Computational Linguistics and Intellectual Technologies:
Proceedings of the International Conference “Dialogue 2018”

Moscow, May 30—June 2, 2018

GENDER, DECLENSION AND STEM-FINAL CONSONANTS: AN EXPERIMENTAL STUDY OF GENDER AGREEMENT IN RUSSIAN

Slioussar N. A. (slioussar@gmail.com)
Higher School of Economics, Moscow, and Saint-Petersburg State University, Saint-Petersburg, Russia

Every adult native speaker of Russian knows that *kon’* is masculine and *lan’* is feminine, although 3rd declension nouns present some difficulties in the first and second language acquisition. However, will the fact that these nouns are less frequent than masculine nouns ending in a consonant or feminine nouns ending in *-a/ja* play a role for online subject-predicate agreement processing? Or will subject-predicate agreement processing be more problematic with subjects of a certain gender? Finally, some final consonants are more characteristic for feminine gender, while the others for masculine gender. Are speakers sensitive to this? We present two experiments addressing these questions. We found that all three factors play a role, but for different tasks (online agreement processing or determining the gender of a novel word) and at different processing stages.

Keywords: grammatical gender, declension, experimental, Russian

РОД, СКЛОНЕНИЕ И КОНЕЧНЫЙ СОГЛАСНЫЙ ОСНОВЫ: ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ СОГЛАСОВАНИЯ ПО РОДУ В РУССКОМ ЯЗЫКЕ

Слюсарь Н. А. (slioussar@gmail.com)
НИУ ВШЭ, Москва, и СПбГУ, Санкт-Петербург, Россия

1. Introduction

The gender of many Russian nouns cannot be determined from their inflectional affixes. But nouns are not distributed evenly among genders and declensions. Firstly, masculine nouns are in general more frequent than feminine and neuter (48% vs. 35% and 17% forms in the grammatically disambiguated subcorpus of the Russian National Corpus (RNC, http://www.ruscorpora.ru)). Secondly, a nominative singular form ending in *-a/ja* or with a zero inflection may be masculine or feminine, but the former is much more likely to be feminine, while the latter is much more likely to be masculine (more details about are given below). Thirdly, if the final consonant of a noun with a zero inflection is taken into account, we will see that, for example, most nouns ending in *-s'* are feminine, while most nouns ending in *-r'* are masculine.

In this paper, we present two experimental studies exploring whether native speakers of Russian are sensitive to the distributional properties outlined above. The first experiment is dedicated to the online processing of gender agreement. In the second pilot experiment, we study how participants determine the gender of real and nonce nouns.

Now let us discuss the relevant properties of nouns in more detail. Table 1 shows the distribution of nouns among genders and declensions in the grammatically disambiguated subcorpus of the RNC. Masculine nouns ending in *-a/ja* and feminine nouns with a zero inflection are termed non-prototypical due to their low relative frequency. These nouns are known to be problematic for the L1 and L2 acquisition (e.g. [Janssen, 2016]; [Rodina & Westergaard, 2012]; [Schwartz et al., 2015]; [Tseitlin 2000]). But [Rusakova 2013] who studied naturally occurring errors in spoken Russian found that adult native speakers do not make more gender agreement errors with such nouns[1]. Still, what does not show up in the number of errors may influence online sentence processing. Experiment 1 is dedicated to this question.

In this paper, we focus on consonant-final nouns. The distributional picture gets more complex if the nature of the final consonant is taken into account. The gender of the words ending in /ž/, /š/, /č'/, /š'č'/ can be determined orthographically, while words ending in other non-palatalized consonant can be only masculine. No such clues are available for the words ending in other palatalized consonants, but, as Table 2 shows, their distribution between the two genders varies greatly (to represent both the relative frequencies of different forms and the number of existing lemmas we relied not only on the RNC, but also on the *Grammatical Dictionary of the Russian Language* [GDRL, Zaliznjak, 1987]).

The majority of nouns ending in labials are feminine. Most nouns ending in *-r'* and *-l'* are masculine, many of them have agentive suffixes like *-tel'* or *-ar'*. Nevertheless, feminine nouns are not just singular cases in this group. Forms of nouns ending in other lingual consonants are either predominantly feminine or are evenly distributed between the two genders (but the number of feminine lemmas is still larger). The majority of nouns ending in *-t'*, which are especially numerous, have the suffix *-ost'*. In our pilot experiment 2, we explored whether adult native speakers are sensitive to these differences.

[1] Cases where gender variation is observed are not taken into account (a corpus-based study of such cases was conducted by [Savchuk 2011]).

Table 1. The distribution of nouns among genders and declensions in the grammatically disambiguated subcorpus of the RNC[2]

Declension and gender	Percentage of nouns in the RNC	Ending in Nom.Sg and prototypicality	Examples
1st decl. feminine	29% nouns	end in *-a/ja,* 'prototypical F'	*zhena* 'wife'
1st decl. masculine	1% nouns	end in *-a/ja,* 'non-prototypical M'	*djadja* 'uncle'
2nd decl. masculine	46% nouns	end in a consonant, 'prototypical M'	*syn* 'son', *gel'* 'gel'
2nd decl. neuter	18% nouns	end in *-o/e,* 'prototypical N'	*pole* 'field'
3rd decl. feminine	5% nouns	end in a consonant, 'non-prototypical F'	*mel'* 'shallow'
irregular and indeclinable	1% nouns		

Table 2. Nouns in the grammatically disambiguated subcorpus of the RNC and in the GDRL

Final consonant	RNC (Nom.Sg forms)		GDRL (lemmas)	
	M	F	M	F
/b'/	34 (24%)	110	1	11
/p'/	3 (2%)[3]	169	—	19
/v'/	13 (1%)	1,448	1	20
/f'/	0	2	—	2
/m'/	0	16	—	3
/d'/	748 (51%)	707	10	55
/t'/	713 (5%)	13,184	17	3414
/z'/	319 (49%)	327	7	34
/s'/	80 (14%)	491	5	57
/n'/	2,354 (45%)	2,842	126	112
/r'/	2,160 (76%)	677	177	34
/l'/	6,648 (71%)	2,653	1083	215

2. Previous experimental studies

Two groups of experimental studies are relevant for the present paper: analyzing gender agreement and nouns with more or less morphologically regular inflections.

[2] The counts are taken from [Slioussar and Samoilova 2015]. Substantivized adjectives were not taken into account.

[3] These three forms are *rup'* (a reduced form of the noun *rubl'* 'ruble').

There are relatively few experimental studies of gender agreement in Russian. In three of them ([Akhutina et al. 1999, 2001]; [Romanova & Gor 2017]) adjectives were presented before nouns audially or visually. In congruent conditions, adjectives agreed with the following nouns, in incongruent ones they did not, and some experiments also included a baseline condition where bare adjective stems without inflections or adverbs were presented. Several methods were employed, including lexical decision (answering whether the presented stimulus is a real word or a nonce word), grammaticality judgment (answering whether the presented fragment is grammatical) and cued-shadowing in which participants must repeat the second presented word (the target noun).

However, the question was always the same: would participants answer significantly faster and more accurately in congruent conditions compared to incongruent ones, and would there be any differences associated with the gender of the nouns? In experiments with a baseline condition, it was also possible to check whether the difference between congruent and incongruent conditions was primarily due to facilitation in the former, or to inhibition in the latter, or both effects were equally prominent. In brief, [Akhutina et al. 2001] observed significant facilitation and inhibition effects for feminine nouns, while for masculine nouns, only inhibition was significant, and for neuter ones, only facilitation was significant[4]. Results from other studies were similar.

The explanations offered in these studies go along the same lines. Masculine gender as the most frequent is assumed to be unmarked, or default, while neuter is considered the most marked. Thus, masculine is expected by default, and strengthening this expectation by a masculine adjective does not produce a big difference (hence no significant facilitation effects). Neuter is the least expected option, so priming a neuter noun with a neuter adjective has the largest effect compared to the baseline condition (hence facilitation effects for neuter nouns are larger than for feminine nouns). Inhibition effects are explained by rechecking, which is especially costly for masculine nouns presented after non-masculine adjectives.

None of these three studies looked at 3rd declension feminine nouns, while the experiments by [Taraban and Kempe 1999] specifically focused on them. Taraban and Kempe selected masculine and feminine nouns ending in a palatalized consonant (opaque condition) and in non-palatalized consonants or in *-a/ja*, which are unambiguously masculine or feminine (transparent condition). They examined the role of such transparency for subject–predicate agreement using word-by-word self-paced reading and forced choice tasks. Participants were asked to read sentence beginnings like (1a) or (2a) and then to select one of the two verb forms in the remaining fragment like (1b) or (2b). In some conditions, sentence fragments contained adjectives. Participants were adult native speakers and L2 learners. For native speakers, transparency and the presence of a gender-marked adjective did not play any role.

(1)	a.	*Daže*	*(obyčnaja)*	*muka/sol'*	*teper'…*
		even	ordinary$_{F}$	flour$_{F.1D}$/salt$_{F.3D}$	now
	b.	*isčez/isčezla*	*iz*	*magazinov.*	
		disappeared$_{M/F}$	from	stores	

[4] This study also involved aphasiac patients, while [Romanova and Gor 2017] compared native speakers to second language learners, but we will not discuss these groups here.

(2)	a.	*Nakanune*	*(otëkšij)*	*palec/lokot'*	*sil'no…*
		the-day-before	swollen$_{M}$	finger$_{M.2D}$/elbow$_{M.2D}$	strongly
	b.	*bolel/bolela*	*ot*	*udara.*	
		hurt$_{M/F}$	from	injury	

[Slioussar and Malko 2016] studied gender agreement attraction. To give an example, an attraction error is present in the English sentence "The key to the cabinets are rusty", where the verb agrees not with the head of the subject phrase, but with another noun, termed *attractor*. In production, such errors are more frequent than agreement errors without attraction. In comprehension, they are missed more often and produce smaller delays in reading times and less pronounced ERP responses.

Number agreement attraction is widely discussed in the literature, while gender agreement has been analyzed only in a few studies so far. Among other things, it was noted that both in production and in comprehension, attraction effects can be observed in the sentences with singular heads and plural dependent nouns (e.g., "The key to the cabinets…"), but not in the sentences with plural heads and singular dependent nouns (e.g., "The keys to the cabinet…"). Almost all proposed explanations appeal to feature markedness, although approaches to markedness may be very different, from representational to frequency-based. Looking for similar asymmetries in gender agreement attraction, several studies of Romance languages obtained controversial results (e.g. [Acuña-Fariña et al., 2014]; [Anton-Mendez et al., 2002]; [Martin et al., 2014]; [Vigliocco & Franck, 1999]). [Badecker and Kuminiak 2007] found that neuter behaves as unmarked in a series of production experiments on Slovak, in which neuter is the least frequent gender, but is used in impersonal sentences, like in Russian.

[Slioussar and Malko 2016] conducted one production and three comprehension experiments. The results of the former were similar to the Slovak study, while in the latter, masculine behaved differently from feminine and neuter. Namely, attraction was observed for all dependent noun genders, but only for neuter and feminine heads. In other words, masculine heads were significantly more resistant to attraction: readers detected agreement errors irrespective of possible attractors' interference[5].

This result can be reconciled with the observations made in [Akhutina et al. 1999, 2001]; [Romanova & Gor 2017]. However, given that different patterns were observed for production and comprehension, we cannot explain them by a particular single property of gender features anymore. This reminds us that the notion of markedness usually invoked to explain all asymmetries between features is problematic because some studies rely on representational markedness (primarily counting the number of positive feature values in formal morphological models), the others consider the most frequent value to be the default etc. From the representational point of view, neuter is the unmarked gender in most accounts, while if we rely on frequency, masculine is. Maybe, these approaches should be seen as complementary, because different properties of features appear to be relevant in different experimental tasks.

[5] It is traditionally assumed that the features of the dependent noun are crucial for attraction, but both this study and some other findings suggest that the features of the head might be more important. We will not discuss this problem here.

Now let us turn to the problem of inflections that are more or less typical for a particular gender. Many experimental studies used different tasks (for example, determining a noun's gender or selecting an article) to show that nouns with more typical inflections are associated with faster and more accurate answers. This was demonstrated for Italian, French, Hebrew, Bulgarian etc. (e.g. [Andonova et al., 2004]; [Bates et al., 1995]; [Gollan & Frost, 2001]; [Spalek et al., 2008]).

However, fewer studies investigated the role of this factor in sentence processing. To give an example, Caffarra et al. (2015) looked at Italian nouns with more and less typical endings presented in the same sentences. They were preceded by articles, which carried gender information. Nouns from the two groups elicited different ERP responses. [Franck et al. 2008] and [Vigliocco and Zilli 1999] demonstrated for Italian, Spanish, and French that heads with regular inflections are more resistant to agreement attraction. The same is true for number agreement attraction (e.g.[Bock & Eberhard, 1993]; [Vigliocco et al., 1995]).

As for Russian, only [Taraban and Kempe's 1999] study discussed above addressed this problem. The authors found no differences between prototypical and non-prototypical subject nouns. However, the task they used required selecting a correctly agreeing verb form, which is a less immediate measure than simple reading times. For this reason, we decided to come back to this factor in the present study.

3. Experiment 1

This experiment was conducted together with Pavel Shilin, an MA student at Saint-Petersburg State University. The goal of the experiment was to find out whether subject–predicate gender agreement is processed differently depending on the declension and gender the subject noun belongs to. 33 native speakers of Russian aged 18–55 (12 women) took part in the experiment.

It is impossible to test all potentially interesting combinations of gender and declension in one experiment, so we selected the following three groups of inanimate nouns as subjects: masculine nouns of the 2nd declension (prototypical masculine, 2DM), feminine nouns of the 1st declension (prototypical feminine, 1DF), and feminine nouns of the 3rd declension (non-prototypical feminine, 3DF). The materials included 36 sets of target sentences in six conditions exemplified in (3a-f). All sentences had the same structure: a subject noun, *byl/byla* 'was$_{M/F}$', an adjective or participle, and a three-word PP.

(3) a. **2DM G(rammatical):** *Xalat* *byl* *potrepannym ot mnogoletnej noski.*
robe$_{M.NOM.SG}$ was$_{M}$ shabby$_{M}$ from years-long wear

b. **2DM U(ngrammatical):** *Xalat* *byla* *potrepannoj…*
robe$_{M.NOM.SG}$ was$_{F}$ shabby$_{F}$

c. **1DF G(rammatical):** *Kurtka* *byla* *potrepannoj…*
jacket$_{F.NOM.SG}$ was$_{F}$ shabby$_{F}$

d. **1DF U(ngrammatical):** *Kurtka* *byl* *potrepannym…*
jacket$_{F.NOM.SG}$ was$_{M}$ shabby$_{M}$

e. **3DF G G(rammatical):** *Šinel'* *byla* *potrepannoj…*
overcoat$_{F.NOM.SG}$ was$_{F}$ shabby$_{F}$
f. **3DF U(ngrammatical):** *Šinel'* *byl* *potrepannym…*
overcoat$_{F.NOM.SG}$ was$_{M}$ shabby$_{M}$
'The robe / jacket / overcoat was shabby from being worn for many years.'

Half of the sentences contained gender agreement errors on the predicate because taking previous studies of agreement into account (primarily agreement attraction experiments), the effects could be expected to be different in grammatical and ungrammatical sentences[6]. Subject nouns in the three declension groups were balanced for frequency and length using the *StimulStat* lexical database (http://stimul.cognitivestudies.ru, [Alexeeva et al., 2018]). Frequency information in this database is taken from the *Frequency Dictionary of Modern Russian Language* [Lyashevskaya & Sharoff, 2009].

Target sentences were distributed into six experimental lists so that each participant saw only one sentence from each set. The lists also contained 80 grammatically correct filler sentences. The sentences were presented on a PC using Presentation software (http://www.neurobs.com). We used the word-by-word self-paced reading methodology. Each trial began with a sentence in which all words were masked with dashes while spaces and punctuation marks remained intact. Participants were pressing the space bar to reveal a word and re-mask the previous one. One third of the sentences were followed by forced choice comprehension questions to ensure that the participants were reading properly.

We analyzed participants' question-answering accuracy and reading times. On average, participants answered 12% questions to target sentences incorrectly, no participants made more than 3 errors. Reading times that exceeded a threshold of 2.5 standard deviations, by region and condition, were excluded [Ratcliff, 1993]. In total, 2.0% of the data were excluded as outliers. Average reading times per region in different conditions are presented in **Figure 1**.

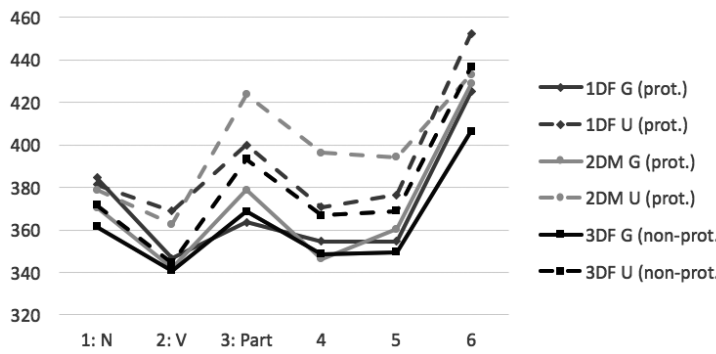

Figure 1. Average reading times per region (in ms) in different experimental conditions

[6] We selected predicates that consisted of a copula and an adjective or participle because such predicates were used in the previous experimental studies of subject—predicate gender agreement in Russian.

For each region, we made pairwise comparisons between the three conditions using a 2 × 2 Repeated Measures ANOVA with grammaticality and declension as factors. Analyses by participants (F_1) and by items (F_2) were performed. In region 1 (the subject noun), there were no significant results, which means that nouns in different conditions were properly balanced and that nouns of a particular gender or declension are not intrinsically more difficult to process.

Region 2 contains the verb *byl / byla* 'was$_{M/F}$'—this is where agreement errors appear in ungrammatical sentences. Figure 1 suggests that participants' reaction to these errors was different depending on the declension of the subject noun. In the conditions 1DF and 2DM (with prototypical feminine and masculine subjects), reading times in ungrammatical sentences are longer than in grammatical ones, while no such difference can be seen in the 3DF conditions (with non-prototypical feminine subjects), which indicates that the error tends to remain undetected in the latter case.

Statistical analyses support this intuition. In the comparison between 1DF and 3DF conditions, grammaticality and the interaction between declension and grammaticality are significant ($F_1(1,32) = 8.13$, $p < 0.01$, $F_2(1,35) = 4.20$, $p = 0.05$; $F_1(1,32) = 7.41$, $p = 0.01$, $F_2(1,35) = 4.05$, $p = 0.05$), while the main effect of declension does not reach significance. This means that the influence of grammaticality depends on the declension of the subject. In the comparison between 2DM and 3DF conditions, grammaticality reaches significance, while the interaction between declension and grammaticality is marginally significant ($F_1(1,32) = 8.01$, $p < 0.01$, $F_2(1,35) = 4.09$, $p = 0.05$; $F_1(1,32) = 3.98$, $p = 0.05$, $F_2(1,35) = 3.17$, $p = 0.08$). When 1DF and 2DM are compared, only the grammaticality factor is significant ($F_1(1,32) = 18.66$, $p < 0.01$, $F_2(1,35) = 10.21$, $p < 0.01$).

In region 3 that contains an adjective or participle, differences between grammatical and ungrammatical sentences become visible in all conditions. Accordingly, the grammaticality factor is significant in all pairwise comparisons ($F_1(1,32) = 15.90$, $p < 0.01$, $F_2(1,35) = 21.24$, $p < 0.01$ for 1DF vs. 2DM; $F_1(1,32) = 11.98$, $p < 0.01$, $F_2(1,35) = 6.20$, $p = 0.02$ for 1DF vs. 3DF; $F_1(1,32) = 9.73$, $p < 0.01$, $F_2(1,35) = 7.83$, $p < 0.01$ for 2DM vs. 3DF). No other factors or interactions reach significance.

Regions 4–6 contain a three-word PP. In region 4, a tendency that can be already detected in region 3 becomes statistically significant: the error-related delay in reading times is more pronounced in the 2DM conditions (with masculine subjects) than in the 1DF and 3DF conditions (with feminine subjects). In the comparison between 1DF and 2DM conditions, grammaticality and the interaction between declension and grammaticality are significant ($F_1(1,32) = 36.95$, $p < 0.01$, $F_2(1,35) = 15.91$, $p < 0.01$; $F_1(1,32) = 9.77$, $p < 0.01$, $F_2(1,35) = 6.45$, $p = 0.02$), while declension is not significant. The same is true for the comparison between 3DF and 2DM ($F_1(1,32) = 50.11$, $p < 0.01$, $F_2(1,35) = 13.17$, $p < 0.01$; $F_1(1,32) = 11.38$, $p < 0.01$, $F_2(1,35) = 5.51$, $p = 0.03$). When 1DF and 3DF are compared, only the grammaticality factor is marginally significant ($F_1(1,32) = 12.34$, $p < 0.01$, $F_2(1,35) = 3.65$, $p = 0.07$).

In region 5, only the grammaticality factor is significant in all pairwise comparisons ($F_1(1,32) = 18.51$, $p < 0.01$, $F_2(1,35) = 17.67$, $p < 0.01$ for 1DF vs. 2DM; $F_1(1,32) = 14.78$, $p < 0.01$, $F_2(1,35) = 6.10$, $p = 0.02$ for 1DF vs. 3DF; $F_1(1,32) = 18.07$, $p < 0.01$, $F_2(1,35) = 10.07$, $p < 0.01$ for 2DM vs. 3DF). In region 6, there are no significant differences.

Finally, let us note that when we planned the experiment, we did not consider assessing the role of the final consonant of 3DF nouns. But 10 out of 36 nouns we selected ended in *-l'* or *-r'*, which is more characteristic for masculine, while other nouns had final consonants characteristic for feminine. So the role of this factor could be estimated, and there were no hints of any relevant differences.

4. Pilot experiment 2

This pilot experiment was included in a study we conducted together with Varvara Magomedova (SUNY, Stony Brook) and Natalia Chuprasova, an MA student at Saint-Petersburg State University. The main goal of the study was to find out how Russian speakers determine the gender of real and nonce nouns with diminutive and augmentative suffixes. However, to make the materials more diverse, other nouns had to be included, and we selected 12 real and 12 nonce nouns ending in palatalized consonants (as well as some indeclinable nouns etc.).

Participants were 30 native Russian speakers (17 women), aged 19–30. They received a list of seven adjectives and then were presented with nouns one by one. They were asked to pick a matching adjective and pronounce the resulting phrase. Adjectives had meanings like 'big', 'small', 'cool', 'bad' etc., to make participants think that the experiment was about semantic connotations of different nouns.

Analyzing the gender of the adjectives selected by the participants, for 12 real nouns ending in palatalized consonants (6 masculine and 6 feminine) we found only 7 errors out of 360 responses. There were three errors with *žen'šen'* 'ginseng$_{M}$', two errors with *stupen'* 'step$_{F}$', and one error with *kisel'* 'starch drink$_{M}$' and with *prorub'* 'ice-hole$_{F}$'. The low number of errors agrees with the previous findings by Rusakova (2013): adult native speakers of Russian do not experience particular difficulties determining the gender of such nouns.

As for 12 nonce nouns, we had two examples with each of the following endings: *-b'*, *-d'*, *-s'*, *-n'*, *-l'* and *-r'*. The number of answers with masculine adjectives in these groups was 22 (out of 60), 26, 38, 30, 44 and 51, respectively. We can see that the nature of the final consonant played a role. To estimate it statistically, we used mixed effects logistic regression with random slopes and random intercepts by participants and by items. Only nonce nouns ending in *-l'* and *-r'* were significantly different from the other groups ($\beta = 2.48$, $SE = 1.15$, $z = -2.15$, $p = 0.03$; $\beta = 4.00$, $SE = 1.54$, $z = -2.60$, $p < 0.01$). Unlike all other palatalized consonants, these final consonants are more characteristic for masculine nouns.

5. Conclusions

In the introduction, we outlined three distributional properties of Russian nouns. Firstly, masculine gender is more frequent than feminine and neuter. Secondly, some combinations of genders and declensions are more frequent than the others (we called them prototypical). Thirdly, for nouns ending in palatalized consonants, some consonants are more characteristic for masculine nouns and the others for feminine nouns.

Previous studies indicate that these factors do not increase the number of naturally occurring gender agreement errors for adult native speakers of Russian [Rusakova 2013]. But speakers might still be sensitive to them, and in the present paper we demonstrated this in two experiments.

Experiment 1 revealed the role of the two first factors. It was not designed to assess the role of the third factor, but, as far as we could estimate from the data, this factor does not play a role in online agreement processing. However, the influence of this factor can be detected when speakers try to guess the gender of a novel noun—like in the pilot experiment 2. In further studies, we plan to use more stimuli, testing speakers' sensitivity to different suffixes etc.

Now let us discuss the results of experiment 1 in more detail. It demonstrated that both gender and declension of the noun influence online processing of the subject–predicate gender agreement in Russian. But, firstly, this influence can be detected only in the sentences with agreement errors, i.e. no gender or declension is intrinsically more difficult to process (at least, in the sentence context[7]). Secondly, declension plays a role at a very early stage and its effect is very short-lived, while the role of gender becomes visible later and its effect is more pronounced.

The fact that a masculine verb form is less readily detected after a 3rd declension subject noun can be explained by the fact that its ending is more typical for masculine nouns that for feminine ones. However, alternative explanations are also possible, for example, all agreement errors (in masculine or in neuter) may be harder to detect after 3rd declension subject nouns, i.e. their gender can be in general harder to retrieve. To exclude this and some other possibilities, other experiments should be conducted. Another line of further research should look at non-prototypical masculine nouns like *papa* 'dad'. The picture may be different not only because of their different gender, but also because all these nouns denote humans, so the gender feature is not semantically empty in this case, which may aid its processing and retrieval.

As for the role of gender as such, we saw that agreement errors with masculine subjects cause a larger delay in reading times compared to errors with feminine subjects, i.e. were costlier for processing. This is in line with the previous findings on gender agreement in comprehension reported in the literature [Akhutina et al., 1999, 2001]; [Romanova & Gor, 2017]; [Slioussar & Malko, 2016]. However, to have a full picture, neuter subjects and predicates should be introduced in further experiments.

So far, several explanations are possible. It is well known that while reading, we generate expectations about the upcoming predicate based on the features of the subject and rechecking is prompted if these expectations are violated (which is associated with increased reading times). Perhaps, the masculine form of the predicate, being the most frequent, causes less disruption if used incorrectly—similarly, using a frequent word instead of an infrequent one provokes less surprise than the opposite mistake. Maybe, these expectations are more robust for masculine subjects, so violating them is more disruptive. Maybe, if an agreement error is detected and rechecking is initiated, masculine subjects are retrieved more readily and reliably— this is what

[7] It is well known that many differences that can be detected in the processing of isolated forms disappear when these forms are embedded in an appropriate context.

[Slioussar and Malko 2016] suggested based on their agreement attraction results where all combinations of genders on subjects, attractors and predicates were examined. All these explanations are compatible with the observed difference between ungrammatical sentences with masculine and feminine subjects. Further experiments are necessary to tease them apart and to gain a better understanding of the patterns observed in previous studies.

The study was supported in part by the grant #16-18-02071 from the Russian Science Foundation.

References

1. *Acuña-Fariña, J. C., E. Meseguer, and M. Carreiras.* (2014). Gender and number agreement in comprehension in Spanish. Lingua 143: 108–128.
2. *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.
3. *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.
4. *Alexeeva, Svetlana, Natalia Slioussar, and Daria Chernova.* (2018). StimulStat: a lexical database for Russian. To appear in: Behavior Research Methods.
5. *Andonova, E., S. D'Amico, A. Devescovi, and E. Bates.* (2004). Gender and lexical access in Bulgarian. Perception and Psychophysics 66: 496–507.
6. *Anton-Mendez, Inés, Janet Nicol, and Merrill F. Garrett.* (2002). The relation between gender and number agreement processing. Syntax 5: 1–25.
7. *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.
8. *Bates, Elizabeth, Antonella Devescovi, Luigi Pizzamiglio, Simona D'Amico, and Arturo Hernandez.* (1995). Gender and lexical access in Italian. Perception and Psychophysics 57: 847–862.
9. *Bock, Kathryn, and Kathleen M. Eberhard.* (1993). Meaning, sound and syntax in English number agreement. Language and Cognitive Processes 8: 57–99.
10. *Caffarra, S., A. Siyanova-Chanturia, F. Pesciarelli, F. Vespignani, and C. Cacciari.* (2015). Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. Psychophysiology 52: 1019–1030.
11. *Franck, Julie, Gabriella Vigliocco, Inés Antón-Méndez, Simona Collina, and Ulrich H. Frauenfelder.* (2008). The interplay of syntax and form in sentence production: a cross-linguistic study of form effects on agreement. Language and Cognitive Processes 23: 329–374.
12. *Gollan, T. H., and R. Frost.* (2001). Two routes to grammatical gender: Evidence from Hebrew. Journal of Psycholinguistic Research 30: 627–651.
13. *Janssen, Bibi E.* (2016). The acquisition of gender and case in Polish and Russian: A study of monolingual and bilingual children. Amsterdam: Pegasus.

14. *Lyashevskaya, Olga, and Sergey Sharov.* (2009). Častotnyj slovar' sovremennogo russkogo jazyka [The frequency dictionary of modern Russian language]. Moscow: Azbukovnik.
15. *Martin, A. E., M. S. Nieuwland, and M. Carreiras.* (2014). Agreement attraction during comprehension of grammatical sentences: ERP evidence from ellipsis. Brain and Language 135: 42–51.
16. *Ratcliff, Roger.* (1993). Methods for dealing with reaction time outliers. Psychological Bulletin 114: 510–532.
17. *Rodina, Yulia, and Marit Westergaard.* (2012). A cue-based approach to the acquisition of grammatical gender in Russian. Journal of Child Language 39: 1077–1106.
18. *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.
19. *Rusakova, Marina.* (2013). Elementy antropocentričeskoj grammatiki russkogo jazyka [Elements of the anthropocentric grammar of the Russian language]. Moscow: Jazyki slavjanskoj kul'tury.
20. *Savchuk, Svetlana.* (2011). Korpusnoe issledovanie variantov rodovoj prinadležnosti v russkom jazyke [A corpus study of morphological variability: variation in gender forms of Russian nouns]. Computer linguistics and intellectual technologies 10, 562–579.
21. *Schwartz, Mila, Miriam Minkov, Elena Dieser, Ekaterina Protassova, Victor Moin, and Maria Polinsky.* (2015). Acquisition of Russian gender agreement by monolingual and bilingual children. International Journal of Bilingualism 19: 726–752.
22. *Sekerina, Irina.* (2012). The effect of grammatical gender in Russian spoken-word recognition. In: Russian language studies in North America. New perspectives in theoretical and applied linguistics, ed. by Veronika Makarova, 107–132. New York: Anthem Press.
23. *Slioussar, Natalia, and Anton Malko.* (2016). Gender agreement attraction in Russian: production and comprehension evidence. Frontiers in Psychology 7: article 1651.
24. *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]. In: Proceedings of the conference 'Dialogue'. http://www.dialog-21.ru/digests/dialog2015/materials/pdf/SlioussarNASamoilovaMV.pdf
25. *Spalek, K., J. Franck, H. Schriefers, and U. H. Frauenfelder.* (2008). Phonological regularities and grammatical gender retrieval in spoken word recognition and word production. Journal of Psycholinguistic Research 37: 419–442.
26. *Taraban, Roman, and Vera Kempe.* (1999). Gender processing in native and non-native Russian speakers. Applied Psycholinguistics 20: 119–148.
27. *Tseitlin, Stella.* (2000). Jazyk i rebenok [Language and the child]. Moscow: Vlados.
28. *Vigliocco, Gabriella, and Julie Franck.* (1999). When sex and syntax go hand in hand: Gender agreement in language production. Journal of Memory and Language 40: 455–478.

29. *Vigliocco, Gabriella, and Tiziana Zilli.* (1999). Syntactic accuracy in sentence production: the case of gender disagreement in Italian language-impaired and unimpaired speakers. Journal of Psycholinguistic Research 28: 623–648.
30. *Vigliocco, Gabriella, Brian Butterworth and Carlo Semenza.* (1995). Constructing subject-verb agreement in speech: The role of semantic and morphological factors. Journal of Memory and Language 34: 186–215.
31. *Zaliznyak, Andrey.* (1987). Grammatičeskij slovar' russkogo jazyka. Slovoizmenenie [The grammatical dictionary of the Russian language. Inflection]. 2nd ed. Moscow: Russkij Jazyk.