

Gender variation in indeclinable inanimate nouns and gender markedness in modern Russian

KIRILL CHUPRINKO^{1*} , VARVARA MAGOMEDOVA²  and
NATALIA SLIOUSSAR^{1,3} 

¹ National Research University Higher School of Economics, Russia

² Independent Researcher, Poland

³ Saint Petersburg State University, Russia

Received: December 1, 2022 • Revised manuscript received: June 22, 2023 • Accepted: June 23, 2023

Published online: August 10, 2023

© 2023 Akadémiai Kiadó, Budapest



ABSTRACT

In this paper, the results of a large web-corpus study on gender of Russian inanimate indeclinable common nouns are presented. In most cases, neuter is assigned to indeclinables as a default. However, morpho-phonological and semantic analogy may lead to feminine and masculine gender assignment. An extensive variation is observed in the whole group of indeclinables and for particular words, which is much larger than anything that can be found in indeclinable nouns. These data support the idea that both masculine and neuter genders have a special status in the Russian gender system (Magomedova & Slioussar 2023). Masculine tends to be chosen in case of conflicting gender cues. When there are no strong cues pointing to any gender, neuter is assigned as the default option. The results of the study are hardly compatible with various structural approaches to gender assignment, but can be accounted for in competition-based models.

KEYWORDS

gender variation, Russian indeclinable nouns, default gender, loanwords, corpus study

1. INTRODUCTION

In this paper, we focus on the grammatical gender of indeclinable nouns in Russian. Russian has three genders: masculine, feminine, and neuter (M, F, and N). Genders closely correlate with

* Corresponding author. E-mail: chuprinko_kirill@mail.ru

inflectional classes, or declensions, and there is an ongoing debate in the literature how to represent this connection (more details are given below). It is also debated which gender should be considered the default, or unmarked, in Russian. Masculine is the most frequent; neuter is the least frequent, but is used in impersonal sentences, i.e. in the syntactic contexts in which the default gender is expected. Various types of data, including data from psycholinguistic experiments, give a controversial picture.

We present the results of the first large corpus study of Russian inanimate indeclinable common nouns. Only a very small portion of Russian nouns are indeclinable. They are interesting for several reasons. Firstly, they demonstrate extensive gender variation. Secondly, they may be used to contribute to both theoretical debates mentioned above. Namely, indeclinable nouns allow us to separate the situation when there are conflicting cues for gender assignment from the situation when there are no (strong) cues. This difference is crucial for the discussion of markedness, as nouns with no cues must get the default gender, at least in the approaches relying on the Elsewhere condition. At the same time, conflicting cues may demonstrate another pattern, depending on the nature of these cues.

The paper is organized as follows. We start by briefly presenting the system of Russian genders and declensions. Then we outline the controversy around gender markedness in Russian. The second section is dedicated to different properties of indeclinable nouns: we briefly overview the previous studies and make some observations of our own. After that we proceed to our own corpus study.

1.1. Genders and declensions in Russian

The majority of Russian nouns are inflected for number and case and are divided into several inflectional classes (declensions) based on their set of inflections. Table 1 presents a widely accepted system with three classes (e.g. Aronoff 1994; Halle 1994; Shvedova 1980). Other approaches have also been proposed,¹ but, since our data do not let us tease them apart, we will not elaborate on this issue. Table 2 shows singular inflections for different declensions (in plural, they have the same inflections in most cases).

As Tables 1 and 2 show, there is a strong correlation between gender and declension, but very often, the gender of the noun cannot be unambiguously determined from its inflection. For example, if a noun ends in *-a* in nominative singular, it is most likely feminine, but can also be masculine; if it ends in *-ju* in instrumental singular, it is definitely feminine etc. Adjectives and participles in singular and verbs in past tense singular show gender agreement, which may help to determine the gender of the noun in case of uncertainty. However, masculine and neuter forms coincide in all cases except for nominative and accusative not only in nouns, but also in adjectives and participles, so even agreement information is not always enough.

The correlation between gender and declension is captured in different ways in several major theoretical frameworks. For example, Kramer (2015, 2020) proposes that declension is a syntactic feature situated lower in the tree than the gender feature. This allows gender to influence declension, but not vice versa. On the other hand, Rice (2005) in the framework of Optimality

¹In brief, they usually separate classes IIa and IIb or make a primary distinction between the I and II ('core') declensions on the one hand and the less frequent III declension on the other.



Table 1. The distribution of Russian nouns by declension and gender²

Declension	Gender	NOM.SG inflection	% in RNC
I	F	(j)a	29% nouns
I	M	(j)a	1% nouns (only animate)
IIa	M	∅	46% nouns
IIb	N	o/e	18% nouns
III	F	∅	5% nouns
indeclinable	different genders	–	1% nouns

Table 2. Singular paradigms in different declensions

	I	IIa	IIb	III
Nominative	(j)a	∅	o/e	∅
Genitive	i/y	(j)a		i
Dative	e	(j)u		i
Accusative	(j)u	= Gen/Nom	= Nom	= Nom
Instrumental	oj/ej	om/em		ju
Locative	e	e		i

Theory or Corbett & Fraser (2000) in the framework of Network Morphology assume that the declension class influences gender assignment.

Importantly, Table 1 does not represent several minor groups. Firstly, some animate nouns in the IIa class denoting professions and social roles allow for feminine gender agreement if their referent is female.³ Given that all masculine nouns in the I class are animate, this means that the connection between gender and declension is even stronger for inanimate nouns than for animate ones. Animate nouns have semantic gender, while inanimate nouns do not have any semantic cues and are left with morphophonological cues only.

Secondly, augmentative and diminutive suffixes are not supposed to change the gender of the base noun (e.g. Vinogradov 1947), but as a result many derivatives do not readily fit the picture in

²Percentages of nouns in the Russian National Corpus, or RNC (www.ruscorpora.ru), are taken from (Slioussar & Samoilova 2015). Their counts were based on the grammatically disambiguated subcorpus and did not take substantivized adjectives into account. Among the nouns with a zero inflection in nominative singular, feminine nouns in the III declension end in a palatalized or alveolo-palatal consonant, while masculine nouns in the IIa declension may end in any consonant.

³A series of formal and functional studies discusses the properties of this group (Panov 1968; Muchnik 1971; Graudina, Icković & Katlinskaja 1976; Corbett 1982, 1991; Asarina 2009; Pesetsky 2013; Steriopolo & Wiltschko 2010; King 2015; Lyutikova 2015; Steriopolo 2019; Landau 2016; Salzmann 2020; Matushansky 2013; Čaha 2019; Privizentseva to appear; Magomedova & Slioussar 2021 among others).



Table 1. This leads to extensive variation: the base noun gender competes with the gender associated with the inflection.⁴ For example, *domina* ‘big house’ is listed in (Zalizniak 1987) as a noun with gender variation (M/F): it is derived from a masculine noun *dom* ‘house’, but inanimate nouns ending in *-a* are feminine. Thirdly, there is a very small group of nouns with irregular inflections.

Now let us look at indeclinable nouns. **Table 3** provides two examples: a declinable noun *leto* ‘summer’ and an indeclinable noun *foto* ‘photo’. They have the same final syllable, the same number of syllables and stress pattern. However, in *leto* the final *-o* is a nominative singular affix, while in *foto* it is a part of the root. Indeclinable nouns do not have overt case marking. As these examples show, often it is not immediately clear why a noun is indeclinable from the morpho-phonological point of view – we will discuss this in more detail in section 2.2.

1.2. The problem of gender markedness

The definition of feature markedness depends on the theoretical framework (see e.g. Haspelmath 2006 for an overview). Structural approaches rely on the so-called representational markedness: a [+a] feature value is more marked than a [–a] value, and no marking is the least marked option. As a result, unmarked gender forms are expected to appear in the structures in which no gender feature is available (e.g. in impersonal sentences).

In Optimality Theory, the unmarked option is the default, which is used when there are no specific requirements to use an alternative option (e.g. C. Rice 2005; K. Rice 2007). Therefore, the unmarked option is expected to be the most frequent and the most productive. Many functional and typological studies hold similar views on markedness.

In spite of these differences, the same feature value is usually selected as unmarked in different approaches. However, this is not true for gender in Russian, which puts it at the centre

Table 3. Paradigms of the indeclinable noun *foto* ‘photo’ and declinable noun *leto* ‘summer’

	Declinable: Singular	<i>leto</i> ‘summer’ Plural	Indeclinable: <i>foto</i> ‘photo’ Singular = Plural
Nominative	<i>let-o</i>	<i>let-a</i>	<i>foto</i>
Genitive	<i>let-a</i>	<i>let-Ø</i>	<i>foto</i>
Accusative	<i>let-o</i>	<i>let-a</i>	<i>foto</i>
Dative	<i>let-u</i>	<i>let-am</i>	<i>foto</i>
Instrumental	<i>let-om</i>	<i>let-ami</i>	<i>foto</i>
Locative	<i>let-e</i>	<i>let-ax</i>	<i>foto</i>

⁴These nouns are briefly discussed in (Vinogradov 1947; Corbett 1982; Hippiusley 1996; Rice 2005; Savchuk 2011; Sitchinava 2011). Steriopolo and colleagues offer a formal analysis (Steriopolo 2008, 2015, 2017; Steriopolo, Markopoulos & Spyropoulos 2021), but it does not take gender variation into account. Corpus and experimental data on gender variation can be found in (Magomedova & Slioussar 2023).



of an ongoing theoretical debate. In Russian, masculine is the most frequent, while neuter is the least frequent (see Table 1), but is used in impersonal sentences.

Experimental studies also provide conflicting results. Several studies of gender agreement processing found that masculine behaves differently from the two other genders (Akhutina et al. 1999, 2001; Romanova & Gor 2017; Slioussar 2018). Slioussar & Malko (2016) discovered that neuter behaves as unmarked in production and masculine in comprehension in their agreement attraction experiments. Badecker & Kuminiak (2007) who studied gender agreement attraction in production in Slovak, in which the gender system is similar to Russian, also observed that neuter behaved as unmarked.

Structural approaches (e.g. Kramer 2015; Nevins 2011) consider neuter to be unmarked in Russian. For example, Nevins encodes feminine as [+FEM], [–MASC], masculine as [–FEM], [+MASC] and neuter as [–FEM], [–MASC]. For Kramer, feminine is [+FEM], masculine is [–FEM] and neuter corresponds to no gender features. Optimality-Theoretic approaches consider masculine to be unmarked and neuter the most marked (e.g. Rice 2005). Many other authors from different frameworks starting from Jakobson (1960) share this position. Finally, to account for the gender frequency distribution and for the neuter use in impersonal sentences at the same time, Corbett & Fraser (2000) assume masculine to be unmarked on the word level and neuter to be unmarked on the sentence level.

Magomedova & Slioussar (2023) suggest a solution to this paradox. Following various structural approaches, they assume that neuter is representationally unmarked, and conclude that it should be chosen when there are no gender assignment cues. If any cues are present, the ones that point to masculine will tend to win over the others (we will come back to this problem in the discussion section). As a result, both masculine and neuter can be seen as ‘default’ options, but in different circumstances and in different senses.

Magomedova & Slioussar (2023) illustrate the defaultness of masculine by their analysis of diminutive and augmentative nouns and nouns ending in a palatalized consonant in nominative singular. In case of conflicting gender cues, masculine is always significantly more likely to be chosen. As for the defaultness of neuter, they note that it can be observed in impersonal sentences and also in indeclinable nouns. They do not analyse these nouns in their paper, but our study confirms and extends their intuition.

2. INDECLINABLE NOUNS IN RUSSIAN

Indeclinable nouns are mostly loanwords ending in *-o*, *-e*, *-(j)a*, *-i* and *-(j)u* vowels.⁵ There are also several indeclinable nouns ending in *-y* and *-ě*, but they are infrequent (e.g. *Jancyz* ‘the river Yangtze’, *Xuanxè* ‘Huang He, the Yellow River’, *kanoé* ‘canoe’). Therefore, we do not consider these nouns in our analysis.

As Table 1 shows, indeclinables constitute about 1% of nouns in the Russian National Corpus. This does not look much, however, in comparison with other Slavic languages, Russian has many more indeclinable loanwords (Muchnik 1971, 256). In Serbian, Croatian, Polish or

⁵There are separate letters for *ja* and *ju* in Russian, but the vowels are the same as those corresponding to the letters *a* and *u* (their use depends on the preceding consonant), so we analyse them together in this paper. The letters *e* and *ě* also encode the same vowel /e/, but their use is less consistent. In particular, *ě* is almost never used in word-final positions.



Slovene, loanwords tend to fully assimilate and incorporate into the declension system, though some marginally stay indeclinable (Thomas 1983; Swan 2002; Shigemori Bučar 2011).

Now let us consider the gender of indeclinable nouns. Most authors modelling the connection between genders and declensions do not take them into account. Corbett (1982) assumes that in inanimate Russian nouns, the gender can be derived directly from the declension. He treats indeclinable nouns as a separate declension and claims that it is associated with the neuter gender. However, while for declinable inanimate nouns the correlation between gender and declension indeed has very few exceptions (almost exclusively diminutive and augmentative nouns), indeclinable nouns show considerable variation in gender assignment even in the dictionaries.

67% of indeclinable nouns are listed in dictionaries as neuter, 17% as masculine, 8% as feminine, and the remaining 7% are listed as having variative gender (Murphy 2000).⁶ One of the goals of our paper is to demonstrate that the variation in real language use is much more extensive. To do so, we conducted the first large corpus study of Russian indeclinable nouns based on a collection of unedited texts.

Masculine and feminine indeclinable nouns are traditionally considered as semantically motivated exceptions (Rozenal' et al. 1998, 203). Thus, *avenju* 'avenue' is feminine due to semantic analogy with the declinable noun *ulica* 'street', *tornado* is masculine due to analogy with *veter* 'wind_M' etc. However, several authors have recently noted that morphophonological factors also play a role: for example, indeclinables ending in *-(j)a* are more likely to be feminine (e.g. Mjakilja 2000; Wang 2014). These two groups of factors, semantic and morphophonological, are discussed in more detail in section 2.1. The third goal of our study, in addition to estimating gender variation and accounting for the prevalence of the neuter, was to analyse their relative importance.

Galbreath (2010) tries to account for the morphophonological factors in his model based on the Optimal Gender Assignment Theory (OGAT). For example, the following constraint is introduced: it is optimal for nouns ending with stressed phonemes /o/ and /e/ to be neuter. This applies both to declinable and to indeclinable nouns. Other constraints take morphology rather than surface forms into account: for instance, it is optimal for a noun of the IIb declension⁷ not to be masculine or feminine. The OGAT also uses gender markedness constraints. For declinable nouns, masculine is assumed to be the default, while for indeclinables, constraints are reranked to make neuter the default. Masculine is indeed the most frequent and productive in declinable nouns, while neuter prevails in indeclinables, but it is not clear in the OGAT model what the underlying reasons for this reranking are. In this paper, we outline a possible solution to this problem based on a new approach to gender markedness in Russian.

Another interesting problem associated with indeclinable nouns is why a certain word actually belongs to this group. In some cases, the answer is straightforward: for example, *-u* is not a nominative singular affix in Russian, so all loanwords ending in *-u* become indeclinable. In the other cases, the reason is not immediately clear: for example, the word *foto* 'photo' in Table 3

⁶For instance, four dictionaries list *kofe* 'coffee' as masculine or neuter. For other indeclinables, gender variation is less widely accepted: M/N variation for *viski* 'whiskey', M/F variation for *biennale* 'biennale' and *vol'vo* 'Volvo', F/N variation for *avenju* 'avenue' is documented only in one dictionary.

⁷The system of declensions adopted by Galbreath (2010) is different from the one used in this paper. In his paper, the IIb declension is class IV.



could in principle decline as the word *leto* ‘summer’ – and such words do decline in other Slavic languages. We discuss this problem in more detail in section 2.2 and hope to address it in our future research.

2.1. Semantic and morphophonological factors in gender assignment in the previous studies

In this paper, we refer to semantically motivated gender assignment to indeclinable nouns as *semantic analogy*. This phenomenon is not well defined: such semantic relations might be similar to hypernymy or to synonymy, in some other cases, they are sporadic conceptual matches. Different terms can be found in literature for this phenomenon: concept association, conceptual link, notional analogy etc. (Murphy 2000, 64–67).

Let us look at semantic and morphophonological factors in more detail. Murphy (2000) found that in contemporary Russian, the impact of these factors may vary for different groups of lexemes. The indeclinables ending in *-o* and *-e* are assigned neuter more often than other nouns regardless of their semantic analogy. She has also found that speakers may use masculine as the default gender for some indeclinable nouns: for unknown words, participants chose masculine agreement almost as frequently as neuter, while feminine was rare. This brings us back to the discussion of gender markedness in section 1.2.

Wang (2014) noted the influence of the final *-(j)a* on feminine gender assignment. Marginal examples of indeclinables ending in *-(j)a* with feminine agreement are also described in (Mjakilja 2000). Both studies also report semantic analogy effects. Some nouns initially demonstrated gender assignment due to semantic analogy, but later fell under the influence of morphophonological cues. For example, *šosse* ‘highway’ used to be feminine due to the semantic analogy with *doroga* ‘road_F’, but later became neuter (Comrie, Stone & Polinsky 1996, 108).

Both morphophonological and semantic analogies are common factors for gender assignment to loanwords cross-linguistically (Corbett 1991, 75–82). For example, a corpus study of Puerto Rican Spanish and Montreal French carried out by Poplack, Pousada & Sankoff (1982) shows that semantic analogy and phonological shape both influence gender assignment to loanwords. However, the size of the effect of each factor can be language-specific as it was found in numerous later works (Rabeno & Repetti 1997; Fuller & Lehnert 2000; Violin-Wigent 2006, among others), including several studies on code-switching in Russian (Leisiö 2001; Pereltsvaig 2004; Chirsheva 2009).

One of the experimental studies on the *Vastry* platform (<https://vastry.ru/>, Dobrushina, Staferova & Belokon 2018) showed that speakers are also sensitive to the prescriptive norm in case of gender assignment to indeclinable nouns. The indeclinable word *kofe* ‘coffee’ is a famous example in Russian where the prescriptive norm (masculine gender) does not coincide with the predominant pattern in real language use (neuter gender). In the experimental study, one group of participants had a reminder that prescriptively *kofe* is masculine, while the other group did not receive that priming. Participants with priming produced more masculine agreement with other indeclinable nouns as well (like *avokado* ‘avocado’ and *gaspacho* ‘gazpacho’) compared to the group without priming.

Another issue that is fundamental for the analysis of semantic factors is the speaker’s familiarity with a noun. It plays a minor role for corpus data (people mostly use words they are familiar with), but is hard to control in the experimental settings. Savchuk (2011)



conducted a study on the Russian National Corpus (www.ruscorpora.ru), which shows that semantic analogy has a strong impact on some groups of indeclinable nouns. For example, the names of car brands, like *Audi* ‘Audi’ and *Ševrole* ‘Chevrolet’, are predominantly used with feminine and masculine agreement, despite being listed in dictionaries as neuter. The analogy with the noun *mašina* ‘car_F’ suggests feminine agreement, while the analogy with *avtomobil* ‘automobile_M, car_M’ triggers masculine agreement. The study also demonstrated that newly borrowed nouns that are not in the dictionaries yet show much larger variation in gender agreement.

To summarize, previous studies show that many factors are at play: semantic analogy, the default status of neuter and maybe masculine, morphophonological factors. In our corpus study, we collected a large dataset to be able to estimate their relative influence.

2.2. What nouns become indeclinable

The distribution of indeclinable nouns discussed in this section stresses not only the strong connection between genders and declensions, but also the role of the stem-final segment and the relative productivity of different declensions. The pattern in the declension IIa (masculine nouns ending in a consonant) is the most productive, the pattern in the declension I (feminine nouns ending in *-(j)a*) is fully productive as well. The classes III and especially IIb get new nouns almost exclusively due to several productive suffixes associated with these declensions accidentally, all these suffixes are used to derive abstract nouns.

Russian nouns ending in a consonant in nominative singular are almost always declinable, with the exception of some female first and last names. Let us look at two examples. The last name of the Russian-born linguist Asya Pereltsvaig does not decline because there are no Russian feminine nouns with a stem-final non-palatalized consonant and a zero inflection. The last name of the Russian linguist Natalia Slioussar, spelled as *Sljuser* in Russian, could go to the III declension, but this never happens to such names – apparently, due to the low productivity of this class. Notably, both last names are declinable in masculine, i.e. when they refer to men.

There are examples of loanwords ending in a consonant like *xevi-metal* ‘heavy metal’ (music genre) that were noted to be indeclinable in the past (Murphy 2000), but now decline normally. All loanwords ending in a non-palatalized consonant go to the IIa declension and become masculine, while loanwords ending in a palatalized or alveo-palatal consonant may be incorporated either in the IIa or in the III declension and become masculine or feminine, respectively. As Savchuk (2011) shows, many of these nouns show gender variation in modern Russian. Importantly, this gender variation always goes hand in hand with declension variation. For example, for the word *šampun* ‘shampoo_{M/F}’ there are two options in genitive singular: *xorošego šampunja* ‘good_{M.GEN.SG} shampoo_{GEN.SG}’ or *xorošej šampuni* ‘good_{F.GEN.SG} shampoo_{GEN.SG}’. Using a masculine adjective with a III declension form or vice versa is absolutely ungrammatical.

Most Russian nouns ending in *-(j)a* are declinable. Some first and last names like *Derrida* are exceptions, although very often, there is variation in real language use. A small number of loanwords do not decline, but mostly monosyllabic ones like *fa* ‘note fa’ or those ending in a hiatus like *fejxoa* ‘fejjoa’, which stresses the importance of the stem for gender assignment (no native Russian words end in a hiatus). However, there are some exceptions like *antraša* ‘entrenchat’.



All nouns ending in *-(j)u* are indeclinable, which is not surprising because *-ju* or *-u* do not correspond to any nominative singular affix. The same is true for the very few nouns that end in *-ě*. Nouns ending in *-i* also do not decline unless *-i* is reanalysed as a nominative plural affix. However, this usually happens only in colloquial Russian (e.g. *dlja xinkalej* ‘for khinkali_{GEN.PL}’ (Georgian dumplings)). Most words ending in *-i* remain indeclinable, but are often used with plural agreement. Among other things, this means that the speaker does not have to decide on their gender because Russian has gender agreement only in singular. There are also a few indeclinable personal names and toponyms ending in *-y* (e.g. *Gurbanguly* (personal name), *Janczy* ‘the river Yangtze’).

Nouns ending in *-o* and *-e* are probably the most interesting. In declinable nouns, these are nominative singular affixes that unambiguously point to neuter (with the exception of some diminutive and augmentative nouns). However, loanwords ending in *-o* and *-e* do not go to the IIb declension and remain indeclinable, which is not the case in other Slavic languages. Moreover, some native proper names and toponyms ending in *-o*, like *Kupčino* (a district in Saint Petersburg), show variation in declinability. This can be associated with the low productivity of this declension in Russian, although it would be great to offer a more precise theoretical explanation in the future.

3. CORPUS STUDY

3.1. Data collection

In this paper, we present the results of a web-corpus study, with a focus on gender markedness and on idiosyncratic noun groups. We used the LiveJournal subcorpus of the General Internet Corpus of Russian (GICR, <http://www.webcorpora.ru/>) as we wanted to study variation which is poorly represented in edited texts. The LiveJournal subcorpus consists of blogs and contains 8,720 million words. It has some morphological annotation (morphological ambiguity, including ambiguity arising from case syncretism, is not resolved). We searched for combination of a preposed attributive adjective as an agreement target and an inanimate indeclinable noun as a head.

For the list of heads, we took all indeclinable nouns from the *Grammatical Dictionary of the Russian Language* (Zaliznjak 1987) and added a few other nouns. Murphy (2000) enumerates over 1,500 indeclinable nouns mentioned in various dictionaries. Firstly, we estimated their frequency subjectively to exclude those that would not be familiar to many Russian speakers (for instance, *kavallo*, a statue of a horseman in the classical Italian art, or *antuka*, a type of umbrella popular in the 19th century). We also did not include several highly frequent nouns: *kofe* ‘coffee’, *kafe* ‘café’, *metro* ‘subway’ and *taksi* ‘taxi’. Using the resulting list, we searched for all available information in the corpus. After that, we excluded some examples that were retrieved by mistake (i.e. did not contain an indeclinable noun with a preposed adjective modifying it) and all nouns with less than two examples of gender agreement.

One instance of gender agreement produced by one author was taken as one observation. If one author produced numerous instances of the same agreement for the same noun, such duplicates were dismissed. However, in 496 cases, one and the same noun was used in different genders by the one and same speaker (in different blog posts or even within one post). These



cases were included as separate observations. They show that the gender of indeclinables can vary even in the grammar of a single speaker.

The resulting dataset contained 66,939 data points. We annotated it for gender, number and case. Firstly, many instances were assigned these features automatically based on the inflections of adjectives. After that, other cases were manually disambiguated. Masculine and neuter adjectives are indistinguishable in cases other than nominative and accusative, so these instances were labelled as NM.

There is no gender agreement in plural, but we were interested in the cases of lexicalized plural agreement. Unfortunately, it is not always clear where an indeclinable noun is used as a pluralia tantum and where we deal with a regular plural formation. Eventually, we marked only those cases where we could be almost certain and left them for future research. In the present study, no instances of plural agreement are discussed, except for one example in Table 5 below.

The GICR also provides some metadata, although they are not always reliable.⁸ Most of the blog posts in our dataset were published between 2010 and 2013. The average birth year of the authors in our dataset is 1980. Most of the users who shared their location were in the Russian Federation, mainly in Moscow. There are also users located in Belarus, Ukraine, and Kazakhstan – the countries where there are many native speakers of Russian.

3.2. General statistics

Our dataset contains 66,939 data points. It is not balanced for frequency, as our goal was to collect all the available data. The most frequent noun in our dataset is *viski* ‘whiskey’ with 4,453 instances, the least frequent noun is *napareuli* ‘napareuli’ (a sort of wine) with 3 instances.

For the purposes of this study we excluded plural forms and monosyllabic nouns from the analysis. Homonyms were treated as two distinct nouns. This left us with 132 nouns (types) and 32,988 observations (tokens) in the main dataset.

Figure 1 presents gender distribution for the nouns that appeared more than 120 times in our dataset. The absolute majority of nouns show gender variation. For some of them, like *xudi* ‘hoodie’, *reno* ‘Renault’ or *pežo* ‘Peugeot’, all three genders have comparable frequencies.

3.3. Semantic analogy and morphophonological patterns

While labelling root-final segments is a very straightforward task, establishing semantic analogies may be trickier. To the best of our knowledge, studies discussing semantic analogy in gender assignment do not report any issues with the procedure (Poplack, Pousada & Sankoff 1982; Fuller & Lehnert 2000; Violin-Wigent 2006). However, in Russian indeclinables, there might be several or no obvious semantic analogies for a given word.

Therefore, we asked 25 native speakers to provide the most obvious word to describe a given indeclinable noun. Relying on the most popular answers, we introduced the following values for the semantic analogy factor. F, M, N were used for the words with one salient semantic analogy and “?” for the absence of any salient semantic analogy. The M/F label was assigned to the nouns

⁸For example, the latest publication year in our dataset is 2037, and the earliest birth year is 1890, which clearly cannot be true.



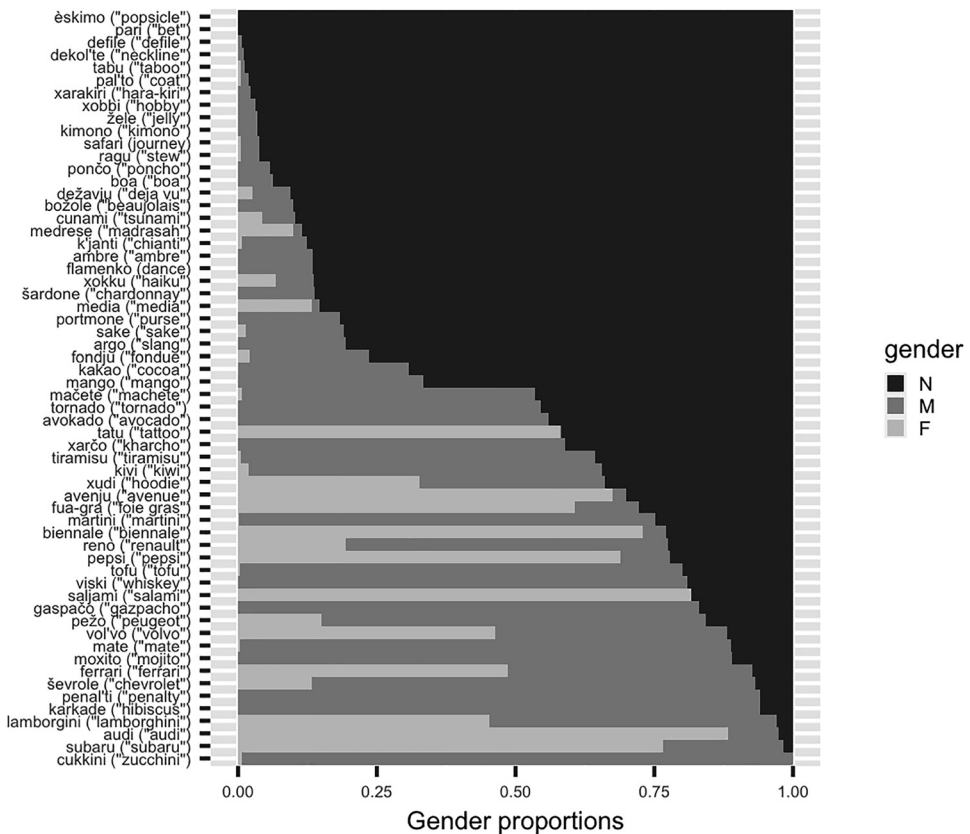


Figure 1. The distribution of gender agreement for frequent nouns in the dataset

with two salient semantic analogies of M and F gender. These are mostly car brands, their semantic analogies are *mašina* 'car_F' and *avtomobil* 'car_M'.

The resulting distribution of target nouns by semantic analogy gender is provided in Table 4. Table 4 also shows the distribution of assigned genders: a total number of tokens found, and the percentages of different genders in each category. Masculine is the most frequent as a semantic analogy gender, while neuter is the least frequent. This reflects the general distribution of genders in Russian nouns (see Table 1). However, the absolute majority of indeclinables with a neuter semantic analogy are neuter, while the same cannot be said about masculine and especially about feminine.

Let us compare the distributions of genders assigned in a group of nouns with the same semantic analogy, but with different root-final segments (Table 5). Here, we also included plural instances into the counts. *Flamenko*, *čačača*, and *sirtaki* are dances, and the word *tanec* 'dance' is masculine in Russian. Most instances of *flamenko* are neuter with zero feminine examples, while *čačača* has 26% of feminine instances. *Sirtaki* has no feminine instances, but has 21% of pluralia tantum forms instead, most probably due to the final *-i*. This may be surprising since *sirtaki* does not implicate any semantic plurality unlike spaghetti or pince-nez. These examples illustrate the influence of the final vowel on gender assignment.



Table 4. The distribution of assigned gender by the gender of semantic analogy (types and tokens)

Semantic analogy	?	n	f	m	m/f
Number of words (types)	30	14	20	52	16
M (tokens)	250 (6%)	131 (9%)	291 (6%)	11,353 (61%)	1,492 (38%)
F (tokens)	208 (5%)	8 (<1%)	1,752 (37%)	62 (<1%)	1,782 (45%)
N (tokens)	3,778 (89%)	1,390 (91%)	2,712 (58%)	7,111 (38%)	678 (17%)

Table 5. Gender assignment to indeclinables denoting dances and to the word *safari* in two meanings

	N	M	F	PL
<i>flamenko</i>	87% (n = 220)	13% (n = 34)	0	0
<i>čaćača</i>	46% (n = 16)	26% (n = 9)	28% (n = 10)	0
<i>sirtaki</i>	33% (n = 14)	45% (n = 19)	0	21% (n = 9)
<i>safari</i> (trip)	96% (n = 177)	3% (n = 6)	1% (n = 1)	0
<i>safari</i> (browser)	7% (n = 2)	79% (n = 23)	14% (n = 4)	0

The role of semantics can be illustrated by gender distribution in homonyms, like *safari* in Table 5. *Safari* can denote a nature trip or a popular web-browser. *Safari* in the first meaning was not labelled as having any salient semantic analogy and shows neuter agreement in the absolute majority of cases. *Safari* in the second meaning usually agrees in masculine due to the analogy with the word *brauzer* ‘browser_M’. Some examples of feminine agreement might be explained by the analogy with *programma* ‘program_F’. Neuter agreement is virtually absent.

Now let us turn to morphophonological factors. Our dataset contains 12 lemmas with the root-final *-(j)a*, 30 lemmas with final *-o*, 26 with final *-e*, 23 with final *-(j)u*, and 41 with final *-i*. The distribution of the assigned gender by the final segment is provided in Table 6. Absolute numbers show how many tokens were found, percentages show the share of the respective gender among all nouns with a particular final segment.

Nouns with the final *-(j)u* are of a particular interest as there is no corresponding declension marker for the nominative case. Thus, *(j)u*-final roots can be used to examine the suggested defaultness of the neuter gender. Table 7 shows gender assignment in these nouns taking the semantic analogy gender into account. It also includes *(j)a*-final and *i*-final nouns that will be discussed below.

Table 6. The distribution of assigned gender by the final segment

Final segment	<i>-(j)a</i>	<i>-e</i>	<i>-o</i>	<i>-(j)u</i>	<i>-i</i>
M	75 (9%)	1,067 (21%)	4,343 (46%)	1,106 (22%)	6,926 (55%)
F	238 (28%)	656 (13%)	263 (3%)	1,012 (20%)	1,643 (13%)
N	534 (63%)	3,306 (66%)	4,931 (52%)	2,956 (58%)	3,942 (32%)



Table 7. The distribution of types and tokens in the subset of nouns with the root-final *-(j)u*, *-i* and *-(j)a*

Semantic analogy	final <i>-(j)u</i>					final <i>-(j)a</i>					final <i>-i</i>				
	?	m	f	n	m/f	?	m	f	n	m/f	?	m	f	n	m/f
Types	10	7	3	1	2	5	5	1	0	1	3	15	9	7	7
Tokens	2,043	1,435	1,312	67	217	467	239	4	0	137	215	7,493	2,061	720	2,022
M	151	894	8	1	52	11	48	0	0	16	18	5,978	234	59	637
F	83	6	780	7	136	121	30	4	0	83	2	13	377	1	1,250
N	1,809	535	524	59	29	335	161	0	0	38	195	1,502	1,450	660	135

Within the *(j)u*-final subset the group with no obvious semantic analogy is the largest. There are nouns like *azu* ‘azu’ (a Tatar dish), *dežavju* ‘deja vu’, *interv’ju* ‘interview’, *kazu* ‘kazoo’ (a musical instrument), *paspартu* ‘passe-partout’, *ragu* ‘stew, ragout’, *randevu* ‘rendezvous’, *ušu* ‘wushu’, *fondju* ‘fondue’, *hokku* ‘haiku’, *džiu-džitsu* ‘jiu-jitsu’. This set of nouns should be the most problematic for the speaker as there are no clear cues for gender assignment. Therefore, the default gender is expected. Indeed, 1,809 of the 2,043 tokens (88.5%) are neuter, 151 are masculine and only 83 are feminine. However, even in the situation of uncertainty we observe some gender variation.

Nouns with the final *-(j)a* look similar to *(j)u*-final nouns in terms of the gender distribution and exactly the opposite in terms of the gender assignment cues. As Table 7 shows, this set mostly consists of nouns with contradictory cues. The root-final *-(j)a* may trigger feminine agreement, but there are almost no words with a feminine semantic analogy. Words with a neuter semantic analogy are also absent (nevertheless, 63% of nouns in this group are assigned neuter).

There are only 12 *(j)a*-final nouns in our data, namely: *amplua* ‘character (theatre)’, *antraša* ‘entrechat’, *boa* ‘boa’, *kinoa* ‘quinoa’, *krema* ‘cream on top of a coffee’, *ljulja* ‘kebab’, *media* ‘media’, *patua* ‘patois’, *tanka* ‘tanka’, *fejxoa* ‘fejjoa’, *fuagra* ‘fois-gras’, *čačača* ‘cha-cha-cha (dance)’. Six of these nouns have hiatus endings, which is very unusual for Russian, and most of them have stress on the root-final vowel. The *(j)a*-final noun set offers another opportunity to test for the defaultness of neuter in indeclinables – they have contradictory cues for gender assignment, but very often end up being neuter.

-i also does not correspond to any nominative singular affix, but the subset of *i*-final nouns differs from *(j)u*-final ones for a number of reasons. Firstly, many *i*-final nouns show lexicalized plural agreement. Secondly, several nouns with a masculine semantic analogy like *viski* ‘whiskey’ are much more frequent than the other words in the dataset, which skews the distribution of genders. A decent share of feminine agreement is also due to a particular group of words: five car brand names (*audi*, *lamborghini*, *micubisi*, *ferrari*, *mazerati*). As we mentioned earlier, these words have *m/f* semantic analogy gender. Feminine agreement instances of these nouns constitute 76% of feminine agreement instances in the whole *i*-final subset. Let us note that neuter agreement is nevertheless highly frequent for the words with feminine and masculine semantic analogy and clearly prevails not only in the *n* group, but also in the ? group.

Finally, we also looked at stress patterns, which can be considered a purely phonological, rather than a morphophonological factor. As we will show in the next section, stress patterns



appear to be significant predictors and improve the predictive power of the model. However, we cannot be completely sure that this significance is actually meaningful. As we have mentioned above, indeclinable nouns are a limited set of words that very often form subgroups with idiosyncratic behaviour, which some of the predictors we used in the models might catch by coincidence. We compared different stress patterns (on the final, prefinal and the second prefinal syllable), but grouping all words with a non-final stress appeared to be more meaningful, so this is what we present in Table 8. Let us add that in declinable nouns, there is no connection between gender and stress patterns.

3.4. Statistical analysis

The statistical analysis was done in the R programming environment (www.r-project.org). We modelled the data with a mixed-effects regression with random intercept by item using the *lme4* package (Bates et al. 2015). For post hoc analyses, we used Tukey’s tests with a Holm-Bonferroni correction using the *glht* function from the *multcomp* package (Bretz, Hothorn & Westfall 2010). We used the same set of independent variables for each model, namely: the root-final segment, the gender of the semantic analogy noun, and stress. We chose neuter as the reference level for semantic analogy gender factor, and *-i* as a reference for the final segment factor.

Masculine gender assignment. Nouns with a feminine, neuter or unclear semantic analogy are less likely to be assigned masculine than nouns with masculine or mixed semantic analogy (see Table 9). Nouns with a mixed semantic analogy are less likely to be assigned masculine than those with masculine analogy. The root-final vowel and the stress pattern are not significant predictors for masculine gender assignment.

Feminine gender assignment. Nouns with a feminine semantic analogy end up with feminine agreement significantly more often than nouns with any other analogy except for the mixed *m/f* (see Table 10). As we mentioned above, this group of nouns mostly includes car brands. The feminine semantic analogy (*mašina* ‘car_F’) is very strong for them, and the other analogy (*avtomobil* ‘automobile_M, car_M’) rather reduces the share of the ‘default’ neuter agreement than undermines the influence of the feminine analogy word. This is probably due to the fact that the word *avtomobil* ‘car_M’ is not as frequent as the word *mašina* ‘car_F’ (156.9 ipm and 490.4 ipm according to Lyashevskaya & Sharov (2009)). There are no significant differences between masculine, neuter and unclear semantic analogy.

Table 8. The distribution of stress patterns in lemmas and tokens and the distribution of the assigned gender by stress patterns

Stress	Final	Non-final
Number of lemmas	58	73
Number of tokens	9,925	23,025
M (tokens)	2,423 (24%)	11,085 (48%)
F (tokens)	1,060 (11%)	2,735 (12%)
N (tokens)	6,442 (65%)	9,205 (40%)



Table 9. Semantic analogy and masculine gender assignment

Semantic analogy	β	SE	<i>P</i>	Significance code
? – n	–0.017	0.609	0.978	
f – n	–0.725	0.643	0.664	
m – n	3.414	0.542	<0.001	***
m/f – n	2.102	0.646	0.006	**
f – ?	–0.708	0.579	0.664	
m – ?	3.431	0.444	<0.001	***
m/f – ?	2.119	0.575	0.001	**
m – f	4.139	0.512	<0.001	***
m/f – f	2.827	0.619	<0.001	***
m/f – m	–1.312	0.508	0.039	*

Table 10. Semantic analogy and feminine gender assignment

Semantic analogy	β	SE	<i>P</i>	Significance code
? – n	1.539	1.072	0.593	
f – n	6.396	1.051	<0.001	***
m – n	0.614	1.017	1.000	
m/f – n	6.643	1.066	<0.001	***
f – ?	4.857	0.704	<0.001	***
m – ?	–0.925	0.640	0.593	
m/f – ?	5.104	0.716	<0.001	***
m – f	–5.782	0.625	<0.001	***
m/f – f	0.2468	0.656	1.000	
m/f – m	6.0287	0.637	<0.001	***

Nouns ending in *-(j)a* are assigned feminine significantly more often than nouns with any other root-final segment (see Table 11). *(j)u*-final nouns are significantly more likely to be feminine than nouns with the root-final *-e* and *-o*; *(j)u*-nouns are more likely to be feminine than *i*-nouns on a tendency level. Nouns with the final stress are significantly less likely to be feminine ($\beta = -1.62$, $SE = 0.57$, $P < 0.005$) than masculine or neuter.

Neuter gender assignment. As for semantic analogy, every value significantly differs from the others (see Table 12), except for *m* vs. *f* and *m* vs. *m/f*, which are different on a tendency level.

Considering the root-final vowel, *-o* is different from *-i*, *-(j)a* and *-(j)u*, but not *-e* (declinable neuter nouns end in *-o* or *-e*). There are no other significant differences (see Table 13). Finally, nouns



Table 11. Root-final segments in feminine gender assignment

Root-final segment	β	SE	<i>P</i>	Significance code
(j)a – i	4.135	0.867	<0.001	***
(j)u – i	1.753	0.730	0.065	.
e – i	–0.708	0.776	0.596	
o – i	–1.542	0.644	0.065	.
(j)u – (j)a	–2.382	0.801	0.015	*
e – (j)a	–4.843	0.881	<0.001	***
o – (j)a	–5.677	0.887	<0.001	***
e – (j)u	–2.461	0.738	0.005	**
o – (j)u	–3.295	0.762	<0.001	***
o – e	–0.834	0.802	0.596	

Table 12. Semantic analogy and neuter gender assignment

Semantic analogy	β	SE	<i>P</i>	Significance code
? – n	–0.105	0.605	0.862	
f – n	–2.250	0.628	0.002	**
m – n	–3.228	0.527	<0.001	***
m/f – n	–4.432	0.645	<0.001	***
f – ?	–2.145	0.565	0.001	***
m – ?	–3.123	0.449	<0.001	***
m/f – ?	–4.326	0.587	<0.001	***
m – f	–0.978	0.497	0.098	.
m/f – f	–2.182	0.616	0.002	**
m/f – m	–1.204	0.521	0.063	.

Table 13. Root-final segments in neuter gender assignment

Root-final segment	β	SE	<i>P</i>	Significance code
o – i	1.677	0.462	0.003	**
o – (j)a	2.267	0.688	0.009	**
o – (j)u	1.719	0.552	0.015	*



with the final stress are significantly more likely to be neuter ($\beta = 1.39$, $SE = 0.42$, $P < 0.01$) than masculine or feminine.

3.5. Discussion. Markedness hierarchy

Data on inanimate indeclinables can complement the discussion on gender defaultness and markedness in the Russian nominal system. Figure 2 shows how often a particular gender is assigned to an indeclinable noun with or without any cues that point to this gender. For example, the root-final *-(j)a* is a cue to feminine. Figure 2 demonstrates that the share of neuter agreement used in the absence of any cues is considerably larger than the corresponding shares of masculine and feminine agreement.

Magomedova & Slioussar (2023) argue that neuter is the default gender in Russian, while the special properties of masculine are explained by the fact that it is most frequent and productive and has the most diverse set of nominative singular forms. Therefore, they predict that neuter should be chosen in the absence of any cues. This prediction can be tested only on indeclinables, because in declinable nouns, declension is a very strong cue for gender assignment, and our data support this prediction. In their own study of diminutive and augmentative nouns, Magomedova and Slioussar show that masculine is significantly more likely to be chosen in case of conflicting gender cues. Thus, both masculine and neuter have a special status in the Russian nominal system, but the nature of this status and its manifestations are very different.

Some effects associated with the special status of masculine can also be observed in our study. Unmotivated masculine and feminine agreement is infrequent in our data. In order to compare them, let us form a subset with 672 (5% of all feminine gender agreement instances) and 127 (3% of all feminine gender agreement instances) entries respectively. When nouns have no salient

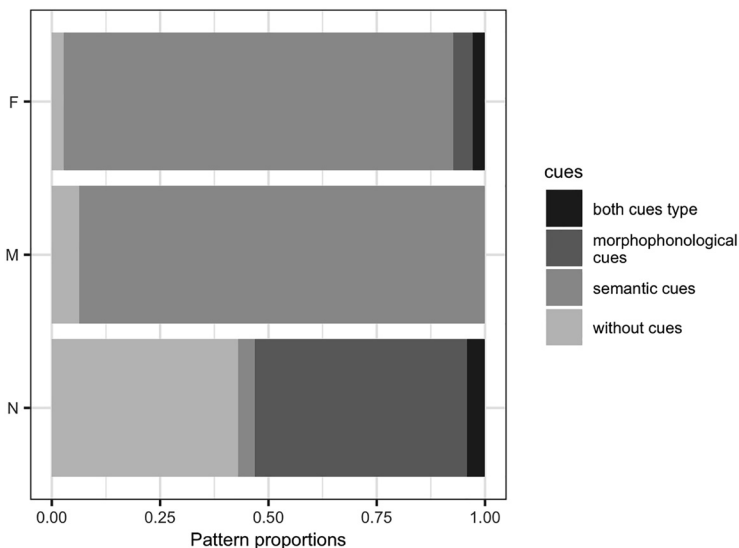


Figure 2. The distribution of gender agreement depending on the presence or absence of different cues



semantic or morphological cues pointing to any gender or when they are assigned an unexpected gender given the morphophonological or semantic factors, they are masculine significantly more often than feminine (according to the two-proportions Z-test, $Z\text{-score} = 4.2$, $P < 0.01$).

In our sample, there are nouns with decent shares of masculine assignment without any cues to it, like *xudi* ‘hoodie’ (33% of masculine respectively). Other nouns have smaller shares of unmotivated masculine agreement: for example, 6% of *pončo* ‘poncho’ instances are masculine, despite its final vowel pointing to neuter. Examples of unmotivated feminine agreement are usually isolated cases, except for the nouns *seppuku* ‘seppuku’ and *xokku* ‘haiku’ (10% and 7% respectively).

3.6. Discussion. Approaches to gender assignment and the challenges posed by indeclinable nouns

Different approaches to gender assignment have been proposed, in the frameworks ranging from Distributed Morphology, as in [Kramer \(2015, 2020\)](#), to Relational Morphology, as in [Jackendoff & Audring \(2020\)](#), and Optimality Theory, as in [Rice \(2005\)](#). Ruth Kramer in her paper on gender assignment claims that it is basically a derivation process and gender is assigned by the Merge operation when a bare root or a stem is merged with a nominal head that contains a gender feature. In simplex nouns, gender is assigned first on the basis of semantics and then the remainder may be assigned gender based on the declension class.

In the analysis of derived nouns, she assumes lexical decomposition: the speaker already knows how to divide a given word into morphemes as gender assignment is done by the Merge operation ([Kramer 2020](#), 58). She also argues for top-down derivation and doubts the existence of phonological gender assignment. Our data do not support this analysis: not only phonological factors matter, the root-final segment is not separable from the root in any formal way.

The same kind of problem holds for the Relational Morphology approach: although this approach allows analogy as a mechanism for assigning features, there is no formal way to separate the root-final vowel from the root and substitute it with a variable. The ‘similar’ relation can be used instead of the ‘same-same’ relation, but we could not find a formal way to do that. Hence, competition-based approaches are optimal for our data. For example, we can specify ‘ends in $-(j)a$ ’ as an OT constraint in a Harmonic Grammar type of approach, which will allow us not only to account for the observed variation, but also to predict the distribution of the genders assigned.

4. CONCLUSION

In this paper, we presented the results of the first large corpus study of Russian inanimate indeclinable common nouns. Our study was based on a corpus of unedited blog posts which reflect real language use rather than the prescriptive norm. There are claims in the literature that most Russian indeclinable nouns are neuter, with the exception of some cases in which a salient semantic analogy plays a role. We demonstrated that the neuter gender indeed prevails in this group, but many instances of masculine and feminine can also be found.

We observed a vast variation both in the group as a whole and in individual nouns, also in one speaker’s production. Nothing similar can be found in declinable nouns, arguably because



their gender choice is restricted by their declension. The only group of declinable nouns that show considerable (but smaller) variation are diminutives and augmentatives, in which the lexical gender and inflectional affixes offer contradictory cues.

We showed that the variation in indeclinable nouns depends on two groups of factors, semantic and morphophonological, and estimated their relative importance. Crucially, these factors do not work in the same way as in declinable nouns and cannot have a similar explanation. In declinable nouns, semantic agreement is found only in words denoting people, in which grammatical gender correlates with biological sex and social gender. If we take the indeclinable noun *avenju* ‘avenue_F’ discussed above, which is feminine because of the semantic analogy with the word *ulica* ‘street_F’, we cannot say that the concept of the street is feminine in any meaningful way.

As for morphophonological factors, in declinable nouns the inflection plays a crucial role, and its role can be explained by the gender-declension connection, in whichever way we represent it in the theory. The final segments of indeclinable nouns are not inflections, so, when they influence the choice of gender, this influence is not morphological, but morphophonological in nature, and other mechanisms must be found to explain it. We believe that some theoretical approaches to morphology are better suited to account for these data than the others and address this topic in a separate paper (Chuprinko, Magomedova & Slioussar to appear).

Finally, our data can contribute to the debate concerning gender markedness in Russian. Some authors assume that masculine is unmarked, while others argue for the default status of neuter. Following Magomedova & Slioussar (2023), we assume that both genders have a special status that may become evident in different contexts. In case of strong conflicting cues pointing to different genders, masculine tends to be chosen significantly more often than feminine or neuter (Magomedova and Slioussar’s data on diminutives and augmentatives confirm this generalization). In the absence of any (strong) cues, neuter becomes the default option. Our data confirm this hypothesis. All cases that do not fit readily the Russian nominal system are more likely to be assigned neuter than any other gender.

ACKNOWLEDGEMENT

The study was partially supported by the Russian Ministry of Science and Higher Education (the research project 075-15-2020-793).

REFERENCES

- Akhutina, Tatiana, Andrei Kurgansky, Maria Polinsky and Elizabeth Bates. 1999. Processing of grammatical gender in a three-gender system: Experimental evidence from Russian. *Journal of Psycholinguistic Research* 28. 695–713.
- Akhutina, Tatiana, Andrei Kurgansky, Marina Kurganskaya, Maria Polinsky, Natalya Polonskaya, Olga Larina, Elizabeth Bates and Mark Appelbaum. 2001. Processing of grammatical gender in normal and aphasic speakers of Russian. *Cortex* 37. 295–326.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.



- Asarina, Alya. 2009. Gender and adjective agreement in Russian. Handout presented at the 4th Annual Meeting of the Slavic Linguistics Society, University of Zadar, Zadar, September 3–6, 2009.
- Badecker, William and Frantisek Kuminiak. 2007. Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language* 56. 65–85.
- Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker. 2015. Fitting linear mixed-effects models using *lme4*. *Journal of Statistical Software* 67. 1–48.
- Bretz, Frank, Torsten Hothorn and Peter Westfall. 2010. Multiple comparisons using R. Boca Raton, FL: CRC Press.
- Caha, Pavel. 2019. Case competition in nanosyntax: A study of numerals in Ossetic and Russian. Berlin: Language Science Press.
- Chirsheva, Galina. 2009. Gender in Russian-English code-switching. *International Journal of Bilingualism* 13(1). 63–90.
- Chuprinko, Kirill, Varvara Magomedova, Natalia Slioussar. To appear. Modelling gender variation in Russian indeclinable nouns: Optimality over structuralism, hierarchical MaxEnt, and degrees of idiosyncrasy. *Journal of Slavic Linguistics*.
- Comrie, Bernard, Gerald Stone and Maria Polinsky. 1996. The Russian language in the twentieth century. Oxford: Clarendon Press.
- Corbett, Greville G. 1982. Gender in Russian: An account of gender specification and its relationship to declension. *Russian Linguistics* 6. 197–232.
- Corbett, Greville G. 1991. Gender. Cambridge: Cambridge University Press.
- Corbett, Greville G. and Norman M. Fraser. 2000. Gender assignment: A typology and a model. In G. Senft (ed.) *Systems of nominal classification*. Cambridge: Cambridge University Press. 293–325.
- Dobrushina, Nina R., Daria A. Staferova and Alexander A. Belokon. 2018. Elektronnaja baza variativnyx javlenij [Electronic database of variation]. *Slověne* 1. 424–436.
- Fuller, Janet M. and Heike Lehnert. 2000. Noun phrase structure in German-English codeswitching: Variation in gender assignment and article use. *International Journal of Bilingualism* 4(3). 399–420.
- Galbreath, Blake Lee Everett. 2010. Gender assignment in contemporary standard Russian: A comprehensive analysis in Optimality Theory. Doctoral dissertation. University of Virginia, Charlottesville, VA.
- Graudina, Ljudmila K., Viktor A. Ickovič and Lija P. Katlinskaja. 1976. Grammatičeskaja pravil'nost' russkoj reči [Grammatical correctness of the Russian speech]. Moscow: Nauka.
- Halle, Morris. 1994. The Russian declension: An illustration of the theory of distributed morphology. In J. S. Cole and C. Kisseberth (eds.) *Perspectives in phonology*. Stanford: CSLI Publications. 29–60.
- Haspelmath, Martin. 2006. Against markedness (and what to replace it with). *Journal of Linguistics* 42. 25–70.
- Hippisley, Andrew. 1996. Russian expressive derivation: A Network Morphology account. *The Slavonic and East European Review* 74. 201–222.
- Jackendoff, Ray and Jenny Audring. 2020. The texture of the lexicon: Relational morphology and the parallel architecture. New York, NY: Oxford University Press.
- Jakobson, Roman. 1960. The gender pattern of Russian. *Studii și Cercetări Lingvistice* 11. 541–543.
- King, Katherine E. 2015. Mixed gender agreement in Russian DPs. M.A. thesis. University of Washington, Seattle, WA.
- Kramer, Ruth. T. 2015. The morphosyntax of gender. Oxford: Oxford University Press.
- Kramer, Ruth. T. 2020. Grammatical gender: A close look at gender assignment across languages. *Annual Review of Linguistics* 6. 45–66.



- Landau, Idan. 2016. DP-internal semantic agreement: A configurational analysis. *Natural Language and Linguistic Theory* 34(3). 975–1020.
- Leisiö, Larisa. 2001. Morphosyntactic convergence and integration in Finland Russian. Doctoral dissertation. University of Tampere, Tampere.
- Lyashevskaya, Olga and Sergey Sharov. 2009. *Častotnyj slovar' sovremennogo russkogo jazyka* [The frequency dictionary of modern Russian language]. Moscow: Azbukovnik.
- Lyutikova, Ekaterina A. 2015. Soglasovanie, priznaki i struktura imennoj gruppy v russkom yazyke [Agreement, features and the structure of the noun phrase in the Russian language]. *Russkij Yazyk v Nauchnom Osveshchenii* 30. 44–74.
- Magomedova, Varvara and Natalia Slioussar. 2021. Gender and case in Russian nouns denoting professions and social roles. *Computational Linguistics and Intellectual Technologies* 20. 483–491.
- Magomedova, Varvara and Natalia Slioussar. 2023. Gender variation and gender markedness in Russian nouns. *Voprosy Jazykoznanija* 2. 7–28.
- Matushansky, Ora. 2013. Gender confusion. In L. L.-S. Cheng and N. Corver (eds.) *Diagnosing syntax*. Oxford: Oxford University Press. 271–294.
- Mjakilja, Kari. 2000. K probleme roda nesklonjaemyx zaimstvovannyx imen naricatel'nyx v sovremennom russkom jazyke [Towards the problem of indeclinable common loans' gender in contemporary Russian]. *Scando-Slavica* 46. 93–103.
- Muchnik, Iosif P. 1971. Grammatičeskie kategorii glagola i imeni v sovremennom russkom literaturnom jazyke [Grammatical categories of verb in contemporary standard Russian]. Moscow: Nauka.
- Murphy, Dianna L. 2000. The gender of inanimate indeclinable common nouns in modern Russian. Doctoral dissertation. University of Ohio, Athens, OH.
- Nevins, Andrew. 2011. Marked targets versus marked triggers and impoverishment of the dual. *Linguistic Inquiry* 42(3). 413–444.
- Panov, Mikhail. 1968. *Russkij jazyk i sovetskoe obščestvo: Morfologija i sintaksis sovremennogo russkogo literaturnogo jazyka* [Russian language and Soviet society: Morphology and syntax of the modern Russian literary language]. Moscow: Nauka.
- Pereltsvaig, Asya. 2004. Gender agreement in American Russian. *Cahiers Linguistiques d'Ottawa* 32. 87–107.
- Pesetsky, David. 2013. Russian case morphology and the syntactic categories. Cambridge, MA: MIT Press.
- Poplack, Shana, Alicia Pousada and David Sankoff. 1982. Competing influences on gender assignment: Variable process, stable outcome. *Lingua* 57(1). 1–28.
- Privizentseva, Mariia. To appear. Mixed agreement in Russian: Gender, declension, and morphological ineffability. *Natural Language and Linguistic Theory*.
- Rabeno, Angela and Lori Repetti. 1997. Gender assignment of English loan words in American varieties of Italian. *American Speech* 72(4). 373–380.
- Rice, Curt. 2005. Optimizing Russian gender: A preliminary analysis. In S. Franks, F. Y. Gladney and M. Tasseva-Kurktchieva (eds.) *Formal Approaches to Slavic Linguistics* 13. Ann Arbor, MI: Michigan Slavic Publications. 265–275.
- Rice, Keren. 2007. Markedness in phonology. In P. de Lacy (ed.) *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press. 79–98.
- Romanova, Natalia and Kira Gor. 2017. Processing of gender and number agreement in Russian as a second language. *Studies in Second Language Acquisition* 39. 97–128.
- Rozental', Ditmar E., Evgenija V. Dzhandzhakova and Natal'ja P. Kabanova. 1998. *Spravočnik po pravopisaniju, proiznošeniju, literaturnomu redaktirovaniju* [Handbook of orthography, pronunciation, editing]. Moscow: CheRo.



- Salzmann, Martin. 2020. The NP vs. DP debate: Why previous arguments are inconclusive and what a good argument could look like: Evidence from agreement with hybrid nouns. *Glossa* 5(1). 83.
- Savchuk, Svetlana. 2011. Korpusnoe issledovanie variantov rodovoj prinadležnosti imen suščestvitel'nyx v russkom jazyke [A corpus study of gender variation in Russian nouns]. *Computer Linguistics and Intellectual Technologies* 10. 562–579.
- Shigemori Bučar, Chikako. 2011. Creative competence in borrowings: Words of Japanese origin in Slovene. *Linguistica* 51(1). 245–262.
- Shvedova, Natalia (ed.). 1980. Russkaja grammatika [Russian grammar], Vol. 2. Moscow: Nauka.
- Sichinava, Dmitri V. 2011. Rod: Materialy dlja proekta korpusnogo opisanija russkoj grammatiki [Gender: Materials for the project of the corpus description of Russian grammar]. Manuscript. Available at: <http://rusgram.ru>.
- Slioussar, Natalia and Anton Malko. 2016. Gender agreement attraction in Russian: Production and comprehension evidence. *Frontiers in Psychology* 7. 1651, 613–632.
- Slioussar, Natalia and Maria Samoilova. 2015. Častotnosti različnyx grammatičeskix xarakteristik i okončanij u suščestvitel'nyx russkogo jazyka [Frequencies of different grammatical features and inflectional affixes in Russian nouns]. Proceedings of the Conference 'Dialogue'. Available at: <http://www.dialog-21.ru/digests/dialog2015/materials/pdf/SlioussarNASamoilovaMV.pdf>.
- Slioussar, Natalia. 2018. Gender, declension and stem-final consonants: An experimental study of gender agreement in Russian. *Computational Linguistics and Intellectual Technologies* 17. 688–700.
- Steriopolo, Olga and Martina Wiltschko. 2010. Distributed gender hypothesis. In G. Zybatow, P. Dudchuk, S. Minor and E. Pshehotskaya (eds.) *Formal studies in Slavic linguistics: Proceedings of the Formal Description of Slavic Languages 7.5*. Bern: Peter Lang. 155–172.
- Steriopolo, Olga, Giorgos Markopoulos and Vassilis Spyropoulos. 2021. A morphosyntactic analysis of nominal expressive suffixes in Russian and Greek. *The Linguistic Review* 38. 1–42.
- Steriopolo, Olga. 2008. Form and function of expressive morphology: A case study of Russian. Doctoral dissertation. University of British Columbia, Vancouver.
- Steriopolo, Olga. 2015. Syntactic variation in expressive size suffixes: A comparison of Russian, German, and Spanish. *SKASE Journal of Theoretical Linguistics* 12. 2–21.
- Steriopolo, Olga. 2017. Nominalizing evaluative suffixes in Russian: The interaction of declension class, gender, and animacy. *Poljarnyj Vestnik: Norwegian Journal of Slavic Studies* 20. 18–44.
- Steriopolo, Olga. 2019. Mixed gender agreement in the case of Russian hybrid nouns. *Questions and Answers in Linguistics* 5(2). 91–105.
- Swan, Oscar E. 2002. A grammar of contemporary Polish. Bloomington, IN: Slavica Publishers.
- Thomas, George. 1983. A comparison of the morphological adaptation of loanwords ending in a vowel in contemporary Czech, Russian, and Serbo-Croatian. *Revue Canadienne des Slavistes* 25(1). 180–205.
- Vinogradov, Viktor V. 1947. Russkij jazyk: Grammatičeskoe učenie o slove [Russian language: A grammatical study of the word]. Moscow: Prosvesčenie.
- Violin-Wigent, Anne. 2006. Gender assignment to nouns codeswitched into French: Observations and explanations. *International Journal of Bilingualism* 10(3). 253–276.
- Wang, Qiang. 2014. Gender assignment of Russian indeclinable nouns. M.A. thesis. University of Oregon, Eugene, OR.
- Zaliznjak, Andrej. A. 1987. Grammatičeskij slovar' russkogo jazyka: Slovoizmenenie [The grammatical dictionary of the Russian language: Inflection], 2nd edn. Moscow: Russkij Jazyk.

