Modeling Gender Variation in Russian Indeclinable Nouns: Optimality over Structuralism, Hierarchical MaxEnt, and Degrees of Idiosyncrasy

Varvara Magomedova, Kirill Chuprinko, and Natalia Slioussar

Abstract: In this paper, we provide an analysis of the grammatical gender of 131 inanimate indeclinable Russian nouns based on the data from the General Internet Corpus of the Russian Language. We demonstrate that most nouns show substantial variation, being used in two or even in all three genders: masculine, feminine, and neuter. We identify several factors affecting this, primarily the gender of the semantic analogy noun and the root-final vowel. We argue that these data can be used to compare several major morphological frameworks and conclude that some approaches, namely optimality-theoretic probabilistic ones, are better suited to account for them. We also compare different models within the chosen set of approaches and show that the hierarchical Maximum Entropy (MaxEnt) models are superior to the classical MaxEnt models.

1. Introduction*

Russian nouns are inflected for number and case and are divided into several declensions or inflectional classes based on their affixes. They are also divided into three genders: masculine, feminine, and neuter (M, F, and N). The gender of a noun closely correlates with its inflectional class, especially in the case of inanimate nouns.

However, there is a small group that does not conform to this generalization: indeclinable inanimate nouns. In this paper, we aim to analyze their grammatical gender based on corpus data. As a result, we reveal extensive variation both in the group as a whole and in individual nouns, and the factors affecting this variation. Our second goal is to show that these data can be used to compare several major morphological frameworks; we argue that some of

* The work on this project was carried out in the framework of the Basic Research Program at the HSE University, Russia. Kirill Chuprinko's research was also partially supported by the Slovenian Research Agency grant no. J6-4615.

them, namely optimality-theoretic probabilistic approaches, are better suited to account for such data than the others. Our third goal is to compare different models within the chosen set of approaches. We argue that the hierarchical Maximum Entropy (MaxEnt) model is superior to the classical MaxEnt models if some idiosyncrasies in the data are properly taken into account.

The paper has the following structure. After providing more information on genders, declensions, and indeclinability in Russian, we briefly formulate the theoretical questions we aim to address, introduce the corpus data we collected, and proceed to discuss them in light of different frameworks. Finally, we compare several optimality-theoretic models of our data.

2. Genders, Declensions, and Indeclinability in Russian

There are several approaches to Russian nominal declensions. Our data do not bear on this question, so we will rely on a widely accepted system with three classes, presented in Table 1 (e.g., Aronoff 1994; Halle 1994; and Shvedova, 1980). Inflections in the classes IIa and IIb are the same in all cases except for nominative and accusative, which is why some authors unite them, as in Table 1, and others do not. On the other hand, inflections in the classes IIa and III are the same only in nominative and accusative (in inanimate nouns). Classes that have distinct inflections in the singular share the same inflections in the plural, and there is no gender agreement, so we do not consider plural forms in this paper.

Table 1. The distribution of Russian nouns by declension and gender[1]

Declension	Gender	NOM.SG inflection	% in RNC	Examples
I	F	-*(j)a*	29% nouns	*mama* 'mom', *zemlja* 'earth'
I	M (animate)	-*(j)a*	1% nouns	*papa* 'dad', *djadja* 'uncle'
IIa	M	∅	46% nouns	*stol* 'table', *gel'* 'gel'
IIb	N	-*o/e*	18% nouns	*vino* 'wine', *pole* 'field'
III	F	∅	5% nouns	*myš'* 'mouse', *mel'* 'shallow'
indeclinable	different genders	—	1% nouns	*kino* 'cinema', *kofe* 'coffee'

[1] Percentages of nouns in the Russian National Corpus, or RNC (www.ruscorpora.ru), are taken from Slioussar and Samoilova (2015). Their counts were based on the grammatically disambiguated subcorpus and did not take substantivized adjectives into account (e.g., *moroženoe* 'ice cream'). Among the nouns with a zero inflection in the nominative singular, feminine nouns in the III declension end in a palatalized or

Table 1 shows that the connection between gender and declension is pervasive but not absolute and does not work at all for indeclinable nouns (i.e., indeclinability is not associated with a particular gender). This table also does not include several minor noun groups that complicate the picture even further. There are different theoretical approaches to the problem of gender and declension (e.g., Caha 2021; Corbett and Fraser 2000; Kramer 2015; and Rice 2005), but we will not discuss them in this paper because our data do not let us tease them apart. In any case, because of this problem and because the affixes in different classes coincide in some forms, the gender of the noun can be unambiguously determined only from agreeing adjectives, participles, and verb forms, or from the choice of pronouns.[2]

All declinable Russian nouns have stems ending in a consonant. Most derivational affixes are associated with a particular declension, but if a noun stem does not have any, it is impossible to decide based on its final consonant to which declension and gender it belongs. There are certain tendencies and restrictions (e.g., see the comments to Table 1), but it is still true that stems in classes I and IIa may end in any consonant. If a noun is inanimate, it is also absolutely impossible to predict its declension and gender based on its semantics.[3] As we show below, these fundamental generalizations do not hold for indeclinable nouns.

Indeclinables are mostly loanwords ending in *-o*, *-e*, *-(j)a*, *-i*, and *-(j)u* vowels. Several indeclinable nouns end in *-y* and *-è*, but they are infrequent—e.g., *Janczy* '(the river) Yangtze', *Xuanxè* 'Huang He, (the Yellow River)'—so we do not analyze them in this paper. If a loanword ends in a consonant, eventually it always becomes a class IIa masculine noun or a class III feminine noun.[4] Most loanwords ending in *-(j)a* become class I feminine nouns, i.e., the final vowel gets reanalyzed as a NOM.SG inflection, but there are exceptions. In some cases, the reason is clear; since all declinable Russian noun stems end in a consonant, loanwords ending in a hiatus, like *kinoa* 'quinoa', become indeclinable,

---

alveolo-palatal consonant, while masculine nouns in the IIa declension may end in any consonant. As for *(j)a,* the Russian alphabet has two separate letters (Я and Ю) that are transliterated in the Latin alphabet as *ja* and *ju*: they convey the same vowels as the letters *a* and *u* but change the articulation of the preceding consonant, introduce the consonant /j/, or are purely a matter of orthographic conventions.

[2] Only the latter option is available to differentiate between masculine and neuter nouns in any form except for the nominative and accusative singular: all affixes coincide not only in nominal, but also in adjectival paradigms.

[3] For animate nouns, especially those denoting people, the grammatical gender is closely connected with the biological sex and/or the social gender of the referent.

[4] The former option is less restricted phonologically and in general much more frequent. In addition to that, many female first and last names that end in a consonant are indeclinable.

unless *j* is appended to the stem. For example, *Ikea*, the brand name, is usually pronounced as *Ikeja* in colloquial Russian and declines as a class I feminine noun. The other cases are more mysterious, like *kannada* 'Kannada (language)'. There is no hiatus, but the word does not decline.

Nouns ending in *-(j)u, -è, -y* and *-i* (like *tabu* 'taboo', *mjusli* 'muesli') do not fit into any existing declension and are doomed to be indeclinable (although *-y/i* can be reanalyzed as a plural affix yielding a pluralia tantum noun). Nouns ending in *-o* and *-e* (like *pal′to* 'coat', *žele* 'jelly') are the most noteworthy: they could be reanalyzed as class IIb neuter nouns, but this almost never happens.[5] Moreover, some native proper names and toponyms ending in *-o,* like *Ivanovo* (an old Russian city), show variation in declinability. In general, although the group of indeclinables is very small in Russian, it is still much larger than in other Slavic languages (Mučnik 1971: 256, Thomas 1983; Swan 2002; and Shigemori Bučar 2011)—an observation that still awaits its explanation and more precise calculations.

Another reason why indeclinables are interesting is that they prefer the neuter gender, but all three genders are represented in this set. According to a dictionary study by Murphy (2000), 67% of indeclinable nouns are listed in various dictionaries as neuter, 17% as masculine, 8% as feminine, and the remaining nouns demonstrate gender variation. Our corpus study shows much more extensive variation in real language use, but neuter is still the most frequent gender. This is remarkable, given that in declinable nouns, masculine is the most frequent and neuter the least frequent, as Table 1 shows. Magomedova and Slioussar (2023) try to explain this, developing a novel approach to the problem of gender markedness in Russian. At the same time, the fact that neuter is not the only choice for such nouns is hard to explain in any model studying the connection between genders and declensions.

Finally, indeclinable nouns are remarkable because their gender assignment was shown to be affected by their semantics and by their stem-final segment. Both factors were noted in the previous studies based on individual examples and some experimental data, while we conducted a corpus study to estimate their role based on a large dataset of naturally occurring examples. Speaking of the semantic factor, there are cases when the gender of a semantically related declinable noun influences gender assignment of an indeclinable (Rozental′ et al. 1998; Murphy 2000; Galbreath 2010; Savchuk 2011, etc.). The nature of these semantic relations is difficult to formalize; they may involve synonymy, hypernymy, or even sporadic conceptual matches. Follow-

[5] For example, the word *èxo* 'echo', clearly borrowed a long time ago, is declinable, although some examples of indeclinable usage are reported by Corbett (2023). At the same time, for many nouns that are indeclinable in modern standard Russian, some declinable forms have been registered in non-standard variants—for example, see Henry (2020) for the history of the word *pal′to* 'coat'.

ing Corbett (1991), we call these relations *semantic analogy*. For example, *avenju* 'avenue' is feminine in the dictionaries and is often used with feminine agreement in our dataset because the synonymous noun *ulica* 'street' is feminine. *Mango* 'mango' and *kivi* 'kiwi' are listed as neuter in the dictionaries, but our dataset contains numerous examples of masculine agreement because their hypernym *frukt* 'fruit' is masculine.

The semantic analogy effect was noted to influence the gender assignment of loanwords cross-linguistically. Besides the case of Russian indeclinables, Corbett (1991: 75–82) provides examples from Archi (Nakh-Dagestanian), Hausa (Afro-Asiatic), and Polish (Slavic). However, as far as we can judge, Russian indeclinables are different in the following way. In all these languages, semantic analogy affects gender assignment when the loanword is adapted to the new language (which affixes and agreement patterns it will be associated with, how its phonological form will be changed), but then its role is the same as for any other noun in this language—in particular, it plays no role for inanimate nouns in Hausa and Polish. Russian also has similar examples—they are briefly discussed at the beginning of §4. However, the situation is different for Russian inanimate indeclinable nouns: semantic analogy affects gender assignment even in the nouns that were borrowed a long time ago and are listed in dictionaries, competing with other factors and resulting in massive gender variation that is not attested for any declinable Russian nouns.[6]

The second major factor is morphophonological analogy with declinable nouns. Some stem-final vowels of indeclinables resemble NOM.SG (-(*j*)*a*, -*o*, -*e*) and NOM.PL (-*i*) affixes. In the dictionary data overview, Murphy (2000: 109–14) discussed how indeclinables ending in -*o* and -*e* are more likely to be neuter, while the ones ending in -(*j*)*a* show a higher share of feminine gender than other indeclinables. This is confirmed in an experimental study by Wang (2014) and by some observations by Mjakilja (2000). Savchuk (2011) discusses the role of morphophonological analogy for nouns like *spagetti* 'spaghetti' and *mjusli* 'muesli' that are mostly used as pluralia tantum. What is especially interesting here is that morphophonological analogy works despite the fact that the nouns remain indeclinable, i.e., their final vowels do not get reanalyzed as inflectional affixes.

The fact that indeclinable nouns consist of a root alone and yet have highly variable gender that appears to depend on the semantic and phonological properties of that root poses several theoretical questions. The main two questions we address in this paper are the following: which theoretical approach can handle these data and how to formalize the analysis? Namely,

[6] Some declinable inanimate nouns show gender variation, but it is limited (e.g., Savchuk 2011). Animate nouns that can be used with masculine and feminine agreement are analyzed in terms of common gender rather than gender variation (the choice of grammatical gender depends on the biological or social gender of the referent).

which formal analysis methods of the selected theoretical approach predict our data distribution the most accurately? In the next section, we present the data in more detail and then turn to these questions.

3. Corpus Study

3.1. Data Collection and Preliminary Analysis

We collected data from the LiveJournal subcorpus (8.72 billion words) of the General Internet Corpus of Russian (GICR, http://www.webcorpora.ru/); it contains blogs that did not undergo editing or proofreading. Our dataset is based on blog posts published between 2010 and 2013. We took the list of inanimate indeclinable nouns from the *Grammatical Dictionary of Russian* (Zaliznjak 1987). We added a few other nouns that we noticed to be in use (both based on our native speaker intuition and on the dictionary study by Murphy 2000) and excluded several very infrequent nouns that would not be familiar to many Russian speakers (for instance, *kavallo*, a statue of a horseman in classical Italian art). We also did not include several highly frequent nouns: *kofe* 'coffee', *kafe* 'café', *metro* 'subway', *kino* 'movie (theatre)', *pal'to* 'coat', and *taksi* 'taxi'.[7]

We searched for combinations of an indeclinable noun and an attributive adjective preceding it as an agreement target. One instance of gender agreement with one noun produced by one speaker was taken as one observation (the corpus has information allowing us to identify different speakers). If one speaker produced numerous instances of the same agreement for the same noun, they were not included in the dataset. However, in 496 cases, one and the same noun was used in different genders by one and the 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 datapoints. After manually cleaning the data, we excluded all nouns with less than two examples of gender agreement and annotated the dataset for gender, number, and case. Many instances were annotated automatically based on the inflections of adjectives; other instances were manually disambiguated. Plural forms were not annotated for gender (there is no gender agreement in plural). In the singular, masculine and neuter adjective forms coincide in all cases except for nominative

[7] All these nouns are much more frequent than the majority of other nouns in our list, so we had a feeling that they could skew the results. They show almost no gender variation in contemporary Russian, being neuter (although see Comrie, Stone, and Polinsky 1996 for some interesting diachronic data), except for the word *kofe* 'coffee'. *Kofe* is masculine according to the conservative literary norm, but many speakers use it with neuter agreement, which provoked one of the biggest orthoepic debates in modern Russian. Gender variation in other indeclinable nouns is not widely discussed.

and accusative, so these instances were labeled as NM. For the analysis in this study, we selected only examples in nominative and accusative singular where gender can be determined unambiguously.[8] We also did not include monosyllabic nouns.[9] As a result, we had 131 nouns (types) and 32,792 observations (tokens) in the final dataset.

We annotated all nouns in this dataset for the root-final vowel, lexical stress (on the final, penultimate, or other syllables), and semantic analogy. The last task was less straightforward. Studies discussing semantic analogy in gender assignment do not report any problems with the procedure (Poplack, Pousada, and Sankoff 1982; Fuller and Lehnert 2000; and Violin-Wigent 2006). But, according to our intuitions, for some Russian indeclinables, one semantic analogy is found easily, while for the others, several variants come to mind, or there is no obvious variant at all. Therefore, we conducted a survey with 25 native speakers of Russian using the IbexFarm platform (https://korpling.german.hu-berlin.de/ibex/). The participants were asked which word they would use to briefly describe an indeclinable noun presented on the screen. Participants could also answer that they did not know the noun or did not know how to explain its meaning briefly. Several indeclinables were frequently described with nouns of different genders, e.g., *bungalo* 'bungalow' as *dom* 'house.M' and *žilišče* 'housing.N'. However, only the nouns in the M/F group were numerous.[10] All other mixed groups contained fewer than five nouns. Therefore, the following labels for semantic analogy were adopted: *F, M, N, M/F,* and *?*. The last label was used when no single noun was present in more than 75% of answers (because our participants struggled to give a short definition or because the definitions they gave were not uniform).

3.2. Statistical Analysis, Results, and Discussion

In this section, we discuss the main results of data analysis, while the following sections focus on fitting them into different morphological frameworks and modeling them. These results are also interesting for the discussion of the gender markedness problem; there is a debate as to whether the masculine or

---

8 It would be interesting to find out how often different indeclinables are used as pluralia tantum nouns, but this is difficult to do based on naturally occurring examples (most indeclinables can be used in plural, so we would have to show that some nouns are used in plural more often than expected).

9 Monosyllabic words with a final vowel are very unusual for the Russian nominal system; almost all of them are names of Latin and Cyrillic letters and musical notes.

10 The M/F group primarily includes car brands. The feminine semantic analogy (*mašina* 'car.F') is very strong for them; the other analogy (*avtomobil'* 'automobile.M, car.M') is less prominent but nevertheless appears regularly and can be shown to affect the behavior of these nouns.

the neuter is the unmarked gender in Russian. We address this question in a separate paper (Chuprinko et al. 2023), which also presents more details on gender variation in indeclinable nouns, providing numerous examples and comparisons.

The first general observation is that gender variation in indeclinable nouns is enormous. Almost every noun in the sample varies in gender to some extent. Individual speakers often select different genders for one and the same noun, even within one text. Secondly, neuter is the most frequent gender in indeclinables (43% of the instances in our dataset), but masculine follows it closely (37% of instances). Feminine examples are clearly a minority (20% of instances). Thus, the picture we get from the corpus differs from the one we get from the dictionaries, although the hierarchy of genders is the same. Notably, many instances of masculine gender assignment are due to semantic analogy. Out of the words that have one salient semantic analogy in our dataset, 52 can be associated with masculine, 21 with feminine, and 15 with neuter (presumably because masculine is the most frequent and neuter the least frequent among declinable nouns). However, the neuter is often assigned in the absence of any clear cues. This pattern is discussed in more detail in Chuprinko et al. (2023).

Before turning to the statistical analysis, a caveat should be mentioned. Semantic and morphophonological features are distributed non-uniformly in our dataset—for example, there are only 12 indeclinable nouns ending in *-(j)a,* and each of them is relatively infrequent. Still, the dataset was large enough to analyze the factors of interest.

The statistical analysis was done in the R programming environment (R Core Team 2021). We tried two approaches to data modeling: mixed-effects logistic regressions with a random intercept by item (*lme4* package; Bates et al. 2015) and logistic regressions with fixed effects only using the built-in *glm* function. We explain the theoretical implications of both approaches and compare the models in §5. The mixed-effects approach turned out to be superior, so below we report the results of mixed-effects models, focusing on the main significant findings (the full description and outputs of all models can be found in the Appendix, available at https://osf.io/wsuaz/). For post hoc analyses, we ran Tukey's tests with the Holm-Bonferroni correction, using the *glht* function from the *multcomp* package (Bretz et al. 2010, and Hothorn et al. 2008). So, all *p*-values reported in tables 2 and 3 are corrected.

We made three mixed-effects models testing the probability of a word to be masculine, feminine, and neuter.[11] We used the same independent vari-

[11] Thus, the dependent variable in our three models is binary (e.g., the word is masculine vs. not masculine depending on the factors included in the model). It is also possible to make a multinominal model with a categorically distributed dependent variable (masculine vs. feminine vs. neuter) using the *mclogit* package (Elff 2022).

ables for each model: the root-final vowel, the gender of the semantic analogy noun, lexical stress (all non-final stresses were grouped), and case (nominative or accusative). We had no specific predictions for the case factor and only wanted to check that it does not influence our results in any unexpected way. We chose neuter as a reference level for the semantic analogy factor and *-i* as a reference for the final vowel factor (the case and stress factors were binary, so no reference level was needed).

The root-final vowel and semantic analogy factors were significant for all genders. Since these factors are not binary, we discuss pairwise comparisons between different values below. Lexical stress was significant only for feminine and neuter models. Nouns with the final stress are significantly less likely to be feminine ($\beta = -1.75$, $SE = 0.59$, $p = 0.003$) than masculine or neuter and significantly more likely to be neuter ($\beta = 1.41$, $SE = 0.43$, $p = 0.001$) than masculine or feminine. This may be due to the fact that the stress makes root-final vowels more salient because Russian has unstressed vowel reduction. The case factor reached significance in all three models, and we specifically checked that it improves their performance using the methods described in §5, but these results are difficult to explain. For instance, it may be that some nouns that tend to be assigned feminine have a relatively higher incidence of accusative forms than those that tend to be assigned neuter, etc. We tried including interactions between different factors in the models, but the resulting models were inferior.

Now let us discuss two main factors in more detail. Figures 1 and 2 on the following page show the role of the semantic analogy. Figure 1 groups the data by the gender of the semantic analogy, while Figure 2 groups them by the assigned gender, which corresponds to our statistical models. This gives us different perspectives. In Figure 1, we can see the general preference for the neuter gender, but masculine and M/F semantic analogies override it, while feminine semantic analogies visibly reduce the share of neuter.

In Figure 2, we can see that the masculine gender group includes mostly nouns with masculine and M/F semantic analogies. The feminine group almost exclusively consists of nouns with feminine and M/F semantic analogies. The neuter group, in contrast, includes a variety of nouns.

However, one value of the dependent variable must be taken as a reference level in such models; for example, if we take the neuter, the model will give us two sets of pairwise comparisons: N vs. M and N vs. F. For the third comparison, another model with a different reference level is needed. We believe that this is not optimal for interpreting the results and for comparing different models, which is one of the goals of this paper. Moreover, as we will show below, it is clear from the data that the factors of interest affect the choice of different genders in different ways. Therefore, we wanted to analyze each gender separately, not comparing it to one of the others but to both of them at once.

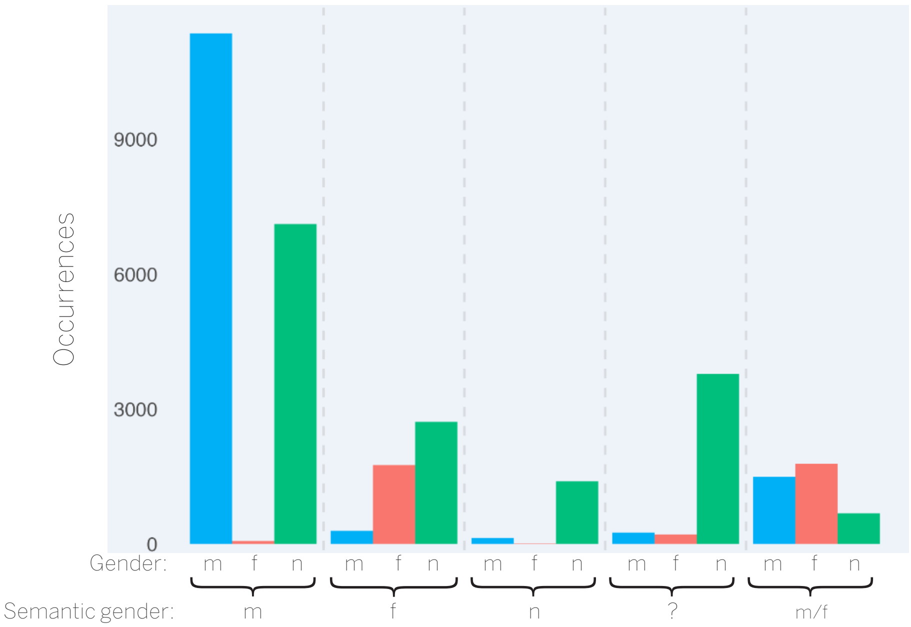

Figure 1. The distribution of genders depending on the semantic analogy[12]

The results of pairwise comparisons can be found in Table 2 (only significant results are reported). Table 2 shows the estimated effect (beta) in log odds, the standard deviation, and the significance for each pair of semantic analogy values for each gender. For example, *gender = masculine, m - n == 0,* $\beta = 3.42$*, SE = 0.53,* $p < 0.001$*** means that when compared to nouns with masculine semantic analogy, nouns with neuter semantic analogy are less likely to end up masculine by approximately 3.42 in log odds with a standard deviation of 0.53, and this difference is significant at the level of $p < 0.001$.

12 "?" indicates that there is no obvious semantic analogy noun, so its gender cannot be determined. "M/F" indicates that there are two salient semantic analogy nouns of masculine and feminine gender.

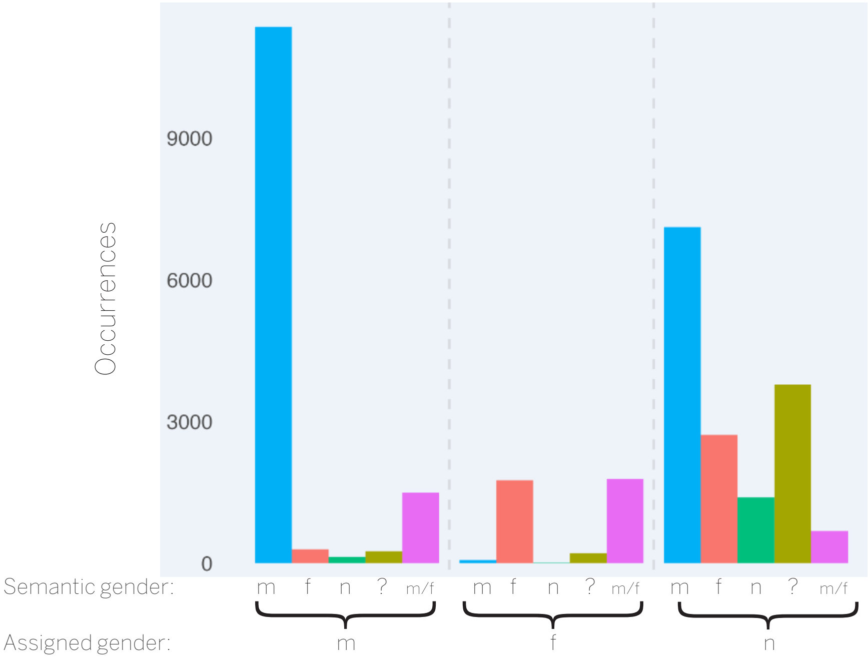

Figure 2. Semantic analogy in each gender group

Table 2. Pairwise comparisons for the semantic analogy factor

Assigned gender	Pair	Beta	SE	*p*	Significance code
masculine	m - n == 0	3.420	0.526	< 0.001	***
masculine	m/f - n == 0	2.107	0.636	0.005	**
masculine	m - ? == 0	3.409	0.450	< 0.001	***
masculine	m/f - ? == 0	2.100	0.579	0.002	**
masculine	m - f == 0	4.140	0.516	< 0.001	***
masculine	m/f - f == 0	2.829	0.622	< 0.001	***
masculine	m/f - m == 0	−1.311	0.508	0.040	*
feminine	f - n == 0	6.444	1.038	< 0.001	***
feminine	m/f - n == 0	6.693	1.046	< 0.001	***
feminine	f - ? == 0	5.016	0.726	< 0.001	***
feminine	m/f - ? == 0	5.265	0.737	< 0.001	***
feminine	m - f == 0	−5.783	0.630	< 0.001	***
feminine	m/f - m == 0	6.032	0.640	< 0.001	***

–continued on next page–

—continued— **Assigned gender**	**Pair**	**Beta**	**SE**	***p***	**Significance code**
neuter	f - n == 0	−2.250	0.643	< 0.001	***
neuter	m - n == 0	−3.230	0.553	< 0.001	***
neuter	m/f - n == 0	−4.431	0.676	< 0.001	***
neuter	f - ? == 0	−2.161	0.571	< 0.001	***
neuter	m - ? == 0	−3.141	0.458	< 0.001	***
neuter	m/f - ? == 0	−4.343	0.598	< 0.001	***
neuter	m - f == 0	−0.980	0.498	0.098	.
neuter	m/f - f == 0	−2.181	0.622	0.002	**
neuter	m/f - m == 0	−1.201	0.530	0.070	.

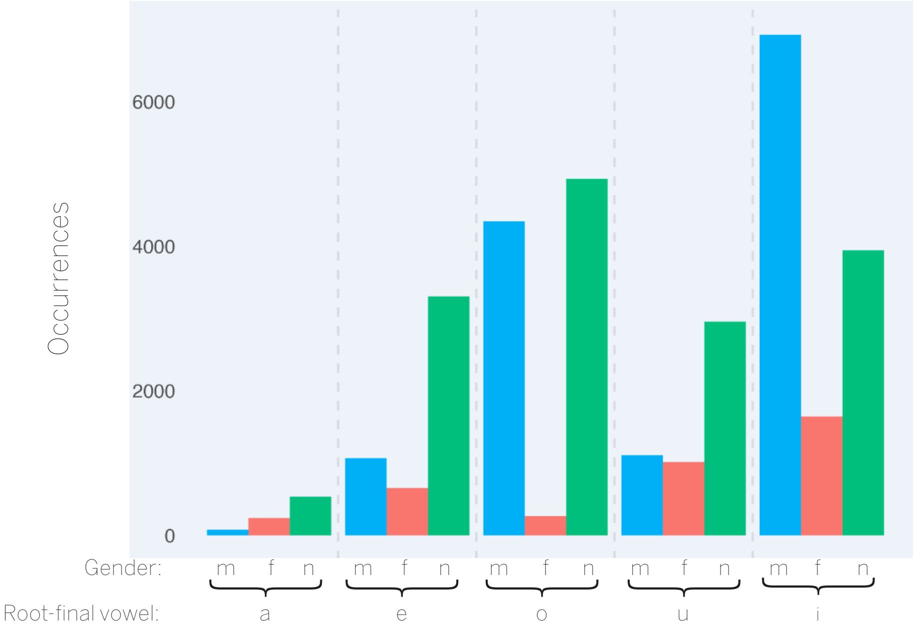

Figure 3. The distribution of genders depending on the root-final vowel

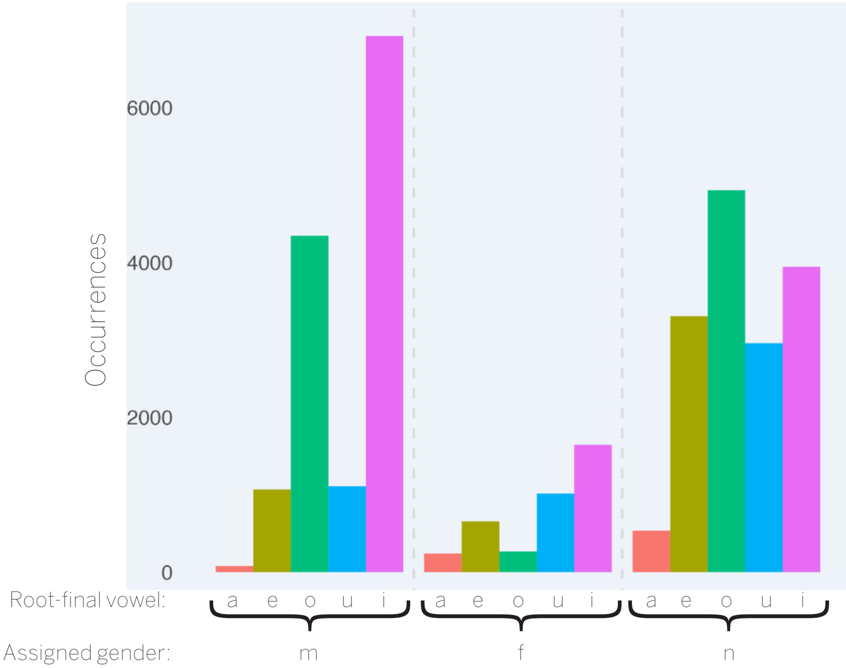

Figure 4. Root-final vowels in each gender group

Figures 3 and 4 show how the morphophonological analogy affects gender distribution. In Figure 3, we can see that the share of feminine agreement is larger for the nouns ending in *-(j)a*, while the share of neuter agreement is larger for the nouns ending in *-e*. For the nouns ending in *-i*, which do not have any morphophonological analogy in the singular, the share of masculine forms is the largest, due to the role of the semantic analogy factor. The semantic factor also explains a large share of masculine forms in the *o*-final group. Still, neuter is the most frequent gender in it, demonstrating the interplay of the two factors.

Figure 4 gives a different perspective on the same data. As we could see above, a word with a masculine semantic analogy has high chances to be masculine, while for a word with a feminine semantic analogy, this is less certain, and the morphophonological factor may affect the distribution. The neuter group is again the most diverse. The results of pairwise comparisons are reported in Table 3. They are significant only in the feminine and neuter models—presumably because the final vowel factor may change the share of these two genders—but not in the masculine, which is selected predominantly based on the semantic factor. In neuter models, the comparisons with *-u* and especially

-e are not significant because these groups are more diverse internally than *-o* (semantic analogy happens to play a stronger role for the nouns in these groups). In the next section, we will discuss these results in light of different theoretical frameworks.

Table 3. Pairwise comparisons for the root-final vowel factor

Assigned gender	Pair	Beta	SE	*p*	Significance code
feminine	a - i == 0	4.287	0.886	< 0.001	***
feminine	o - i == 0	−1.653	0.651	0.044	*
feminine	u - i == 0	1.903	0.750	0.044	*
feminine	e - a == 0	−4.885	0.890	< 0.001	***
feminine	o - a == 0	−9.940	0.924	< 0.001	***
feminine	u - a == 0	−2.384	0.807	0.016	*
feminine	u - e == 0	2.501	0.747	0.005	**
feminine	u - o == 0	3.556	0.802	< 0.001	***
neuter	o - i == 0	1.689	0.470	0.003	**
neuter	o - a == 0	2.300	0.703	0.010	**
neuter	u - o == 0	−1.745	0.566	0.017	*

4. Comparing Different Frameworks

In this section, we show why Russian indeclinable nouns are an interesting dataset to test different theoretical approaches to morphological gender. We compare three approaches representing opposing views on the grammatical architecture: Distributed Morphology (DM), Relational Morphology (RM), and Optimality Theory (OT). We compare both classic and probabilistic versions of these approaches. The crucial properties of the theories relevant for our data are (i) what kind of information is available when gender assignment takes place and (ii) whether the model predicts one outcome or several possible outcomes of different probability.

First, let us summarize some interesting properties of Russian indeclinable nouns. Crosslinguistically, relying on semantic analogy and morphophonology in assigning gender to loanwords is not uncommon (e.g., Corbett 1991: 75–82). The case of Russian indeclinables is interesting because most other examples discussed in the literature are either properly incorporated into the native inflectional system or come from languages with poor inflectional morphology in which this problem does not arise. Russian also has numerous loanwords of the first type. For example, the English word *bug* (in the meaning 'a software error') was adopted into Russian in two versions: *bag* (class IIa masculine) and *baga* (class I feminine). The first version, which is more frequent, is closer to the original in the nominative singular, while the second version probably appeared due to semantic analogy with the word *ošibka* 'error.F'. The choice of the feminine gender predetermined the choice of the inflectional class. Examples like *baga* with an added affix are rare in Russian, while examples like *vendetta* 'vendetta.F' borrowed from Italian are frequent; in such loanwords, *-a* is reanalyzed as a class I inflectional affix, so they become feminine.

Indeclinables are different in three respects. Firstly, most lexemes demonstrate gender variation, which may be extensive: some nouns, like *xudi* 'hoodie', have comparable frequencies of all three genders in our dataset. Secondly, once an inanimate word is borrowed and becomes declinable, its semantics does not influence its gender, while in indeclinables, speakers continue to rely on this information. Some inanimate declinable nouns show gender variation, but it is never due to semantics—rather, class III loanwords like *šampun'* 'shampoo' migrate to the IIa class, and a gender change accompanies the change of declension.[13] Thirdly, as we noted above, loanwords like *vendetta* become feminine because *-a* is reanalyzed as a class I inflectional affix, and all inanimate nouns in this class are feminine. This pattern can be accounted for in most major morphological theories. In indeclinables, no reanalysis takes place, the last vowel belongs to the root, but the morphophonological analogy affects gender assignment nevertheless. In this section, we argue that its role is difficult to explain in several major frameworks.

4.1. Distributed Morphology

Structural theories mostly assume modularity and sequential processing, which restricts the availability of certain information at different derivation stages. Being a syntactic feature, gender must receive its value at an early derivation stage—when the syntactic tree is constructed. But phonological information is not available at this stage. This raises a problem for our results as we have shown that both phonological and semantic factors influence gender assignment and give rise to variation, which is especially difficult to account for. This problem is relevant for all sequential approaches; let us consider the most influential one, Distributed Morphology, as an example.

13 Another small group of nouns that show gender variation are nouns with diminutive and augmentative affixes: their gender depends on the gender of the base noun as well as on their inflectional class (e.g., Steriopolo et al. 2021 and Magomedova and Slioussar 2023). Thus, this variation is also not due to semantics.

Kramer (2020) claims that gender is assigned by the Merge operation when a bare root or a stem is merged with a nominal head already containing a gender feature. Declension is also a syntactic feature situated lower in the structure. Therefore, it is easy to explain how declension can influence the choice of gender—like in Russian declinable loanwords and many other cases—but very difficult to explain the behavior of Russian indeclinables.

Several other authors working in the Distributed Morphology framework aim to "decompose" Russian declensions, reducing them to combinations of several features. For example, Privizentseva (2023) suggests the following system: class I is [+fem][−α] (the small group of animate masculine nouns in this class receives a separate treatment), class IIa is [−fem][+α], class IIb is [−fem][−α], and class III is [+fem][+α], where α is a declension feature and *fem* a gender feature. Importantly, in all approaches, formal or functional, including Kramer (2020) and Privizentseva (2023), gender features are realized by inflections and not by roots.[14] Therefore, there is no way to explain why an indeclinable noun with a root-final *-a* is more likely to be feminine—unless the process can refer to the surface form.

4.2. Relational Morphology

Relational Morphology (e.g., Jackendoff and Audring 2020) postulates a continuum between the lexicon and the grammar: a unified space of schemas, where a schema is a multilayer representation of a real or possible word. Different layers contain semantic, syntactic, and phonological information. Different parts of this information are connected by indices. An example of a schema for the noun *reader* is given in (1).

(1) Semantics: $[\text{PERSON}_a; [\text{READ (a, Y)}]_1]_2$
Morphosyntax: $[\text{N } [\text{V}]_1 \text{ aff}_3]_2$
Phonology: $/\text{ri:d}_1 \text{ ər}_3/_2$

Schemas are connected to each other by the "same-same" relations: relations point to a part of a given schema that can also be found in other schemas. One of the basic principles of Relational Morphology is the availability of all information in a schema and its relations at any given moment. This allows us to account for the influence of semantic and phonological factors at once.

However, to be able to predict gender distribution, it is still necessary to look inside the root and to connect the root-final segment in indeclinables to

[14] In Caha's (2021) approach, different roots spell out different sets of features—he uses this to explain why they select different inflections (in order to get rid of arbitrary declension features). Unfortunately, this approach also does not predict that roots ending in particular vowels would be more likely to be assigned a particular gender.

inflectional affixes. We cannot do this (or properly formalize such operations) as each part of the schema must be substitutable with a variable—this guarantees the possibility to generate new words in this approach. It is possible to substitute a number of entities with a single variable, but it is not clear what would be the theoretical basis to divide a root into smaller structural pieces.

Notably, Doleshal (2000) developed a model of Russian gender that is similar in spirit to Relational Morphology. Her model even accounts for variation, although the distribution of variants cannot be predicted. However, Doleshal does not analyze indeclinables, so the problems outlined above are also relevant for her approach.

4.3. OT, Harmonic Grammars, and Maximum Entropy

In optimality-theoretic (OT) approaches, all kinds of information are available at the same time and thus can be used for gender assignment. It is also crucial for our data that these approaches do not require morphological reanalysis of a loanword to make morphophonological analogies possible. The derivation in this framework is a process of selecting the best candidate from a theoretically infinite set of possible realizations. All candidates are tested against a set of constraints. In the classic version of OT, these constraints are ranked, while in OT-based probabilistic theories, such as Harmonic Grammars and Maximum Entropy models, they are weighted. OT models of Russian gender not taking variation into account were developed by Rice (2005, 2006) and Galbreath (2010). In this study, we turn to probabilistic models because they allow for analysis of variation.

5. Modeling Our Data: Comparing Different OT Approaches

In the previous section, we demonstrated that our data can be analyzed only using OT-based approaches. To account for variation, we opted for probabilistic Maximum Entropy grammars over a strict OT grammar. Since it is possible to build statistical models based on these approaches, to choose the optimal one, we used them to model our data and determined which models have the largest predictive power.

Maximum Entropy grammars are essentially logistic regressions (Hayes and Wilson 2008), where betas (or estimates, see the comments for Table 2) are the weights of constraints and binary independent variables are the constraints themselves. These models were shown to capture many wider grammatical generalizations but were not ideal for predicting the idiosyncratic behavior of different lexemes: to do so, we would need to introduce different weights of the same constraint for specific words. Zymet (2019) used mixed-effects regressions to capture this idiosyncratic behavior, which he introduces to the theory as Hierarchical MaxEnt.

The predictive power of the models is generally the reason why we trust statistics, but choosing between different statistical methods can also be seen as a part of the theoretical analysis in the case of probabilistic approaches. In this section, we compare the predictive power of mixed-effects and fixed-effects regressions and conclude that Zymet's Hierarchical MaxEnt approach accounts for our data more precisely. We also discuss the levels of the Hierarchical MaxEnt: what kind of behavior should be accounted for using constraints and what should be treated as idiosyncratic?

Our mixed-effects models treated each indeclinable noun as a random effect because each lexeme may have a different predisposition for gender assignment. For example, semantic analogy is more salient for some nouns than for others, etc. Speakers were not included as random effects, as we mostly had only one instance per lexeme from a speaker in our dataset. To separate language-wide factors (fixed effects) from idiosyncratic behavior (random effects) we excluded feature values that could be found only in a very few nouns (five or less). Namely, we did not introduce *m/n* and *f/n* values for the semantic analogy factor—they were included in the *?* group. Our motivation was simple: if we only have a couple of words with a certain value, it is impossible to tell whether they are different because of this feature value or because of some idiosyncrasy.[15]

We also tried including several minor factors as fixed effects. For example, we discussed in footnote 10 that nouns denoting car brands are more likely to be used with feminine agreement than any other group of nouns, so we included this as a separate semantic factor. Other minor factors we considered were word length and word-final hiatus. However, these factors did not improve the models, which indicates that such properties of smaller noun sets are better treated as idiosyncratic.[16]

Now let us focus on the comparison between two sets of models: three mixed-effects models and three models with fixed effects alone. In both sets, only four major factors were included as fixed effects: the semantic analogy gender, root-final vowel, lexical stress, and case. As we already mentioned in §3, we tried including interactions between different factors, but the resulting models were inferior according to the tests. All information on these six models, as well as pairwise comparisons for non-binary factors, can be found in

[15] To test this intuition, we built the models including all possible feature values, but they had smaller predictive power. We did not include them in the comparisons reported below for reasons of space.

[16] We tested this using the methods discussed in this section, but we cannot report all the results for reasons of space.

the Appendix (https://osf.io/wsuaz/), while below we discuss different estimates of their performance.

Indeclinable nouns are a limited set of words, so we tested the models on the same data that they were trained on. In particular, we could not afford to divide the dataset into test and training sets because mixed effects were calculated by word, and one and the same word could not appear in both sets. Unfortunately, this approach has its pitfalls: we cannot detect overfitting. As we show below, to control for this, we relied on the Akaike and Bayesian Information Criteria.

First of all, we used the pROC package (Robin et al. 2011) for ROC-AUC metrics. The ROC (Receiver Operating Characteristic) curve and the AUC (Area Under the Curve) metrics are widely used for evaluating the performance of binary classification models. The ROC curve is a graphical representation of the true positive rate (sensitivity) against the false positive rate (specificity) for a binary classifier system as its discrimination threshold is varied. The AUC refers to the area under the ROC curve. It provides a scalar value that characterizes the overall performance of the model across all thresholds. Figure 5 shows the ROC-AUC metrics for our models. Mixed-effects models are better than fixed-effects models for neuter (89% vs. 79%), slightly better for feminine (97% vs. 94%), and the same for masculine (82% vs. 82%).

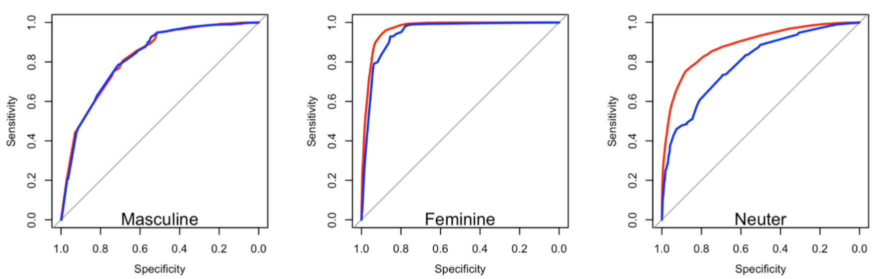

Figure 5. The performance of fixed- and mixed-effects models (shown in blue and in red, respectively)

To estimate the performance of our models in terms of the maximum likelihood, errors, and complexity, we used the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the adjusted, marginal and conditional r-squared (R^2), and the root mean squared error (RMSE), calculated with the help of the performance package (Lüdecke et al. 2021). The information-theoretic indicators are based on maximal likelihood (ML) models because restricted maximal likelihood (REML) models are not available for binomial logistic regressions in the *lme4* package. As Table 4 shows, mixed-effects models are superior for all three genders according to all these criteria.

Table 4. Performance estimates for mixed- and fixed-effects models

Model	AIC	BIC	RMSE	marginal R^2	conditional R^2	adjusted R^2
Masculine: mixed effects	26202	26337	0.353	0.360	0.666	
Masculine: fixed effects	33159	33251	0.412			0.299
Feminine: mixed effects	9598	9732	0.212	0.571	0.796	
Feminine: fixed effects	12178	12271	0.243			0.414
Neuter: mixed effects	27969	28103	0.365	0.301	0.644	
Neuter: fixed effects	36019	36111	0.430			0.252

The RMSE criterion shows the square root from the squared sum of residuals divided by the number of observations, where a residual is a distance from a datapoint to the regression curve. The smaller this number is, the better the model. As we can see in Table 4, mixed-effects models are more precise by a large margin. R^2 is a more sophisticated criterion based on the same principle; it is calculated as the sum of squared differences between each actual datapoint and the model prediction divided by the sum of squared differences between each actual datapoint and the mean of all datapoints, all subtracted from one. Therefore, the bigger this number is, the better the model. Marginal R^2 shows how fixed effects alone account for variance in a mixed-effects model, while conditional R^2 shows an estimate for both fixed and mixed effects. As we can see in Table 4, conditional R^2 is larger than marginal R^2 for every model. In addition to that, adjusted R^2 calculated for fixed-effects models are always smaller than both marginal and conditional R^2 calculated for mixed-effects models.

However, both RMSE and R^2 criteria are known to prefer more complex models and do not penalize for overfitting. To take this problem into account, we used the Akaike and Bayesian Information Criteria. The AIC and BIC are based on the likelihood function and are almost the same, but BIC penalizes complex models more. The absolute value of both criteria is of no importance for the model selection; these numbers are used relatively—the smallest number among all the models trained on the same data indicates the best model. As we can see in Table 4, both criteria strongly prefer mixed-effects models for all three genders. There may be some further complications in comparing mixed- and fixed-effects models (e.g., see McNeish and Kelley 2019), which makes us interpret the results in Table 4 with a certain caution. However, this does not apply to the ROC-AUC metrics in Figure 5, which show the superiority of mixed-effects models very clearly.

We can conclude that our results support the Hierarchical MaxEnt approach (Zymet 2019) over the ordinary MaxEnt one. The feminine gender is predicted more successfully than masculine and neuter, which is probably because it is assigned to indeclinables only in the presence of very strong cues (a stem-final *-(j)a* or an obvious semantic analogy word of feminine gender). This makes the feminine gender much less frequent in this group of nouns than masculine and neuter, and at the same time, easier to predict. Interestingly, if we look at pairwise comparisons for the semantic analogy and final vowel factors in the Appendix (https://osf.io/wsuaz/), we can see that fixed-effects models give a larger number of significant results than mixed-effects ones. Since the performance of fixed-effects models was inferior according to all used criteria, we can conclude that this significance is in fact misleading. This shows once again how important it is to compare different approaches to choose an optimal one.

6. Conclusions

We analyzed grammatical gender assignment in a set of 32,792 instances of 131 indeclinable inanimate nouns from the General Internet Corpus of the Russian Language. We found a substantial variation in gender for the majority of nouns and identified several contributing factors: the gender of the semantic analogy noun and the root-final vowel, as well as the lexical stress and case. The first two factors were noted in earlier studies, but their role has never been estimated on corpus data or discussed in the context of different morphological frameworks.

We believe that Russian indeclinables are noteworthy because in declinable inanimate Russian nouns, the gender does not depend on semantics or on the properties of the stem—rather, it strongly correlates with the inflectional class. When a new loanword becomes declinable, its semantics can sometimes affect the choice of declension and gender, and its phonology very often does (for example, the final *-a* is usually reanalyzed as a class I nominative singular affix). But once the loanword has been incorporated in the Russian nominal system, these factors stop playing a role, unlike in indeclinable nouns. It is especially remarkable that the final vowel in indeclinables influences gender assignment although it does not get reanalyzed as an affix and remains part of the root.

We argued that indeclinable nouns are an interesting dataset for comparing different morphological frameworks. We selected three approaches representing opposing views on the grammatical architecture: Distributed Morphology, Relational Morphology, and Optimality Theory, for which classic and probabilistic versions were compared. We identified two properties of these approaches as crucial for our dataset: (i) what kind of information is available when gender assignment takes place and (ii) whether the model predicts one outcome or several possible outcomes of different probability. As for (i), in DM, no phonological information is available when gender assignment takes place. In RM, phonological information may be accessed, but not inside the root. In OT, constraints can take any information into account. As for (ii), only probabilistic OT approaches can deal with multiple outcomes of different probability.

Finally, within the family of probabilistic OT approaches, we aimed to select the one that would be optimal for our data, accounting both for some general tendencies and for some idiosyncratic properties of small noun groups or individual words. Since it is possible to build statistical models based on these approaches, we used them to model our data and determined which models have the largest predictive power. As a result, we demonstrated that the hierarchical Maximum Entropy (MaxEnt) models (Zymet 2019) were superior to the classical MaxEnt models.

References

Aronoff, Mark. (1994) *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT press.

Bates, Douglas, Martin Maechler, Ben Bolker, and Steve Walker. (2015) "Fitting linear mixed-effects models using lme4". *Journal of statistical software* 67(1): 1–48.

Bretz, Frank, Torsten Hothorn, and Peter Westfall. (2010) *Multiple comparisons using R*. Boca Raton, FL: CRC Press.

Caha, Pavel. (2021) "Modeling declensions without declension features: The case of Russian". *Acta linguistica academica* 68(4): 385–425.

Chuprinko, Kirill, Varvara Magomedova, and Natalia Slioussar. (2023) "Gender variation in indeclinable inanimate nouns and gender markedness in modern Russian". *Acta linguistica academica* 70(3): 317–38.

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

Corbett, Greville. (1991) *Gender*. Cambridge, UK: Cambridge University Press.

———. (2023) "The dog didn't bark, the noun didn't inflect: A typology of significant absences". Paper presented at DGfS (45th annual conference of the German Linguistic Society), Cologne, Germany, March 2023.

Corbett, Greville, and Norman M. Fraser. (2000) "Gender assignment: A typology and a model". Gunter Senft, ed. *Systems of nominal classification*. Cambridge, UK: Cambridge University Press, 293–325.

Doleschal, Ursula. (2000) "Gender assignment revisited". Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen, and Mirja Saari, eds. *Gender in grammar and cognition. Part I: Approaches to gender*. Berlin: Mouton de Gruyter, 117–66.

Elff, Martin. (2022) "mclogit: Multinomial logit models, with or without random effects or overdispersion". R package version 0.9.6. Available at: https://CRAN.R-project.org/package=mclogit. Last accessed 10 December 2022.

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*. PhD dissertation, University of Virginia, Charlottesville, Virginia.

Halle, Morris. (1994) "The Russian declension: An illustration of the theory of Distributed Morphology". Jennifer Cole and Charles Kisseberth, eds. *Perspectives in phonology*. Stanford, CA: CSLI Publications, 29–60.

Hayes, Bruce, and Colin Wilson. (2008) "A maximum entropy model of phonotactics and phonotactic learning". *Linguistic inquiry* 39(3): 379–440.

Henry, Louise. (2020) "Morfologičeskij status slova *pal′to* v russkom jazyke: Korpusnoe issledovanie". *Movna osobistist′: Lingvistika i lingvodidaktika* 5: 350–56.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. (2008) "Simultaneous inference in general parametric models". *Biometrical journal* 50(3): 346–63.

Jackendoff, Ray, and Jenny Audring. (2020) *The texture of the lexicon: Relational Morphology and the Parallel Architecture*. Oxford: Oxford University Press.

Kramer, Ruth. (2015) *The morphosyntax of gender*. Oxford: Oxford University Press.

———. (2020) "Grammatical gender: A close look at gender assignment across languages". *Annual review of linguistics* 6: 45–66.

Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, Philip Waggoner, and Dominique Makowski. (2021) "performance: An R package for assessment, comparison and testing of statistical models". *Journal of open source software* 6(60): 3139. DOI 10.21105/joss.03139.

Magomedova, Varvara, and Natalia Slioussar. (2023) "Gender variation and gender markedness in Russian nouns". *Voprosy jazykoznanija* 2: 7–28.

McNeish, Daniel, and Ken Kelley. (2019) "Fixed effects models versus mixed effects models for clustered data: Reviewing the approaches, disentangling the differences, and making recommendations". *Psychological methods* 24(1): 20–35.

Mjakilja, Kari. (2000) "K probleme roda nesklonjaemyx zaimstvovannyx imen naricatel′nyx v sovremennom russkom jazyke". *Scando-Slavica* 46: 93–103.

Mučnik, Iosif. (1971) *Grammatičeskie kategorii glagola i imeni v sovremennom russkom literaturnom jazyke*. Moscow: Nauka.

Murphy, Dianna L. (2000) *The gender of inanimate indeclinable common nouns in modern Russian*. PhD dissertation, Ohio University, Athens, Ohio.

Poplack, Shana, Alicia Pousada, and David Sankoff. (1982) "Competing influences on gender assignment: Variable process, stable outcome". *Lingua* 57(1): 1–28.

Privizentseva, Mariia. (2024) "Semantic agreement in Russian: Gender, declension, and morphological ineffability." *Natural language and linguistic theory* 42: 767–814.

R Core Team. (2021) "R: A language and environment for statistical computing". Vienna, Austria: R Foundation for Statistical Computing. Available at https://www.R-project.org/.

Rice, Curt. (2005) "Optimizing Russian gender: A preliminary analysis". Steven Franks, Frank Gladney, and Mila Tasseva-Kurktchieva, eds. *Formal Approaches to Slavic Linguistics #13: The South Carolina Meeting.* Ann Arbor, MI: Michigan Slavic Publications, 265–75.

———. (2006) "Optimizing gender". *Lingua* 116(9): 1394–417.

Robin, Xavier, Natacha Turck, Alexandre Hainard, Natalia Tiberti, Frédérique Lisacek, Jean-Charles Sanchez, and Markus Müller. (2011) "pROC: An open-source package for R and S+ to analyze and compare ROC curves". *BMC bioinformatics* 12, article 77. DOI 10.1186/1471-2105-12-77

Rozental′, Ditmar, Evgenija Džandžakova, and Natal′ja Kabanova. (1998) *Spravočnik po pravopisaniju, proiznošeniju, literaturnomu redaktirovaniju*. Moscow: ČeRo.

Savchuk, Svetlana. (2011) "Korpusnoe issledovanie variantov rodovoj prinadležnosti imen suščestvitel′nyx v russkom jazyke". *Computer linguistics and intellectual technologies* 10: 562–79.

Shigemori Bučar, Chikako. (2011) "Creative competence in borrowings: Words of Japanese origin in Slovene". *Linguistica* 51(1): 245–62.

Shvedova, Natalia, ed. (1980) *Russkaja grammatika.* Moscow: Nauka.

Slioussar, Natalia, and Maria Samoilova. (2015) "Častotnosti različnyx grammatičeskix xarakteristik i okončanij u suščestvitel′nyx russkogo jazyka". In *Proceedings of the conference "Dialogue"*. Available at: www.dialog-21.ru/digests/dialog2015/materials/pdf/SlioussarNASamoilovaMV.pdf.

Steriopolo, Olga, Giorgos Markopoulos, and Vassilis Spyropoulos. (2021) "A morphosyntactic analysis of nominal expressive suffixes in Russian and Greek". *The linguistic review* 38(4): 645–86.

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". *Canadian Slavonic papers* 25(1): 180–205.

Violin-Wigent, Anne. (2006) “Gender assignment to nouns codeswitched into French: Observations and explanations”. *International journal of bilingualism* 10(3): 253–76.

Wang, Qiang. (2014) “Gender assignment of Russian indeclinable nouns”. Master’s thesis, University of Oregon, Eugene, OR.

Zaliznjak, Andrej A. (1987) *Grammatičeskij slovar' russkogo jazyka: Slovoizmenenie*. 2nd ed. Moscow: Russkij Jazyk.

Zymet, Jesse. (2019) “Learning a frequency-matching grammar together with lexical idiosyncrasy: MaxEnt versus Hierarchical Regression”. Katherine Hout, Anna Mai, Adam McCollum, Sharon Rose, and Matthew Zaslansky, eds. *Proceedings of the 2018 Annual Meeting on Phonology*. Available at: https://journals.linguisticsociety.org/proceedings/index.php/amphonology/article/view/4495/4198.

Varvara Magomedova
Independent researcher
Warsaw, Poland
varya.magomedova@gmail.com

Kirill Chuprinko
Center for Cognitive Science of Language
University of Nova Gorica
Nova Gorica, Slovenia
chuprinko_kirill@mail.ru

Natalia Slioussar
School of Linguistics
HSE University
Moscow, Russia
slioussar@gmail.com