Norwegian compounds and corresponding constructions in Russian: The case of nouns with deverbal heads

Anonymous author

Abstract: This article presents a corpus study of Norwegian compounds with deverbal heads (e.g., papirproduksjon 'paper production' from produsere 'produce') and corresponding constructions in Russian, such as the genitive (proizvodstvo bumagi 'paper production'), the adjective (bumažnoe proizvodstvo 'paper production'), the preposition (priglašenie na užin 'dinner invitation'), and compound constructions (zemlevladelec 'landowner'). Test of the "Non-Head Function Hypothesis" (Mezhevich 2002) indicates that the genitive construction is the most frequent equivalent of Norwegian compounds where the non-head functions as an internal argument (object). However, the adjective and compound constructions represent important competitors, while the preposition construction is more marginal. The genitive construction is shown to be particularly frequent for non-agentive nouns. A number of generalizations about the use of compounds are proposed, and it is argued that the adjective construction involves "typification", which is an example of the general cognitive process "construal" (Langacker 2008). Finally, an "Extended Non-Head Function Hypothesis" is proposed, according to which the choice of a Russian construction depends on the closeness of the relation between head and non-head of the Norwegian compound. The closer the relation, the more likely is the use of the genitive. The more distant the relation, the more likely is the use of the adjective construction.

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

1. Problem and data

As pointed out in Nesset 2018a–b, compounds represent an interesting problem in Germanic-Slavic contrastive grammar, since a Germanic compound may correspond to a number of constructions in Slavic languages.¹ Consider the following examples from Norwegian and Russian (excerpted from the The RuN corpus):²

- (1) a. The genitive construction: *papirproduksjon-proizvodstvo bumagi* 'paper production'
 - b. The adjective construction: *papirproduksjon-bumažnoe proizvodstvo* 'paper production'
 - c. The compound construction: *jordeier-zemlevladelec* 'landowner'
 - d. The preposition construction: *middagsinvitasjon-priglašenie na užin* 'dinner invitation'

"One-to-many situations" of this type where one construction in the Germanic source language has more than one corresponding construction in the Slavic target language, are notoriously difficult for L2 learners and translators. The task that faces linguists working on contrastive grammar is to formulate generalizations that can help L2 learners and

¹ I use "construction" in the way this term is used in Construction Grammar, i.e. about "conventional, learned form-function pairings at varying levels of complexity and abstraction" (Goldberg 2013, 17). In the same way as Booij (2010) I treat morphological patterns such as compounds as constructions.

² The RuN corpus is available at http://tekstlab.uio.no/glossa2/run.

translators to choose the right construction in the target language. Nesset (2018a-b) proposed a number of generalizations about Norwegian and Russian; the present article is a follow-up study focusing on a particularly complex group of Norwegian compounds, viz. compound nouns with deverbal heads.

Although the exact definition of compounds in Norwegian has been debated (Johannessen 2001), for the purposes of the present study it is sufficiently precise to say that a compound is a word whose stem consists of more than one stem. Thus, *papirproduksjon* in (1a) is a compound, since its stem consists of the two stems *papir* 'paper' and *produksjon* 'production'.

In Norwegian, the second stem of a compound is the head, because this stem determines the grammatical categories of the compound as a whole. Thus, *rødvin* 'red wine' is a noun of masculine gender, since its head *vin* 'wine' is a masculine noun, while *vinrød* 'burgundy, wine red' is an adjective, inasmuch as its head *rød* 'red' is an adjective.³

A compound has a deverbal head if the head noun is derived from a verb. I adopt a traditional definition of "derived from", whereby X is derived from Y if the meaning of Y is included in the meaning of X. By way of example, consider the noun *læring* 'learning' as in *motivasjon er viktig for læring* 'motivation is important for learning'. We can define *læring* as *det å lære* '(the process of) learning', and since the meaning of the verb *lære* 'learn, teach' is included in the meaning of the noun *læring*, we can conclude that the noun is derived from the verb, i.e. that we are dealing with a deverbal noun.

Although the definition itself is clear, there are nevertheless cases where it is not straightforward to determine whether X is derived from Y or the other way around (Faarlund et al. 1997, 123). For instance, is *dikt* 'poem' derived from *dikte* 'write poems', or is it the other way around? The paraphrase *skrive dikt* 'write poems' suggests that the meaning of the noun is included in the meaning of the verb, and therefore that the verb is derived from the noun. At the same time, however, one might paraphrase the meaning of the noun as *resultatet av å dikte* 'the result of *dikte*', in which case the meaning of the verb is included in the meaning of the noun. In such cases, I have included the relevant nouns in my database in order not to exclude potentially interesting examples.

It is worth mentioning that the category of deverbal nouns as defined above is not limited to action nouns, i.e. nouns like *læring* 'learning' that denote the same action as the corresponding verb *lære* 'learn, teach'. An important subgroup of deverbal nouns are agentive nouns, i.e. nouns like *lærer* 'teacher' that denote the person who carries out the action denoted by the relevant verb.⁴ Since *lære* in Norwegian covers both 'learn' and 'teach', we can say that the meaning of the verb is included in the meaning of the agentive noun *lærer*, which we may define as 'a person who teaches'. In other words, agentive nouns like *lærer* meet the definition of "deverbal noun", and they are therefore included in the present study. Notice that I use the term "agentive" in a broad sense so as to include nouns denoting inanimate objects, such as *båndopptaker* 'tape recorder' and *tannpirker* 'toothpick'. Some words may have both animate and inanimate interpretations. Thus, *brødskjærer* 'bread slicer' could conceivably denote both a person who slices bread, and an instrument used for slicing bread. For a discussion of "agent" and other semantic roles, see Frawley (1992, 201-229).

In order to facilitate an empirical study of Norwegian compound nouns with deverbal heads and their Russian counterparts, I used the RuN corpus, which is a parallel corpus of

2

³ Admittedly, there exist scattered exceptions to this rule, e.g. *brennvin* 'hard liquor', which is a neuter noun despite the fact that *vin* is masculine (see Nesset 2016: 95 for discussion).

⁴ Note that the Norwegian word for 'learner' is *elev*, which I will not discuss in this article.

Norwegian and Russian texts.⁵ The corpus contains fiction and non-fiction, and includes texts that are translated from Norwegian to Russian and from Russian to Norwegian, as well as some texts that are translated into Russian and Norwegian from a third language. I searched for all Norwegian deverbal nouns with the suffixes mentioned in the authoritative grammar by Faarlund et al. (1997, 97-104) and manually weeded out all non-compounds. Thus, all my searches took Norwegian as their point of departure, regardless of which language the relevant texts were originally written in. The suffixes, for which I found relevant examples, are listed in Table 1.

Suffix	Example of deverbal noun	Corresponding verb
-ade	promenade 'promenade'	promenere 'walk'
-ant	representant 'representative'	representere 'represent'
-ari	havari 'accident'	havarere 'have an accident'
-ens	residens 'residence'	residere 'reside'
-ent	student 'student'	studere 'study'
-er	lærer 'teacher'	<i>lære</i> 'teach, learn'
-eri	<i>lureri</i> 'trickery'	lure 'cheat'
-ing/-ning	<i>åpning</i> 'opening'	<i>åpne</i> 'open'
-ium	studium 'studies'	studere 'study'
-nad	søknad 'application'	søke 'apply'
-sel	ferdsel 'traffic'	ferdes 'travel'
-sjon	operasjon 'operation'	operere 'operate'
-t	fart 'traffic'	fare 'travel'
-tør	ekspeditør 'shop assistant'	ekspedere 'serve (customer in shop)'

Table 1: Norwegian suffixes included in the present study (from Faarlund et al. 1997, 97-104)

In addition to suffixed deverbal nouns, Norwegian also have two types of deverbal nouns without derivational suffixes. These types are referred to as "conversion", where "the form of the base remains unaltered" (Haspelmath 2002, 24).6 The first type, which I will call "infinitive conversion", involves nouns that are identical to the infinitive form of the verb, such as *reise* 'journey' which is related to the verb *reise* 'to travel'. In order to identify such nouns in the corpus, I searched for all nouns ending in *-e*, and then weeded out irrelevant examples manually. In the second group of non-suffixed deverbal nouns we find examples like *hjelp* 'help (noun)', which equals the stem of the corresponding verb *hjelpe* 'to help'. This type I refer to as "stem conversion". Unfortunately, there is no systematic way to extract all nouns of the stem conversion type from the corpus; since there is no suffix involved, these nouns can have a wide variety of shapes. In order to find as many examples as possible, I searched for all relevant stems from the database of Nesset (2018a–b).

The corpus searches yielded a database of 5,708 tokens (text examples) and 2,987 types (lemmas).⁷ It is important to keep in mind that the corpus is dominated by fiction

⁵ Corpus searches were performed in May and June 2019.

⁶ An alternative analysis of conversion is in terms of zero suffixes. For arguments against zero suffixes, see Anderson (1992) and Nesset (1998).

⁷ When I count tokens (text examples) and types (lemmas), I count pairs of Norwegian and Russian words, such as *dørhåndtak – ručka dveri* 'door handle' and *dørhåndtak – dvernaja ručka* 'door handle'. Notice that these two pairs are counted as different, since one of the members of each pair is different. The total number of types (lemmas) is the number of unique pairs. The total number of tokens is the number of all pairs in the database. The database is available at the Tromsø Repository of Language and Linguistics (TROLLing): https://doi.org/10.18710/YRIO2V.

and that it is exclusively based on written language. However, despite these limitations the database covers all relevant types of deverbal nouns, and it is large enough to facilitate a number of generalizations about the relationship between Norwegian deverbal compounds and the corresponding constructions in Russian.

Table 2 shows the distribution of the Norwegian suffixes and conversion patterns in the database. While most suffixes have few examples, four patterns dominate: -er, -ing/-ning, infinitive conversion, and stem conversion. Together, these patterns constitute more than 90% of the database (91.5% for tokens and 90.5% for types). Notice that even if stem conversion is the only pattern for which it is not possible to extract all relevant examples from the corpus, this pattern is nevertheless by far the largest group in the database.

	Toke	Tokens		Types		
	Raw numbers	Percent	Raw numbers	Percent		
-ade	4	0,1	3	0,1		
-ant	11	0,2	10	0,3		
-ari	2	0,0	1	0,0		
-ens	5	0,1	3	0,1		
-ent	31	0,5	21	0,7		
-er	1022	17,9	462	15,5		
-eri	38	0,7	26	0,9		
-ing/-ning	830	14,5	577	19,3		
-ium	2	0,0	2	0,1		
-nad	6	0,1	5	0,2		
-sel	82	1,4	33	1,1		
-sjon	79	1,4	71	2,4		
-t	203	3,6	95	3,2		
-tør	20	0,4	14	0,5		
Infinitive conversion	848	14,9	240	8,0		
Stem conversion	2525	44,2	1424	47,7		
Total	5708	100,0	2987	100,0		

Table 2: Distribution of Norwegian morphological patterns in the database

The contribution of the present study can be summarized as follows. First, in section 2 I will test the "Non-Head Function Hypothesis" (Mezhevich 2002). This test indicates that the genitive construction is the most important Russian counterpart to Norwegian compounds where the non-head is an internal argument (grammatical object). However, the adjective and especially the compound constructions represent serious competitors, while the preposition construction is more marginal. In section 3, it is shown that the genitive construction is particularly frequent for non-agentive nouns. Section 4 proposes a number of generalizations about the use of compounds, before in section 5 it is argued that the adjective construction involves "typification" and more generally "construal" (Langacker 2008). Finally, in section 6, an "Extended Non-Head Function Hypothesis" is proposed, which states that the choice of a Russian construction depends on the closeness of the relation between head and non-head of the Norwegian compound. The closer the relation, the more likely is the use of the genitive. The more distant the relation, the more likely is the use of the adjective construction.

2. Testing the "Non-Head Function Hypothesis": do internal arguments yield genitive constructions in Russian?

In a contrastive study of English and Russian, Mezhevich (2002, 96) explores English compounds where the "head noun is derived from a verb and the non-head is interpreted as an internal argument of the head noun." By "internal argument" she means an argument that functions as a grammatical object of the corresponding verb. By way of example, consider *physics teacher*, where the non-head *physics* represents the internal argument of the corresponding verb *teach*: *teach physics*. Note that for the purposes of the present study I include both examples with direct objects and so-called preposition objects in the category of internal arguments. Consider *bjørnejakt* 'bear hunt', which corresponds to a verb followed by a prepositional phrase: *jakte på bjørn* 'hunt bear'. Such prepositional phrases are traditionally analyzed as grammatical objects in Norwegian, since the prepositions cannot easily be substituted by other prepositions (Faarlund, Lie & Vannebo 1997, 697; Enger and Kristoffersen 2000: 245, and Haugen 2006, 187). Since we are interested in how Norwegian compounds are translated, the definition of "internal argument" refers to Norwegian words, not to their Russian counterparts.

For nouns of the relevant type, Mezhevich argues that they are "usually expressed by means of a genitive construction" in Russian. Since the hypothesis focuses on the syntactic function of the non-head, we may refer to the hypothesis as the "Non-Head Function Hypothesis", although Mezhevich herself does not use this name. The hypothesis may be formulated explicitly as follows:

(2) The Non-Head Function Hypothesis (Mezhevich 2002)

Deverbal compounds where the non-head is an internal argument usually correspond to genitive constructions in Russian.

Focusing on theoretical issues, Mezhevich does not test this hypothesis against corpus data. Her study is concerned with English and Russian, but since English and Norwegian are similar in their use of compounds, it makes sense to test the hypothesis against Norwegian data. To this end, I excerpted all nouns in my database where the non-head represents an internal argument. I limited myself to Norwegian compounds that correspond to one of the four Russian constructions mentioned in section 1:

- (3) a. Genitive: høyberging-uborka sena 'hay harvesting'
 - b. Adjective: bjørnejakt-medvež'ja ochota 'bear hunt'
 - c. Compound: hærfører-polkovodec 'military commander'
 - d. Preposition: ballongjakt-ochota za šarami 'balloon hunt'

These constructions represent a good testing ground since in all of them both the head and the non-head are expressed in both languages. As shown in Nesset 2018b, Norwegian compounds sometimes correspond to single words (non-compounds), such as *huseier*—

-

⁸ In Norwegian, there is variation between *jakte på bjørn* with the preposition *på* 'on' and *jakte bjørn* without a preposition. I will not discuss this variation in the present article, although it suggests that preposition objects are closely related to direct objects without a preposition.

⁹ Notice that I do not claim that English and Norwegian are *identical* with regard to compounding. Already Jespersen (1942, 38) observed that there are differences among the Germanic languages, but compounding is nevertheless a central word-formation process in all Germanic languages, and English and Norwegian are sufficiently similar to make test of Mezhevich's hypothesis against Norwegian data meaningful.

chozjain 'landlord'. In such cases, the non-head is not expressed in Russian, and examples of this type therefore were not included in the test of Mezhevich's hypothesis.

The results of the test are summarized in Table 3 and visualized in Figure 1. Do the results match the predictions of the hypothesis? Let us first consider type frequency. Mezhevich (2002, 96) included the word "usually" in her hypothesis, and if we understand it as "the most frequent option", we may say that the hypothesis is confirmed, insofar as the genitive represents more than half the examples (51.7%), as shown in Table 3. Needless to say, however, we are not dealing with a particularly strong tendency in favor of the genitive.

	Type freque	ency	Token frequency		
	Raw numbers	Percent	Raw numbers	Percent	
Adjective	37	15.3%	54	13.4%	
Genitive	125	51.7%	172	41.8%	
Preposition	7	2.9%	10	2.4%	
Compound	73	30.1%	174	42.3%	
Total	242	100.0%	410	100.0%	

Table 3: Distribution of constructions where the non-head represents an internal argument

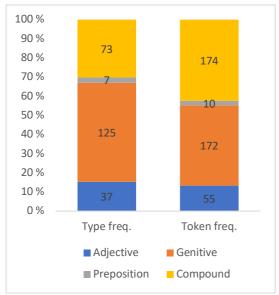


Figure 1: Distribution of constructions where the non-head represents an internal argument

For token frequency, the results match the predictions of Mezhevich's hypothesis to a lesser degree. As shown in the table, the compound and genitive constructions display nearly identical frequencies. In other words, while compounds are attested in fewer lemmas, some of these lemmas show higher token frequency than genitive constructions. We will return to this point in section 4.

Both type and token frequency indicate that the adjective construction is less frequent than the compound and genitive constructions. The preposition construction is the least frequent option in my database with less than 3% for both type and token frequency.

It is interesting to compare the findings in Table 3 to the results reported in Nesset 2018a, 247. The data from Nesset 2018a in Table 4 concern type frequencies of constructions corresponding to all kinds of Norwegian compounds, i.e. not only deverbal nouns where the non-head is an internal argument. Comparison with the type frequencies in Table 3 indicate that the genitive construction receives a much higher frequency for deverbal compounds of the relevant type (51.7%) than for compounds in general

(20.6%). The difference is statistically significant with a modest, but robust and reportable effect size.¹⁰ This speaks in favor of Mezhevich's hypothesis in (2). Clearly, a deverbal head and a non-head that functions as an internal argument are factors that increase the likelihood of using a genitive construction in Russian.

	Raw numbers	Percent
Adjective	1,531	62.5%
Genitive	505	20.6%
Preposition	223	9.1%
Compound	190	7.8%
Total	2.449	100.0%

Table 4: Distribution of constructions corresponding to Norwegian compounds (type frequency, not only deverbal nouns where the non-head is an internal argument). Adapted from Nesset 2018a, 247.

To sum up, the data under scrutiny offer some support to the Non-Head Function Hypothesis, although the situation appears to be more complex than stated in the hypothesis, since compounds and, to a lesser degree, the adjective construction are also well attested in my database. Two questions arise. First, will the data match the hypothesis better if we adopt a more fine-grained classification of the nouns in question? Second, are there generalizations that motivate the use of the compound and adjective constructions instead of the genitive construction? As we will see in the following three subsections, the answers to both questions are in the affirmative.

3. Are agentive nouns different?

We arrive at a more fine-grained classification if we distinguish between agentive nouns, such as *hærfører-polkovodec* 'military commander' and non-agentive nouns, such as *høyberging-uborka sena* 'hay harvesting'. As we will see, the genitive construction is by far the most frequent option for non-agentive nouns, whereas for agentive nouns compounds are as frequent as the genitive construction.

As mentioned in section 1, I adopt a broad definition of "agentive noun", insofar as both persons (e.g. *fysikklærer* 'physics teacher') and objects (e.g. *båndopptaker* 'tape recorder') are included. Even more fine-grained classifications of these nouns are indeed conceivable, and the same holds for non-agentive nouns. However, my dataset is too small to make quantitative analysis feasible for a more detailed classification, and in the following we will therefore stick to the distinction between agentive and non-agentive nouns.

It is worth mentioning that some agentive compounds are of the so-called parasynthetic type, whereby compounding is combined with suffixation (Bisetto and Melloni 2008, Johannesen 2017). A case in point is *ženonenavistnik–kvinnehater* 'woman hater, misogynist'. Notice that the base is arguably a verb phrase (*nenavidet' ženščin–hate kvinner* 'hate women'), rather than a compound, since there are no verbs **ženonenavidet'-*kvinnehate*. My database does not include enough examples to facilitate detailed investigation of parasynthetic compounds.

The data for agentive and non-agentive nouns are summarized in Tables 5 (type frequency) and 6 (token frequency), and visualized in Figure 2. The following

_

 $^{^{10}}$ I compared the type frequency of genitive vs. other constructions in Tables 3 and 4. Pearson's Chisquared test with Yates' continuity correction (X-squared = 116.56, df = 1) returned a p-value < 2.2e-16. The Cramer's V-value is 0.2, which indicates a robust and reportable effect size between small and medium (King and Minium 2008: 327–329).

observations can be made. First, the data reinforces the conclusion from the previous section that the genitive is widely used, while compounds and, to a lesser extent, adjective constructions remain serious competitors. Second, the Non-Head Function Hypothesis fares better for non-agentive than for agentive nouns. Third, the hypothesis has better predictions for type than for token frequency. For token frequency of agentive nouns, the genitive construction receives a lower percentage (41.2%) than compounds (51%), as shown in Table 6.

	Agentive no	ouns	Non-agentive nouns		
	Raw numbers	Percent	Raw numbers	Percent	
Adjective	8	8.0	29	20.4	
Genitive	44	44.0	81	57.0	
Preposition	4	4.0	3	2.2	
Compound	44	44.0	29	20.4	
Total	100	100.0	142	100.0	

Table 5: The distribution of constructions where the non-head is an internal argument: agentive vs. non-agentive nouns (type frequency)

	Agentive no	Non-agentive nouns		
	Raw numbers	Percent	Raw numbers	Percent
Adjective	9	4.6	45	20.7
Genitive	80	41.2	93	42.9
Preposition	6	3.2	4	1.8
Compound	99	51.0	75	34.6
Total	194	100.0	194	100.0

Table 6: The distribution of constructions where the non-head is an internal argument: agentive vs. non-agentive nouns (token frequency)

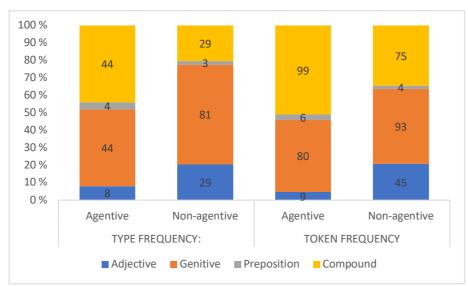


Figure 2: The distribution of constructions where the non-head is an internal argument: agentive vs. non-agentive nouns

The available data suggest that Mezhevich's hypothesis represents a valuable generalization, especially for type frequency of non-agentive nouns. At the same time, the hypothesis is far from being a categorical rule, so it is necessary to explore the competition between the genitive and its two most important rivals, viz. compounds and the adjective construction. Compounds will occupy us in section 4, before we turn to adjectives in section 5. The third competitor, the preposition construction, will not be discussed in the

following, since it is quite marginal in my dataset, as shown in this and the previous sections.

4. Genitive vs. compound: four constraints

On the basis of my data, I propose four generalizations about the rivalry between genitive constructions and compounds. These generalizations place constraints on compounding as a word-formation pattern in Russian. While there are virtually no restrictions on compounding in Norwegian, compounding is more constrained in Russian, and this explains why the Russian genitive frequently corresponds to compounds in Norwegian.

The first generalization concerns productivity, i.e., the ability of a pattern to be extended to new lexical items:¹¹

(4) The productivity constraint
In Russian, compounding is less productive than in Norwegian.

Norwegian has no such constraint, insofar as compounds are formed freely, even for highly unconventional combinations of concepts (Eiesland and Lind 2017). The following examples from my database illustrate this:¹²

- (5) a. håndklefinner 'towel finder' otyskivatel' polotenec
 - b. lekepåfinner 'person who comes up with games to play' vydumščik igr
 - c. papirsluker 'paper swallower' požiratel' bumagi
 - d. tallerkenkaster 'thrower of plates' kidal'ščik bljud
 - e. tuppemaler 'chicken painter' risoval'ščik petuchov

If the productivity constraint in (4) is on the right track, we would expect such examples to correspond to genitive constructions, rather than compounds in Russian. This is indeed the case, as shown in (5).

This is not to say that compounding is unproductive as a word-formation pattern in Russian. As is well known, certain types of compounds are productive in the sense that they attract new members (Švedova (ed.) 1980, 247, 251 et passim).¹³ However, even if compounding displays some degree of productivity in Russian, "occasionalisms" (words created on the fly for the purposes of communication) are not formed as freely as in Norwegian compounds, as illustrated in (5). In section 2, we saw that Russian compounds scored higher for token frequency than for type frequency. This suggests that a relatively large proportion of the Russian compounds occur frequently in texts, which in turn testifies to their conventionality in the speech community.

Another constraint concerns the structure of the non-head:

 $^{^{11}}$ As pointed out by an anonymous reviewer, the term "constraint" is used in a wider sense in (4) than elsewhere in the article. As opposed to the constraints in (6), (8) and (10), the "productivity constraint" in (4) does not map a structure onto an (un)acceptability value. However, (4) is a constraint in the sense that it describes a limitation on the use of compounds in Russian.

¹² As a measure of unconventionality I used two criteria: (i) the relevant item should not be present in Landrø and Wangesteen (1986), a standard dictionary of Norwegian, and (ii) the item in question should not have more than one single attestation in the NoWaC corpus, a web corpus containing 700 million words (https://www.hf.uio.no/iln/tjenester/kunnskap/sprak/korpus/skriftsprakskorpus/nowac/).

¹³ A more detailed discussion of the complex concept of "productivity" in language is beyond the scope of

¹³ A more detailed discussion of the complex concept of "productivity" in language is beyond the scope of the present study. The interested reader is referred to Baayen (1993), Bybee (1995), as well as Barðdal 2008 and 2012.

(6) The simplex non-head constraint Russian compounds normally have simplex non-heads.

"Simplex" is here used to rule out compounds as non-heads. In Norwegian, compounds with complex non-heads can be formed freely, as shown by the following examples:

- (7) a. *barnekulturforsker* 'researcher of children's culture' *issledovatel' detskoj kul'tury*
 - b. husbukktemmer 'beetle tamer' ukrotitel' domomučitel'nic
 - c. jernbaneanlegg 'railroad construction' postrojka železnych dorog
 - d. nybilsalg 'sale of new cars' prodaža novych mašin
 - e. olje- og gassutvinning 'oil and gas extraction' dobyča nefti i gaza

As we would expect from the simplex non-head constraint in (6), the Norwegian examples in (7) do not correspond to compounds in Russian, and the genitive construction is used instead.

	Raw numl	Percent	
	Complex non-head	All compounds	complex non-head
Norwegian	159	2103	8
Russian	0	206	_

Table 7:Distribution of compounds(lemmas) with complex non-heads in Norwegian and Russian

Table 7 summarizes the distribution of compounds with complex non-heads in my database. While such examples constitute 8% of the Norwegian compounds (lemmas) in the dataset, I have no corresponding examples for Russian. It might therefore be tempting to omit the word "normally" in the constraint in (6) and treat the absence of complex non-heads as a categorical rule in Russian. However, Švedova (ed. 1980, 246) mentions a couple of exceptions: *korneklubnemojka* 'washer for roots and tubers' and *plodoovoščesušilka* 'fruit and vegetable drier'. Examples of this kind appear to be marginal in Russian, and they also seem to be restricted to cases where the non-head consists of two coordinated (paratactically related) elements. Thus, in *plodoovoščesušilka* the two elements *plod* 'fruit' and *ovošč* 'vegetable' are coordinated, as suggested by the use of *and* in the gloss 'fruit and vegetable drier'. Likewise, in *korneklubnemojka koren*' 'root' and *kluben*' 'tuber' are coordinated, insofar as we are dealing with a machine the washes roots *and* tubers. The suggestion of the su

Notice that example (7e) is of the paratactic type where *olje* 'oil' and *gass* 'gas' are coordinated. Even if compounds are in principle possible in cases of this type in Russian, as shown by examples such as *korneklubnemojka* and *plodoovoščesušilka*, in (7e) the genitive construction was used in Russian: *dobyča nefti i gaza*.

A potential constraint on Russian compounds concerns the referential properties of the non-head:

(8) The individual reference constraint

-

¹⁴ An anonymous reviewer provides two additional examples of paratactic compounds with a complex non-head (*neftegazoxrana* 'oil and gas protection' and *neftegazoprovod* 'oil and gas pipeline'), which show that such compounds are possible in Russian, albeit marginal. The reviewer suggests that such compounds may be less marginal for adjectives (e.g., *neftegazodobyvajuščaja promyšlennost* 'oil-gas-extracting industry'. Testing this hypothesis is beyond the scope of the present study.

The non-head of Russian compounds normally does not have individual reference.

In Norwegian compounds, the non-head may refer to unique individuals, as illustrated by examples where the non-head is a proper name:

(9) a. *Auschwitz-kommandant* 'commandant of Auschwitz' – *komendant Osvencima* b. *Conan Doyle-leser* 'Conan Doyle reader' – *čitatel' Konana-Dojla*

While compounds with proper names as non-heads are attested in Russian, they do not seem very widespread. For instance, toponyms form relative adjectives where Norwegian may have compounds, as illustrated by *tul'skij samovar* 'samovar from Tula', which may be translated as the compound *Tula-samovar* into Norwegian. Likewise, where Norwegian forms compounds with person names, Russian uses relative adjectives or a genitive construction, as in *Mendeleevgaten* 'Mendeleev street', which corresponds to *Mendeleevskaja ulica* or *ulica Mendeleeva* in Russian.

I hasten to add that I do not claim that proper names are *impossible* as non-heads of Russian compounds. Examples like *Moskva-reka* 'Moscow river' and *šekspiroved* 'Shakespeare specialist' show that such compounds exist. With regard to the latter example, it is worth mentioning that the non-head is used metonymically about a field of research; a Shakespeare specialist is not first and foremost an expert on Shakespeare's person, but rather a scholar who studies Shakespeare's literary work. While compounds with proper names as non-heads are possible in Russian, it is interesting to see that in the two examples in (9) a genitive construction is used in Russian. This even applies to the metonymic example in (9b), where Conan Doyle refers to the books by Conan Doyle, not the person. Unfortunately, my database does not contain enough relevant examples for a rigorous test of the "individual reference constraint" in (8), so this must be left for future research.

A final constraint refers to the prosodic properties of Russian compounds:

(10) The prosody constraint
Russian compounds typically have monosyllabic non-heads.

Consider the following examples where a hyphen separates the non-head, head, as well as the "linking morph" /o/ between non-head and head:

- (11) a. *ljud-o-ed* 'cannibal'
 - b. les-o-rub 'lumberjack'
 - c. zub-o-čistka 'toothpick'
 - d. zeml-e-delie 'agriculture'
 - e. blag-o-detel' 'benefactor'
 - f. zdrav-o-chranenie 'health care'

These examples comply with the prosodic constraint in (10), insofar as the non-heads are monosyllabic.

Although the examples in (11) are suggestive, it is not difficult to find counterexamples, such as *literatur-o-vedenie* 'literature studies' with a long non-head. However, examination of my database as a whole yields a clear picture. Table 8 shows that 150 out of 206 compounds (agentive and non-agentive) have monosyllabic non-head

(i.e., 73%), and that only 14 compounds have non-heads with three or more syllables (i.e., 7%). 15

It is instructive to compare the numbers for compounds in Table 8 with the corresponding numbers for genitive constructions. While for compounds the majority of non-heads are monosyllabic, a majority of genitive constructions have non-heads with two or more syllables (127 monosyllabic vs. 199 di- and polysyllabic non-heads). The difference between compounds and genitive constructions is statistically significant with a robust and moderate effect size. ¹⁶

There is evidence for the prosodic constraint in (10) from beyond my database. On the list of frequently occurring non-heads provided in the Russian Academy Grammar (Švedova ed. 1980, 761-762) 98 out of 151 items are monosyllabic.¹⁷ Circumstantial evidence also comes from the fact that the non-heads often have short metathesis forms of Church Slavic origin (e.g., *zdrav*- 'health'), rather than the corresponding longer East Slavic pleophony forms like *zdorov*- 'health' (Švedova (ed.) 1980, 243).

	Number of syllables in non-head								
	0	1	2	3	4	5	6	7	8
Agentive compounds	0	60	12	5	4	0	0	0	0
Non-agentive compounds	1	90	29	5	0	0	0	0	0
Genitive constructions	0	127	104	42	32	14	3	0	4

Table 8: Length of non-head measured in number of syllables. For compounds, the linking /o/ is not counted, since it is arguably not part of the non-head. For genitive constructions, the inflectional ending is not counted. The numbers are based on type frequency.

The upshot of this section is simple. Russian has a number of constraints on compounding, and it is therefore natural that Russian often prefers other constructions where Norwegian uses compounds. Three conclusions can be drawn about the constraints discussed above. First, a single compound often complies with several constraints, and it is thus difficult to assess the relative importance of the constraints. For instance, *čàepítie* 'tea drinking' not only meets the prosodic constraint, but also represents has a simplex non-head that does not refer to a unique individual. Second, the constraints describe tendencies, not categorical rules. Finally, the empirical evidence for the constraints I have proposed is relatively limited, and they may therefore be regarded as hypotheses to be tested more thoroughly in future research.

5. Genitive vs. adjective: "Typification" and construal

The rivalry between the genitive and adjective constructions was discussed in detail in Nesset 2018a, and the conclusions from that analysis also apply to compounds with deverbal heads and their Russian equivalents. In particular, the adjective construction involves what we may refer to as "typification" and highlights the role of different construals of a situation, i.e. the ability of language users to view a situation in different ways (Langacker 2008, 4).

 $^{^{-15}}$ Notice that a non-head may contain no syllables, as in $\it zl\mbox{-}o\mbox{-}radstvo$ 'malevolence', where I analyze the

vowel in the first syllable as a "linking vowel" that is not part of the non-head. Since there is only one example with a non-syllabic non-head in my dataset, I do not discuss such examples in detail. 16 Pearson's Chi-squared test with Yates' continuity correction (X-squared = 60.838, df = 1) returned the

¹⁶ Pearson's Chi-squared test with Yates' continuity correction (X-squared = 60.838, df = 1) returned the p-value 6.197e-15. Cramer's V-value ≈ 0.34 (a moderate effect size, King and Minium 2008: 327-329).

¹⁷ The Academy Grammar includes the linking morph /o/ in the representation of the non-head, but I have not counted this syllable in order to arrive at numbers that are comparable to those in Table 8.

Examples like *okno kontory* and *kontorskoe okno*, which may both be glossed as 'office window', illustrate the subtle difference between the genitive and adjective constructions. While *okno kontory* may refer to a window in a particular office, *kontorskoe okno* denotes a *type*, namely the kind of window you may come across in an office (Nesset 2018a, see also Rakhilina 2008).

To what extent does this difference apply to constructions with deverbal heads? In the relevant constructions, we have an action that is directed towards an argument. This is commonly expressed in a genitive construction, such as *proizvodstvo bumagi* 'production of paper':¹⁸

(12) V 1720 godu Pëtr I izdal ukaz o stroitel'stve fabrik po proizvodstvu bumagi po vsej strane. (*Nauka i žizn'* 2009)

'In 1720, Peter I issued a decree about building factories for the production of paper all over the country.'

Here, the focus is on what the relevant factories produced, i.e. paper as opposed to other commodities. However, a slightly different construal is possible, whereby paper production is a type of production:

(13) Takova tenevaja storona burnogo razvitija chimii (v tom čisle neftepererabotki, bumažnogo proizvodstva, metallurgii i dr.), kotoruju my nabljudaem v poslednie gody. (*Chimija i žizn'* 1966)

'Such is the shadowy side of the booming development of chemistry (including oil refining, paper production, metallurgy etc.), that we have observed in recent years.'

In this example, the author provides a list of different types of chemical industry, and since the focus is on types, the adjective construction is chosen. I suggest referring to this construal as "typification", since it takes an action with an argument and views it as a type of action.

I use the term "construal" in order to emphasize how subtle the difference between the genitive and adjective constructions is. "Construal" as a term is used about different ways of viewing the same situation (Langacker 2008, 4). The choice between the genitive and adjective constructions brings out slightly different aspects of the situation, as illustrated in (12) and (13), but the nuances are subtle, and in many cases both constructions are possible.

The Russian constructions corresponding to Norwegian *avislesing* 'reading of newspapers' illustrate the subtle semantic difference between the genitive and adjective constructions, but bring in two other factors, viz. language change and idiolectal preferences. Since we are dealing with an action and an internal argument, the genitive construction is the first option that comes to mind: *čtenie gazet* 'reading of newspapers'. In the Russian National Corpus, *čtenie gazet* is found in 160 examples in 137 different documents. At the same time, the adjective construction is also attested, albeit only in nine examples.¹⁹ Here is one of them:

(14) Gazetnoe čtenie polnosť ju ubedilo menja v sobstvennoj nikčemnosti. (Terechov 1997-2008)

.

¹⁸ This and the following examples in this section are from the Russian National Corpus, available at www.ruscorpora.ru.

¹⁹ Corpus searches were carried out in January, 2020.

'Newspaper reading completely convinced me of my own worthlessness.'

This example is arguably about a type of reading and the impact this reading has had on the author. In other words, this is a good example of "typification", although a genitive construction would be possible in the same context. In addition to the fact that the adjective construction may be replaced by the genitive, which testifies to the very subtle semantic difference between the two constructions, two things are striking. First, seven out of nine examples with the adjective construction come from one text - the novel "Peterburg" by the symbolist Andrej Belyj. This fact suggests that the choice between the two constructions, i.e., between two construals of the situation, to some extent may depend on the personal preferences of the language users – i.e., their idiolects. A second striking point is that eight out of nine examples are more than one hundred years old. Although the examination of one single case does not enable us to draw general conclusions, it illustrates that the preferences for one or the other construal may change over time. As shown in the previous sections, in present day language the adjective construction enjoys a relatively modest role as the Russian equivalent of Norwegian deverbal compounds, and it is tempting to suggest that this may be the result of diachronic change whereby the adjective construction has become less frequent over time. However, my dataset does not make it possible to test this hypothesis empirically, and the hypothesis therefore must be left for future research.

6. A broader perspective: the Extended Non-Head Function Hypothesis

So far, we have been concerned with constructions where the non-head corresponds to the internal argument (object) of the verb from which the head is derived. We now turn to other non-heads. I propose that the choice of construction is sensitive to the relationship between verb and non-head. The closer the relationship between verb and non-head, the more likely is the use of the genitive. Conversely, the more distant the relationship between verb and no-head, the more likely is the use of the adjective construction.

For the purposes of this study, I distinguish between four broad classes of non-heads. "Object" covers cases of the type we have been concerned with in the previous sections, where the head is derived from a transitive verb and the non-head represents the internal argument (object) of that verb. I use "transitive subject" about examples of the following type where the head is derived from a transitive verb and the non-head corresponds to the subject of that verb:

- (15) a. jentekyss devičij poceluj 'girl's kiss'
 - b. embedsmannsuttrykk vyraženie činovnika 'bureaucrat's expression'
 - c. enehersker samoderžec 'autocrat'
 - d. ansiktsuttrykk vyraženie na lice 'facial expression'

Notice that my classification takes the Norwegian member of each pair as its starting point, since my research question is based on Norwegian: what are the Russian equivalents of Norwegian compounds with deverbal heads?

In some cases, the Norwegian and Russian members of one pair are not completely parallel. A case in point is *ansiktsuttrykk* 'facial expression' in (15d) above. The Norwegian non-head *ansikt* 'face' corresponds to the subject of a transitive verb (*ansiktet uttrykker noe* '(someone's) face expresses something'), while the Russian construction *vyraženie na lice* 'expression on (someone's) face' rather suggests an adverbial interpretation of the

non-head as the place where the expression of some emotion occurs. It is worth pointing out that the most natural way to translate *ansiktsuttrykk* into Russian is by means of the genitive construction *vyraženie lica*. In my database, the genitive is attested in 46 example (tokens), while the preposition construction occurs in only four examples. The preposition construction appears to require an overt possessor (e.g. *Kitty* in (16)) or a description of what is expressed on the face (e.g. *cholodnoe otčajanie* 'cold despair' in (17)):

- (16) Levin dumal o tom, čto označala èta peremena vyraženija na lice Kiti [...]. 'Levin was wondering what the change in the expression on Kitty's face meant.' (Tolstoj, RuN Corpus)
- (17) Ni slova bol'še, povtorial ona, i s strannym dlja nego vyraženiem cholodnogo otčajanija na lice ona rasstalas' s nim.
 - '- Not one more word. she repeated, and with what he thought was a strange expression of cold despair on her face, she parted with him.' (Tolstoj, RuN Corpus)

"Intransitive subject" is used in cases where the head corresponds to the subject of an intransitive verb, as in the following examples:

- (18) a. hundegjøing-sobačij laj 'dog's barking'
 - b. hestevrinsk- fyrkan'e lošadej 'horse whinny'
 - c. isgang ledochod 'drifting of ice'
 - d. *stormaktkappløp sorevnovanie meždu velikimi deržavami* 'race between superpowers'

Finally, I use "adverbial" as a cover term for cases where the non-head expresses the time, location, means, source, goal, or topic of the action:

- (19) a. aftenssamtale vežernjaja beseda 'evening conversation'
 - b. kjærlighetsdans tanec ljubvi 'love dance'
 - c. hushjelp domrabotnica 'housemaid'
 - d. soppsamtale razgovor o gribach 'conversation about mushrooms'

Table 9 and Figure 3 show the distribution of the various types of non-heads for the four rivaling constructions: adjective, genitive, preposition, and compound. The non-heads are mentioned according to how closely they are related to the relevant verb. Internal arguments are governed by the verb and are traditionally analyzed as parts of the verb phrase, while transitive subjects are outside the verb phrase and not governed by the verb. Intransitive subjects occupy an intermediate position between internal arguments (objects) and transitive subjects; as is well known from typological studies, in ergative languages intransitive subjects group with objects, whereas in nominative-accusative languages intransitive subjects behave like transitive subjects with regard to case marking. Adverbial non-heads are least closely related to the verb, since adverbials are traditionally not analyzed as arguments and are generally not obligatory parts of a sentence.

	Adjective	Genitive	Preposition	Compound
Internal argument (object)	37	125	7	75
Intransitive subject	79	131	4	16
Transitive subject	26	9	1	5

Adverbials 219 52 42 11

Table 9: The distribution of non-heads with various functions for four constructions (type frequency)

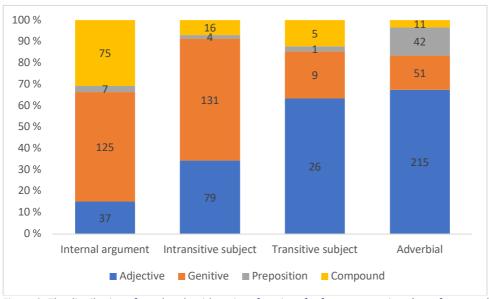


Figure 3: The distribution of non-heads with various functions for four constructions (type frequency)

Table 9 and Figure 3 show that the choice between the adjective and the genitive constructions is sensitive to the type of non-head.²⁰ The following generalization emerges:

(20) The Extended Non-Head Function Hypothesis

- a. The closer the relationship between verb and non-head, the more likely is the use of the genitive.
- b. The more distant the relationship between verb and no-head, the more likely is the use of the adjective construction.

The hypothesis in (20) is an extended version of Mezhevich's "Non-Head Function Hypothesis" discussed in section 2. While Mezhevich only considered non-heads that function as internal arguments (objects), the hypothesis in (20) has a broader scope, in that it covers non-heads with different syntactic functions.

The picture is less clear for compounds and constructions with prepositions, but two patterns are noteworthy:

- (21) a. Compounds are most commonly attested in cases where the non-head corresponds to an internal argument (transitive object).
 - b. Prepositions are most common when the non-head is an adverbial.

-

 $^{^{20}}$ A series of chi-squared tests indicates that the differences in Table 9 and Figure 3 are statistically significant and involve strong effect sizes. I first compared adjective and genitive, for which Pearson's Chi-squared test (X-squared = 169.71, df = 3) returned a p-value < 2.2e-16. Cramer's V-value is 0.5. Comparison of genitive and preposition (X-squared = 91.606, df = 2) gave a p-value < 2.2e-16. Cramer's V-value is 0.5. Finally, I compared prepositions and compounds. Here Pearson's Chi-squared test (X-squared = 73.586, df = 2) returned a p-value < 2.2e-16. Cramer's V-value is 0.7. For the comparisons of genitive, preposition and compound, the numbers for transitive subjects are too small to permit statistical analysis, and these numbers were not included in the relevant chi-squared tests. Cramer's V larger than 0.5 is regarded as a strong effect size (King and Minium 2008: 327–329 and Levshina 2015: 209).

Although these generalizations are suggestive, it should be kept in mind that they are based on a limited dataset. However, they may be regarded as hypotheses to be tested in future research.

7. Concluding remarks

The contribution of my analysis can be summarized as follows. First, it has been shown that four morphological types dominate Norwegian compounds with deverbal heads: infinitive conversion (båtreise 'boat trip'), stem conversion (fiolinspill 'violin playing'), and suffixations with -er (fysikklærer 'physics teacher') and -ing/-ning (bringebærsylting 'making jam from raspberries'). Second, four Russian constructions correspond to the relevant Norwegian compounds: the adjective, genitive, preposition, and compound constructions. Third, an empirical test of Mezhevich's (2002) Non-Head Function Hypothesis has confirmed that the genitive construction is used extensively, although compounds and, to a lesser degree, the adjective and preposition constructions are also attested. Fourth, it has been argued that the genitive construction is more frequently attested for non-agentive nouns than for nouns of the agentive type. Fifth, four generalizations that constrain the use of Russian compounds have been proposed. Sixth, it has been argued that the use of the adjective construction is related to the process of "typification" and, more generally, to construal, i.e. to our ability to portray a situation in different ways. Last, but not least, an Extended Non-Head Function Hypothesis has been proposed, according to which the choice of Russian construction depends on the syntactic relationship between the non-head and the head of the Norwegian compound.

These generalizations represent valuable tools for L2 learners and scholars interested in contrastive grammar and translation, but since the generalizations are based on a limited data material they should not be considered definite conclusions, but rather hypotheses to be confirmed or refuted in future research. Germanic-Slavic contrastive studies of word-formation represent a fertile field that deserves more attention from linguists in the future.

References

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.

Baayen, R. Harald. 1993. On frequency, transparency and productivity. *Yearbook of morphology* (1992), 181–208).

Barðdal, Jóhanna. 2008. *Productivity: Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamins.

Barðdal, Jóhanna. 2012. "Predicting the productivity of argument structure constructions." In Zhenya Antić, Charles B. Chang, Emily Cibelli, Jisup Hong, Michael J. Houser, Clare S. Sandy, Maziar Toosaravandani & Yao Yao (eds.), *Proceedings of the Thirty-second Annual Meeting of the Berkeley Linguistics Society (2006)*: 467–478.

Bybee, Joan B. 1995. Regular morphology and the lexicon. *Language and cognitive processes* 10(5), 425–455.

Bisetto, Antonietta and Chiara Melloni. 2008. "Parasynthetic Compounding". *Lingue e linguaggio* 7 (2), 233-260.

Booij, Gert. 2010. Construction morphology. Oxford: Oxford University Press. Eiesland, Eli Anne & Marianne Lind. 2017. "Hva er en 'grevlingby'? En undersøkelse av hvordan nylagde sammensetninger fortolkes av personer med og uten afasi". *Logpeden* 63(2), 6–14.

- Enger, Hans-Olav and Kristian Emil Kristoffersen. 2000. *Innføring i norsk grammatikk. Morfologi og syntaks.* Oslo: Cappelen Akademisk Forlag.
- Faarlund, Jan T., Svein Lie and Kjell I. Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Frawley, William. 1992. *Linguistic semantics*. Hillsdale NJ and London: Lawrence Erlbaum Associates.
- Goldberg, Adele E. 2013. Constructionist Approaches. In *The Oxford Handbook of Construction Grammar*, edited by Thomas Hoffmann and Graeme Trousdale. Oxford: Oxford University Press. DOI:10.1093/oxfordhb/9780195396683.013.0002
- Haspelmath, Martin. 2002. *Understanding Morphology*. London: Arnold.
- Haugen, Tor Arne. 2006. "Preposisjonsobjekt frå norrønt til moderne norsk. Ei undersøking av valens i leksikalsk nærskylde verb." *Maal og Minne* 2: 187–207.
- Jespersen, Otto. 1942. A Modern English Grammar on Historical Principles. Part VI: Morphology. Copenhagen: Munksgaard.
- Johannessen, Janne B. 2001. "Sammensatte ord." *Norsk Lingvistisk Tidsskrift* 19 (1): 59–91.
- Johannessen, Janne B. 2017. "Puzzling parasynthetic compounds in Norwegian." *In Order and Structure in Syntax II: Subjecthood and Argument Structure*, edited by Michelle Sheehan and Laura R. Bailey (eds.), 257–273. Berlin: Language Science Press.
- King, Bruce M. and Edward Minium. 2008. *Statistical reasoning in the behavioral sciences*. Hoboken, NJ: John Wiley & Sons.
- Landrø, Marit Ingebjørg and Boye Wangensteen. 1986. *Bokmålsordboka*. Bergen, Oslo, Stavanger, Tromsø: Universitetsforlaget.
- Langacker, Ronald W. 2008. *Cognitive Grammar*. Oxford: Oxford University Press.
- Levshina, Natalia. 2015. *How to do linguistics with R*. Amsterdam and Philadelphia: John Benjamins.
- 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, 95–114. Vancouver: Simon Fraser University.
- Nesset, Tore. 1998. Russian Conjugation Revisited. Oslo: Novus Press.
- Nesset, Tore. 2016. "Spøkelsesfiske, makrellfotball og traktoregg: norske sammensetninger og konseptuell integrasjon". *Maal og Minne* 2: 85–110.
- Nesset, Tore. 2018a. "How to Translate Compounds into Russian?" *Scando-Slavica* 64.2: 242–263.
- Nesset, Tore. 2018b. "When a single word is enough: Norwegian compounds and their Russian counterparts." *Slovo* 59: 61–72.
- Rakhilina, Ekaterina V. 2008. "Semantika russkich imennych konstrukcij s genitivom: 'ustojčivost'." In Festschrift in Honour of Professor Arto Mustajoki, edited by Jouko Lindstedt. Helsinki: University of Helsinki, 338–49.
- Švedova, Natalija Ju. (ed.) 1980. Russkaja grammatika 1. Moscow: Nauka.