Chapter 8

God, the Devil, and Christ: A corpus study of Russian syntactic idioms and their English and Finnish translation correspondences

Mikhail Mikhailov
Tampere University

In any language, phrases like *holy Christ/God/cow* can be found. They are sometimes called syntactic idioms, because they are identified in the first place by their syntactic structure and only in the second place by their variable lexical elements. Such expressions are difficult to present in dictionaries, and for this reason they are problematic for language learners. In this paper, the structure, meanings, and use of the Russian construction *Nominative + s 'with' + Instrumental* (*bog s toboj* 'god with you', *čërt s nim* 'the devil with him', etc.) as well as its equivalents in other languages are studied. The construction has four main meanings: 'blessing', 'disagreement', 'permission', and 'acceptance with disapproval'. These meanings are determined by context, and in many cases the expressions are ambiguous. A large web corpus of Russian, ruTenTen11, was used for studying the composition of the construction, its obligatory and optional components, and its functioning in speech. To study the English and Finnish correspondences of the construction, data from parallel corpora of literary texts were used. Parallel concordances demonstrated the absence of direct equivalents for the construction in both English and Finnish. The data also show that this construction is often misunderstood by translators. This phenomenon is obviously connected to insufficient information supplied by monolingual and bilingual dictionaries. The use of the CxG methodology helps to make syntactic idioms more visible and provide better descriptions for them.

Mikhail Mikhailov. 2021. God, the Devil, and Christ: A corpus study of Russian syntactic idioms and their English and Finnish translation correspondences. In Aleksandar Trklja & Łukasz Grabowski (eds.), *Formulaic language: Theories and methods*, 191–223. Berlin: Language Science Press. DOI: 10.5281/zenodo.4727675

1 Introduction

Only the core part of a language consists of free sequences of elements that are combined according to the language's basic rules. The remaining – quite substantial – part consists of so-called "exceptions", for which no clear-cut rules can be suggested. While some rules can be worked out, they are so complicated that it is extremely difficult to use them. This is one of the reasons why learning languages is difficult.

Among those non-free sequences of components are expressions that cannot be interpreted from their constituent parts because of a certain added meaning. Such units are called idiomatic expressions. Many of them are registered at the end of dictionary entries after the basic meanings of the main lexical element are explained. For example, the expression *to kick the bucket* would be probably found at the end of the entry on the noun *bucket*, or, less likely, at the end of the entry on the verb *to kick*.

Some idiomatic expressions pretend to be free expressions. Is the phrase *How do you do?* in English idiomatic? Evidently, it is, although it does look like a normal English phrase. A person who says this phrase is not really interested in the health or personal problems of the addressee, nor is the phrase a question. The phrase should be uttered exactly in this form when one is introduced to a person and has the same meaning and function as *Nice to meet you.* Any changes to the phrase (*How are you? How are you doing? How did you do*? etc.) may lead to a communicative failure. Hence, there are good reasons to treat the speech formula *How do you do?* as an idiom.

Other borderline cases are combinations of a noun or a verb with a preposition or an adverb, and a good example of this would be English phrasal verbs like *put on*, *show off*, *cut in*, *run out*, etc. Some of these phrases can be registered in dictionaries as idioms, while some are believed to be free expressions. In any case, it is clear that all of them are difficult for non-native speakers and they often cause misinterpretation.

A good example of such a mistake caused by the misunderstanding of an idiomatic expression is a passage from the adaptation of John Wilson's tragedy *The City of the Plague* (1816) by the Russian poet Alexander Pushkin (*Pir vo vremja čumy* [The feast in the time of the plague], 1832).

Here is the quotation from the original English text:

(1) Priest. O impious table! Spread by impious hands!
Mocking with feast and song and revelry
The silent air of death that hangs above it,

<...>
I could have thought that hell's exulting fiends
With shouts of devilish laughter dragged away
Some harden'd atheist's soul unto perdition.
Several voices. How well he talks of hell! Go on, old boy!

The Russian translation of the last line looks like this:

(2) Несколько голосов. Он мастерски об
several.PRON voice.NOUN.GEN.PL he.PRON.3.NOM skillfully.ADV about.PREP
аде говорит! Ступай, старик! ступай
hell.NOUN.LOC talk.PRES.3SG. go.IMP 'old_man'.NOUN.NOM go.IMP
своей дорогой!
own.PRON.INS way.NOUN.INS.SG
'Several voices. He skilfully talks about hell. Go away, old man!'

In the original text of the play, the audience mockingly encourages the priest to continue his speech. Pushkin evidently understood *go on* as 'continue on your way' and the reaction of the priest's audience in the Russian translation is the opposite.[1] Pushkin read in the original and translated many English authors – Shakespeare, Byron, Milton – and his translations show a very good understanding of the source text. The error in the translation of Wilson is most likely caused by a lack of knowledge of the spoken language and the possible scarceness of information in the dictionaries of that time[2].

The modern world is more open, there are more language manuals and dictionaries, the methods of learning languages have improved, and people speak foreign languages much better than in Pushkin's time. Besides, online dictionaries, text corpora, and encyclopaedias make it possible to make very complicated queries. Does this mean that idiomatic expressions do not present problems for learners and translators nowadays?

In any language, there can be found idiomatic expressions that have idiomaticity programmed into their syntactic structure; they are a kind of frame into which variable lexical components can be inserted. For example, there is an English tautological expression *N-Pl will be N-Pl*, which is most often realized as *boys will be*

[1]The matter was discussed on Russian social media in 2019 with many Russian scholars participating, Yakov Testelets and Dmitri Sitchinava among them.

[2]The opinions on Pushkin's command of English are very contradictory; some researchers believe he spoke the language fluently, while others think he could barely read, see Zaharov 2008 for more information.

boys (enTenTen15: 878 occurrences, 0.05 ipm[3]), but one can coin other phrases based on that pattern: *men will be men* (enTenTen15: 29 occurrences), *women will be women* (enTenTen15: 3 occurrences), *students will be students* (enTenTen15: 7 occurrences), etc. Such expressions are sometimes called syntactic idioms or phraseoschemes (see, e.g. Baranov & Dobrovol′skij 2008: 16), and usually they are not registered in dictionaries of idioms, partly due to technical issues (e.g. where to place the entry) and partly because of their very complicated semantics. However, such idioms often become topics for linguistic publications, for example Wierzbicka 1987 on *boys will be boys*.

In this paper, I will study the Russian syntactic idiom *N-Nom s 'with'-N/Pron-Instr* (hereafter, I will use a shorter version *N-s-N*, although it is less precise), which can be realized in expressions like *bog s toboj* 'god with you', *čërt s rabotoj* 'devil with work', etc. I will study the idiom with the help of corpus data and describe its structure and meaning using the formalisms of Construction Grammar (CxG, see Fried & Östman 2004). I will check parallel corpora for possible correspondences of this idiom in other languages and ascertain whether translators understand it correctly.

My main sources of data will be ruTenTen11, Russian-English and English-Russian parallel corpora at the Russian National Corpus (RNC) and the Russian-Finnish and Finnish-Russian parallel corpora ParRus and ParFin compiled at Tampere University (Mikhailov & Härme 2015, Härme & Mikhailov 2016).

2 The construction *N-s-N*: An overview

Let us start with usage examples from ruTenTen11, a Russian language corpus hosted at SketchEngine (sketchengine.eu).

(3) a. Пожалел псарь хорошенькую
pity.PAST.MSG dog-trainer.NOUN.NOM.SG pretty.ADJ.ACC
девочку и сказал: «Ну и
girl.NOUN.ACC.SG and.CONJ say.PAST.PTCP and.PTCP go.IMP
ступай. Бог с тобой,
god.NOUN.NOM.SG with.PREP you.PRON.2 poor.ADJ.NOMFSG
бедная девочка!»
girl.NOUN.NOM.SG
'The dog-trainer took pity on the pretty girl and he said: Go. God be with you, poor girl!'

[3]ipm = instances per million words.

b. Путин работает. И Бог
Putin.NOUNPROPER.NOM work.PRES.3SG and.PTCP god.NOUN.NOM
с ним!
with.PREP he.PRON.INS
'Putin is working, and let him be!'

c. -Наверное, вы правы. Бог
maybe.ADV you.PRON.2.NOM.PL right.ADVerbial god.NOUN.NOM
с ней, с политикой.
with.PREP she.PRON.3F.INS.SG with.PREP politics.NOUN.F.INS.SG
'Maybe you are right and we should not discuss politics.'

d. Однако все дела закончить не
however.ADV all.PRON thing.NOUN.PL.ACC finish.INF not.PTCP
удастся. И бог с
succeed.REFL.FUT.3SG and.PTCP god.NOUN.NOM with.PREP
ними!
it.PRON.3.PL.INS
'However, we won't be able to finish all the jobs, and we need not do them!'

Example (3a) is different from the remaining three. In spite of the word *bog* 'god' being the headword of the expression, *bog s X* 'god with X', only this example really relates to piety. The meaning can be described as wishing somebody success, prosperity, or other achievements – later on, I group these meanings together as 'blessing'. In the remaining three examples, the expression has the meaning of acceptance of an inevitable state of affairs the speaker (probably) does not approve of – later on, I refer to this meaning as 'acceptance'. The findings of this chapter will demonstrate that the meaning of 'acceptance' is much more frequent in present-day Russian than the meaning of 'blessing'.

The expression *bog s X* 'god with X' is not unique, and many expressions can be found of a similar structure with different nouns in the initial position.

(4) a. Честное слово,
honest.ADJ.NEUT.NOM.SG word.NOUN.NEUT.NOM.SG
чёрт с ней, с
devil.NOUNM.NOM.SG with.PREP she.PRONF.INS.SG with.PREP
такой работой!
such.PRON.INS job.NOUN.F.INS.SG
'I am serious, to hell with such a job!'

b. Правда, я застрял на Ибице,
however.ADV I.PRON stick.PAST.MSG on.PREP Ibiza.NOUNPROPER.LOC
ну и хрен с ним
so.PTCP and.PTCP horseradish.NOUN.NOM with.PREP it.PRON.INS.SG
<...>

'Although I am stuck in Ibiza, it is no big deal…'

c. Ладно. Пёс с ними, с
allright.ADV dog.NOUNM.NOM.SG with.PREP it.PRON.PL.INS with.PREP
высокими идеями.
high.ADJ.PL.INS idea.NOUN.F.PL.INS
'OK, I do not care about these high ideas.'

The construction is flexible, and it has two variable components. The first component should be a noun in the Nominative case, the second is the preposition *s* 'with', and the third can be a noun or a pronoun in the Instrumental case. Additionally, the construction has optional elements. It can be introduced with particles or particle combinations *da*, *i*, *da i*, *nu*, *nu i*, and *nu da i*. If the third component is a pronoun, it can be explicitated (i.e. be made more explicit) with a propositional group headed with preposition *s* 'with' and a noun, sometimes with an attribute, like in examples (4a) and (4c).

The expressions are very typical in spoken Russian and rather misleading for non-native speakers, as many colloquialisms are. The meaning often depends on the context and the intonation. The construction is used in the written language as well, and many examples can be found in fiction, mass media, and letter exchange.

The two most frequent of them, *bog s X* and *čërt s X*, are occasionally registered in dictionaries of the Russian language. The Ozhegov-Shvedova Dictionary (Ožegov & Švedova 1992) has both (and even the *pës s X* 'dog with X'), while the Concise Academic Dictionary of Russian (Evgen′eva 1984) has only *čërt s X*. The Efremova's Dictionary of Russian language does not register any of these idioms (nor does it seem to register syntactic idioms at all).

Russian phraseological dictionaries, even the latest and the most complete Academic Dictionary of Russian phraseology (Baranov & Dobrovol′skij 2015), register only *bog s X* and ignore *čërt s X*.

The bilingual Russian-English Phraseological Dictionary by Sophia Lubensky (1995) is the most accurate with this group of idioms: it registers both *bog s X* and *čërt s X* and mentions that the first element can be replaced by other words, *bog*

'god' > *gospod'* 'Lord', *Hristos* 'Christ', *čërt* 'devil' > *shut* 'clown', *pës* 'dog', *prah* 'ashes', and *hren* 'cock, vulg.'.

In *Constructicon for Russian* (Borin et al. 2012), a repository of Russian constructions[4], only the construction *čërt s X* is registered with the following definition: "This construction expresses consent with [a participant or situation] Theme imposed on the speaker. The speaker negatively evaluates the participant or situation, and contrary to their will, accepts these conditions". According to Constructicon, the X component can only be a pronoun (which is not true, cf. e.g. a quite acceptable phrase *čërt s karantinom* 'to devil with the quarantine'). The article does not mention the possibilities of changing *čërt* 'devil' to other nouns, and *bog s X* was not registered, at least at the time this paper was written.

However, in spite of the fact that an average Russian native speaker is very likely to connect the expressions *bog s X*, *čërt s X*, *hren s X*, etc., these relations are not shown in Russian monolingual dictionaries and only partly registered in Lubensky's Russian-English Dictionary.

In linguistic literature, the construction *N-s-N* has not yet been a subject of special study, although Dobrovol'skij et al. (2019: 12) mention in their paper that this construction is productive and deserves a separate study.

Thus, neither dictionaries and lexical databases nor current linguistic research provide a thorough analysis of this syntactic idiom and give a concise picture of its structure, meanings, and functioning in speech. In this publication, I will therefore try to fill this gap.

3 Obtaining the corpus data

As it has already been mentioned, the construction *N-s-N* belongs to language for general purposes, and it is not likely to be found in specialist discourse. It can be used in posts on social media and other informal messages, in mass media texts, and in fiction. Thus, to collect data on this construction, we need a corpus of language for general purposes, and this corpus must be very large, because frequencies of multiword expressions are much lower than frequencies of single words. We also need a concordancing tool with the capacity to make complicated queries to look up syntactic constructions. Currently, the most suitable resource is SketchEngine, which uses its own ruTenTen11, currently the largest corpus of the Russian language available (18.2 G running words). The service permits the download of search results in convenient formats, which was very important for

[4]https://spraakbanken.gu.se/karp/#?mode=konstruktikon-rus

the current study. Therefore, the choice to use SketchEngine and ruTenTen11 was easy.

The service supports the CQL query language, and this makes it possible to run very complicated search queries. However, in this particular case, it was problematic to obtain the data in one step. The problem is that the sequence *Nominative + s + Instrumental* is very common in the Russian language, and searching for it directly would produce an immense amount of noise like *kofe s molokom* 'coffee with milk', *obed s drugom* 'lunch with friend', *kniga s kartinkami* 'book with pictures', etc. Of course, one can always search for particular words in particular forms, but it was necessary to find out first what lexemes serve as the first component of the construction, the noun in the Nominative case. For this reason, I decided to start with a search on the sequences *Noun.Nominative + s + Personal_pronoun.Instrumental* forming a sentence, i.e. delimited with end-of-sentence punctuation marks. Of course, such a search would not yield all the relevant data, and there might still be some noise in the results (e.g. *obed so mnoj*[5] 'lunch with me' or the above-mentioned *kofe s molokom* 'coffee with milk' as separate sentences). Still, the task of this particular query was not to find all the data with 100% precision, but to get a list of candidates for the headword of our construction.

The first query therefore had the following form:

(5) Query 1.

```
[word="\." | word="!" | word="\?"][word="ну"]?[word="и"]?[tag="
    N..sn.*"]
[word="со?"][tag="P....i.*"][word="\." | word="!" | word="\?" |
     word=";"]
```

I will give here only a very brief explanation of the query: for more details, see the manual of the CQL language on the website[6] of SketchEngine. Each token of the search phrase is put in square brackets. A full stop means any character. A question mark after any element (token, character) means that it is optional.[7] An asterisk means that the preceding element can be repeated from zero to an

[5]The Russian preposition *s* 'with' has a phonetic variant *so* that is used if the next word starts with a combination of consonants, e.g. *so mnoj*, *so stakanom*, *so zvonom*, etc. This variant is included in the search queries of this study, for example in Query 1 below.

[6]https://www.sketchengine.eu/documentation/cql-basics/

[7]To include in a query "real" full stops, question marks, asterisks, and other characters with special meaning, they should be preceded with a backslash ("\.", "\?", "*", etc.).

indefinite number of times. The tag "word" is used for querying by tokens (running words and punctuation marks), "lemma" by dictionary form, and "tag" by grammatical features. Different tags of the same token can be combined by the logical operators: | ("or"), & ("and") and ! ("not"). So, Query 1 can be read as follows: "a full stop, an exclamation mark, or a question mark – optional particle *nu* – optional particle *i* – a noun in the Nominative singular[8] – a preposition *s* 'with' or its phonetic variant *so* – a pronoun in the Instrumental case – a full stop, an exclamation mark, a question mark, or a semicolon".

This query running on a Gigacorpus would have produced a vast concordance that I did not need, so I ordered 10,000 random examples. After loading the concordance into R, separating the first noun into a separate column and creating a frequency list of these nouns, I obtained a table with 1,280 lines. To be on the safe side, I decided to check the whole frequency list, even the single occurrences. As it has been already mentioned, the combination N.NOM+S+PRON.INS is very common in Russian, and even after restricting the sequence to a separate sentence, many nouns on the list had nothing to do with the construction in question. After removing the noise, the list was dramatically reduced to what can be seen in Table 8.1.

Having the list of headword candidates, it was easy to run the queries to collect all usage examples for the construction *N-s-N* with the words from the list. All the queries run on the second stage of the search were formed like Query 2 below. This particular query looks up the constructions with *bog* 'god' as the headword. The construction does not have to be a separate sentence (commas added to initial and final tokens), and the third element can be a noun or a pronoun.

(6) Query 2.

```
[word="\." | word="!" | word="\?" | word=","][lemma="ну"]?[
    lemma="и"]?[lemma="бог"][word="с" | word="со"][tag="P....i
    .*" | tag="N...i.*"][word="\." | word="!" | word="\?" |
    word=";" | word=","]
```

The queries for all headwords from the list in Table 8.1 were done by replacing the headword in Query 1 with relevant words: `[lemma="бог"]` ⟶ `[lemma="хрен"]`, `[lemma="леший"]`, etc. In some cases for the words that have different spellings or could have been lemmatized incorrectly, matching with regular expressions was

[8]The codes for grammatical forms are explained in the tagsets for each language; the Russian tagset can be found at https://www.sketchengine.eu/russian-tagset/.

Table 8.1: The frequency list of the headwords.

Head			Freq
бог	bog	‘god’	2,622
черт	čërt	‘devil’	1,802
хрен	hren	‘horseradish’	907
господь	gospoď	‘Lord’	520
фиг	fig	‘fig’	391
шут	šut	‘fool’	105
христос	hristos	‘Christ’	100
хер	her	‘prick’	92
пес	pës	‘dog’	85
аллах	allah	‘Allah’	30
леший	lešij	‘forest imp’	16
хуй	huj	‘prick’	15
дьявол	ď̂avol	‘devil, satan’	13
бес	bes	‘devil’	8
шайтан	šajtan	‘devil for muslims’	5
иисус	iisus	‘Jesus’	4
демон	demon	‘demon’	3
сатана	satana	‘satan’	3
будда	budda	‘buddha’	2
холера	holera	‘cholera’	2
госдеп	gosdep	‘Department of State’	1
зевс	zevs	‘zeus’	1
перун	perun	‘Perun, Slavic god of thunder’	1
фюрер	fjurer	‘führer, Hitler’	1
член	člen	‘organ’	1

used or a *lemma* tag was replaced with a *word* tag, e. g., `[lemma="ч[е|о|ё]рт"]`, `[lemma="х.й"]`, `[word="[г|Г]осподь"]`.

The second search produced a more exact picture, because this time all examples and not a random sample were collected, and all the variants of the construction were looked up.

To check the precision of the search, a random sample of 1,000 examples was generated from the concordance and manually checked. Only 9 examples were wrong and the precision was therefore $991/1000 * 100 = 99.1\%$.

Evaluating the recall is more difficult due to the size of the corpus. The search was more or less accurate concerning high-frequency nouns detected with the help of Query 1. However, there might have been a number of low-frequency words used in the construction, and they may not have been detected with the query. Let us assume that there were around 500 examples with the names of people, gods, and mythological creatures that had a low frequency and passed unnoticed. Besides, there are always misspelled words and typos. With an error rate of 5%, about 1,600 headwords or other important components of the construction could have contained a typo, and these contexts would not have been found with the query. Another issue is the parsing accuracy. According to Nivre & Fang (2017), the accuracy of Russian Universal Dependency parsers is currently on the level of 79.79%. An accuracy of 79.79% for 30,000 examples means about 6,500 examples might have been incorrectly annotated and not found by the query. Thus, the recall of the search would be

$$\left(1 - \frac{500 + 1600 + 6500}{30000}\right) * 100 = 71.3\%.$$

The estimation is rough, but it is clear that one cannot expect a high recall rate in a very large and noisy corpus.

The results of the second search are presented in Table 8.2. The words after *šajtan* had very low frequencies and were removed from the table. The absolute (F) and relative (ipm) frequencies are given for each headword, along with the log-likelihood index (LL) (Dunning 1993; Xiao 2015: 111; Levshina 2015: 223–239). The values of LL are significant for all headwords at the $p > 0.0001$ level.

Grammatical constructions are certain lexemes that occur in speech in certain forms and in a certain order, and therefore collocation searches can provide additional information on the composition and use of constructions. Collocation searches on large Russian language corpora can be performed with the online collocator CoCoCo[9] (Kopotev et al. 2016; Kormacheva 2020). Unfortunately, gram-

[9]http://cococo.cosyco.ru/

Table 8.2: Headwords of the construction N-s-N: ruTenTen11.

Word	Translit	Meaning	F	ipm	Connotation	LL
бог	bog	'god'	10706	0.586	pos	7160.01
черт	čërt	'devil'	6981	0.382	neg	5694.86
хрен	hren	'horse-radish'	4054	0.222	neg	6173.16
фиг	fig	'fig'	2585	0.141	neg	3158.73
господь	gospodʹ	'Lord'	164	0.09	pos	8029.45
шут	šut	'fool'	1087	0.059	neg	1278.75
хуй	huj	'prick'	923	0.05	neg	511.64
пес	pës	'dog'	907	0.05	neg	76.94
хер	her	'prick'	473	0.026	neg	456.73
христос	hristos	'Christ'	304	0.017	pos	8190.28
аллах	allah	'Allah'	123	0.007	neg	2826.67
леший	lešij	'forest imp'	85	0.005	neg	58.79
дьявол	d'âvol	'devil, satan'	81	0.004	neg	1459.43
бес	bes	'devil'	37	0.002	neg	1011.68
шайтан	šajtan	'devil for muslims'	33	0.002	neg	26.24

matical searches are not available in the current version,[10] and therefore one can only submit queries on the concrete lexical realizations of constructions. I tested the headwords from Table 8.2 in combination with the word *s* 'with' (*bog + s*, *fig + s*, etc.) and no third collocation could be found for the words *gospod'*, *šut*, *hren*, *allah*, *lešij*, *bes*, and *šajtan*. For the remaining headwords, only pronoun collocations were found. The search for the noun preceding the phrases *s nim*, *s nej*, and *s nimi* yielded the collocates *bog*, *fig*, *hren*, *čërt*, and *shut*. The CoCoCo service performs searches on three Russian corpora: *Taiga*, the Russian National Corpus, and *I-ru*. Taiga (Shavrina & Shapovalova 2017) is the largest corpus and the only one suitable for our searches (and even in Taiga there was not enough data for some words). This shows that for studying syntactic idioms, one needs very large data sets, and the existing manually collected corpora are too small. Evidently, this is the reason why only part of my findings was confirmed with CoCoCo. Sadly, webcorpora like ruTenTen11 are also problematic in terms of data quality.

The total number of examples collected with Query 2 was 30,019 and the relative frequency of the construction was 1.64 ipm. This is a low frequency, e.g.

[10]To be more precise, it is possible to compose a query with a grammatical form and no lexeme, but the search does not work.

the Frequency Dictionary of Russian by Ljashevskaja & Sharov (2009) includes 20,000 words with a relative frequency higher than 2.6 ipm. An additional problem is that unlike lexemes, constructions cannot be detected by means of tokenization and lemmatization.

After studying Table 8.2, we can define the semantic restrictions for the head of the *N-s-N* construction: 'divine force', 'dark force', or 'masculine sexual organ' (obviously serving as a euphemism for dark forces, other swear words will not work in this construction).

In most cases, the collected examples have only obligatory components without optional particles at the beginning and the nominal group at the end. Still, 11,645 examples have emphasizing particles in the initial position of the construction: *da* (2,634), *da i* (1,948), *da nu i* (84), *i* (2,936), *nu* (461), and *nu i* (3,582). About 23% of the examples (6,897) have the optional nominal group with explicitation of the pronoun of the construction's third obligatory element.

It is impossible to analyse in detail the usage examples that were collected, as the size of the concordance was more than 30,000 items. Still, some of the most typical meanings could be found by studying random examples and collocations. The number of meanings has grown from the two meanings detected at the beginning of section 2 of this paper to the following four meanings:

Blessing: X gives Y a blessing to perform Z (headwords: *bog* 'god', *gospod'* 'Lord', *Hristos* 'Christ')

Disagreement, disbelief, surprise: X disagrees with Y/does not believe Y/is surprised with what Y says (headword: *bog* 'god', *gospod*' 'Lord', *Hristos* 'Christ')

Permission: X allows Y to perform Z (headwords: *bog* 'god', *gospod'* 'Lord', *Hristos* 'Christ')

Acceptance with disapproval: X is reluctant that Y is planning Z, but cannot prevent it (headwords: *fig* 'fig', *hren* 'horseradish', *her* 'prick')

The meanings can be connected: on the one hand, a positive attitude to Z (blessing, permission, see example 3) can gradually turn into a negative (acceptance with disapproval, see example 4). On the other hand, a blessing can transform into a disagreement (see example 7).

One might think that blessing and disagreement have nothing in common. However, disagreement can be expressed by blessing the existence of another point of view and in this way express that the speaker's point of view differs

from that of the interlocutor. The nouns in the first position should be *bog/gospod'/Hristos*. The "evil forces" are not fit for expressing the respect that is necessary for this meaning. The third component must be the second-person pronoun *ty/vy*. An exclamation mark is very typical for such contexts.

(7) a. —Ну, какая коррупция в
well.PTCP what.PRON corruption.NOUN.F.NOM in.PREP
Англии, да Бог с
England.NOUN.LOC and.PTCP god.NOUN.NOM with.PREP
вами!
you.PRON.2.INS.PL
'What corruption in England are you talking of? I don't believe you!'

b. Что вы, господь с вами!
what.PRON you.PRON Lord.NOUN.NOM with.PREP you.PRON.INS.PL
это не он.
this.PRON not.PTCP he.PRON.3M.NOM
'What are you talking about, you are wrong! It is not he!'

c. —Что ты! Христос с
what.PRON you.PRON.2SG Christ.NOUN.NOM with.PREP
тобой! — воскликнул я,
you.PRON.2.INS.SG shout.PAST.MSG I.PRON.NOM slightly.ADV
несколько испуганный.
frighten.PTCP.PASSMSG.NOM
'What are you talking about! You are wrong!' I shouted, a little frightened.'

The connections of the meanings are shown schematically in Figure 8.1.

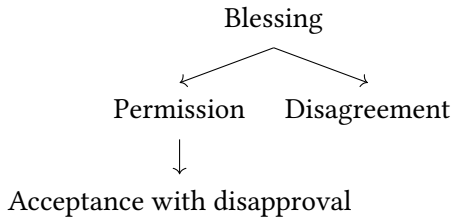

Figure 8.1: The meanings of the construction N-s-N.

Although *bog* 'god' occupies the first row of Table 8.2 and is almost twice as frequent as *čërt* 'devil', the words with negative connotations clearly dominate the

list, and the sums of frequencies of the expressions with the negatively coloured headwords outmatch the positively coloured. The result is 17,369 versus 12,650, which makes 58% against 42%. Even in cases where the headword is positively coloured, there might be contexts with negative connotations (e.g. example 7).

To sum up, the meanings of the construction are interrelated and have many borderline cases. For this reason, it is practical not to treat them as separate homonymous constructions, but rather as a single construction.

4 Constructing the construction

To present the *N-s-N* construction as a whole, I will take advantage of the box notation used in Construction Grammar (CxG). This notation is "a convenient way of organizing all the information needed to give an adequate account of linguistic structure" (Fried & Östman 2004: 13). The result of summing up the findings from the concordances is presented in the following box diagrams (Figures 8.2–8.4).

In section 3, the existence of semantic and structural variation in the research data was demonstrated. The easiest way to handle this heterogeneity is to define three variants of the construction *N-s-N*. Still, it is better to treat them as variants of the same construction rather than as independent constructions. The first variant (*N-s-N_a*, Figure 8.2) covers the meaning 'blessing'; the second one (*N-s-N_b*, Figure 8.3) has the meaning 'disagreement', 'surprise', or 'disbelief'; and the last one (*N-s-N_c*, Figure 8.4) handles the remaining meanings.

The construction *N-s-N_a* (Figure 8.2) is the simplest. The choice of the first noun is limited to three: *bog* 'god', *gospod'* 'Lord', and *Hristos* 'Christ'. The second nominal component is always a second-person pronoun. This pronoun can be explicitated by the optional noun phrase with the noun in the Nominative (or, rather, Vocative) case. The beneficiary of this construction is always a person (or a personified animal/artefact, etc).

(8) a. Здоровья Вам и Вашим
health.NOUN.GEN you.PRON.2.DAT.PL and.CONJ you.PRON.POSS.DAT.PL
близким, терпения! Бог с
relative.NOUN.DAT.PL, patience.NOUN.GEN! God.NOUN.NOM with.PREP
Вами!
you.2.INS.PL
'I wish you and your relatives health and patience. God be with you!'

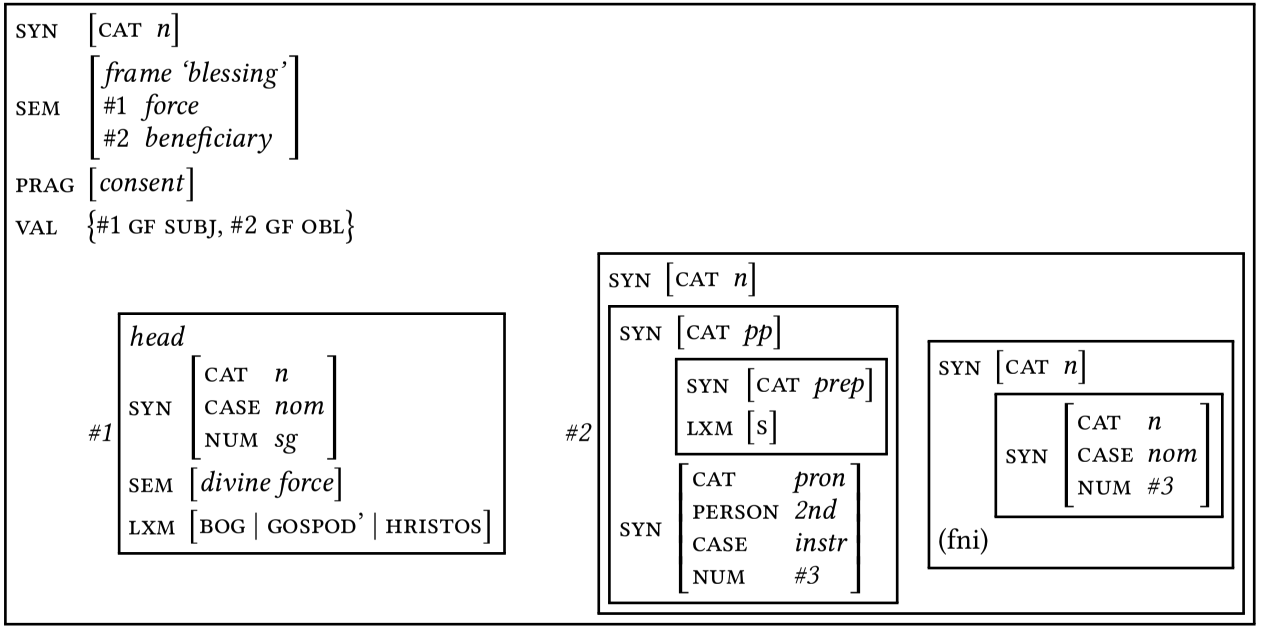

Figure 8.2: The representation of the N-s-N_a construction

b. Прощай, Оля, господь с
farewell.ADV Olja.NOUNPROPER.NOM, Lord.NOUN.NOM with.PREP
тобой.
you.PRON.2.INS.SG
'Farewell Olja, the Lord be with you.'

c. Прощайте! Христос с вами!
farewell.ADV! Christ.NOUN.NOM with.PREP you.PRON.2.INS.PL
'Farewell! Christ be with you!'

The variant *N-s-N_b* (Figure 8.3) looks very similar to the previous one and can be easily confused with it, as we will see in section 5 of this chapter. When spoken, the intonation of this variant is different from *N-s-N_a*, with a phrasal stress on the first nominal element; graphically it may be expressed with the exclamation mark. Besides, there is a structural difference: an optional particle *nu* or *da* in the beginning.

(9) a. Я ведь тебя убить хотел...
I.PRON.1.NOM.SG but.PTCP you.PRON.2.ACC.SG kill.INF want.PAST.MSG.
– Что ты, Бог
what.PRON you.PRON.2.NOM.SG. god.NOUN.NOM with.PREP
с тобой!
you.PRON.2.INS.SG!
'I wanted to kill you... – What are you talking about? I don't believe you.'

b. Откуда это вы взяли, что
where.PRON this.PRON you.PRON.2.NOM.PL. take.Past.PL that.PRON
я отрицаю ваше
I.PRON.1.NOM.SG negate.PRES.1SG you.PRON.2.POSS.ACC.SG
существование. Да Бог с
existence.NOUN.ACC.SG. and.PTCP god.NOUN.NOM with.PREP
вами!
you.PRON.2.INS.PL
'How did you decide that I do not believe in your existence? This is not true!'

c. ... говорят, что повально вся
... say.PRES.3.PL that_Pron "in_mass".ADV all.PRONF.NOM.SG
молодежь идет в
"young_people".NOUN.NOM.SG go.PRES.3SG to.PREP

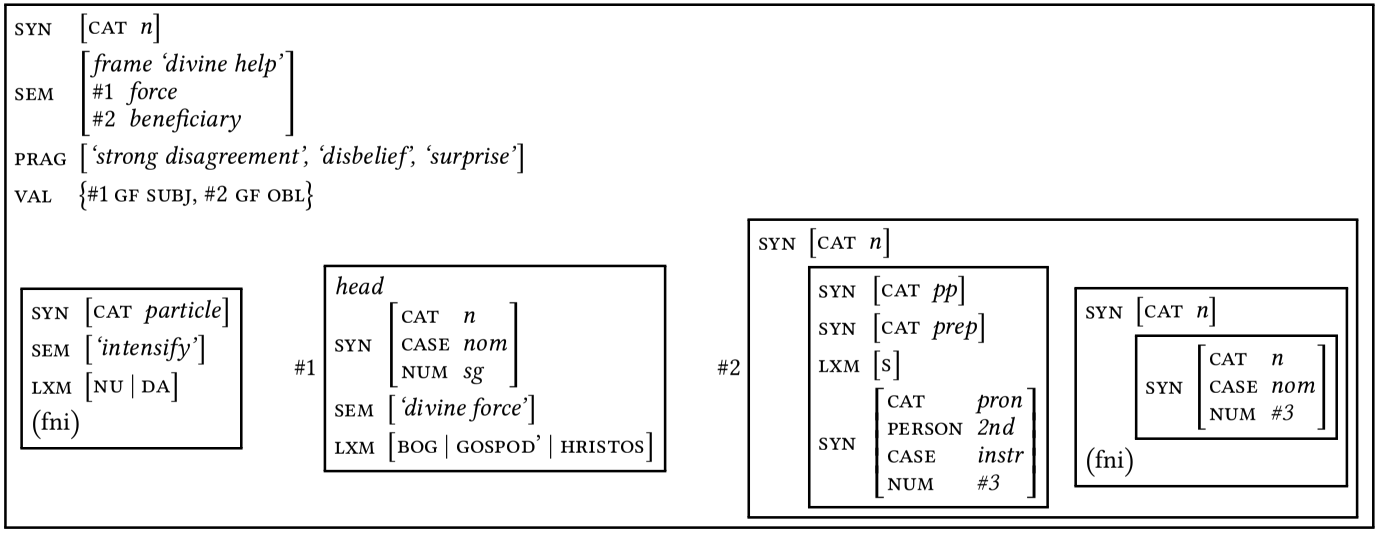

Figure 8.3: The representation of the N-s-N_b construction.

вузы. Да господь с
university.NOUN.ACC.PL. and.PTCP Lord.NOUN.NOM with.PREP
вами!
you.PRON.2.INS.PL
'They say all young people go to universities. No way!'

d. Лапшинов с нарочитым
Lapshinov.NOUNPROPER.NOM with.PREP faked.ADJ.INS.SG
негодованием успокаивал: -Что ты,
indignation.NOUN.INS.SG pacify.PAST.MSG. what.PRON you.PRON.2SG,
Никита! Христос с
Nikita.NOUNPROPER.NOM! Christ.NOUN.NOM with.PREP
тобой!
you.PRON.2.INS.SG
'Lapshinov was pacifying him with a faked indignation. Nikita, what are you (doing/saying)? You don't mean it!'

The construction *N-s-N_c* (Figure 8.4) gives more freedom to choose the first nominal element. Any noun from the list of Table 8.2 can be used, including the three nouns from the *N-s-N_a* and *N-s-N_b* variants. The list is open and other nouns with the semantics of 'superhuman force' are applicable (see section 3 for details). The second nominal element can be a noun or a pronoun, and there are no semantic restrictions: it can be a person, a thing, an activity, a situation, etc.

This variant can have an optional initial element: a particle or a combination of particles that work as an intensifier. The palette is richer than in *N-s-N_b*, which has only two options. At least the following combinations are used quite frequently: *da, i*, *da i*, *nu i*, and *nu da i*. The most frequently used is the combination *nu i* (3,582 examples in the concordance). The expression *nu i* is used in other contexts as well (e.g. *Nu i durak* 'what a fool you are'), and Dobrovol′skij et al. (2019) claim that it is a separate construction.

(10) a. Диска с ПО, естественно,
disk.NOUN.GENSG with.PREP software.NOUN_Abr, "of_course".ADV,
тоже никакого нет, ну и
also.PTCP none.PRON.GENSG no.Pred, well.PTCP and.PTCP
аллах с ним.
Allah.NOUN.NOM with.PREP he.PRON.3.INS.SG
'Of course there is no software included, well, I don't care.'

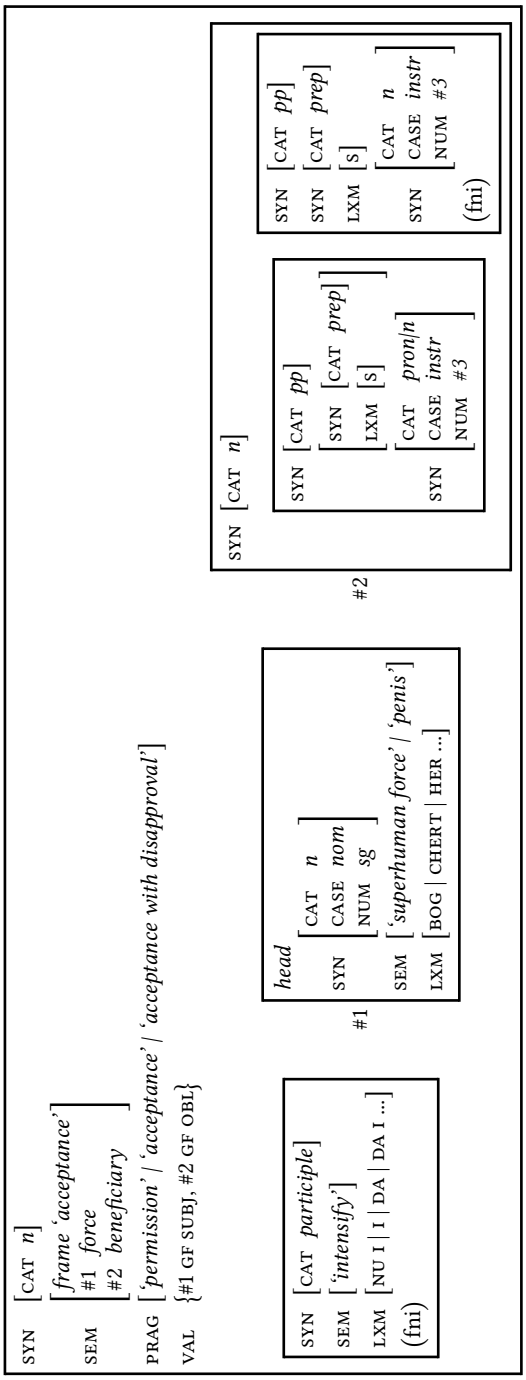

Figure 8.4: The representation of the N-s-N_c construction.

b. Да бес с ними,
and.PTCP devil.NOUN.NOM.NOUN.NOM.SG with.PREP he.PRON.INS.3.PL
с британцами.
with.PREP Brit.NOUN.INS.PL
'I don't care about the Brits.'

c. Радуйтесь, и Господь с вами.
enjoy.IMP.PL and.PTCP Lord.NOUN.NOM with.PREP you.PRON.2.INS.PL
'Be happy and the Lord be with you.'

Another optional component is the prepositional phrase, which can be used for the explicitation of the second nominal element if the latter is a pronoun. Unlike the construction N-s-N_a, this element is not in the Nominative case, but it repeats the structure of the second element: preposition *s* 'with' + noun phrase in the Instrumental case.

This prepositional phrase can be a combination of a preposition with a single noun (11a), or it can have quite a complicated structure (11b and 11c).

(11) a. Черт с ним, с
devil.NOUN.NOM with.PREP he.PRON.INS.SG, with.PREP
народом!
people.PRON.INS.SG
'I don't care about the people!'

b. Бог с ней, с
god.NOUN.NOM with.PREP she.PRON.3F.INS.SG with.PREP
этой конкретной передачей.
this.PRON.INS.SG concrete.ADJ.INS.SG programme.NOUN.F.INS.SG
'I don't care about this particular programme.'

c. Впрочем, аллах с ними, с
anyway.ADV, Allah.NOUN.NOM with.PREP it.PRON.INS.3.PL with.PREP
нашими смешными читательскими
our.PRON.POSS.INS.PL funny.ADJ.INS.PL reader.ADJ.INS.PL
проблемами!
problem.NOUN.F.INS.PL
'Anyway I do not care about our funny readers' problems!'

In rare cases, the use of two pronouns is possible, like in (12), where the speaker is the beneficiary.

(12) Бог с ним, со мной.
god.NOUN.NOM with.PREP he.PRON.3M.INS.SG with.PREP I.PRON.1.INS.SG
'I do not care about myself.'

If the context is limited, the construction may become ambiguous, as in example (13), which can be interpreted as a blessing, as an acceptance, and even as disagreement.

(13) –Друже Чумак, <…> я
friend.NOUN.VOC Chumak.NOUNPROPER.NOM I.PRON.NOM.SG NO.PTCP
не требую никаких
demand.PRES.1SG NO.PRON.NEG.GEN.PL explanation.NOUN.GEN.PL.
объяснений. Бог с вами.
god.NOUN.NOM with.PREP you.PRON.INS.PL
'Friend Chumak, I am not demanding any explanations. God be with you/I do not care/Not at all.'

The most frequently used is the construction N-s-N_c. A check of the same random sample of 1,000 examples that was used for calculating the precision of the search (see section 2) confirmed this: among the 991 correct examples, only 49 (4.9%) belonged to N-s-N_a. The construction N-s-N_b occurred about the same number of times, in 47 examples (4.7%), and all the remaining 895 (90.3%) examples belonged to N-s-N_c.

5 The challenges of parallel concordancing

To check the equivalents that are used when translating contexts containing a certain construction, one needs many examples from parallel texts. This becomes a problem when studying multiword expressions, because their frequencies are low, and therefore large amounts of text are needed to get enough examples. As it has already been mentioned, the best source of data for studying idiomatic expressions are fiction and mass media texts. Such texts are available in parallel corpora, but the sizes of parallel corpora of literary texts are quite modest compared to gigaword monolingual corpora. The data I used for this study were as follows:

1. Parallel corpora at the Russian National Corpus (RNC)
    - Russian-English subcorpus (6.5m running words)
    - English-Russian subcorpus (18m running words)

2. Parallel corpora at Tampere University
    - ParRus, the Russian-Finnish corpus of fiction texts (6m running words)
    - ParFin, the Finnish-Russian corpus of fiction texts (3m running words)

It is obvious that the amounts of data from these parallel corpora are microscopic in comparison with ruTenTen11. Besides, the Russian-English subcorpus of the RNC is not well-balanced: works by Vladimir Nabokov clearly dominate over all other authors and periods. However, there were no other data available. Parallel corpora at SketchEngine are larger, but their composition is unclear, and it is impossible to filter out indirect translations and pseudotranslations. Hence, our data will be suitable only for detecting general tendencies for some of the expressions.

Table 8.3: Frequencies of the headwords N-s-N construction in the parallel corpora.

Word			RuEn F	ipm	EnRu F	ipm	RuFi F	ipm	FiRu F	ipm
бог	bog	'god'	65	9.86	23	1.27	73	23.09	9	5.03
господь	gospoď	'Lord'	8	1.21	13	0.72	0	0	0	0
пес	pes	'dog'	1	0.15	0	0	0	0	0	0
Христос	Hristos	'Christ'	10	1.52	0	0	0	0	0	0
черт	čert	'devil'	42	6.37	28	1.55	52	16.44	6	3.35
шут	šut	'clown'	2	0.30	1	0.06	0	0	0	0
хрен	hren	'horseradish'	0	0	9	0.50	0	0	0	0

It is easy to observe in Table 8.3 that the normalized frequencies of headwords are much higher than in ruTenTen11, although not all expressions were found (only seven of fifteen). This can be explained by the structure of ruTenTen11, which contains many genres in which the construction *N-s-N* is never used. The causes of the differences in frequencies between parallel corpora are the corpora's imbalance and their construction from whole texts, and therefore a couple of very long texts could skew the whole collection.

The comparison of the frequencies of *N-s-N* in ruTenTen11 and the parallel corpora demonstrates that the frequencies of expressions are much less stable than those of single words, and it is problematic to obtain reliable statistics from the observations. For example, the frequency of the expression *bog s X* is 9.8 ipm in Russian-English RNC and 23.1 ipm in ParRus, although both corpora are collections of Russian fiction texts.

Regardless, one important observation can be made from the frequencies: the construction *N-s-N* is much more frequent in the original Russian texts than in the translations from English and Finnish into Russian. This is the sign of the evident absence of matching constructions in both English and Finnish. The findings are also in line with Tirkkonen-Condit's (2004) hypothesis about the underrepresentation of unique items of the source language in translated language.

The statistics from the parallel concordances give the impression that something is not right. As it was shown in the previous sections, the construction *N-s-N* is polysemous, and the actual meaning depends on the context. The most misleading is the construction with *bog* 'god' as a headword: it can be used in all three variants of the construction described in section 4 of this paper. The variant *N-s-N_a* is not very frequent: I demonstrated this by the study of random examples. Still, in the Russian-English data, 28 contexts out of 65 were translated into English with expressions containing the word *god*. In the Russian-Finnish data, there are 73 contexts with *bog* 'god', and 48 of them are translated with the expressions containing *jumala* 'god', *herra* 'Lord', or *luoja* 'Creator'. From the above-mentioned study of random examples, I would have expected that only about 7% of the contexts of *bog s X* would belong to the *N-s-N_a* variant, while the statistics from the parallel corpora show a much higher rate in both the Russian-English and Russian-Finnish data.

It is true that the data are not balanced, and that the frequencies of the expressions in our data vary greatly. It is therefore quite possible that the data from the parallel corpora might contain far more *N-s-N_a* contexts than the ruTenTen11 data. For this reason, it is necessary to check the actual contexts to confirm the statistical observations.

The checking of the Russian-English concordance with *bog* 'god' on the Russian side and *god* on the English side confirmed my suspicions: 19 cases out of 28 show an obvious misunderstanding of the source text.

(14) «Ну, бог с тобой, оставайся
"well.PTCP god.NOUN.NOM with.PREP you.PRON.2.INS.SG stay.IMP
уж», — решила в тоске
well.PTCP", — decide.PASTFSG in.PREP melancholy.NOUN.LOCSG
Грушенька, сострадательно ему
Grushenka.NOUNPROPER.NOM, compassionately.ADV he.PRON.DATSG
улыбнувшись.
smile.GERUND
"OK, I don't care, you can stay, decided Grushenka in her melancholy and smiled at him compassionately."

> "Well, God bless you, you'd better stay, then," Grushenka decided in her grief, smiling compassionately at him.' (F. Dostoevsky. 1878. *Brat´â Karamazovy* [The Brothers Karamazov], transl. C. Garnett, 1912)

In example (14), the speaker reluctantly gives the interlocutor her permission to stay, while the translator obviously understood the expression as a blessing or at least as a demonstration of piety (which is strange for Grushenka, who, as we know, was not a very pious person).

In the Russian-Finnish data, 44 contexts with an obvious misunderstanding were found. An additional factor for misinterpreting is Russian-Finnish dictionaries, some of which register the phrase *bog s X* only with the meaning of blessing (see, e.g. Kuusinen & Ollikainen 1984).

(15) Господин Разумихин не то-с,
mister.NOUN.NOM.SG Razumihin.NOUNPROPER.NOM not.PTCP that.PRON,
да и человек посторонний,
and.PTCP and.Particle person.NOUN.NOM.SG stranger.ADJ.NOMMSG,
прибежал ко мне весь такой
run.PAST.MSG to.PREP I.PRON.1.DAT all.PRONM.NOM such.ADJ.NOM.SG
бледный ... Ну да бог с
pale.ADJ.NOMMSG. but.PTCP and.PTCP god.NOUNM.NOM with.PREP
ним, что его сюда мешать.
he.PRON.3.INS.SG, what.PRON he.PRON.3.ACC.SG here.ADV involve.INF
'Mister Razumihin is a stranger, but he ran to me so pale. Never mind, why shall we involve him in this.'

(16) Herra Razumihinhan on vallan
mister.NOUN.NOM Razumihin.NOUN.NOM be.PRES power.NOUN.GEN
toista maata, sivullinen
another.ADJ.PART country.NOUN.PART stranger.ADJ.NOM
ihminen, vaikka hän juoksi silloin
person.NOUN.NOM although.CONJ he.PRON.NOM run.PAST then.ADV
kasvot kalpeina luokseni ... Luoja
face.NOUN.NOM white.ADJ.ESSIVE "to me".ADV. Lord.NOUN.NOM
hänen kanssaan, eihän hänellä ole tässä
he.PRON.GEN with.POSTP not.PTCP he.PRON.3.ALL be.PRES here.ADV.ADESS
osaa eikä arpaa.
part.NOUN.PART not.PTCP lot.NOUN.PARTSG
'Mister Razumihin is like from another country, a stranger, still he ran to me with a white face. God be with him, he has nothing to do in this

business.' (F. Dostoevsky. 1866. *Prestuplenie i nakazanie* [Crime and Punishment], transl. J. Konkka, 1970)

The expression *čërt s X* 'devil with X' also contains a trap: it can be interpreted as swearing and blasphemy, although in many cases it has a different meaning and belongs to the construction *N-s-N_c*.

(17) Об чем? Ну, да черт
about.PREP what.PRON.LOC well.PTCP and.PTCP devil.NOUN.NOM
с тобой, пожалуй, не сказывай.
with.PREP you.PRON.INS.SG maybe.ADV not.PTCP tell.IMPSG
'What about? Well, do not tell, I don't mind.'
'What about? Confound you, don't tell me then.'
(F. Dostoevsky. 1866. *Prestuplenie i nakazanie* [Crime and Punishment], transl. C. Garnett, 1914)

One might think that such things take place only in very old translations of even older source texts. However, this is not so: in (18) is an example of a relatively recently published translation from Russian into Finnish.

(18) Черт с ним! – сердито подумала
devil.NOUN.NOM with.PREP he.PRON.INS.SG angrily.ADV think.PASTFSG
Вероника.
Veronika.NOUNPROPER.NOM.
'I don't care, thought Veronika angrily.'

(19) "Hitto", Veronika mietti
devil.NOUN.NOM, Veronika.NOUNPROPER.NOM think.PAST.3SG
vihaisena.
angry.ADJ.ESS.SG
'Devil, thought Veronika angrily.' (A. Marinina. 1995. *Za vse nado platit´* [You have to pay for everything] transl. O. Kuukasjärvi, 2005)

It should be mentioned that the parallel concordance also provided enough examples with interesting solutions for this construction. I will give here only two examples from the Russian-English data. In (20a) an English expression *all right* is used, while in (20b) the meaning of expression is explicitated (*I will take it*).

(20) a. Ну что ж, бог с
well.PTCP what.PRON but.PTCP, god.NOUN.NOM with.PREP
вами, пусть пять рублей
you.PRON.2.INS.PL, let.PTCP five.NOUN.NUM.NOM rouble.NOUN.GEN.PL
будет. Только деньги попрошу
be.V.FUT.3SG. only.ADV money.NOUN.ACC.PL ask.V.FUT.1SG
вперед.
forward.ADV
"OK, let it be five roubles, but I would like to have the money in advance."
"Well, all right, make it five roubles. Only I want the money in advance, please."
(Ilya Ilf, Evgeny Petrov. 1927. Двенадцать стульев (*Dvenadcat´ stul´ev*) [The Twelve Chairs], transl. J. Richardson, 1961).

b. Ну, бог с вами, —
well.PTCP god.NOUN.NOM with.PREP you.PRON.2.INS.PL say.PAST
сказал Махин, кладя на
Mahin.NOUNPROPER.NOM, put.GERUND on.PREP counter.NOUN.ACC.SG
витрину купон.
coupon.NOUN.ACC.SG
"'OK, I agree,' said Mahin, putting the coupon on the counter".
"Well, I will take it," said Mahin, and put the coupon on the counter.
(Leo Tolstoy. 1889–1904. *Fal´šivyj kupon* [The Forged Coupon] 1889–1904, transl. Louise and Aylmer Maude, 1911)

To sum up the findings from the parallel concordances, the main problem of the data obtained from translations from Russian into other languages is the possibility of misunderstanding the source texts by translators. Hence, translations from other languages into Russian quite unexpectedly become a very useful source of reference data. Translators into Russian write in their native language and their work is addressed to other native speakers of Russian. As a result, the expression that served as a stimulus for the Russian expression may be with a few reservations used as an equivalent for translating in the opposite direction. Of course, in this case there is an issue of the correct understanding of the source text in language X.

The RNC's English-Russian subcorpus is larger and richer than the Russian-English one. In spite of this, the construction *N-s-N* features in it much less frequently (see Table 8.3). Still, the parallel concordance produces some interesting solutions that seem suitable for translating from Russian into English as well.

(21) a. I've been told I ought to have a salon, whatever that may be. Never mind. Go on, Badger.
Мне частенько говорили, что мне
I.PRON.1.DATSG often.ADV say.PAST.PL that.PRON I.PRON.1.DATSG
надо бы завести салон, что
need.Pred would.PTCP start.INF salon.NOUN.ACC.SG, what.PRON
бы это там ни значило. Ну,
would.PTCP this.PRON there.ADV not.PTCP mean.PAST. well.PTCP
бог с ним! Продолжай,
god.NOUN.NOM.SG with.PREP he.PRON.3.INS.SG. continue.IMP,
Барсук.
Badger.NOUN.NOM.
'They often said to me that I should start a salon, whatever it may mean. Continue, Badger.' (Kenneth Grahame. 1908. *The Wind in the Willows*, transl. I. Tokmakova, 1988)

b. "You still have half your balls there." "I don't care. This will set my game back a month."
–У тебя еще осталась
by.PREP you.PRON.2.GENSG still.PTCP remain.PASTFSG
половина мячей. –И черт
half.NOUN.NOM.SG ball.NOUN.GEN.PL. and.PTCP devil.NOUN.NOM
с ним. Это отбросит
with.PREP he.PRON.3.INS.SG. It.PRON throw.V.FUTure.3SG
мою технику на
my.PRON.POSSF.ACC.SG technique.NOUN.ACC.SG on.PREP
месяц назад.
month.NOUN.ACC.SG back.ADV
'You still have half of the balls. I don't care. It will throw my technique a month back.' (Michael Connelly. 2002. *City Of Bones*, transl. D. Vozniakevitch, 2006)

The same can be observed in the Finnish-Russian parallel concordance obtained from the ParFin corpus.

(22) Lukeneilla ihmisillä sellanen on
study.PTCP.ALL.PL person.NOUN.ALL.PL such.ADJ.NOM.SG be.V.3PRESSG
ja hyvä niin.
and.CONJ good.ADJ.NOM.SG so.ADV
'Educated people have this and this is good.'

(23) У тех, кто учился, есть,
by.PREP he.PRON.GEN.PL, who.PRON.NOM study.PAST.MSG, be.PRES.3,
и бог с ними.
and.PTCP god.NOUN.NOM with.PREP he.PRON.INS.PL.
'Those who studied have it and let it be' (Kari Hotakainen. *Juoksuhaudantie*, transl. I. Uretski)

Strangely, although the English stimuli *never mind* and *I don't care*, as well as the Finnish stimulus *hyvä niin* 'OK' can be considered as very good variants for conveying the meaning of the Russian construction *N-s-N_c*, they are not very typical for translations from Russian. The expression *never mind* occurs only 7 times in the Russian-English concordance, and the verb *care* only three times. In the Russian-Finnish parallel concordance, there is not a single example of *hyvä niin* used as an equivalent for *N-s-N*.

6 Discussion

The case study performed in this paper demonstrates the usefulness of monolingual and parallel corpora for studying constructions. Corpora provide information on the variability of constructions and statistics. Monolingual concordancing is helpful in the study of the components of the construction, the lexemes used for its realization, and even semantic issues. The analysis reveals that the construction *N-s-N* can be implemented in the form of ready-made phrases (like *bog s nim, čërt s nim*, etc.) that are used very frequently, as well as in the form of *hapax legomena* constructed with the same template. As a result, some phrases may be registered in dictionaries, while occasionalisms remain outside both dictionaries and grammar descriptions due to their rarity and specificity. Evidently, the best way of describing and storing such units would be databases like FrameNet or Constructicon.

To study the links of the construction with other languages, parallel corpora were used. However, the usability of this resource was limited. Parallel corpora did not help so much in looking up translation equivalents as one might have expected. The first reason was that the search did not return enough usage examples; one would have needed much larger data collections to obtain a parallel concordance at least comparable with the monolingual concordance from ruTenTen11. The data that were available were sufficient only for demonstrating the fact that the *N-s-N* construction in Russian does not have corresponding constructions in English or Finnish, and that this absence causes difficulties for translators.

The second reason was the rather high rate of errors in the translations. Of course, one might expect errors in any language data – this is quite natural – but in this case the errors were repeated, and their main cause was misinterpretation of the source text. On the one hand, this is a challenge to modern statistical and neural machine translation technologies, which are based on parallel corpora and use human translations for modelling MT. The developers of MT presume that there might be errors and mistakes in the data, but are they ready for errors on such a scale? On the other hand, this is a challenge to the belief that the translation of a literary work into another language is the *same* story told in other words. The real data show that literary translators sometimes do not understand the source text well enough.

Why does this happen? The first priority of a literary translator is to produce a good target text, one that meets the standards of a literary text. The correspondence of the translation to the source text comes second, and it is not likely that every passage of the translation is compared to the original text. Of course, the translation should not be very different, but how correct should it be? There is also some evidence that the literary translators' command of the source language is not as advanced as one might expect. For example, Nikolai Čukovskij, one of the leading Russian literary translators working from the 1920s to the 1960s, was very critical of his own proficiency in English (Čukovski & Čukovskij 2004), and there existed writers (and especially poets) who "translated" by editing earlier translations or literal translations produced by other people (see, e.g. Kamovnikova 2019).

These issues make the use of parallel corpora of literary texts a specific resource. They cannot be, for example, the main source of data for bilingual dictionaries, but rather reference data for rechecking translation equivalents. Parallel corpora also demonstrate that even nowadays, proficiency in non-native languages is limited and needs to be improved. The data from parallel corpora might be of great help in finding such weak points.

List of dictionaries and corpora

Academic dictionary of Russian phraseology : See Baranov & Dobrovol′skij 2015.
Constructicon for Russian : See Borin et al. 2012 and https://spraakbanken.gu.se/karp/#?mode=konstruktikon-rus
Dictionary of Russian language : See Evgen′eva 1984.
enTenTen15 : English corpus from the web, see https://www.sketchengine.eu/ententen-english-corpus/

Parallel corpora at the RNC : See http://ruscorpora.ru/new/search-para-en.html
ParFin : Russian–Finnish parallel corpus of literary texts, Tampere University, htts://puolukka.uta.fi/texthammer
ParRus : Finnish–Russian parallel corpus of literary texts, Tampere University, htts://puolukka.uta.fi/texthammer
ruTenTen11 : Russian corpus from the web, see
The explanatory dictionary of Russian language : See Ožegov & Švedova 1992.
The Frequency dictionary of the Russian language (based on RNC) : See Ljashevskaja & Sharov 2009 and http://dict.ruslang.ru.
The large Russian–Finnish dictionary : See Kuusinen & Ollikainen 1984.
The Random House Russian-English dictionary of idioms : See Lubensky 1995.

References

Baranov, Anatolij & Dmitrij Dobrovol′skij. 2008. *Aspekty teorii frazeologii.* [Aspects of the phraseological theory]. Moskva: Znak.

Baranov, Anatolij & Dmitrij Dobrovol′skij (eds.). 2015. *Akademičeskij slovar' russkoj frazeologii [Academic dictionary of Russian phraseology].* Moscow: LEKSRUS.

Borin, Lars, Markus Forsberg, Leif-Jöran Olsson & Jonatan Uppström. 2012. The open lexical infrastructure of Språkbanken. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 3598–3602. Istanbul: ELRA.

Čukovski, Kornej & Nikolaj Čukovskij. 2004. My očen′ raznye literatory… [We are very different writers]. *Znamja* 1. http://www.chukfamily.ru/kornei/prosa/pisma/my-ochen-raznye-literatory-ki-chukovskij-nk-chukovskij-perepiska.

Dobrovol′skij, Dmitrij, Mihail Kopotev & Ludmila Pöppel. 2019. Gruppa konstrukcij nu i x: Semantika, pragmatika, sočetaemost′ [The group of constructions nu i X: Semantics, pragmatics, combinability]. *Scando-Slavica* 65. 5–25. DOI: 10.1080/00806765.2019.1586578.

Dunning, Ted. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics* 19(1). 61–74. https://www.aclweb.org/anthology/J93-1003/.

Evgen′eva, A. P. (ed.). 1984. *Slovar′ russkogo âzyka. [Dictionary of Russian language].* Moskva: Russkij âzyk.

Fried, Mirjam & Jan-Ola Östman. 2004. Construction grammar: A thumbnail sketch. In M. Fried & J. Östman (eds.), *Construction grammar in a cross-language perspective*, 11–86. Amsterdam/Philadelphia: John Benjamins.

Härme, Juho & Mikhail Mikhailov. 2016. From Russian to Finnish and back: Compiling Russian–Finnish–Russian parallel corpora. In Lubica Medvecká (ed.), *Translation from/into languages of limited diffusion*, vol. 3, 139–147. Bratislava: The Slovak society of translators of scientific & technical literature.

Kamovnikova, Natalia. 2019. *Made under pressure: Literary translation in the soviet union, 1960-1991 (studies in print culture and the history of the Book)*. Amherst, Massachussets: University of Massachusetts Press.

Kopotev, Mihail, Lidia Pivovarova & Daria Kormacheva. 2016. Constructional generalization over Russian collocations. In Begoña Sanromán Vilas (ed.), *Collocations cross-linguistically* (Mémoires de la Société néophilologique de Helsinki Tome C), 121–140. Helsinki: Modern Language Society.

Kormacheva, Daria. 2020. *Collocations on the way: How words come together in Russian*. University of Helsinki. (Doctoral dissertation). http://urn.fi/URN:ISBN:978-951-51-6083-6.

Kuusinen, Martti & V. Ollikainen (eds.). 1984. *Venäläis-Suomalainen suursanakirja [The large Russian-Finnish dictionary]*. Helsinki: WSOY.

Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins.

Ljashevskaja, Olga & Sergej Sharov. 2009. *Častotnyj slovar´ russkogo âzyka. (Na materialah Nacional´nogo korpusa russkogo âzyka). [The frequency dictionary of the Russian language (based on Russian National Corpus)]*. Moskva: Azbukovnik.

Lubensky, Sophia. 1995. *Random House Russian-English dictionary of idioms*. New York: Random House.

Mikhailov, Mikhail & Juho Härme. 2015. Parallelnyje korpusa hudožestvennyh tekstov v Tamperskom universitete [Parallel corpora of fiction texts at the University of Tampere]. *Russkij jazyk za rubežom* (Spetsvypusk). 16–19.

Nivre, Joakim & Chiao-Ting Fang. 2017. Universal dependency evaluation. In *Proceedings of the NoDaLiDa 2017. Workshop on Universal Dependencies (UDW 2017)*, 86–95. Gothenburg.

Ožegov, S. I. & N. Û. Švedova. 1992. *Tolkovyj slovar´ russkogo âzyka [The explanatory dictionary of Russian language]*. Moskva: Az″.

Shavrina, Tatiana & Olga Shapovalova. 2017. *To the methodology of corpus construction for machine learning: "taiga" syntax tree corpus and parser*. In Trudy meždunarodnoj konferencii "Korpusnaja lingvistika–2017". Sankt-Peterburg: SPbGU & ILI, 78.

Tirkkonen-Condit, Sonja. 2004. Unique items: Over or underrepresented in translated language? In Anna Mauranen & Pekka Kujamäki (eds.), *Translation universals: Do they exist?*, 177–184. Amsterdam/Philadelphia: John Benjamins.

Wierzbicka, Anna. 1987. Boys will be boys: 'Radical semantics' vs 'radical pragmatics'. *Language* 63(1). 95–114.

Xiao, Richard. 2015. Collocation. In Douglas Biber & Randy Reppen (eds.), *The Cambridge handbook of English corpus linguistics*, 106–124. Cambridge, UK: Cambridge University Press. DOI: 10.1017/CBO9781139764377.

Zaharov, N. V. 2008. Anglijskij âzyk v tezauruse puškina [The English language in the Pushkin's thesaurus]. *Znanie. Ponimanie. Umenie* 9.