In which case are Russians afraid? *Bojat'sja* with genitive and accusative objects

Abstract

The present article investigates case use with the verb *bojat'sja* 'be scared' in Russian. Many verbs with *-sja* never combine with objects in the accusative case. The verb *bojat'sja* historically was among such verbs, but this verb is undergoing a shift and is currently used with both genitive and accusative objects. This study examines the parameters that motivate this change. Using data from the Russian National Corpus and an experimental study, we show that the accusative case is more likely to appear when the object is individuated. We furthermore demonstrate that the use of accusative objects depends on register: less restricted registers such as newspaper texts and answers in the experiment show more uses of accusative objects.

1. Introduction

The speakers of Old Russian used *sja* as a clitic accusative reflexive pronoun. However, the clitic underwent grammaticalization and has developed into a verbal affix in Contemporary Standard Russian (Zaliznjak 2008, Nesset 1998a, 1998b).[1] Since the slot of the accusative object was originally filled with *sja*, verbs in *–sja* normally do not combine with accusative objects. Instead, objects of such verbs appear in the genitive, dative or instrumental cases. In this paper we discuss the verb *bojat'sja* 'be scared' which traditionally is described as one of the verbs that combine exclusively with the genitive case as in (1), but sometimes is also attested with the object in the accusative case as in (2).[2]

(1) On   **bo-it-sja**   **žen-y**   i   stara-et-sja
     he.NOM.SG   be.scared-PRAES.3SG-REFL wife-GEN.SG   and   try-PRAES.3SG-REFL
     byt'   točn-ym.
     be.INF   precise-INS.SG

---

[1] Many researchers refer to *–sja* as a "postfix", since it occurs at the right edge of the word, after inflectional endings.

[2] Other verbs with *–sja* that govern the genitive are *bereč'sja* 'be careful', *čuždat'sja* 'shun', *deržat'sja* 'hold on to', *dičit'sja* 'be shy of', *dobivat'sja* 'strive for', *dožidat'sja* 'wait for', *domogat'sja* 'seek after', *gnušat'sja* 'have an aversion to', *kasat'sja* 'touch', *opasat'sja* 'be afraid of', *osteregat'sja* 'beware of', *pugat'sja* 'be frightened of', *slušat'sja* 'obey', *stesnjat'sja* 'feel shy', *storonit'sja* 'shun', *strašit'sja* 'be afraid of', *stydit'sja* 'be ashamed of', *udaljat'sja* 'move away', *užasat'sja* 'be horrified', *lišat'sja* 'be deprived of', and *xvatit'sja* 'notice the absence of'. However, in the present article we limit ourselves to discussing *bojat'sja*. All our examples are from the Russian National Corpus available at www.ruscorpora.ru. It is interesting to notice that the accusative is used in (2) even though the verb is negated, since negation generally favors the use of the genitive (Timberlake 1985).

'He is afraid of his wife_{GEN} and tries to be precise.' [Vs. V. Ivanov. Dnevniki (1940-1948)]

(2) Poèt.∅   ne   **bo-it-sja**   **žen-u**,   s   kotor-oj
     poet-NOM.SG   NEG   be.scared-PRAES.3SG-REFL wife-ACC.SG   with   which-INS.SG
     ne   živ-et,   bespoko-it-sja   ob   izdani-i
     neg   live-PRAES.3SG   be.worried-PRAES.3SG-REFL   about   publishing-LOC.SG
     knig.∅,   svo-ix   i   druz-ej,   i   mečta-et   o
     book-GEN.PL   own-GEN.PL   and   friend-GEN.PL and   dream-PRAES.3SG about
     večn-om   sojuz-e   «dv-ux   istinnyx   xudožnikov»,
     everlasting-MASC.LOC.SG union-LOC.SG   two-GEN   true-GEN.PL   artist-GEN.PL
     ne   zameč-aja   nača-vš-ego-sja   otdaleni-ja.
     NEG   notice-GER   begin-PARTCP.ACT.PAST-NEUT.GEN.SG-REFL estrangement-GEN.SG
     'The poet is not afraid of his wife_{ACC}, who he doesn't live with, he worries about publishing his own and his friends' books, he dreams about the everlasting union of "two true artists", and he does not notice the beginning of the estrangement.' [Aleksandr Pjatigorskij. Vspomniš' strannogo čeloveka (1997)]

There is no general agreement concerning the status of accusative objects for verbs like *bojat'sja*. The academy grammar (Švedova (ed.) 1980 vol. 2: 35) and major dictionaries (Evgen'eva 1999, Ožegov/Švedova 1992 and Ušakov 1935-40/2008) analyze *bojat'sja* as combining only with genitive objects. However, some grammars (Švedova/Lopatin 1989: 364, Timberlake 2004: 319) and literature on normative speech (Butorin 1966, Černyšev 1911, Gorbačevič 1971: 237, Ickovič 1982: 35-37, Ljustrova et al. 1982, Prokopovič et al. 1975, Rozental' 1986, 1988, Rozental' and Telenkova 1984) mention that *bojat'sja* occasionally combines with the accusative. Comrie, Stone and Polinsky (1996: 144-147) and Nichols (1993: 82) also briefly mention that *bojat'sja* may appear with the accusative case.

Among the scholars who acknowledge the existence of accusative objects with *bojat'sja* the assessments of the situation are different. For instance, Butorin (1966: 130) concludes that "[i]n 19th century literature and contemporary usage examples with *bojat'sja* + the accusative are sometimes ("*inogda*") encountered." Krys'ko (1997: 240-245) and Ferm (2004) offer two larger and more recent diachronic studies. Krys'ko (1997: 245) claims that the accusative is used "massively", while Ferm analyzes a small sample (300 examples) from the Russian National Corpus (RNC) and finds only few examples with the accusative.

Since the status of accusative objects for verbs like *bojat'sja* is controversial, we decided to carry out an empirical investigation of the situation. Our study, which is based on data from the Russian National Corpus and a psycholinguistic experiment, investigates the factors motivating the use of the accusative. Our contribution can be summarized as follows. First, we show that accusative objects are still relatively infrequent in Contemporary Standard Russian. Second, it is demonstrated that the use of the accusative is more frequent for highly individuated objects. Third, our findings indicate that the use of the accusative depends on

register; we find more accusative objects in newspaper texts and in our experimental data than in more formal registers.[3]

The remainder of the article is structured as follows. After a brief discussion of the individuation and register hypotheses in section 2, we present our corpus study in section 3, where we test these hypotheses. In section 4, we turn to the experimental data, which enable us to assess the relative importance of the factors favoring the accusative. The contribution of our study is summarized in section 5.

2. Hypotheses: Individuation and Register

Variation between objects in the genitive and accusative cases is well studied in three constructions: the genitive of negation, the partitive construction and the construction with so-called intensional verbs, such as *ždat'* 'wait'. In the genitive of negation construction, the accusative and genitive cases are in competition when a transitive verb is negated – the object of the transitive verb is marked with genitive (3). The object of the non-negated transitive verb appears in accusative case (4). A highly individuated animated object tends to appear in the accusative even under negation (Babby 1980: 154–158, Padučeva 2006: 31–32), see (5).

(3) Ja   ne   **čita-ju**   **gazet**-∅   i   ne   smotr-ju
    I   NEG   read-PRAES.1SG   newspaper-GEN.PL   and   NEG   watch-PRAES.1SG
    televizor,   no   ne   sta-l   ob   èt-om
    television-ACC.SG but   NEG   begin-PAST.MASC.SG   about   this-LOC.SG
    rasprostranja-t'-sja.
    extend-INF-REFL
    'I don't read newspapers and don't watch TV, but I decided not to talk about it.' [Mariam Petrosjan. Dom, v kotorom... (2009)]

(4) **Čita-jte**   **gazet-y**:   mož-et   obnaruž-it-sja   ves'ma
    read-IMP.PL   newspaper-ACC.PL   can-PRAES.3SG   appear-PRAES.3SG-REFL   very
    polezn-aja   informaci-ja.
    useful-FEM.NOM.SG   information-NOM.SG
    'Read the newspapers: perhaps useful information will appear.' [Čto nas ždet v avguste (2003) // «Kriminal'naja xronika», 2003.06.24]

(5) Ja   ni   raz-u   ne   **vide-l**   **Gal-ju**   s
    I   NEG   time-PART.SG   NEG   see-PAST.MASC.SG Galya-ACC.SG   with
    podrug-ami.
    friend-INS.PL

---

[3] For the purposes of the present study we use the term "register" in a relatively loose sense as a language variety defined according to how it is used in social situations (see Hudson 1980: 48). In particular, we are interested in registers reflecting different levels of formality,

'I have never seen Galya with friends'. [Vasilij Aksenov. Zvezdnyj bilet // «Junost'», 1961]

Krasovitsky et al. (2011) investigate case assignment of direct object of Russian negated verbs in a diachronic corpus study. In earlier periods only genitive case was possible for direct objects of verbs under negation, but from the beginning of the seventeenth century isolated instances of accusative direct objects appear. By the second half of the twentieth century the amount of accusative objects reaches 49% of all direct objects of negated verbs. Krasovitsky et al. (2011) test the effect of referentiality, individuation and definiteness using corpus data. Comparing abstract and concrete direct objects they show that in the second half of the twentieth century 59% of concrete direct objects of negated verbs appear in the accusative, while only 25% of abstract direct objects under negation have accusative case.

The partitive construction appears with perfective verbs when an indefinite amount of a homogeneous substance is described as in (6). When the object is specified, the accusative case is preferred, see (7).

(6) My   obmy-li   nov-uju   poem-u,   to   est'
     we   bathe-PAST.PL new-FEM.ACC.SG poem-ACC.SG   this   be.PRAES.3SG
     vypi-li   **vodk-i**   i   zakusi-li   kopčen-oj   ryb-oj.
     drink-PAST.PL   vodka-GEN.SG   and   eat.with-PAST.PL   smoked-INS.SG   fish-INS.SG
     'We have celebrated the new poem, that is we drank some vodka and ate
     some smoked fish.' [V. P. Kataev. Almaznyj moj venec (1975-1977)]

(7) Lapov.∅   prinja-l   stakan   i   medlenno,   ne
     Lapov-NOM.SG   take-PAST.MASC.SG   glass-ACC.SG   and   slowly   NEG
     pomoršči-vši-s',   **vypi-l**   **vodk-u**.
     wince-GERUND-REFL   drink-PAST.MASC.SG   vodka-ACC.SG
     'Lapov took the glass and slowly, without wincing, drank the vodka.' [Grigorij
     Saburov. Peškom po volnam // «Zvezda», 2002]

Nonreferential objects of weak intensional verbs (cf. Kagan 2013) are marked with the genitive (see (8)), while referential objects of such verbs are marked with the accusative (see (9)).

(8) **Žda-l**   **avtobus-a**   pod   fonar-em,
     wait-PAST.MASC.SG   autobus-GEN.SG   under   street.lamp-INS.SG
     raskry-l   knig-u —   na   stranic-y   sta-l
     open-PAST.MASC.SG   book-ACC.SG   on   page-ACC.PL   begin-PAST.MASC.SG
     pada-t'   sneg.∅.
     fall-INF   snow-NOM.SG
     [He] was wating for a bus under the street lamp, [he] opened a book – the
     snow started to fall on the pages. [Mixail Šiškin. Venerin volos (2004) //
     «Znamja», 2005]

(9) Na   drug-oj   storon-e   dorog-i   t[o]   e[st']
     on   other-FEM.LOC.SG   side-LOC.SG   road-GEN.SG   this   be.PRAES.3SG

naprotiv	tak	nazyvaem-oj	gostinic-y	ja	**žd-u**
across	so	called-FEM.GEN.SG	hotel-GEN.SG	I	wait-PRAES.1SG

avtobus.∅   nomer.∅   22…
autobus-ACC.SG   number-NOM.SG   twenty-two

'On the other side of the road, across the street from the so-called hotel, I am waiting for the bus number twenty-two… [Pis'mo sestre (1982)]

We see that in all three cases above, choice between genitive and accusative objects depends on the degree of individuation, understood as the conventional likelihood of viewing something as an individual (cf. Timberlake 1985). In view of this, it would be logical to assume that individuation also plays a role in the distribution of the genitive and accusative with *bojat'sja*. Indeed, a number of researchers mention individuation as a factor that increases the probability of the accusative case with *bojat'sja* (Comrie, Stone and Polinsky 1996: 145, Ljustrova et al. 1982: 95, Maier 2010: 144, Nichols 1993: 82, Timberlake 2004: 319). Another frequently mentioned factor is register: substandard or colloquial speech is often described as a factor favoring the accusative (Gorbačevič 1971: 237, Janko-Trinickaja 1962: 60, Vinogradov 1947: 623, 1986: 505, Comrie, Stone and Polinsky 1996: 145, Ickovič 1982: 36, Ljustrova et al. 1982: 95f., Rozental' and Telenkova 1984: 54, and Timberlake 2004: 319).

Although the factors of individuation and register are frequently mentioned in the scholarly literature, we are not aware of any large-scale empirical studies based on corpus data or data from psycholinguistic experiments. Our goal is to carry out such an investigation in the present article. The hypotheses we test can be made explicit as follows:

(10)   The Individuation Hypothesis: A high degree of individuation favors the use of the accusative in the object of *bojat'sja*.

(11)   The Register Hypothesis: Less restricted registers favor the use of the accusative in the object of *bojat'sja*.

In our study we operationalize individuation in terms of three parameters: animacy, use of proper names vs. common nouns and word order. We expect that objects with *bojat'sja* would show the following distribution: animate nouns would be more likely to appear in the accusative than inanimate nouns, proper names would be more likely to appear in the accusative than common nouns, and preverbal objects would be more likely to appear in the accusative than postverbal objects. It is clear why animacy and proper names are related to individuation, but the relationship between word order and individuation might need more explanation. This relationship occurs because in Russian word order is associated with information structure. While the Verb-Object (VO) word order is neutral, the Object-Verb (OV) word order marks the object as thematic/topicalized, i.e. the object represents given information. Since an object representing given information (OV word order) is more likely to be viewed as an individual than a new object (VO word order), we expect more examples with the accusative case in sentences with OV word order.

In order to test the Register Hypothesis, we investigate two corpora within the RNC: the main corpus and the newspaper corpus. The former corpus mainly includes literary texts, and the majority of the texts are from the second half of the twentieth century and the first decade of the twenty-first century. The latter corpus only contains newspaper articles written between 2000 and 2011. Although there is no one-to-one relationship between corpus and register, a number of socio-linguists have commented on the rapid and radical changes in the language of the press in the post Soviet period, whereby elements of colloquial and substandard language became pervasive in the press. For instance, Wade and Ryazanova-Clarke (1999) mention the "free use of obscenities" and "and the coarse kind of language" used by journalists:

(12)   Journalists from some newspapers criticised each other's Russian, at the same time making free use of obscenities [...]. Journalists, especially in papers aimed at young people, cultivated a coarse kind of language designed to 'win hearts and minds by its realism' 'expand the circle of readers' [...]. (Wade and Ryazanova-Clarke 1999: 308)

In a similar vein, Kakorina (1996) discusses the frequent use of elements from substandard (*prostorečie*) and jargon in the press:

(13)   Nekotorye issledovateli sklonny konstatirovat' novyj vitok vul'garizacii literaturnogo jazyka, i dlja ètogo imejutsja vse osnovanija. Prostorečnye i žargonnye èlementy svobodno vključajutsja v gazetnye teksty, stanovjatsja privyčnymi sredstvami publičnogo obščenija. (Kakorina 1996: 80)

Gorham (2014) summarizes the development as follows:

(14)   The new market for freer, more open discourse also inevitably meant more spontaneous language production that tended to be more colloquial in style, a trend exacerbated by the influx of a new generation of journalists and TV personalities lacking traditional Soviet training. On one level, this meant a general "coarsening" of the language of the print and electronic media (Gorham 2015: 80).

As an illustration of how grammatical elements from colloquial Russian make their way to newspaper prose, consider the rivalry between the zero genitive plural form (-Ø) and *-ov* with words of measurement such as *gramm* and *kilogramm.* Glovinskaja [1996: 240] notes that in the end of the 20th century the zero form experienced a revival, and started expanding in colloquial Russian. In order to find out if this change was reflected in the Russian National Corpus, we explored quantitative constructions that contain a numeral together with the measurement words *gramm* and *kilogramm* in the newspaper subcorpus of the RNC and compared the more standard construction with the *-ov* ending (e.g. *sto kilogrammov* '100 kilograms') with the corresponding construction with zero genitive plural form (e.g. *sto kilogramm*). Our results indicate a decrease of the first construction and an

increase of the second in the first decade of the 21ˢᵗ century. While in year 2001 only 2% (5 out of 208 examples) of the two quantitative constructions with *gramm* and *kilogramm* used zero genitive plural form, already in 2010 these constructions constitute 15% (115 out of 651 examples) of all uses of the two quantitative constructions[4]. Thus, we see that journalists pick up and actively use innovative grammatical constructions and these processes can be observed using the newspaper subcorpus of the RNC.

A detailed study of the language of the post-Soviet press is well beyond the scope of the present study. However, if we take claims like the ones in (12), (13) and (14) seriously, as well as our small-scale corpus investigation of measurement constructions, the Register Hypothesis predicts a high proportion of accusative objects with *bojat'sja* in the Newspaper corpus compared to texts from the same period in the main corpus.

3. Corpus study

3.1.　Methodology

The contrast between the accusative and genitive cases in Russian is noticeable in the following nouns in the singular: first-declension neuter nouns (*okno* 'window'), first-declension inanimate nouns (*stol* 'table'), second-declension nouns (*mama* 'mom' and *papa* 'dad'), and third-declension nouns (*myš'* 'mouse'). Due to syncretism of the Russian declension system, animate nouns of the first declension have the same forms in the accusative and genitive cases (*otc-a* 'father_ACC' or 'father_GEN'). Thus, this last class of nouns is excluded from our investigation. Nichols (1993: 82) claims that "[t]he accusative with these verbs [*bojat'sja* and *slušat'sja*] is possible only with second-declension nouns[…]". As opposed to that, Israeli (1997: 44) and Krys'ko (1997: 243) argue that the accusative is also possible in third and first declensions. The results of our study support Israeli's and Krys'ko's claims: both third-declension nouns like *myš'* 'mouse' (15) and inanimate first-declension nouns like *narod* 'nation' (16) appear in the accusative with *bojat'sja*.

(15)　— Razve　　možno　　**boja-t'-sja**　　**myš'**-Ø? —
　　　really　　　possible　　be.scared-INF-REFL　　mouse-ACC.SG
　　　udivi-l-sja　　　　　　Birjukov-Ø.
　　　be.surprized-PAST.MASC.SG-REFL　　Birjukov-NOM.SG
　　　'"Is it possible to be afraid of a mouse?" – Birjukov remarked in surprise.'
　　　[Jurij Petkevič. Živye cvety zimoj (2001)]

(16)　Naš-i　　　pravitel-i,　　poxože,　　ne　　tol'ko　nenavid-jat,　no
　　　our-NOM.PL　　leader-NOM.PL　probably　　NEG　　only　　hate-PRAES.3PL　but
　　　i　　　paničeski　　**boja-t-sja**　　**sv-oj**　　　**narod**
　　　and　　anxiously　　be.scared-PRAES.3SG-REFL　own-MASC.ACC.SG　nation-ACC.SG

---

[4] Pearson's Chi-squared test: X-squared = 29.29, df = 1, p-value = 6.225e-08. Cramer's V-value = 0.2 (small to moderate and robust effect size).

'It seems that our leaders don't just hate, but dread in panic their own people.' (Pravda 1993, 12.5.)

However, although accusative objects are attested for first and third declension nouns, examples are too few for statistical analysis. For instance, we found only two examples with accusative objects in the third declension (one example in the main corpus and one in the newspaper corpus). For the purposes of the corpus study, we therefore decided to focus on the second declension. We will return to the third declension in our experimental study in section 4.

The corpus study of accusative and genitive objects with *bojat'sja* presented us with a methodological problem: the two phenomena we are interested in are unequally distributed in the RNC. There are few examples with the accusative case and an abundance of examples with the genitive case. If we restricted our search to a small sample of the corpus, we would not have enough accusative examples for analysis. Ferm (2004) investigated a sample of 300 examples with *bojat'sja* and found only three examples with the accusative. Such a small number does not enable the researcher to find out much about the accusative construction apart from the fact that it exists and is relatively low frequent in the corpus. However, if we chose to study all examples of the verb *bojat'sja*, we would "drown" in thousands of examples with genitive objects, and it would be infeasible to find the needle (the accusative) in the haystack of examples with the genitive. In order to solve the methodological problem, we have chosen to combine two methods.

For the accusative case, we investigated all available examples in both corpora and weeded out the noise manually. Four types of objects were searched for: 1) second declension feminine 2) second declension masculine 3) third declension 4) first declension masculine inanimate. However, as mentioned above, only the second declension returned enough examples in the accusative to make statistical analysis possible.

For the genitive, we collected random samples of 300 examples for the same four types of objects. The noise was weeded out manually within those samples, and then we extrapolated the amount of relevant examples in the corpus, based on the number of relevant examples in the samples. For example, in the main corpus we found 2644 attestations of *bojat'sja* with second declension feminine nouns like *žena* 'wife' in the genitive. We randomly chose a sample of 300 examples and tagged those examples. Among those 300 examples we found twenty-seven examples that were not relevant to our search and had to be excluded. The remaining 273 examples were distributed as shown in the third column of Table 1: we found twenty-one animate objects, 252 inanimate objects, seven proper names, 266 common nouns, eight examples of OV word order and 265 examples of VO word order. We expect that 2644 attestations of second declension feminine nouns follow the same distribution. Thus, we expect that we will find approximately 238 ((27/300)*2644) non-relevant examples among those 2644 attestations. We further expect to find that the remaining 2406 will be distributed between animate and inanimate objects as follows: we will find approximately 185 ((21/300)*2644) examples with animate objects and approximately 2221 examples with inanimate objects. The remaining approximations for the second declension feminine nouns

are shown in the fourth column of Table 1, which therefore summarizes our approximations for genitive objects in the second declension.

Parameter	Value	Sample	Approximation
Animacy	Animate	21	185
	Inanimate	252	2221
Proper vs. common	Proper names	7	62
	Common nouns	266	2344
Word order	OV	8	71
	VO	265	2336

Table 1: Distribution among relevant parameters in random sample and approximation for the main corpus: second declension feminine nouns.

We made such extrapolations separately for each type of objects (first declension masculine, second declension masculine, second declension feminine, third declension) and then summarized the results. Thus, at the end of this procedure we received a table similar to Table 1, but with approximations for all relevant declensions. These numbers were then compared to the numbers of examples with accusative objects. The obtained results are shown and discussed in the next section.

3.2. Results

Table 2 presents overall results: we found the accusative in 1% of all examples in the main corpus and in 7% of examples in the newspaper corpus. This difference is statistically significant [5]. This distribution supports the Register Hypothesis advanced in section 2. As we mentioned above, newspaper articles of the 21st century are less restricted in terms of register, and we therefore expect more uses of the accusative with *bojat'sja* in the newspaper corpus. The paragraphs below explain in more detail how the main corpus and the newspaper corpus differ in terms of semantic parameters related to individuation.

	#ACC	#GEN	#ACC+GEN	%ACC
Main	30	3437	3467	1
Newspaper	53	757	810	7

Table 2: Accusative/genitive variation for the verb *bojat'sja* in two corpora.

Let us first consider the difference between animate and inanimate nouns. As shown in Figure 1 and Table 3, in both corpora the accusative case is more frequent among animate nouns, which supports the Individuation Hypothesis. [6] However, the

---

[5] Pearson's Chi-squared test: X-squared = 108.27, df = 1, p-value < 2.2e-16. Cramer's V-value = 0.2 (small to moderate and robust effect size).

[6] Main corpus: Pearson's Chi-squared test with Yates' continuity correction: X-squared = 231.82, df = 1, p-value < 2.2e-16. Cramer's V-value = 0.3 (moderate effect size). Newspaper corpus: Pearson's Chi-squared test with Yates' continuity correction: X-squared = 386.16, df = 1, p-value < 2.2e-16. Cramer's V-value = 0.7 (large effect size).

distribution in the two corpora is different. While in the main corpus 9% of animate objects appear in the accusative, in the newspaper corpus 90% of animate objects are found in the accusative. Similarly, while almost no examples of inanimate objects in the accusative are attested in the main corpus, 3% of inanimate objects in the newspaper corpus are in the accusative. In short, animacy is a factor favoring the accusative, and in the newspaper corpus this factor plays a much more important role than in the main corpus.
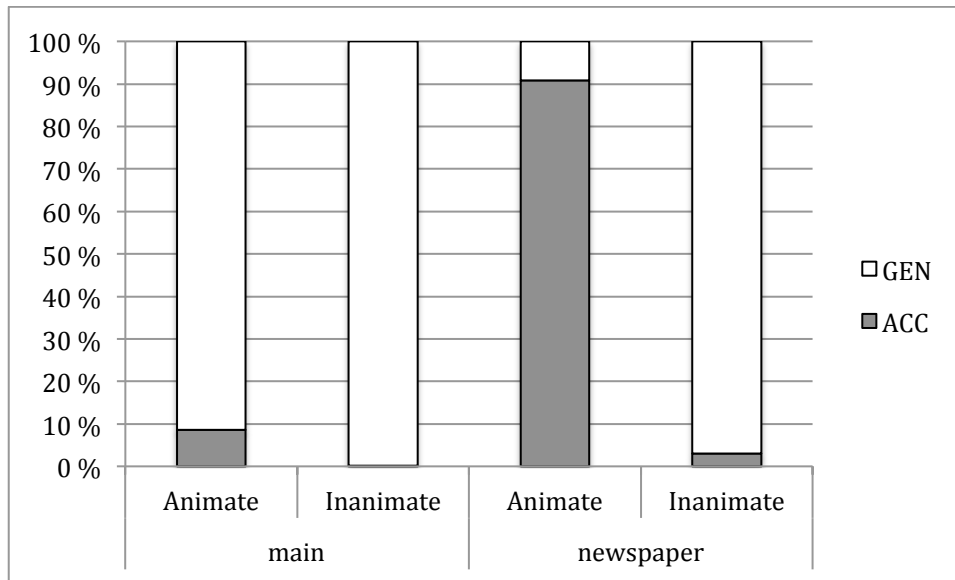
Figure 1: Animacy – comparison of attestations of accusative and genitive in the RNC.

		#ACC	#GEN	#ACC+GEN	%ACC
Main	Animate	27	286	313	9
	Inanimate	3	3152	3155	0
Newspaper	Animate	30	3	33	91
	Inanimate	23	754	777	3

Table 3: Animacy – comparison of attestations of accusative and genitive in the RNC.

Word order also plays an important role in the distribution of the two cases, as shown in Figure 2 and Table 4. Both corpora show statistically significant differences between preverbal and postverbal objects.[7] As mentioned in section 2, the preverbal (topicalized or thematic) position is characteristic for more individuated objects, and we therefore expect such objects to appear in the accusative case more frequently. While data from both corpora confirms this, the effect is more prominent in the newspaper corpus: here the accusative is 4.7 times more likely to appear in preverbal position as opposed to postverbal position. For

---

[7] Main corpus: Pearson's Chi-squared test with Yates' continuity correction: X-squared = 6.48, df = 1, p-value = 0.01. Cramer's V-value = 0.04 (non-reportable effect size). Newspaper corpus: Pearson's Chi-squared test with Yates' continuity correction: X-squared = 28.67, df = 1, p-value < 0.001. Cramer's V-value = 0.2 (small to moderate and reportable effect size).

the main corpus, the accusative is only 3 times more likely to appear in preverbal position.
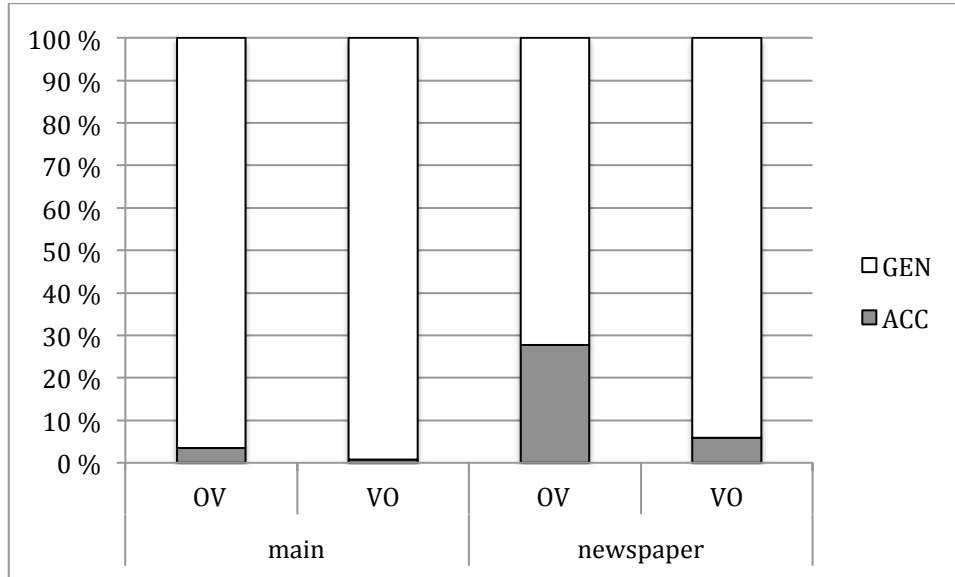

Figure 2: Word order – comparison of attestations of accusative and genitive in the RNC.

		#ACC	#GEN	#ACC+GEN	%ACC
Main	OV	4	112	116	3
	VO	26	3326	3352	1
Newspaper	OV	13	34	47	28
	VO	44	710	754	6

Table 4: Word order – comparison of attestations of accusative and genitive in the RNC.

Proper names usually refer to highly individuated objects. We therefore expect a more frequent use of the accusative for such objects. Again, data from both corpora confirms this: percent of the accusative case is higher for proper names as opposed to common nouns.[8] In the main corpus we see almost no uses of the accusative with common nouns and 15% of accusatives among proper names. In the newspaper corpus the difference is even more striking (4% accusative for common nouns as opposed to 44% accusative for proper names).

---

[8] Main corpus: Pearson's Chi-squared test with Yates' continuity correction: X-squared = 336.27, df = 1, p-value < 2.2e-16. Cramer's V-value = 0.3 (moderate effect size). Newspaper corpus: Pearson's Chi-squared test with Yates' continuity correction: X-squared = 145.45, df = 1, p-value < 2.2e-16. Cramer's V-value = 0.4 (moderate effect size).
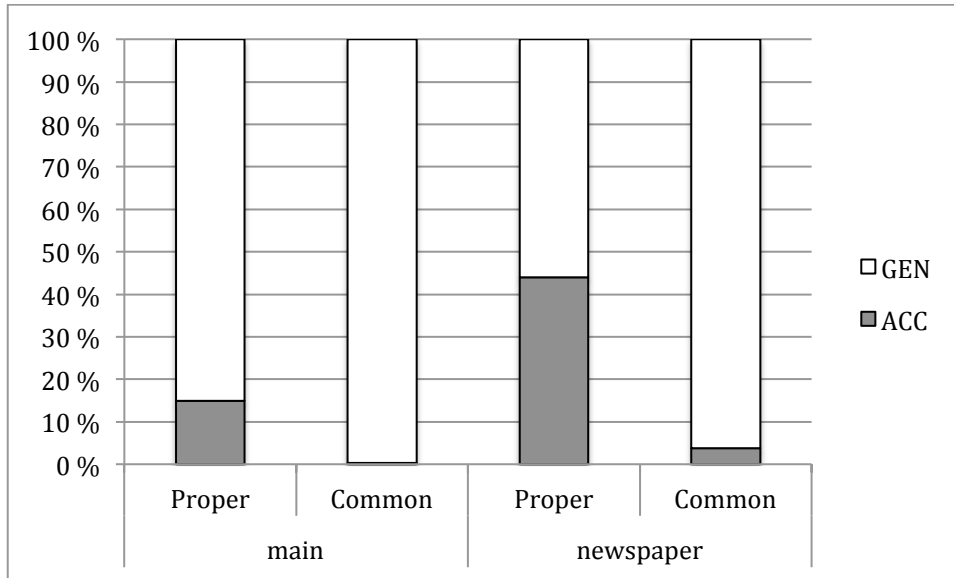
Figure 3: Proper vs. common – comparison of attestations of accusative and genitive in the RNC..

		#ACC	#GEN	#ACC+GEN	%ACC
Main	Proper	22	126	148	15
	Common	8	3310	3318	0
Newspaper	Proper	29	37	66	44
	Common	28	725	753	4

Table 5: Proper vs. common – comparison of attestations of accusative and genitive in the RNC.

To sum up, all three parameters related to individuation – animacy, word order and proper names vs. common nouns – play an important role in the distribution of accusative vs. genitive case with *bojat'sja*. In both corpora more individuated objects are more likely to appear in the accusative. This supports the Individuation Hypothesis proposed in the literature and discussed in section 2.

In addition, the two corpora differ in the distribution of the accusative and the genitive, insofar as the accusative construction occurs in the newspaper corpus more frequently. Two corpora differ with regard to time periods, and in order to eliminate this difference we conducted an additional search and compared data from main corpus and newspaper corpus from year 2000 onwards. We restricted this search to animate nouns, since we have shown that these nouns are more likely to undergo a shift from genitive to accusative. We also excluded examples with negation in order to eliminate this confounding factor. Texts of the main corpus created since year 2000 contain twenty-eight examples of the verb *bojat'sja* with animate nouns, ten of which appear in accusative case (36%), while newspaper texts produced during the same period of time contain thirty-three examples with this verb, thirty of which are in accusative (91%)[9]. Thus, difference between main and newspaper corpus with regard to the case of an object of the verb *bojat'sja* is

[9] Pearson's Chi-squared test with Yates' continuity correction: X-squared = 18.06, df = 1, p-value < 2.129e-05. Cramer's V-value = 0.54 (large effect size).

12

due to register, and not diachronic change. The accusative construction is characteristic of substandard and colloquial speech, which as mentioned in section 2 has become pervasive in the press in the post Soviet era. In this way, our findings appear to corroborate the Register Hypothesis, which was proposed in the literature and discussed in section 2 of the present article. We hasten to add that the relationship between corpus and register is complex, and a more detailed study would be required in order to support stronger conclusions. However, such a study is beyond the scope of the present article.

Since some combinations of the parameters are never attested in the corpus and some are attested in only a few examples, we cannot compare the impact of each parameter. In order to include the infrequent combinations of the relevant parameters, we need to investigate accusative and genitive objects via an experimental study with balanced conditions. This will allow us to weigh the relative importance of each parameter.

4. Experiment

4.1. Methodology

Based on the corpus data analyzed in the previous section we decided to focus on three parameters for the experimental study: animacy, word order, proper vs. common noun. In addition to these three parameters relating to individuation, we decided to investigate whether the declension a noun belongs to has an impact on the choice between accusative and genitive objects. For this purpose, we compare nouns of the second and third declensions, which are the declensions where accusative objects are most likely to occur. In our experiment we have twelve different conditions. For the second declension we have eight conditions that represent every possible combination of animacy, word order and proper vs. common nouns. For each of these eight conditions we included two examples in our questionnaire. However, for the third declension we found that both animate proper names and animate common nouns are infrequent. There is only one animate proper name, *Ljubov'*, that is perceived as neutral by native speakers of Russian. All other proper names suggested by dictionaries (e.g. *Ninel', Ruf'*) are infrequent and are perceived as foreign or unfamiliar. Since we did not want to add frequency and familiarity of a proper name as an additional parameter we chose to decrease the number of targets in the third declension. As a result the third declension is only represented by four conditions. We chose to include only conditions with the more neutral VO word order. We have also reduced the number of examples per condition: only one example for each of the third-declension conditions was used. Thus, the questionnaire for the experiment consists of 20 questions and their distribution among the parameters is shown in Table 6. The full questionnaire[10] also

---

[10] The questionnaire for this study is available online at the Tromsø Repository of Language and Linguistics (TROLLing: http://opendata.uit.no/).

contained 40 fillers with other Russian verbs with non-standard argument structure, i.e. with objects in the dative, genitive and instrumental.

Second declension (–a)

	VO			OV	
	Proper name	Common noun		Proper name	Common noun
Animate	2	2	Animate	2	2
Inanimate	2	2	Inanimate	2	2

Third declension (–ь)

				OV	
				Proper name	Common noun
			Animate	1	1
			Inanimate	1	1

Table 6: Distribution of examples per condition in the questionnaire.

The survey was administered on-line. All participants in the experiment are non-linguists, since it has been shown that the linguist's judgments differ significantly from those who do not have linguistic background, see Dąbrowska 2010. In total, 409 speakers of Russian participated in the experiment. Of the participants 68% are female, and 32% are male. This gender imbalance might have affected the final results, since we found that women are more innovative in the use of the accusative case with *bojat'sja* than men. The average year of birth is 1976, so at the time of the experiment the average age of the participant is 37 years. In our sample there are many more participants with college education than there would be if the sample were balanced for education (80% with college education, 8% with only high school education, 11% with unfinished college education). This might have affected the results, since the data shows that less educated speakers are more prone to innovation. However, despite the imbalance with regard to gender and education the experiment yielded robust results, as we will see in sections 4.2 and 4.3.

4.2.   The impact of the parameters in the experiment

The experimental data supports the individuation hypothesis as concerns animacy: the object is three times more likely to be marked with accusative case when the object is animate, as shown in Figure 4 and Table 7.[11]

---

[11] Pearson's Chi-squared test with Yates' continuity correction: X-squared = 890.72, df = 1, p-value < 2.2e-16. Cramer's V-value = 0.33 (moderate effect size).
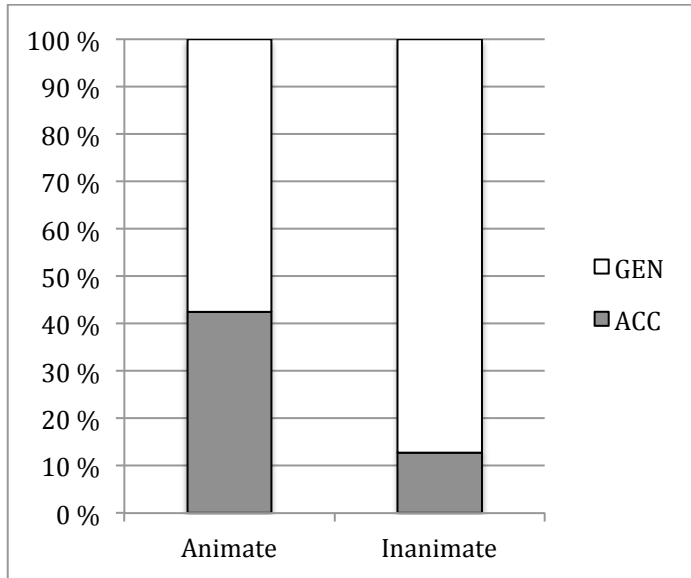
Figure 4: Animacy – comparison of attestations of accusative and genitive in the experiment.

	#ACC	#GEN	%ACC
Animate	1726	2336	42
Inanimate	518	3527	13

Table 7: Animacy – comparison of attestations of accusative and genitive in the experiment.

Note here that in 13% of all our examples inanimate nouns are marked with the accusative. This number is higher than the corresponding numbers for inanimate nouns in both corpora we have studied (0.001% accusatives in the main corpus and 3% accusatives in the newspaper corpus). The answers to our questionnaire represent less standardized speech than texts in written corpora, either literary or newspaper. This difference suggests that the use of accusative case with *bojat'sja* is more prominent in spoken language than in written texts, which is in harmony with the Register Hypothesis discussed in section 2.

As can be seen from Figure 5 and Table 8, the experimental data also supports the Individuation Hypothesis with regard to word order: objects preceding the verb are two times more likely to be in the accusative than objects following the verb (40% for OV as opposed to 20% for VO).[12]

---

[12] Pearson's Chi-squared test with Yates' continuity correction: X-squared = 377.04, df = 1, p-value < 2.2e-16. Cramer's V = 0.2 (small to moderate effect size).
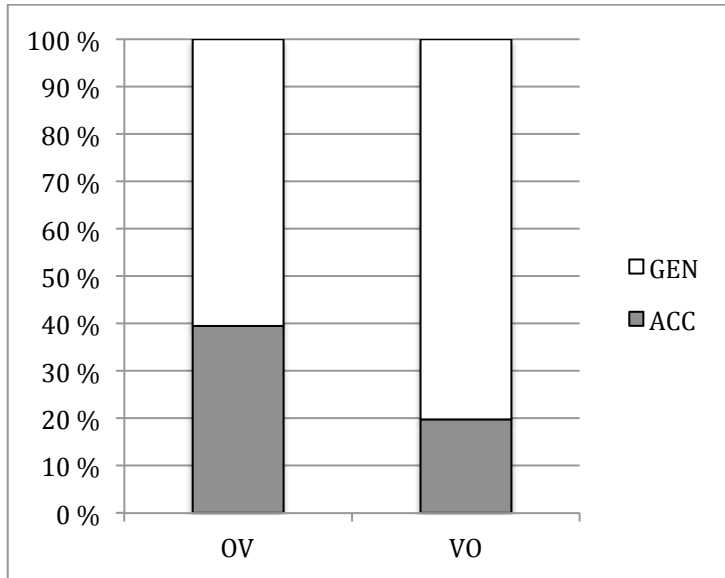
15

Figure 5: OV vs. VO word order – comparison of attestations of accusative and genitive in the experiment.

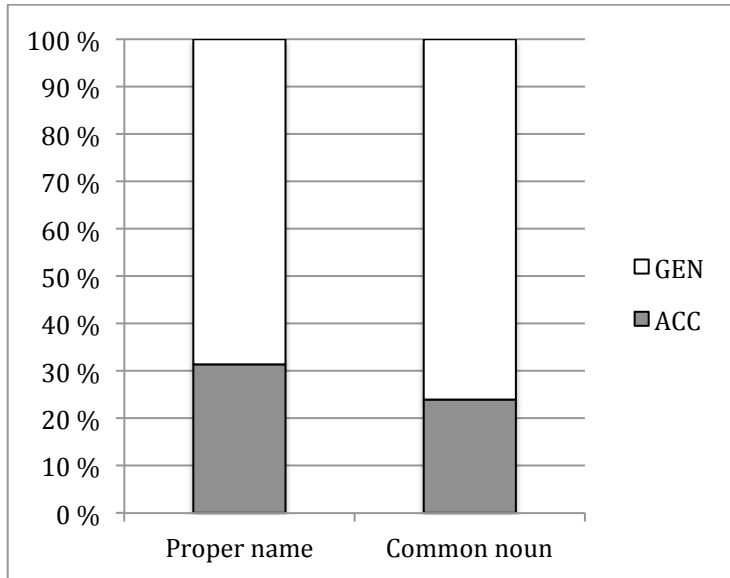|	#ACC	#GEN	%ACC
OV	1279	1956	40
VO	965	3907	20

Table 8: OV vs. VO word order – comparison of attestations of accusative and genitive in the experiment.

Similar to the distribution in the corpus, the experiment shows a significant difference between the behavior of proper names and common nouns.[13] We see that the presence of a proper name triggers the use of the accusative in 31% of all responses, while common nouns yield the accusative case in only 24% of all responses. However, unlike the corpus results, the experiment results show that the effect size of this difference is Cramer's V = 0.08, which does not indicate a reportable difference. In the corpus data set, proper names are often used for animate subjects and in sentences with OV word order. Both these features favor the accusative. In the experimental data set we are able to control the use of different parameters and thus to pinpoint how each of the parameters affects the choice of the case. It seems likely that the difference between proper and common nouns may be explained by their correlation with other important parameters.

_____

[13] Pearson's Chi-squared test with Yates' continuity correction: X-squared = 54.78, df = 1, p-value = 1.346e-13. Cramer's V = 0.08 (non-reportable effect size).

Figure 6: Proper name vs. common noun – comparison of attestations of accusative and genitive in the experiment.

	#ACC	#GEN	%ACC
Proper name	1272	2783	31
Common noun	972	3080	24

Table 9: Proper name vs. common noun – comparison of attestations of accusative and genitive in the experiment.

In the corpus data set we found only two examples of third declension nouns in the accusative case (one in the main corpus and one in the newspaper corpus). There are two conceivable explanations for this pattern. It is possible that the accusative is less likely to be used with third declension nouns, or it might be the case that this effect is just a conspiracy between two other preferences: (1) The accusative case is more likely to be used with animate nouns; (2) There are few animate nouns in the third declension. In the experiment, one of our goals was to check whether the third declension plays an independent role in the distribution of the genitive vs. accusative case with *bojat'sja*, or whether the third declension effect is just a side effect of accusative's preference for animacy. We included four sentences with third declension nouns in the questionnaire. These sentences represented animate and inanimate proper and common nouns, and all sentences represented the more neutral VO word order. For this reason, it is appropriate to compare them to the second declension examples with the same word order.

Figure 7 and Table 10 show that nouns in the second declension are seven times more likely to appear in the accusative (28% as opposed to 4%).[14] This indicates that declension serves as a distributional parameter in its own right in addition to the individuation parameters discussed above. Native speakers of Russian are less likely to use the accusative case for third declension nouns with the verb *bojat'sja*, and this does not seem to be a mere side effect of the lack of animate

---

[14] Pearson's Chi-squared test with Yates' continuity correction: X-squared = 381.5876, df = 1, p-value < 2.2e-16. Cramer's V = 0.28 (moderate effect size).

nouns in the third declension. In the experiment, we were able to keep other relevant factors constant (animacy, word order, proper name vs. common noun), and therefore clarify the impact of the declension as such. For proper names, the participants in our experiment used the accusative in 52% of all responses for the second declension names *Ol'ga* and *Nina*, whereas for the third declension name *Ljubov'* the corresponding number was 8%. For common nouns, participants used accusative in 43% of all responses for the second declension kinship terms *mama* 'mom' and *babuška* 'grandmother' as opposed to only 7% of accusatives for the third declension kinship term *svekrov'* 'mother-in-law (husband's mother)'.

This result suggests that native speakers are sensitive to the differences between declension classes. When a pattern spreads to animate nouns in one declension (in our case the second), the pattern does not necessarily extend to animate nouns in another declension (in our case the third), even if both declensions are dominated by nouns of the same gender (in our case feminine).
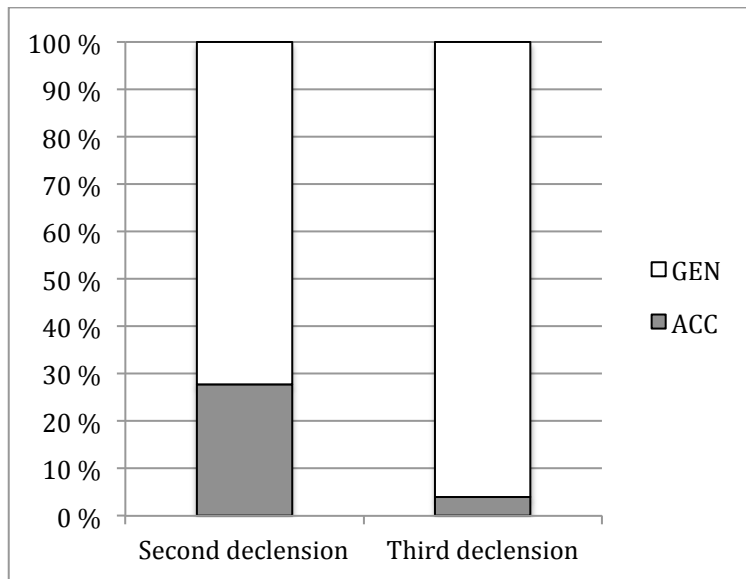

Figure 7: Declension – comparison of attestations of accusative and genitive in the sentences with VO word order

	#ACC	#GEN	%ACC
Declension II	900	2348	28
Declension III	65	1559	4

Table 10: Declension – comparison of attestations of accusative and genitive in the sentences with VO word order.

4.3.   Interaction of variables in the experiment

The interaction of the four variables animacy, word order, proper vs. common noun and declension can be seen in Table 11 below. Note that second declension animate proper nouns in OV word order have the highest percent of accusative case (57%) and are more likely to appear in the accusative than in the genitive, while third declension inanimate common nouns in VO word order have the lowest percent of accusative case (0%), and are always marked with the genitive.

18

Second declension (–a)					
VO			OV		
	Proper name	Common noun		Proper name	Common noun
Animate	57	53	Animate	52	43
Inanimate	29	18	Inanimate	14	2

Third declension (–ь)					
			OV		
				Proper name	Common noun
			Animate	8	7
			Inanimate	2	0

Table 11: Interaction of animacy, proper vs. common nouns, word order, declension – percent of accusatives among the attestations in the experiment

We have employed the random forests method to rank the importance of the four variables in our experimental study. Random forests (proposed by Leo Breiman and Adele Cutler, see Breiman 2001) are a method for classification, which constructs multiple decision trees. One of the important features of this algorithm is that it can estimate the relative importance of the input variables: the more trees in the forest indicate a particular variable as important for the classification, the larger is the importance of this variable overall. In figure 8 we can see that animacy, appearing on the left, is ranked highest. This means that animacy has the largest impact on the choice of case for the object of *bojat'sja*. Declension is ranked next, which again shows that declension serves as a separate parameter in the choice of the case: declension effects cannot be explained by animacy. Word order is ranked next, and the proper vs. common nouns parameter is ranked last. This ranking supports the hypothesis that was offered in the previous section, viz. that the effect of proper nouns is less evident once we fix all other parameters. Thus, the effect of proper nouns in our corpus data is most likely affected by the concurrence of proper nouns with animate nouns and OV word order. However, note that the importance of proper vs. common parameter is ranked above 0 and is similar to the importance of word order; therefore this parameter is relevant for the case distribution for the object of the verb *bojat'sja*.
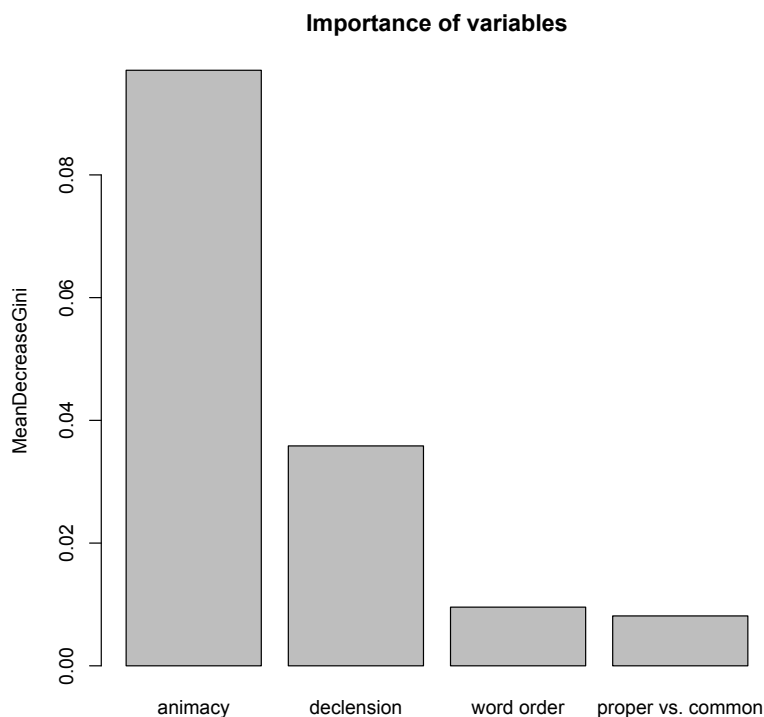
Importance of variables

Figure 8: Importance of variables according to the random forest method

5. Conclusions

In our corpus and experimental studies we have investigated the case of the object of the Russian verb *bojat'sa*. Originally *bojat'sja* was used exclusively with genitive case objects, but in Contemporary Standard Russian the accusative is also used.

Both our corpus study and our experimental study revealed that the accusative appears more frequently with individuated objects: three parameters related to individuation – animacy, word order and proper names vs. common nouns – significantly affected the distribution of cases with *bojat'sja*. The corpus study also revealed a significant difference between the main corpus and the newspaper corpus. We suggest that this difference is due to the fact that the accusative is more likely to appear in less restrictive registers, which have become pervasive in the press in the post Soviet period.

The controlled setting of the experiment allowed us to study combinations of parameters that are underrepresented in the corpus. The experiment results show that (1) the parameter of animacy has most impact on the distribution; (2) the behavior of third declension nouns differs significantly from the second declension nouns; (3) the difference between proper and common nouns in the corpus is most likely the result of other correlated parameters.

The fact that the accusative is more used with *bojat'sja* in less restricted registers may suggest that accusative objects will become more widespread in the future, since such registers are likely to indicate the direction of ongoing change.

However, more data is needed in order to make substantial predictions about the use of accusative objects with *bojat'sja* in the future.

Both object of the reflexive verb *bojat'sja* in our study and object of the transitive verb under negation in Krasovitsky et al. (2011) study demonstrate similar patterns. Genitive case, which is less commonly used to mark an object, is replaced with a default case for direct objects – accusative. In opposition between genitive and accusative cases, genitive is more likely to be used for non-individuated nouns, while accusative demonstrates preference for highly individuated nouns. Both diachronic shifts are governed by the same tendencies. Nouns that are more compatible with preferences of the accusative case (concrete nouns, proper names, animate nouns) are more likely to undergo shift first, while nouns that are more compatible with genitive case preferences (abstract nouns, common nouns, inanimate nouns) linger and are less likely to undergo the shift. Thus, this case study together with Krasovitsky et al.'s corpus data illustrates how semantics governs morphological change.

References

Babby, Leonard H. (1980). *Existential sentences and negation in Russian* (Linguistica Extranea Studia, 8). Ann Arbor: Karoma.

Breiman, Leo (2001). Random Forests. *Machine Learning* 45 (1): 5–32.

Butorin, D. I. (1966): Ob osobyx slučajax upotreblenija vinitel'nogo prjamogo ob''ekta v sovremennom russkom literaturnom jazyke. In G. A. Kačevskaja and K. S. Gorbačevič (eds.): *Normy sovremennogo russkogo literaturnogo slovoupotreblenija*. Moscow and Leningrad: Nauka.

Černyšev, Vasilij I. (1911): *Pravil'nost' i čistota russkoj reči: opyt russkoj stilističeskoj grammatiki*. St. Petersburg: Imperatorskaja akademija nauk.

Comrie, Bernard, Gerald Stone and Maria Polinsky (1996): *The Russian language in the twentieth century*. Oxford: Clarendon Press.

Dąbrowska, Ewa. (2010). Naive v. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27 (2010), 1–23.

Evgen'eva, Anastasija P. (ed.) (1999): *Malyj Akademičeskij Slovar'.* Moscow: Russkij jazyk.

Ferm, Ludmila (2005): *Variativnoe bespredložnoe glagol'noe upravlenie v russkom jazyke XVIII veka*. Stockholm: Södertörns högskola.

Glovinskaja, Marija Ja. (1996): Aktivnye processy v grammatike (na materiale innovacij i massovyx jazykovyx ošibok). In E. A. Zemskaja (ed.): *Russkij jazyk konca XX stoletija (1985-1995).* Moscow: Jazyki russkoj kul'tury.

Gorbačevič, Kirill S. (1971): *Izmenenie norm russkogo literaturnogo jazyka*. Leningrad: Prosveščenie.

Gorham, Michael S. (2014): *After Newspeak: Language Culture and Politics in Russia from Gorbachev to Putin.* Ithaca, NY, USA: Cornell University Press.

Hudson, Richard A. (1980): *Sociolinguistics*. Cambridge: Cambridge University Press.

Ickovič, Viktor A. (1982): *Očerki sintaksičeskoj normy*. Moscow: Nauka.

Israeli, Alina (1997): *Semantics and Pragmatics of the "Reflexive" Verbs in Russian*. Munich: Verlag Otto Sagner.

Janko-Trinickaja, Nadija A. (1962): *Vozvratnye glagoly v sovremennom russkom jazyke*. Moscow: Izdatel'stvo Akademii Nauk SSSR.

Kagan, Olga (2013): *Semantics of genitive objects in Russian: A study of genitive of negation and intensional genitive case*. Dordrecht: Springer.

Kakorina, E.V. (1996): Transformacii leksičeskoj semantiki i sočetaemosti. In E. A. Zemskaja (ed.): *Russkij jazyk konca XX stoletija (1985-1995).* Moscow: Jazyki russkoj kul'tury.

Krasovitsky, Alexander, Matthew Baerman, Dunstan Brown, Greville. G. Corbett (2011). Changing semantic factors in case selection: Russian evidence from the last two centuries. *Morphology* 21, 573-592.

Krys'ko, Vadim B. (1997): *Russkij istoričeskij sintaksis. Ob'ekt i perexodnost'*. Moscow: Indrik.

Ljustrova, Zoja N., Lev I. Skvorcov and Viktor Ja. Derjagin (1982): *Druz'jam russkogo jazyka*. Moscow: Znanie.

Maier, Ingrid (2010): Recension – Tore Nesset: Russiskstudentens beste venn: Elementær innføring i kasuslære. Oslo 2010: Novus forlag. *Slovo* (Uppsala), No. 51:141–148.

Nesset, Tore (1998a): *Russian Conjugation Revisited. A Cognitive Approach to Aspects of Russian Verb Inflection.* Oslo: Novus Press.

Nesset, Tore (1998b): Affiks eller klitikon?, *Norsk Lingvistisk Tidsskrift* 16: 185-206.

Nichols, Johanna (1993): Transitive and Causative in the Slavic lexicon: Evidence from Russian. In Bernard Comrie and Maria Polinsky (eds.): *Causatives and transitivity*. Amsterdam and Philadelphia: John Benjamins.

Ožegov, Sergej I. and Natalija Ju. Švedova (eds.) (1992): *Tolkovyj slovar' russkogo jazyka*. Moscow: Az'' Ltd.

Padučeva, Elena V. (2006). Genitiv dopolnenija v otricatel'nom predloženii. *Voprosy jazykoznanija* 6, 21–43.

Prokopovič, Nikolaj N., Lija A. Deribas and Elena N. Prokopovič (1975): *Imennoe i glagol'noe upravlenie v sovremennom russkom jazyke*. Moscow: Russkij jazyk.

Rozental', Ditmar È. (1986): *Upravlenie v russkom jazyke. Slovar'-spravočnik dlja rabotnikov pečati*. Moscow: Kniga.

Rozental', Ditmar È. (1988): *Punktuacija i upravlenie v russkom jazyke*. Moscow: Kniga.

Rozental', Ditmar È. and Margarita A. Telenkova (1984) *Slovar' trudnostej russkogo jazyka*. 3rd edition. Moscow: Russkij jazyk.

Ryazanova-Clarke, Larisa and Terence Wade (1996): *The Russian Language Today.* London and New York: Routledge.

Švedova, Natalija Ju. (ed.) (1980): Russkaja grammatika I. Moscow: Nauka.

Švedova, Natalija Ju. and Vladimir V. Lopatin (1989): *Kratkaja russkaja grammatika*. Moscow: Russkij jazyk.

Timberlake, Alan (1985): Hierarchies in the genitive of negation, in: Richard D. Brecht and James S. Levine, (eds.), Case in Slavic. Columbus, Ohio: Slavica Publishers, 338–360.

Timberlake, Alan. 2004. *A Reference Grammar of Russian*. Cambridge: Cambridge University Press.

Ušakov, Dmitrij N. (1935-40/2008): *Bol'šoj tolkovyj slovar' sovremennogo russkogo jazyka*. Moscow: Alta Print.

Vinogradov, Viktor V. (1947): *Russkij jazyk (grammatičeskoe učenie o slove).* Moscow and Leningrad: Gosudarstvennoe učebno-pedagogičeskoe izdatel'stvo ministerstva prosveščenija RSFSR.

Vinogradov, Viktor V. (1986): *Russkij jazyk (grammatičeskoe učenie o slove)* (3rd edition). Moscow: Vysšaja škola.

Zaliznjak, Andrej A. (2008): *Drevnerusskie ènklitiki*. Moscow: Jazyki slavjanskix kul'tur.