russian constructions differ in frequency, with simplex word as the most widely used option and compound displaying the lowest frequency. Third, it is proposed that although the choice between the Russian constructions is not fully predictable, it is possible to state statistical tendencies, and to this end five prototypical patterns are discussed in detail. Finally, it is argued that the Norwegian-based prototypical patterns must be supplemented with Russian-based generalizations that capture the meanings of the Russian constructions that correspond to Norwegian compounds.

Keywords: compound, contrastive linguistics, word-formation, Russian, Norwegian, genitive, adjective, prepositional phrase

1.0 Why compounds?

Of particular interest to contrastive linguistics are what we may call “one-to-many situations”, i.e. situations where one structure in one language corresponds to more than one structure in another language. Take aspect as an example. While Russian and other Slavic languages have an aspectual distinction between perfective and imperfective verbs, this distinction is absent in Scandinavian and other Germanic languages. This creates a problem for L2 learners and translators, since for each Scandinavian verb (e.g. Norwegian *skrive* ‘write’) there are two corresponding verbs in Russian, viz. imperfective *pisat’* and perfective *napisat’*. Conversely, Russian L2 learners and translators face a similar “one-to-many situation” with regard to definiteness. Since Russian lacks the distinction between indefinite and definite articles, each Russian noun (e.g. *kniga* ‘book’) has more than one corresponding structure in Germanic (e.g. Norwegian (*en*) *bok* ‘(a) book’ and *boken* ‘the book’). “One-to-many” situations are challenging for L2 learners, as anyone teaching Russian aspect to Scandinavian students or definiteness to Russian L2 learners of Scandinavian languages can testify to.¹

The present study addresses a “one-to-many situation” that has received less attention than aspect and definiteness. Scandinavian compounds, i.e. words whose stem consists of more than one stem, confront L2 learners and translators with several corresponding structures in Russian.² While some Scandinavian compounds (e.g. Norwegian *dampskip* ‘steamboat’) can be rendered as compounds in Russian (*paroxod* from *par* ‘steam’ and *xod* ‘going, motion’), in other cases the Russian corresponding structures involve an adjective

¹ While it seems clear that “one-to-many situations” are challenging for L2 learners, it is a debated question how important this factor is compared to other factors that influence language learning (Collins et al. 2009). I will not discuss this question in the present article.

² The exact definition of compounds in Norwegian is controversial (Johannessen 2001). Are compounds in Norwegian best described as words with stems consisting of two *stems* or words with stems consisting of two *words*? I will not discuss this question which is tangential to my study.

+ a noun (e.g. *kuchonnyj stol* for Norwegian *kjøkkenbord* 'kitchen table') or a genitive construction (e.g. *prodavec mašin* for Norwegian *bilselger* 'car salesperson'). The following questions arise: Which Russian constructions do most frequently correspond to Scandinavian compounds? To what extent is the choice between the Russian constructions predictable? The present study is an attempt to answer these questions.

In order to investigate these questions empirically, I created a database of all compound nouns from three Norwegian novels and their Russian translations, as well as three Russian works of fiction and their Norwegian translations:

- (1) a. Erlend Loe: *Naiv super/Naivno. Super*
- b. Herbjørg Wassmo: *Dinas bok/Kniga Diny*
- c. Nikolaj Frobenius: *Latours katalog/Katalog Latura*
- d. Anton Čechov: *Rassказы/Kvinneskjebner og andre fortellinger*
- e. Ljudmila Ulickaja: *Medeja i eë deti/Medeia og hennes barn*
- f. Andrej Kurkov: *Piknik na l'du/Døden og Pingvinen*

The texts were chosen because they are part of the RuN corpus of parallel texts; the coding of the texts in the RuN corpus made it possible to create a spreadsheet file where each row contained one Norwegian sentence and the corresponding sentence in Russian.³ Based on this document, all Norwegian compound nouns and their Russian equivalents were then identified and annotated manually. The resulting database contains 8,390 Norwegian tokens that represent a total of 4,636 different lexemes.

The contribution of the present study can be summarized as follows. First, in section 2 it is demonstrated that Norwegian compounds have five relatively frequent corresponding constructions in Russian (here mentioned according to decreasing frequency): simplex word, adjective + noun, noun + noun in the genitive, noun + prepositional phrase, and compound. Second, in section 3 it is argued that although the choice between the Russian constructions cannot be predicted with complete accuracy, a number of generalizations can be made. I propose five prototypical patterns. Third, in section 4 I show that in addition to Norwegian-based generalizations, generalizations about Russian grammar are also helpful, especially generalizations about the genitive case in Russian. My findings are summarized in section 5.

Before we turn to the analysis, two points need clarification. First, my data come from Norwegian and Russian, and my generalizations are therefore strictly speaking only valid for Norwegian compounds and their Russian equivalents. However, in view of the close relations among the Scandinavian languages, it is likely that most of my conclusions carry over to Danish and Swedish. To what extent the generalizations are relevant for other Germanic languages is less clear. Already Jespersen noted that although compounding is a central word-formation mechanism across Germanic, there are considerable differences among the Germanic languages: "numerous German compounds must be rendered in English by an adjective plus a substantive or some similar phrase" (Jespersen 1942: 38). I will furthermore not discuss to what extent the generalizations I propose are relevant for Slavic languages other than Russian.

A second note of clarification pertains to the relationship between contrastive linguistics and translation studies. Although the word *translate* occurs in the title, the

³ The RuN corpus is available at www.hf.uio.no/ilos/english/research/projects/run/corpus/. Notice that it was not possible to use the RuN corpus directly, since Norwegian compounds and their Russian equivalents have no formal properties that would make it possible to identify them through regular corpus searches. It was therefore necessary to use the texts the corpus is based on, and merge and annotate these texts.

present study is not meant as a contribution to translation studies. I limit myself to generalizations about Norwegian compounds and their Russian counterparts, but at the same time hope that this contribution to the contrastive study of Scandinavian and Slavic languages will be of interest also for specialists in translation studies.

2.0 Distribution: Norwegian compounds and their Russian counterparts

The following five constructions are the most frequent equivalents of Norwegian compounds in my dataset:

- (2) a. Adjective + noun: *kuchonnyj stol – kjøkkenbord* ‘kitchen table’,
- b. Noun + noun in the genitive: *prodavec mašin – bilselger* ‘car salesperson’
- c. Noun + prepositional phrase: *kol’co dlja ključej – nøkkelring* ‘key chain’
- d. Compound: *paroxod – dampskip* ‘steamboat’
- e. Simplex word: *kladbišče – kirkegård* ‘cemetery’

Notice that “simplex word” in (2e) is used in a broad sense so as to encompass all single words that are not compounds. In other words, the category includes nouns with derivational affixes (*kladbišče – kirkegård* ‘cemetery’, *konjux – stallgutt* ‘horseman’ and *vdovec – enkemann* ‘widower’), as well as underived nouns (*vokzal – togstasjon* ‘railway station’, *lif – kjoleliv* ‘bodice’, *svet – dagslys* ‘daylight’ and *sejf – jernskap* ‘safe’). Nominalized adjectives are also included in the “simplex word” category (e.g., *verdensrom – vselennaja* ‘cosmos’), and the same holds for a small number of cases where a Norwegian compound corresponds to a pronoun in Russian (for instance, the Norwegian noun *konjux* ‘horseman’ may be rendered as *on* ‘he’ in the relevant Russian text). In addition to the five constructions in (2), a number of less frequent patterns were attested. Abbreviations, for instance, are sometimes used in the Russian texts, while the Norwegian texts have compounds (e.g. *OMON – spesialpoliti* ‘special police force’ and *ZAGS – vielseskontor* ‘Civil Registry Office’),⁴ and on some occasions the translators have chosen to rephrase:

- (3) Norwegian: Vært på reisefot til stadighet. (Loe)
Russian: Ja by vsë vremja putešestvoval.
‘I’d be traveling all the time.’

Here, the Norwegian noun *reisefot* can be analyzed as a compound based on *reise* ‘journey’ and *fot* ‘foot’, but there is no corresponding noun in the Russian sentence, since the translator has chosen to rephrase and use the past tense form of the verb *putešestvovat* ‘travel’ instead.

Table 1 (raw numbers) and Figure 1 (percentages) summarize the distribution of the relevant patterns. Three conclusions can be drawn. First, it is shown that more than a third of all Russian examples (37%) involve simplex words (in the sense described above). Second, only in 4% of the examples do Norwegian compounds correspond to compounds in Russian. Third, among the three syntactic constructions (i.e. the constructions that involve combinations of words), adjective + noun (33%) is three times as frequent as noun + noun in the genitive (11%), which in turn is more than twice as frequent as noun + prepositional phrase (5%). Due to considerations of space, detailed analysis of all five constructions is not feasible in the present article. In the following, we will therefore focus on the three syntactic constructions and ask to what extent the choice between them is predictable.

⁴ Whether this indicates that Russian in general uses more abbreviations than Norwegian is a question that must be left open for further research.

	#lemmas
Adjective + noun	1531
Noun + noun in the genitive	505
Noun + prepositional phrase	223
Compound	190
Simplex word	1727
Other	460
Total	4636

Table 1: Distribution of Russian constructions (number of lemmas)

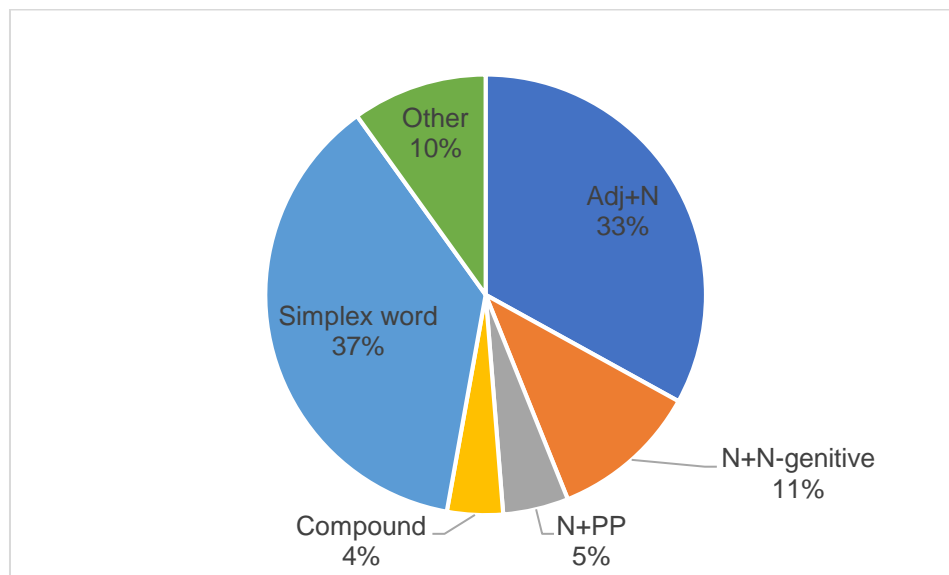


Figure 1: Distribution of Russian constructions (in percent)

3.0 Norwegian-based generalizations: Five prototypical patterns

Although we will see that it is not possible to create a deductive rule system that completely predicts the choice between the Russian constructions, based on statistical analysis I will propose five prototypical patterns representing typical correspondences between Norwegian compounds and Russian grammatical constructions.

3.1 Classification

In order to assess the predictability of the Russian constructions, I constructed a random sample of 500 Norwegian noun + noun compounds and their Russian equivalents from my database. Since the aim was to investigate the relationship between the three syntactic constructions in (2a–c), the random sample did not contain other Russian constructions. The relative frequencies of the three constructions in the random sample mirrors the relative frequencies in the dataset as a whole: 339 examples (68%) of adjective + noun, 112 examples (22%) of noun + noun in the genitive, and 49 examples (10%) of noun + prepositional phrase. Two limitations deserve mention: First, as mentioned, all Norwegian members of the random sample are so-called noun + noun compounds, i.e. compound nouns based on two nouns, such as *kjøkkenbord* ‘kitchen table’, which is based on the nouns *kjøkken* ‘kitchen’ and *bord* ‘table’. Second, only what we may call “parallel structures” are included. *Kjøkkenbord* – *kuchonnyj stol* is “parallel” insofar as both the concepts of ‘kitchen’ and ‘table’ are represented in both languages. “Non-parallel” are those cases where this condition is not fulfilled, e.g. *kaffedoktor* ‘(lit.) coffee doctor’, a mix of coffee and hard liquor that corresponds to *kofo s vodkoj* ‘(lit.) coffee with vodka’.

Evidently, there is nothing in the Russian construction that corresponds to *doctor* ‘doctor’, and this is therefore a non-parallel case.

Compounds have three elements that may form the basis of classification: the non-head constituent to the left (e.g. *kjøkken* in *kjøkkenbord*), the head constituent to the right (e.g. *bord* in *kjøkkenbord*), and the relationship between the two constituents (e.g. “place”; *kjøkkenbord* may be interpreted as a table located in a kitchen). All three elements were subjected to semantic classification in order to establish to what extent they facilitate predicting the choice between the three Russian syntactic constructions.

For the non-head and head constituents I adopted the following classification based on the so-called animacy hierarchy (see e.g. Sasse 1993 and Enger and Nettet 2011):

- (4) a. Human, e.g. *tannlege* ‘dentist’ in *tannlegestol* ‘dental chair’ and *eier* ‘owner’ in *fabrikkeier* ‘factory owner’
b. Animal, e.g. *bjørn* ‘bear’ in *bjørnejakt* ‘bear hunt’ and *fisk* ‘fish’ in *gullfisk* ‘goldfish’
c. Body part, e.g. *hjerte* ‘heart’ in *hjertesykdom* ‘heart disease’ and *hode* ‘head’ in *hestehode* ‘horse’s head’
d. Other concrete countable, e.g. *fly* ‘airplane’ in *flybillett* ‘airplane ticket’ and *hjelme* ‘helmet’ in *sykkelhjelme* ‘bike helmet’
e. Concrete mass, e.g. *gull* ‘gold’ in *gullkjede* ‘gold chain’ and *syltetøy* ‘jam’ in *blåbærsyltetøy* ‘blueberry jam’
f. Abstraction, e.g. *økonomi* ‘economy’ in *økonomiavsnitt* ‘economy department’ and *instinkt* ‘instinct’ in *jaktinstinkt* ‘hunting instinct’
g. Temporal notion, e.g. *vår* ‘spring’ in *vårblomst* ‘spring flower’ and *dag* ‘day’ in *avreisedag* ‘departure day’
h. Spatial notion, e.g. *sør* ‘south’ in *sørhimmel* ‘southern sky’ and *mark* ‘field’ in *slagmark* ‘battlefield’

In addition, I classified the non-head and head constituents as relational vs. sortal nouns (Barker and Dowty 1993, Barker 2011, Eckhoff 2011: 12-17, Stefanowitsch 2003). As suggested by the term, relational nouns describe relationships between arguments, while sortal nouns can be described in terms of inherent properties. Thus, while the relational noun *friend* can only be defined in terms of a relationship between people, the sortal noun *man* can be described in terms of inherent properties, such as “male” and “adult”. Typical examples of relational nouns are deverbal nouns such as *investigation*, which involves a relationship between an investigator (agent argument) and what is investigated (patient argument). The relational/sortal distinction was included in the annotation because Mezhevich (2002) has argued that is relevant for the translation of English compounds into Russian.

With regard to the relationships between the constituents of a compound, I have adopted the classification proposed in Eiesland (2015), a doctoral dissertation that classifies a large number of Norwegian compounds:

- (5) a. Place: the head is located in the non-head (*kjøkkenbord* ‘kitchen table’ is a table located in a kitchen)
b. Time: the head is used, takes place etc. at a time described by the non-head (*høstfrakk* ‘autumn coat’ is a coat used in the autumn)
c. Part of whole: the head is part of the non-head (*fiskekjeft* ‘fish mouth’ is a mouth that is part of a fish)
d. Source: the head comes from the non-head (*geitost* ‘goat cheese’ is cheese that comes from (the milk of) the goat)

- e. Material: the head is made from the non-head (*gullkjede* ‘gold chain’ is a chain that is made from gold)
- f. Used by: the head is used by the non-head (*mannesminke* ‘mens’ makeup’ is makeup used by men)
- g. Belonging: the head belongs to/is owned by the non-head (*kongsgård* ‘king’s farm’ is a farm owned by the king)
- h. Purpose: the head has the non-head as its purpose or goal (*bokhylle* ‘bookshelf’ is a shelf with the purpose of carrying books)
- i. Identity: the compound is both an example of the head and the non-head (*baby-pingvin* ‘baby penguin’ is both a baby (infant) and a penguin)
- j. Participant: the head involves a relation/action that the non-head participates in (*bjørnejakt* ‘bear hunting’ is an action where the bear participates (as the patient argument))
- k. Characteristic part (*fløtesaus* ‘cream sauce’ is a sauce that has cream as a characteristic part)
- l. Caused by: the head is caused by the non-head (*brannfare* ‘fire hazard’ is a hazard caused by fire)
- m. Comparison: the head is compared to/resembles the non-head (*sekkebukse* ‘baggy pants’ are pants that resemble a bag)
- n. Instrument: the head is an action that is carried out by means of the non-head (*telefonsamtale* ‘telephone conversation’ is a conversation carried out by means of telephone)
- o. Topic: the head is something that the non-head is about (*nissehistorie* ‘story about a gnome’ is a story about a gnome)

As Eiesland (2015) discusses extensively, there are borderline cases that are hard to classify. For instance, while as noted in (5a) *kjøkkenbord* ‘kitchen table’ may be described as a table located in a kitchen (“place”), one might also define it as a table designed for use in a kitchen, in which case the “purpose” relation could be invoked. Borderline cases like these make it impossible to carry out a completely objective classification, a fact that needs to be taken into consideration while interpreting the results.

3.2 Statistic analysis: Random Forest

In order to find out which attributes of the Norwegian compounds (described in the previous section) are the best predictors of the Russian constructions, I carried out a Random Forest analysis (Strobl et al. 2009). This is a statistical technique that divides the data material and predictors into subsets and creates a decision tree for each subset that tries to predict an outcome (in our case the choice between the Russian constructions) based on the available information (in our case the classification of Norwegian compounds). Each tree in the “forest” votes about the relative importance of the predictors, and Random Forest analysis therefore gives a reliable assessment of the situation as a whole, although caution should be exercised since we are dealing with a small dataset. The analysis returned the variable importance plot in Figure 2, which indicates the relative importance of the relevant predictors. In other words, Random Forest suggests that the relation between the constituents (classified in (5)) is a better predictor than the meaning of the non-head and head constituents. Among the properties pertaining to the non-head and head, the relational/sortal distinction is most important, while the meaning (as classified in (4)) of the non-head is more important than the meaning of the head. However, while Random Forest provides important information, the

analysis leaves a number of questions open. Some of these questions are addressed in the following section.

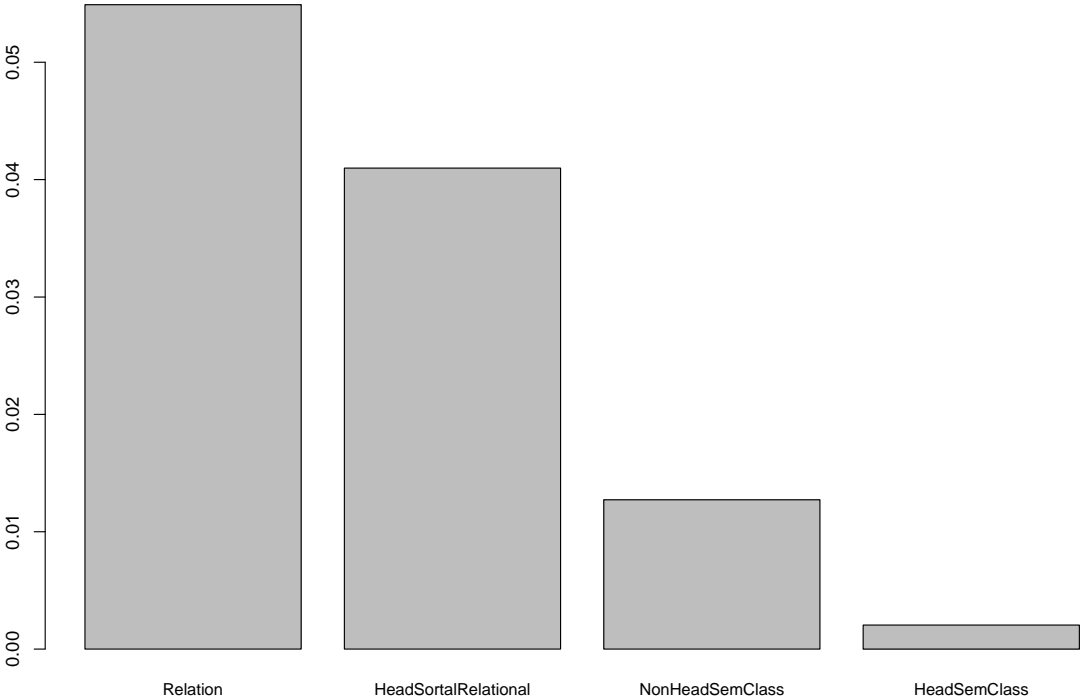


Figure 2: Variance importance plot from Random Forest analysis of 500-word sample

3.3 Statistics: Classification And Regression Tree

I furthermore carried out a Classification And Regression Tree (CART) analysis of the dataset (Strobl et al. 2009). In the same way as Random Forest, CART produces a decision tree that tries to predict the outcome based on the available information. However, CART creates one tree for the whole database, where each non-terminal node names the relevant predictor and gives a p-value indicating statistical significance. The predictions take the form of binary splits, as shown in Figure 3. First, the model selects the relation as a predictor (node 1) and makes a binary split between the “participant”, “part of whole” and “source” relations (node 2) and other relations (node 3). Under node 3, which takes the relational/sortal distinction for the head noun as its starting point, a further binary split is carried out, whereby non-relational (sortal) nouns and certain relational nouns (deverbal relational nouns and relational nouns denoting body parts and garments, node 4) are opposed to all other relational nouns (node 5).

The terminal nodes of the decision tree are bar diagrams where the leftmost bar represents the adjective + noun construction, the middle bar the noun + noun in the genitive construction, and the rightmost bar the noun + prepositional phrase construction. The y-axis of the bar diagrams represents the proportion of examples with each construction. Thus, node 2 indicates about 40% adjective + noun, 50% noun + noun in the genitive, and 10% noun + prepositional phrase for compounds with the “participant”, “part of whole” and “source” relations. These proportions must be compared to the random distribution in the dataset, which, as mentioned above, consists of 68% adjective + noun, 22% noun + noun in the genitive, and 10% noun + prepositional phrase. Since the actual distribution in node 2 has 40% adjective + noun instead of the

expected 68%, it is clear that the adjective + noun construction is underrepresented, while noun + noun in the genitive is overrepresented (50% instead of the expected 22%) and noun + prepositional phrase is as expected (10% of the examples). In other words, the “participant”, “part of whole” and “source” relations favor the noun + noun in the genitive constructions while slightly disfavoring the adjective + noun construction.

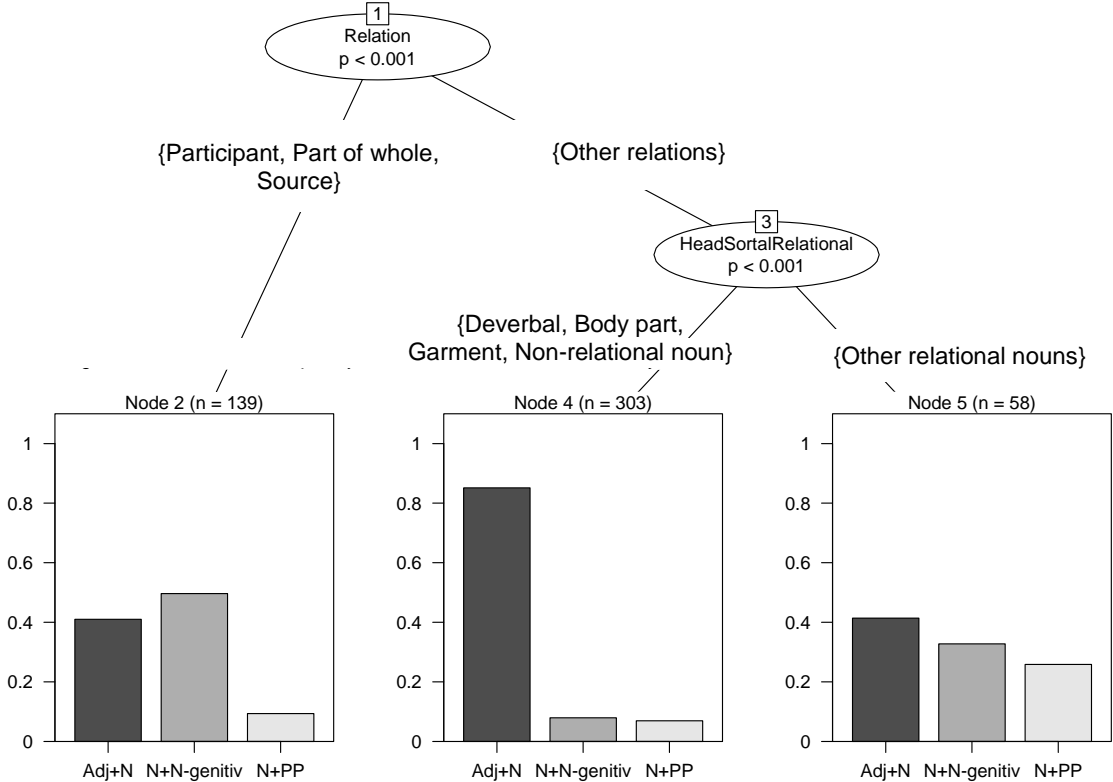


Figure 3: CART analysis of 500-word sample

We now turn to node 4, which is more important than node 2, insofar as node 4 covers 303 examples (n = 303), while node 2 only accounts for 139 examples. As shown, node 4 favors the adjective + noun construction for deverbal nouns, body parts, garments and non-relational nouns that do not involve the participant, part of whole or source relations. It may come as a surprise that deverbal nouns, which are typical examples of relational nouns, group together with non-relational nouns. However, this is presumably because numerous deverbal nouns involving the participant relation have been sifted out and thus end up under node 2. Node 5 only concerns deverbal nouns with other relations, and they seem to favor the adjective + noun construction in the same way as non-relational nouns.

Node 5, which covers all other constellations of predictors (a total of 58 examples), disfavors the adjective + noun construction (40% instead of the expected 68%). At the same time, noun + noun in the genitive and especially noun + prepositional phrase constructions are overrepresented.

What can we learn from the CART analysis in Figure 3? Since the analysis is based on a relatively small dataset and since, as mentioned above, the classification encounters problematic borderline cases, it is clear that the results must be interpreted with caution. Nevertheless, two conclusions can be drawn. First, the differences between the distributions in nodes 2, 4 and 5 indicate that we can formulate generalizations based on

relations and the relational/sortal noun distinction.⁵ Second, these generalizations are not categorical “all or nothing” rules, but rather statistical tendencies. It is, of course, possible that we have not identified the best predictors, and that other predictors would produce clearer tendencies or even categorical rules. However, at least until such predictors have been discovered, we will have to stick to statistical tendencies.

3.4 Prototypical patterns

A number of generalizations emerge from the statistical analysis presented in sections 3.2-3.3. In the following, I will zoom in on five such generalizations. The generalizations take the form of prototypes, since the statistical analysis has shown that categorical rules are not a realistic goal. While the prototypical patterns are based on the statistical analysis, in the following I will go beyond the 500-word sample analyzed in the previous section and include examples from my entire database.

Each prototype is a constellation of up to four components: (a) Russian construction, (b) semantics of the non-head of the Norwegian compound, (c) relation between non-head and head of the Norwegian compound, and (d) semantics of the head of the Norwegian compound.⁶ In addition, I provide four examples for each pattern (under (e) for each type). The first prototypical pattern has the following attributes:

- (6) Prototype 1:
- a. Russian construction: Adjective + noun
 - b. Non-head: concrete mass noun
 - c. Relation: material
 - d. Head: concrete count noun
 - e. Examples: *gullkjede – zolotaja cepočka* ‘gold chain’, *stråhatt – solomennaja šljapa* ‘straw hat’, *steinpir – kamennyj mol* ‘stone pier’, *mursteinshus – kirpičnyj dom* ‘brick house’

As shown in (6), the first prototype captures the generalization that a Norwegian concrete count noun whose non-head is a mass noun frequently corresponds to the adjective + noun construction in Russian. Presumably, the head can be any type of count noun, including jewelry, garments, as well as larger structures such as buildings. A variety of materials are possible, including metal, straw, rock, brick, etc.

The second prototype involves time:

- (7) Prototype 2:
- a. Russian construction: Adjective + noun
 - b. Non-head: temporal noun
 - c. Relation: time
 - d. Head: Inanimate
 - e. Examples: *natteskodde – nočnoj tuman* ‘night fog’, *høstfrakk – osennee pal'to* ‘autumn coat’, *morgengymnastikk – utrennjaja zarjadka* ‘morning exercise’, *førskolealder – doškol'nyj vozrast* ‘preschool age’

The generalization is that Norwegian compounds where the non-head is a temporal noun and the relationship is time tend to correspond to the adjective + noun construction in

⁵ With regard to the fifteen relations listed in (5a-o), it is interesting to notice that only three of the relations (participant, part of whole, and source) turn out to be relevant for the CART analysis. Whether this indicates that a less fine-grained classification with fewer relations may be more useful for the analysis of compounds, is a question that must be left open for future research.

⁶ The statistical analysis has shown that these four components are of different value as predictors, and as we will see, some components are left unspecified in some of the prototypical patterns.

Russian.⁷ Seemingly, there are no strong restrictions on the head noun, which in the attested examples range from concrete countable nouns like *frakk* ‘coat’ to processes like *gymnastikk* ‘exercise’ and abstract notions such as *alder* ‘age’. However, animate nouns are atypical, hence the specification “inanimate” in (7d).

The next prototype takes us from time to space:

(8) Prototype 3:

- a. Russian construction: Adjective + noun
- b. Non-head: spatial noun
- c. Relation: place
- d. Head: —
- e. Examples: *elvebåt – rečnoe sudno* ‘river boat’, *hagebenk – sadovaja skamejka* ‘garden bench’, *hjørnerom – uglovaja komnata* ‘corner room’, *gatefotograf – uličnyj fotograf* ‘street photographer’

This pattern accommodates the generalization that Norwegian compounds where the non-head is a spatial term that (as one would expect) relates the head to a certain place frequently correspond to the adjective + noun construction in Russian. I have not included any specifications for the head; although animate heads may be less common than inanimate heads, animate heads do occur, as shown by *gatefotograf* ‘street photographer’.

We now turn to prototypes concerning the noun + noun in the genitive construction:

(9) Prototype 4:

- a. Russian construction: Noun + noun in the genitive
- b. Non-head: —
- c. Relation: participant
- d. Head: Deverbal relational noun
- e. Examples: *sykkelprodusent – proizvoditel’ velosipedov* ‘bike manufacturer’, *solnedgang – zachod solnca* ‘sunset’, *skiveutglidning – smeščenje diskov* ‘disk prolapse’, *folkeforflytning – pereselenie narodov* ‘resettlement’

This prototype involves compounds with the participant relation, which as shown in the CART analysis in section 3.3 favors the noun + noun in the genitive construction (although adjective + noun is admittedly also frequent). The participant relation is typically found in compounds with deverbal heads, as specified in (9d). Since there appear not to be any clear restrictions on the meaning of the non-head noun, this slot is left open in (9b).

The final prototype involves compounds with the part of whole relation, which as shown in the CART analysis in section 3.3 frequently involves the noun + noun in the genitive construction:

(10) Prototype 5:

- a. Russian construction: Noun + noun in the genitive
- b. Non-head: —
- c. Relation: part of whole
- d. Head: relational noun

⁷ It may seem superfluous to specify that the relation is “time” when the non-head is a temporal noun. However, there are a few examples where temporal non-heads arguably involve other relations than “time”. A case in point is *minuttviser* ‘minute hand (on a watch)’ where the relation can be analyzed as “participant”, since *minutt* ‘minute’ is the patient argument (what is shown by the minute hand).

- e. Examples: *flyvinge – krylo samolëta* ‘airplane wing’, *guttehode – golova mal’čika* ‘boy’s head’, *dørklinke – ručka dveri* ‘door handle’, *jakkeslag – lackan pidžaka* ‘lapel’

This pattern represents the generalizations that compounds with a relational noun and the part of whole relation often correspond to the noun + noun in the genitive construction in Russian. In these compounds, the head describes something that is a part of something else, such as body parts or parts of garments or objects. No specification is given in (10b), since apparently there are no restrictions on the meaning of the head noun.

As shown, the five patterns discussed in this section concern two Russian constructions, viz. adjective + noun and noun + noun in the genitive. The former construction typically corresponds to three types of Norwegian compounds involving the relations material, time and space, whereas the latter construction corresponds to Norwegian compounds with relational nouns as heads.

Before we leave the five prototypical patterns, it is worth pointing out again that they are not categorical rules. Therefore, it is not difficult to come up with counterexamples. However, at the same time the prototypes represent frequent patterns, and as such they could be valuable for L2 learners and translators. The value increases, I suggest, when the knowledge about the prototypes is coupled with knowledge about the relevant constructions in Russian. This is the topic of the next section.

4.0 Russian-based generalizations

All generalizations stated so far have been based on the semantics of the Norwegian compounds. However, L2 learners and translators also have knowledge of the target language, Russian. How does this knowledge add to the picture presented in the previous section? In the following, I discuss a number of generalizations that may help us choose the correct construction corresponding to Norwegian compounds. The aim is not to give an exhaustive analysis of the relevant constructions, but rather by means of some well-known examples illustrate the interaction between Norwegian-based and Russian-based generalizations.

4.1 Stative vs. dynamic meanings

Janda and Clancy (2002: 111) describe the use of the Russian genitive in terms of a network with four nodes representing the following meanings:

- (11) a. A source, e.g. *uechat’ iz Moskvy* ‘leave Moscow’
 b. A goal, e.g. *dochat’ do doma* ‘drive up to the house’
 c. A whole, e.g. *čast’ tela* ‘part of a body’
 d. A reference point, e.g. *tufel’ka Diny* ‘Dina’s shoe’

The “source” and “goal” meanings are self-explanatory, while the “whole” meaning refers to what can be called “partitive constructions” in a broad sense, and the “reference point” category encompasses possessive constructions, as indicated by the example in (11d). While detailed discussion of Janda and Clancy’s analysis is beyond the scope of the present study, one important generalization emerges from the classification in (11):

- (12) a. The dynamical meanings “source” and “goal” use genitive *with* a preposition (noun + prepositional phrase).
 b. The stative meanings “whole” and “reference point” use genitive *without* preposition (noun + noun in the genitive).

This generalization has implications for Norwegian compounds. Since the noun + noun in the genitive construction is largely confined to the stative meanings “whole” and “reference point”, it follows that compounds involving dynamic relationships such as movement or direction will prefer the noun + prepositional phrase construction. Compounds describing physical motion or direction are not frequent in my dataset, but attested examples include *Reinsnes-jekt – karbas iz Rejnsnesa* ‘boat from the Reinsnes village’ and *trondhjemsfrue – skiperša iz Trondchejma* ‘madame from Trondheim’, where a person or object coming from a particular place is referred to.

The dataset also contains examples of food or drink, where some ingredient is described as the source: *viltsaus – sous iz diči* ‘game sauce’, *solbærtoddy – punš iz černoj smorodiny* ‘hot blackcurrant drink’ and *ripsgelé – žele iz krasnoj smorodiny* ‘redcurrant jelly’. While as expected in my database the noun + noun in the genitive construction is not attested for examples of this type, there is competition between the noun + prepositional phrase and the adjective + noun constructions (cf. prototype 1 in section 3.4). Data from the Russian National Corpus (main corpus, www.ruscorpora.ru) also illustrates this competition. For instance, the corpus has 18 attestations of “food/drink *iz černoj smorodiny*” and the same number of examples with the adjective *černosmorodinovyj* immediately followed by a noun denoting food or drink. Further discussion of this competition is beyond the scope of the present study.

Other constructions involving (metaphorical) dynamic or directional relationships use the preposition *dlja* ‘for’ and *iz-pod* ‘(lit.) from under’: *nøkkelring – kol’co dlja ključej* ‘key chain’, *crocketsett – prinadležnosti dlja kroketa*, *(tom) hatteeske – kartonka iz-pod šljap* ‘(empty) hat box’, and *(tomt) majonesglass – banočka iz-pod majoneza* ‘(empty) mayonnaise jar’. The examples with *dlja* involve a purpose, which can be analyzed as a metaphorical goal, while *iz-pod* indicates a source.

4.2 Constructions with measure nouns

In a number of publications, Partee and Borschev (2003, 2012a, 2012b, see also Borschev and Partee 2001 and 2004) have explored the semantics of Russian genitive constructions with measure nouns (e.g. *stakan moloka* ‘glass of milk’) and provided comparisons with related constructions (e.g. *stakan s molokom* ‘glass with milk’). A brief discussion of two of their observations is in order since these observations have consequences for the translation of Norwegian compounds.

One of their observations is that the noun + noun in the genitive construction in examples like *stakan moloka* ‘glass of milk’ implies that the glass is full, while no such implication follows from *stakan s molokom* ‘glass with milk’ with the noun + prepositional phrase construction. The full/not full difference is relevant for compounds, as shown by the following example from my dataset:

- (13) Norwegian: Hun støtte punsjeglasset ut av lensmannens hånd. (Wassmo)
 Russian: Ona vyrvala iz ruki lensmana stakan s punšem.
 ‘She kicked the punch glass out the police officer’s hand.’

Both *stakan punša* and *stakan s punšem* may correspond to the Norwegian compound *punsjeglass*, but the translator has chosen the second option, which is compatible with a glass that was not full. This makes sense since it is likely that the police officer had already drunk some of his punch before the glass was kicked out of his hands.

Other compounds translated by means of the preposition *s* ‘with’ in my dataset include *flaskereol – polka s bytylkami* ‘shelf with bottles’ and *bokholderskap – škaf s buxgalterskimi knigami* ‘closet with accounting books’. Why is the noun + prepositional phrase construction used instead of a noun + noun in the genitive? As pointed out by Borschev

and Partee (2004: 37), genitive constructions tend to be reserved for cases where a container can be used as a measure of its contents. This is not natural in the two examples with *polka* ‘shelf’ and *škaf* ‘closet’ since shelves and closets are not normally used to measure the quantity of bottles and books. The noun + prepositional phrase construction, on the other hand, does not have an affinity to measurement and is therefore a natural choice as Russian translations of the compounds *flaskereol* and *bokholderskap*.

4.3 Type descriptions vs. individual reference

One important difference between the adjective + noun and the noun + noun in the genitive constructions is that the former is a type description that involves a permanent relationship between the property (adjective) and the noun, whereas no such entailment holds for the noun + noun in the genitive construction (Rakhilina 2000: 88). By way of example, consider two possible equivalents of *loftsvindu* ‘attic window’, discussed in Rakhilina 2008: 342: *čerdačnoe okno* and *okno čerdaka*. The former describes a type of window, while the latter may refer to a window of an individual attic, as shown by the possibility of adding the determiner to the genitive noun (*okno ètogo čerdaka* ‘the window of this attic’). To what extent does this difference affect the translation of Norwegian compounds?

The non-head of Norwegian compounds are generally not inflected, and the bare form of the noun is compatible with type descriptions.⁸ Thus, *dameblad* ‘women’s magazine’ is a type of journal designed for women in general (not for any individual woman). As one would expect, *dameblad* corresponds to an adjective + noun: *damskij žurnal*. At the other extreme, the non-head of a Norwegian compound may refer to a unique individual, e.g. when the non-head is a proper name. An example from my dataset is Wassmo’s *Lorch-cello* ‘Mr. Lorch’s cello’, the cello of the protagonist’s teacher. For this compound, the translator has chosen the noun + noun in the genitive construction: *violinčel’ Lorka*.

As an illustration of the differences between the two constructions, consider two compounds from my database with *kontor* ‘office, as the non-head: *kontorgutt* – *kontorskij mal’čik* ‘office boy’ and *kontorvindu* – *okno kontory* ‘office window’. The former is used as a description of a type of job:

- (14) Russian: [Samonja] byl otpravlen nakonec v Odessu, gde i načal trudovuju dejatel’nost’ v kačestve kontorskogo mal’čika. (Ulickaja)
 Norwegian: [Samonja] ble omsider sendt til Odessa, der han tok fatt på sitt yrkesliv som kontorgutt.
 ‘Samonja was finally sent to Odessa where he started his career as an office boy.’

Since we are dealing with a type description, the use of the adjective + noun construction is expected. *Kontorvindu* occurs in the following context:

- (15) Norwegian: Lensmannen ble forskrekket. Han gløttet med nærsynte øyne ut gjennom kontorvindet da han hørte hesten. (Wassmo)
 Russian: Lensman ispugalsja Blizoruko priščurivšis’, on vygljanul v okno kontory, kogda uslyšal, čto kto-to pod”echal.
 ‘The police officer was alarmed. He looked out of the window with shortsighted eyes when he heard the horse come.’

Here, we are dealing with a window in a particular office, a situation that is adequately rendered in Russian with the noun + noun in the genitive construction. The alternative

⁸ For the purposes of the present study, I disregard exceptional examples like *fedreland* ‘fatherland’ where the non-head is an inflected form, in this case the plural form *fedre* of *far* ‘father’ (Faarlund et al. 1997: 62).

kontorskoe okno would be possible in Russian (and is attested in the Russian National Corpus) but would emphasize what kind of window the police officer looked out of – a less likely translation. Examples like these indicate that the difference between description of types and individual reference is relevant for the translation of Norwegian compounds, but that it is difficult to establish clear rules.

5.0 Concluding remarks

Based on an empirical investigation of Norwegian compounds and their Russian equivalents from six works of fiction I have argued that compounds represent a “one-to-many situation” that is challenging in the same way as aspect is problematic for Scandinavian L2 learners of Russian and definiteness presents challenges for Russian students of Scandinavian and other Germanic languages. The contribution of my study can be summarized as follows. First, it has been shown that Norwegian compounds frequently correspond to five Russian constructions: adjective + noun, noun + noun in the genitive, noun + prepositional phrase, compound, and simplex word. Second, it is shown that the five Russian constructions differ in frequency; the most frequent option is a simplex word, while the compound construction is least frequent; taken together, these five constructions constitute 90% of the examples in my database. Third, it is demonstrated that the choice of Russian constructions is not completely predictable, but that statistical tendencies exist. Fourth, I have proposed five prototypical patterns that represent correspondences between Norwegian compounds and Russian syntactic constructions. Fifth, I have shown that the Norwegian-based generalizations must be supplemented with Russian-based generalizations capturing the semantic subtleties of the Russian syntactic constructions.

A number of issues are left open for future research. Due to considerations of space, the present study has not explored Russian compounds and simplex words, although they function as Russian equivalents of Norwegian compounds. In-depth studies of these options are therefore called for. A limitation of the present study is the fact that it is based on a relatively small dataset from works of fiction. A broader investigation of a larger dataset from different genres would therefore be welcome. A final alley for future research is to find out to what extent the conclusions drawn in the present study are valid for compounds in Germanic languages other than Norwegian and corresponding structures in Slavic languages other than Russian. However, while much remains to be done, the present study represents a first step towards a systematic analysis of Norwegian compounds and the corresponding constructions in Russian – a field that has much to offer.

References

Primary sources

- Čechov, Anton. 1974-1983. *Polnoe sobranie sočinenij i pisem*. Moscow: Nauka.
Čechov, Anton. 1993. *Kvinneskjebner og andre fortellinger*. Oslo: Cappelen.
Frobenius, Nikolaj. 1996. *Latours katalog*. Oslo: Gyldendal.
Frobenius, Nikolaj. 2002. *Katalog Latura*. St. Petersburg: Azbuka.
Kurkov, Andrej. 2013. *Piknik na l'du*. Char'kov: Folio.
Kurkov, Andrej. 2013. *Døden og pingvinen*. Oslo: Cappelen Damm.
Loe, Erlend. 1996. *Naiv super*. Oslo: Cappelen.
Lu, Èrlend. 2008. *Naivno. Super*. St. Petersburg: Azbuka klassika.
Ulickaja, Ljudmila. 1996. *Medeja i eë deti*. Moscow: Vagrius.

Ulickaja, Ljudmila. 2002. *Medea og hennes barn*. Oslo: Bazar Forlag.
Vassmu, Xerb'erg. 2002. *Kniga Diny*. St. Petersburg: Azbuka klassika.
Wassmo, Herbjørg. 1989. *Dinas bok*. Oslo: Gyldendal.

Secondary sources

- Barker, Chris. 2011. "Possessives and relational nouns." In: *Semantics: An international handbook of Natural Language Meaning, volume 2*, edited by Klaus von Heusinger, Claudia Maienborn and Paul Portner, 1109-1113. Berlin: De Gruyter Mouton.
- Barker, Chris and David Dowty. 1993. "Non-verbal Thematic Proto-Roles." *Proceedings of North-Eastern Linguistics Conference (NELS-23)*, edited by Amy Schafer, 49-62. Amherst, MA: GLSA Publications.
- Borschev, Vladimir and Barbara H. Partee. 2001. "Genitive Modifiers, Sorts, and Metonymy." *Nordic Journal of Linguistics* 24 (2): 140-160.
- Borschev, Vladimir and Barbara H. Partee. 2004. "Genitives, Types and Sorts: The Russian Genitive of Measure." In *Possessives and Beyond: Semantics and Syntax*, edited by Ji-Yung Kim, Yury A. Lander and Barbara H. Partee, 29-43. Amherst, MA: GLSA Publications.
- Collins, Laura, Pavel Trofimovich, Joanna White, Walcir Cardoso and Marlise Horst. 2009. "Some Input on the Easy/Difficult Grammar Question: An Empirical Study." *The Modern Language Journal* 93: 336-353.
- Eckhoff, Hanne M. 2011. *Old Russian Possessive Constructions*. Berlin: De Gruyter Mouton.
- Eiesland, Eli Anne. 2015. *The Semantics of Norwegian Noun-Noun Compounds: A corpus-based study*. PhD dissertation: University of Oslo.
- Enger, Hans-Olav and Tore Nessel. 2011. "Constraints on diachronic development: the Animacy Hierarchy and the Relevance Constraint." *STUF* 64 (3): 193-212.
- Faarlund, Jan T., Svein Lie and Kjell I. Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Janda, Laura A. and Steven J. Clancy. 2002. *The case book for Russian*. Bloomington: Slavica.
- Jespersen, Otto. 1942. *A Modern English Grammar on Historical Principles. Part VI: Morphology*. Copenhagen: Ejnar Munksgaard.
- Johannessen, Janne B. 2001. "Sammensatte ord." *Norsk Lingvistisk Tidsskrift* 19 (1): 59-91.
- King, Bruce M. and Edward W. Minium. 2008. *Statistical reasoning in the behavioral sciences*. Hoboken: Wiley.
- Mezhevich, Ilana. 2002. "English Compounds and Russian Relational Adjectives." In *Proceedings of the North West Linguistics Conference 2002*, edited by Geoffrey S. Morrison and Les Zsoldos. Vancouver: Simon Fraser University: 95-114.
- Partee, Barbara H. and Vladimir Borschev. 2003. "Genitives, relational nouns, and argument-modifier ambiguity." In *Modifying Adjuncts*, edited by E. Lang, C. Maienborn and C. Fabricius-Hansen, 67-112. Berlin: De Gruyter.
- Partee, Barbara H. and Vladimir Borschev. 2012a. "Sortal, Relational, and Functional Interpretations of Nouns and Russian Container Constructions." *Journal of Semantics*, 29: 445-486
- Partee, Barbara H. and Vladimir Borschev. 2012b. "Dva stakana moloka: Substances and Containers in Genitive of Measure Constructions in Russian." *Russkij jazyk v naučnom osveščanii* 24 (2): 140-166.
- Rakhilina, Ekaterina V. 2000. *Kognitivnyj analiz predmetnyh imen: semantika i sočetaemost'*. Moscow: Russkie slovari.

- Rakhilina, Ekaterina V. 2008. "Semantika russkich imennykh konstrukcij s genitivom: 'ustojčivost'." In *Festschrift in Honour of Professor Arto Mustajoki*, edited by Jouko Lindstedt. Helsinki: University of Helsinki, 338-349.
- Sasse, Hans-Jürgen. 1993. "Syntactic phenomena in the world's languages I: categories and relations." In: *Syntax. Ein internationales Handbuch zeitgenössischer Forschung*, edited by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann, 646–686. Berlin: Mouton de Gruyter.
- Stefanowitsch, Anatol. 2003. Constructional semantics as a limit to grammatical alternation: The two genitives of English. In *Determinants of Grammatical Variation in English*, edited by Günter Rohdenburg and Britta Mondorf, 413-441. Berlin and New York: Mouton de Gruyter.
- Strobl, Carolin, James Malley and Gerhard Tutz. 2009. "An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests." *Psychological Methods* 14 (4): 323–348.