Advances in formal Slavic linguistics 2016

Edited by

Denisa Lenertová

Roland Meyer

Radek Šimík

Luka Szucsich

Draft
of January 21, 2019, 13:51

Open Slavic Linguistics

Editors: Berit Gehrke, Denisa Lenertová, Roland Meyer, Radek Šimík & Luka Szucsich

In this series:

1. Lenertová, Denisa, Roland Meyer, Radek Šimík & Luka Szucsich (Eds.). Advances in formal Slavic linguistics 2016.

Advances in formal Slavic linguistics 2016

Edited by

Denisa Lenertová

Roland Meyer

Radek Šimík

Luka Szucsich

Lenertová, Denisa, Roland Meyer, Radek Šimík & Luka Szucsich (ed.). 2019. *Advances in formal Slavic linguistics 2016* (Open Slavic Linguistics 1). Berlin: Language Science Press.

Freie Universität Berlin

Contents

Preface

The present volume, *Advances in Formal Slavic Linguistics 2016*, marks a delectable double premiere: It initiates both the book series *Open Slavic Linguistics* as a whole, and its sub-series of collective volumes on formal Slavic linguistics.

Open Slavic Linguistics aims at publishing high quality books with a focus on Slavic languages on the empirical side, which at the same time reflect the state of the art and current developments in general linguistics. Its core principles are strict adherence to a genuine Open Access policy and to quality control through double-blind peer review. The series takes a broad linguistic perspective and invites monographs and topical collective volumes from virtually all subdisciplines. This may include theoretically oriented work on Slavic linguistic phenomena, advanced empirical/experimental work on Slavic languages, as well as handbooks, introductions and companions to the linguistic analysis of a given language. The defining characteristics of the series is that it seeks a solid grounding in up-to-date theoretical and empirical methods, fosters mutual understanding of linguists across object languages and subdisciplines, and seeks to contribute both to narrowly defined Slavic linguistics and to general linguistics and linguistic typology.

Advances in Formal Slavic Linguistics 2016 presents a selection of high quality papers authored by young and senior linguists from around the world and contains both empirically oriented work, underpinned by up-to-date experimental methods, and more theoretically based contributions. The volume covers all major linguistic areas, including morphosyntax, semantics, pragmatics, phonology, and their mutual interfaces. The particular topics discussed range from argument structure, word order, case, agreement, tense, aspect, and the left clausal periphery to segmental phonology. The thematic breadth and analytical depth of the contributions reflect the vitality of the field of formal Slavic linguistics and testify to its relevance for the global linguistic endeavor.

Early versions of the papers included in this volume were presented at the conference on Formal Description of Slavic Languages 12 or at the satellite Workshop on Formal and Experimental Semantics and Pragmatics, which were held in Berlin on 7–10 December 2016 – the year referred to in the title of the volume. Half of the submitted abstracts made it into the 44 presentations of the

conference. The 21 papers in the present volume were developed from these contributions in the course of a further thorough reviewing process. Neither the original conference nor the present volume would have been possible without the readiness of so many experts to devote their time and thoughts to the critical evaluation and helpful commenting of their colleagues' research papers. We wish to express our gratitude both to the 75 anonymous reviewers of the original conference abstracts, and to the more than 50 external reviewers for the present volume. Their commitment testifies to the liveliness and ambition of the field of Slavic linguistics. This book would have also been impossible without our student assistants, Bella Badt, Justina Bojarski, Andrei Koniaev and Jake Walsh, and the invaluable help of the Language Science Press editors Sebastian Nordhoff and Felix Kopecky. We gratefully acknowledge their efforts and support. Finally, we would like to acknowledge the authors themselves. Open Access publishing is a collective endeavor and we appreciate the authors' willingness to collaborate with us closely not just on linguistic and scientific issues, but also on editorial matters. We sincerely hope that the authors and readers of this volume will share our conviction that it has been worthwhile.

<div align="right">

Denisa Lenertová, Roland Meyer, Radek Šimík & Luka Szucsich
Berlin, 14 December 2018

</div>

Chapter 1

Russian case inflection: Processing costs and benefits

Maria D. Vasilyeva

Lomonosov Moscow State University

Mechanisms underlying the processing and storage of morphological case are still debatable in psycholinguistics. The key questions concern the nature of the special status of the nominative, the homogeneity/heterogeneity of oblique case forms, the impact of case syncretism and paradigmatic relations on nominal processing and the organization of the mental lexicon. We investigate these issues turning to Russian nominal processing. We performed two experiments with feminine and masculine nouns in different cases (experiment 1: nouns in singular, experiment 2: nouns in plural) using the visual lexical decision task. In this task, we measure the speed and accuracy with which the participant classifies sequences of letters as words or non-words. Evidence from both experiments indicates that differences in processing exist not only between the nominative and the other case forms, but also among the obliques. Experiment 1 points to the influence of wordform and exponent ambiguity, while experiment 2 reveals effects that are specific for case per se. We discuss the role of zero vs. overt phonological form, grammatical features, (non-)accidental homonymy, context, frequency, inflectional and relative entropy in case recognition.

1 Introduction

The role of frequency and regularity in processing of inflectional morphology has for long been of utmost concern for psycholinguists. Meanwhile, it is still not clear whether grammatical features that an inflectional marker conveys play an additional role in wordform processing. For instance, if we are speaking about nouns, a natural question to ask is how case influences nominal recognition.

Studies of isolated wordform processing suggest that nominative wordforms are processed faster than other case forms (see, e.g. Lukatela et al. 1978 for Ser-

bian; Niemi et al. 1994 for Finnish; Abulizi et al. 2016 for Uyghur; Gor et al. 2017 for Russian). Yet, there is no uniform explanation of this fact. Likewise, it is debatable whether oblique cases entail equal processing costs or not.

Finnish and Uyghur researchers provide only the pooled mean for all the inflectional variants, comparing it to the nominative and do not inspect contrasts between oblique forms, though they usually use more than one oblique case in their experiments (Niemi et al. 1991; 1994; Hyönä et al. 1995; Laine & Koivisto 1998; Laine et al. 1999; Abulizi et al. 2016). As the nominative has zero inflection in these languages, oblique processing cost is attributed to a morphological decomposition procedure that is obligatory for inflected obliques, but absent in the non-inflected nominative.

This explanation is unsatisfactory for several reasons. Firstly, the nominative advantage disappears when case forms are embedded in context (Bertram et al. 2000; Hyönä et al. 2002). Bertram et al. (2000) and Hyönä et al. (2002) suggest that oblique processing disadvantage in a context-less environment arises not due to the decomposition cost, but precisely due to the lack of an appropriate context. Yet, they do not examine if all oblique cases suffer from the lack of context or benefit from its presence to the same extent. Secondly, processing of zero inflection receives a benefit in recognition speed only if the zero is associated with the nominative, but not with an oblique case (Gor et al. 2017). Finally, phonologically zero and overt nominatives appear not to differ in processing speed (see, e.g., Lukatela et al. 1980 for Serbian; Gor et al. 2017 for Russian). Thus, it is not the zero inflection that makes Finnish and Uyghur nominative wordforms special, but the the nominative case itself.

Early Serbian studies did compare processing of oblique cases, but mainly failed to find significant differences in response latencies (Lukatela et al. 1978; 1980; 1987; Katz et al. 1987; Kostić & Katz 1987; Feldman & Fowler 1987). These results, starting with Lukatela et al. (1980), were analyzed within the satellite model. The nominative form represents the nucleus of the nominal paradigm, while oblique case forms surround it as satellites. Satellites are assumed to be equidistant from the nucleus (Feldman & Fowler 1987). Deviations from the predictions of this model were attributed to specific experimental settings in case of nouns (Feldman & Fowler 1987; Todorović 1988); differences in adjectival case processing were assumed to rely on different mechanisms (Kostić & Katz 1987). However, not more than three case forms belonging to one number were compared at once. It is likely that some effects that could show up in a more elaborate design were obscured. Moreover, ambiguity of case forms that is present in Serbian declension did not receive enough attention.

Draft of January 21, 2019, 13:51

According to subsequent Serbian studies, processing speed of a wordform correlates positively with the number of syntactic functions/meanings that its inflectional ending encompasses (Kostić 1991; 1995; Kostić et al. 2003; Filipović Đurđević & Kostić 2003; Ševa & Kostić 2003), which hints at oblique processing differences. The proposed methodology of calculating syntactic functions/meanings is not flawless, since the authors bring under this umbrella term both syntactic notions such as subject or complement and semantic notions such as instrument or goal. Furthermore, it is not taken into account that lexemes in the same case have different probabilities of expressing the same thematic role, e.g. 'girl-INS' is less likely to be an instrument than 'hammer-INS'. Likewise, as all homonymous forms are treated equally, differences between accidental and non-accidental ambiguity is disregarded.

Later on, this problematic measure was abandoned, and the focus was shifted to paradigmatic relations between wordforms captured by inflectional and relative entropy measures (see, e.g. Milin et al. 2009). The inflectional entropy $H(P)$ reflects the amount of information associated with the inflectional paradigm of the target lexeme (see (1)), where f stands for frequency, the wordform w_i belongs to the paradigm P of a lexeme w) and correlates negatively with response latencies: when a lexeme has a higher value of the inflectional entropy, its wordforms are processed faster, and vice versa (Moscoso del Prado Martín et al. 2004). The relative entropy $D(IP\|IC)$ captures the divergence between the frequency distribution of the target lexeme w and the frequency distribution of its inflectional class IC (see (2)), where e_i stands for inflectional exponent), and it correlates positively with response latencies: wordforms belonging to paradigms with higher values of relative entropy are processed more slowly (Milin et al. 2009). When surface and lemma frequency combined with entropy measures are taken into account, case differences appear to play no additional role (Milin et al. 2009); yet, this claim was made on a small subset of wordform: wordforms in *-u* 'ACC.SG' and *-e* 'GEN.SG'/'NOM/ACC.PL' for feminine nouns, wordforms in *-om* 'INS.SG' and *-u* 'DAT/LOC.SG' for masculine nouns.

(1)

$$H(P) = - \sum_{w_i \in P} \frac{f(w_i)}{f(w)} \log_2 \frac{f(w_i)}{f(w)}$$

(2)

$$D(IP\|IC) = \sum_{w_i \in P} \frac{f(w_i)}{f(w)} \log_2 \frac{f(w_i)/f(w)}{f(e_i)/f(e)}$$

Another viewpoint predicting differences in oblique case processing and paying attention to wordform ambiguity was developed primarily by Clahsen et al. (2001). They adopted minimalist principles in morphology (see, e.g., Wunderlich 1996), suggesting that in the mental lexicon, the meaning of a case exponent is represented as a set of binary features. Non-accidental ambiguous inflectional markers receive underspecified representations. Along with this "natural" underspecification, radical underspecification is assumed to be present as well: only positive values are stored in the mental lexicon, while negative ones are deduced from paradigmatic oppositions. Hence, a direct implication for the psycholinguistic models of wordform processing arises. The number of specified (positive) features should determine the processing ease: the more information a form carries, the longer it takes to be recognized.

Main evidence supporting this claim comes from studies on German adjectival declension. Adjectival case forms with more specified representations are recognized slower in the lexical decision task (Clahsen et al. 2001). Such case forms show reduced priming effects under cross-modal priming if the adjective serving as a prime does not share all the positive features with the target (Clahsen et al. 2001). Similar priming effects are to a certain extent replicable even with highly proficient L2-German speakers (Bosch & Clahsen 2016; Bosch et al. 2017). As far as sentence processing is concerned, when an ungrammatical sentence contains an adjective or a determiner that is compatible with the context by its feature specification, this does not lead to an ungrammaticality effect in a sentence matching task, observed for ungrammatical sentences where specificity is violated, i.e. when the feature set of the wordform mismatches context requirements (Penke et al. 2004). These two types of ungrammatical sentences result in distinct ERP responses (Opitz et al. 2013).

If the radical underspecification hypothesis is true, the same principles should hold for nominal case inflection in other languages as well. Yet, prior studies on case processing shed doubts on its tenability, and additional evidence is needed.

2 Present study

The present study aims to verify whether case form processing is determined by the grammatical features, nominative vs. oblique dichotomy, or context. We addressed this issue in two lexical decision task experiments employing Russian data: experiment 1 with singular nouns and experiment 2 with plural nouns.

Russian was not chosen incidentally, but due to its particular pattern of case syncretism (convergence of inflectional exponents in different paradigmatic cells)

Russian has six major cases and several inflectional classes of nouns. We will restrict ourselves to inanimate nouns and discuss only two most productive inflectional classes (Wiese 2004): feminine nouns with the nominative ending *-a* and masculine nouns with the nominative ending *-∅*. As is evident from Table 1, the two classes of nouns choose uniform endings in plural (except for genitive), but behave differently in singular.

Table 1: Russian case endings for the two most productive inflectional classes

		Nom	Acc	Gen	Dat	Loc	Ins
Singular	Feminine	-a	-u	-y	-e	-e	-oj
	Masculine	-∅	-∅	-a	-u	-e	-om
Plural	Feminine	-y	-y	-∅	-am	-ax	-ami
	Masculine	-y	-y	-ov	-am	-ax	-ami

Context . Presenting case forms in isolation, we can test whether all oblique case forms rely equally on the context. Russian data is particularly suitable for resolving this issue, as there is a special case in Russian, namely locative (also called prepositional), which, unlike other cases, is always governed by a preposition. If the context is crucial for efficient oblique case recognition, locative wordforms should be processed longer compared to other oblique cases, since the latter do not need any preceding context on the left (e.g., if they occur at the beginning of a sentence). This hypothesis is partly supported by Vasilyeva et al.'s (2014) finding: masculine locative singular wordforms are processed as slowly as pseudowords with the same syllabic structure, and they are often qualified as nonwords. However, in the singular form, this processing cost could be caused by the homonymy of *-e* 'LOC.MASC' with *-e* 'DAT/LOC.FEM'. If the effect is induced by the lack of prior preposition activation, locative plural processing should also be impaired. If locative plural processing is not more difficult than processing of other obliques, difficulty of masculine locative singular can not be explained by the absence of an appropriate preposition alone and, in general, the context-based hypothesis is not tenable.

Plural: case features vs. exponent frequency . As oblique plural exponents are non-ambiguous, these data are fruitful for exploring the role of case in nominal

processing. If surface and lemma frequency are accounted for and differences in oblique case processing still arise, they could be attributed either to exponent frequencies or to the set of grammatical features associated with the particular case. Frequency counts provided in Samojlova & Slioussar (2014) suggest a hierarchy in (3a). Different approaches to Russian declension employ different sets of features and, thus, give conflicting predictions see (3b)–(3e), where we arrange oblique cases according to the number of positive features they express (as suggested by Clahsen et al. 2001).

(3) a. frequency: Gen < Ins < Loc < Dat (34% < 11% < 10.3% < 4.7%)
 b. Müller 2004: Loc < Dat ≈ Ins < Gen (⟨+obl⟩ < ⟨+obl, +gov⟩ ≈ ⟨+obl, +subj⟩ < ⟨+obl, +gov, +subj⟩)
 c. Wiese 2004: Loc < Ins ≈ Dat ≈ Gen (⟨+obl⟩ < ⟨+obl, +inst⟩ ≈ ⟨+obl, +dat⟩ ≈ ⟨+obl, +gen⟩)
 d. Wunderlich 1996: Gen < Dat < Loc < Ins (⟨+hr$_N$⟩ < ⟨+hr, +lr⟩ < ⟨ +hr & additional semantic features ⟩ < ⟨ semantic features⟩))
 e. Caha 2008: Gen < Loc < Dat < Ins

Zero oblique inflection . Gor et al.'s (2017) study demonstrated that oblique overt and zero inflection trigger similar processing costs. But their conclusion was based on the comparison of feminine -∅ 'ɢᴇɴ.ᴘʟ' to masculine -*a* 'ɢᴇɴ.sɢ'. A comparison with masculine -*ov* 'ɢᴇɴ.ᴘʟ' is needed to support their claim.

Nominative ambiguity . Feminine -*y* 'ɴᴏᴍ.ᴘʟ' coincides with 'ɢᴇɴ.sɢ'. According to the approach advocated by Kostić (1991), etc., such ambiguous wordforms should benefit from their wider syntactic distribution and be recognized faster than their unambiguous masculine counterparts -*y* 'ɴᴏᴍ.ᴘʟ'.

Singular: case syncretism . Even if we obtain no significant differences in plural oblique processing, differences in singular oblique processing might arise due to ambiguity. Comparing instrumental wordforms, which are non-ambiguous, to other obliques, we can determine how interparadigmatic and intraparadigmatic syncretism influences wordform recognition. Furthermore, comparisons of wordforms with the same exponents, but belonging to different inflectional classes might help to resolve the debates concerning accidental vs. non-accidental homonymy in Russian singular declension.

If the two -*u*-s are accidentally homonymous (Wiese 2004), 'ᴀᴄᴄ.ꜰᴇᴍ' is expected to be processed faster than 'ᴅᴀᴛ.ᴍᴀsᴄ'. If the two -*e*-s are accidentally

homonymous (Müller 2004), 'DAT./LOC.FEM' is expected to be processed faster than 'LOC.MASC'.

We also decided to compare -*e* 'DAT/LOC.FEM' to -*u* 'DAT.MASC'. If the dative reading is dominant for -*e*, there should be no difference between these two conditions. Finally, we compared feminine and masculine instrumental wordforms. These endings are also used in adjectives of the respective gender, but their distribution is different: feminine -*oj* covers all oblique cases, while masculine -*om* is used in locative only. This difference might lead to an advantage of feminine instrumental over masculine instrumental.[1]

Zero nominative inflection. Russian overt and non-overt nominative inflection (-*a* 'NOM.FEM' and -∅'NOM.ACC.MASC') was already compared in Gor et al. (2017), and no difference was observed. However, their study employed an auditory lexical decision task, and it is unclear whether their results are modality-neutral.

3 Method

3.1 Participants

Ninety-six Russian native speakers, all right-handed (aged 17–25 years) were tested. Half of them participated in experiment 1, the other half in experiment 2.

3.2 Stimuli

We used all six case forms of inanimate nouns belonging to two declensional classes (54 feminine nouns ending in -*a* and 54 masculine nouns ending in -∅ matched for lemma frequency). All stimuli were base nouns, they did not undergo any stem alternations and had fixed stress on the stem (the 1a inflectional class according to Zaliznyak 1977). Length in nominative differed from 4 to 6 (each group comprised one third of words with each length). 108 nouns with pseudoendings and 108 inflected pseudostems served as nonwords. In experiment 1, nouns were presented in singular; in experiment 2, in plural. Latin-square design was employed with the number of lists corresponding to the number of case forms.

---

[1]Since feminine genitive singular -*y* is homonymous with nominative plural, we did not compare it to the masculine genitive singular.

3.3 Procedure

Each participant was assigned to one of the six experimental lists and was tested individually. Experiments were run using DMDX software (Forster, Forster, 2003). Before the test phase (324 trials), participants received written instructions and performed a practice phase (20 trials). In each trial, participants had to decide whether the string of letters presented on the screen was a real Russian word or not. They were instructed to respond as fast and accurately as possible. Each trial started with a fixation sign (+) that was displayed on the screen for 600 ms. The stimulus remained on the screen until response or time-out (2500 ms). The interstimulus interval was set to 2500 ms.

3.4 Data analysis

We used linear mixed-effects modeling for the analysis of reaction times and logistic mixed regression for the accuracy data (Baayen 2008). Statistical analysis was implemented in the package lme4 (Bates et al. 2014) in the statistical software R (R Core Team 2014). *T*-values, *z*-values, *p*-values, and standard errors were determined using the package lmerTest (Kuznetsova et al. 2017). Fixed and random effects were included only if they significantly improved the model fit in a backward stepwise model selection procedure. Models were selected using Chi-square log-likelihood ratio tests with regular maximum likelihood parameter estimation.

Subject and lexeme were treated as random effects. Lemma and wordform frequency, length in letters and syllables, mean Levenstein distance to the nearest 20 lexeme-neighbors, inflectional and relative entropy measures were additionally included as covariates.[2,4,5] Trial order (*z* transformation on log numbers) was included to control for longitudinal task effects such as fatigue or habituation. All these covariates were log-transformed. To avoid multicollineriarity, all counts except for trial were transformed into 5 principal components, explaining 93.5% of variance (Baayen 2008). The first principal component (PC1) captured

---

[2]Lemma frequency was taken from the frequency dictionary (Lyashevskaya & Sharoff 2009). Wordform frequency was manually extracted from the main undisambiguated subcorpus of Russian national corpora[3]; for ambiguous endings the cumulative frequency was taken, relying on Milin et al.'s (2009) experience. To avoid zero frequencies, one was added to all counts, as suggested by Brysbaert & Diependaele (2013).

[4]The Levenstein distance was calculated in the vwr package (Keuleers 2013) in the R software (R Core Team 2014).

[5]In order to calculate relative entropy, frequency of exponents was taken form the database created by Samojlova & Slioussar (2014).

orthographic characteristics of the stimulus. The second component (PC2) was inversely related to frequency. The third component (PC3) was inversely related to relative entropy and positively related to inflectional entropy. Paired contrasts were carried out in the package lsmeans (Lenth 2016). For planned comparisons, FDR adjusted p-values are reported (Benjamini & Hochberg 1995).

As our words were presented without context, case labels for ambiguous endings (feminine -*y* and -*e*, masculine -*y* and -∅) are somewhat arbitrary. Hence, we do not expect any differences between feminine locative and dative -*e*, nor between masculine nominative and accusative -∅. However, this is needed for counterbalancing issues, as patterns of syncretism do not coincide across our two noun groups. In the statistical analysis, the mean pooled over the two "conditions" will be used.

4 Results

Two participants in experiment 1 and two participants in experiment 2 gave fewer than 75% correct answers to word stimuli, so we recruited four additional people to replace them. We excluded from further statistical analysis two lexemes in experiment 1 and one lexeme in experiment 2 due to low mean accuracy score.

Reaction time (RT) data were analyzed as follows. Incorrect responses were removed from the analysis (7.1% of all data in experiment 1, 7.9% in experiment 2). Too fast (< 300 ms) or too slow responses (> 1 500 ms) were likewise excluded from further analysis. We applied log-transformation to reduce the positive skew. After that, remaining outliers were cut off via interquartile trimming.[6] In sum, 3.9% of correct responses were removed in experiment 1 and 5.7% in experiment 2.

Raw RTs and error rates (ER) are presented in Table 2.

Final models for RTs and accuracy included the following factors: PC2, PC3, case, gender, a case by gender interaction and a PC3 by gender interaction. The model accounting for RTs in experiment 1 also included trial. All other predictors and interactions turned out to be insignificant. Full model specifications are presented in the Appendix (see Table 5 for experiment 1 and Table 6 for experiment 2).

---

[6]We kept only those RTs, which satisfied the following formula Q1 – (2.5 × IQR) < RT < Q3 + (2.5 × IQR), by participants, items (lexemes), gender and case (Q1 stands for first quartile, Q3 for third quartile, and IQR = Q3 – Q1 for interquartile range).

Table 2: Mean RT (in ms) and ER (in %) to feminine and masculine nouns in different cases and numbers (*SD* is provided in brackets)

		Nom	Acc	Gen	Dat	Loc	Ins
				Experiment 1: singular			
f	RT	738 (181)	782 (198)	764 (195)	805 (203)		770 (194)
	ER	1.7 (12.9)	4.8 (21.4)	4.6 (20.9)	9.3 (29)		3.9 (19.3)
m	RT	719 (176)		838 (228)	799 (203)	850 (231)	803 (219)
	ER	2.3 (15)		8.3 (27.7)	6.3 (24.2)	20.8 (40.7)	2.3 (15.1)
				Experiment 2: plural			
f	RT	818 (202)		894 (225)	900 (226)	873 (216)	894 (230)
	ER	4.6 (21)		22.4 (41.7)	7.5 (26.4)	8.5 (27.9)	7.3 (26.1)
m	RT	821 (196)		875 (209)	904 (235)	879 (212)	932 (234)
	ER	5.1 (22)		4.2 (20)	7.4 (26.2)	7.4 (26.2)	7.2 (25.8)

4.1 Experiment 1: Singular

Trial had a facilitative effect on RTs ($B = -0.012$, $t(4541) = -4.45$, $p < .001$). PC2 (inversely related to frequency) had a facilitative effect on RTs and accuracy rate ($B = 0.018$, $t(121) = 6.03$, $p < .001$, respectively and $B = -0.364$, $z = -5.86$, $p < .001$, respectively).

PC3 (entropy meausures) affected differently the two types of nouns ($B = 0.024$, $t(103) = 3.2$, $p = .002$ and $B = -0.472$, $z = -3.628$, $p < .001$, respectively): there was a facilitation for feminine nouns ($B = -0.022$, $t(103) = -3.68$, $p < .001$ and $B = 0.509$, $z = 4.89$, $p < .001$, respectively) and no effect on masculine nouns ($B = 0.002$, $t(103) = 0.48$, $p = .633$ and $B = -0.11$, $z = 0.481$, $p = .631$, respectively).

Paired contrasts for different cases are summarized in Table 3 (for statistical details see Table 7). Apart from nominative vs. oblique differences, we observe differences between oblique case forms. Instrumental and the -*u* form ('ACC.FEM' and 'DAT.MASC') are "easy" obliques with faster responses and higher accuracy scores. The -*e* forms ('DAT/LOC.FEM' and 'LOC.MASC') constitute the "difficult" oblique▉ group with slower responses and lower accuracy scores. Feminine genitive falls into the "easy" group, while masculine genitive patterns with the difficult -*e* 'LOC'

form.[7]

Table 3: Experiment 1 (singular nouns): summary of paired contrasts analysis for feminine and masculine singular nouns in different cases

	RTs	accuracy
f	Nom < Ins ≈ Gen ≈ Acc < -*e*	Nom < Gen ≈ Acc < -*e*,
		Nom ≈ Ins, Gen ≈ Ins ≈ Acc, Ins < -*e*
m	-∅ < Ins ≈ Dat < Gen ≈ Loc	-∅ ≈ Ins < Dat ≈ Gen < Loc

In the analysis of accuracy, in contrast to the RT data, we fail to observe the nominative superiority over instrumental in any noun group. What is more, according to the accuracy analysis, masculine genitive yields higher accuracy rate than the masculine locative.

Paired contrasts for gender   (for statistical details see Table 7). There was no general gender effect either in the RT or accuracy analysis. Feminine -*e* forms ('DAT/LOC.FEM') are recognized faster and more accurately than masculine -*e* forms ('LOC.MASC'). Feminine instrumental wordforms are recognized faster than masculine ones, but there is no effect in the accuracy analysis. Feminine nominative is responded to slower than masculine -∅ forms ('NOM/ACC.MASC'); no effect shows up in the accuracy analysis. Feminine -*e* forms ('DAT/LOC.FEM') are recognized less accurately than masculine -*u* forms ('DAT.MASC'), but there is no significant difference in the RT analysis. There is no significant difference between feminine and masculine -*u* forms ('ACC.FEM' and 'DAT.MASC', respectively).

4.2 Experiment 2: Plural

PC2 (inversely related to frequency) had a facilitative effect on RTs and accuracy rates ($B = 0.019$, $t(126) = 4.98$, $p < .001$ and $B = -0.285$, $z = -4.21$, $p < .001$, respectively).

PC3 (entropy measures) affected differently the two types of nouns ($B = 0.021$, $t(107) = 2.28$, $p = .025$ and $B = -0.423$, $z = -2.727$, $p = .006$, respectively): there was a facilitation for feminine nouns ($B = -0.022$, $t(111) = -2.91$, $p = .004$ and $B = .442$, $z = 3.7$, $p < .001$, respectively), but no effect on masculine nouns ($B = -0.0004$, $t(101) = -0.06$, $p = .949$ and $B = 0.019$, $z = 0.19$, $p = .847$, respectively).

---

[7]In Table 3 and further "<" stands for significantly faster or significantly more accurate responses, "≈" stands for no significant difference.

Paired contrasts for case are summarized in Table 4 (for statistical details see Table 8). According to the RT analysis, we observe a tripartite division of oblique forms: locative as the easiest, dative in the middle and instrumental as the most difficult. Genitive is recognized significantly faster than instrumental, but differs neither from locative, nor from dative.

In the accuracy analysis, only two contrasts are retained: between the -*y* form ('NOM/ACC') and instrumental and the difference between the -*y* form ('NOM/ACC') and genitive.

Table 4: Experiment 2 (plural nouns): summary of paired contrasts analysis for plural nouns in different cases

RTs	accuracy
-*y* < Loc < Dat < Ins,	-*y* < Gen ≈ Ins
Gen < Ins,	-*y* ≈ Loc ≈ Dat
Gen ≈ Loc, Gen ≈ Dat	Loc ≈ Dat ≈ Gen ≈ Ins

Paired contrasts for gender (see Table 7 for statistical details). There was no difference between two groups of nouns either in the RT or accuracy analysis. In genitive, feminine nouns have higher odds to be recognized incorrectly than masculine nouns; no significant difference shows up in the RT analysis. Feminine and masculine -*y* forms ('NOM/ACC') differ neither in the RT analysis, nor in the accuracy analysis.

5 Discussion

Results of our two experiments replicate the nominative/oblique dichotomy effect, previously reported for Russian and other languages (see §1). Apart from this trivial finding, we obtained several significant differences between oblique case processing both in singular and in plural. As we took into account lemma and surface frequency of a wordfom, such oblique case processing differences should stem from the properties of the inflectional exponents.

Inflectional and relative entropy. We considered inflectional and relative entropy among potential covariates in the statistical analysis, as these factors were

assumed to be highly predictive of nominal processing in Serbian (Milin et al. 2009). Prior to the analysis, we transformed our counts into principal components. PC3 capturing these two measures emerged in the statistical analysis of RTs and accuracy in both experiments. Unfortunately, the influence of PC3 was attested for feminine nouns only. The effect lies in the same direction as reported by Milin, Filipović Đurđević & Moscoso del Prado Martín (2009), but in the Serbian study masculine and feminine nouns were equally sensitive to entropy measures. However, to calculate the entropy values, they used frequencies of feminine exponents, as this inflectional class is assumed to be dominant in Serbian. In Russian, masculine -∅ nouns are slightly more frequent than feminine -*a* nouns (Samojlova & Slioussar 2014) and, thus, might be considered dominant. However, as the patterns of syncretism in these two noun groups do not coincide, we decided against using dominant class frequencies and employed feminine frequencies for feminine nouns and masculine frequencies for masculine nouns. This decision might be a possible reason for the observed discrepancies with Milin et al. (2009), but a more refined study is needed in order to make more solid conclusions.

Context: the locative issue . Initially, we hypothesized that if absence of context is an important source of oblique processing cost, preposition-less locatives should suffer the most, both in singular and in plural. Although masculine locative singular was one of the most difficult forms to recognize, plural locative was processed faster than all other obliques. Thus, we conclude that context-based explanations do not receive support at least for wordforms with non-ambiguous case markers. Influence of context on forms with ambiguous case exponents will be discussed below.

Plural. The hierarchy of plural case processing speed (4a) does not follow the order of exponent frequency: otherwise, instrumental would have been the easiest to process. So we can conclude that exponent frequency does not play a major role in the case form recognition. Nor does this hierarchy agree with the predictions derived from the frequency of exponents and feature sets proposed by Müller (2004); Wiese (2004); Wunderlich (2004). Interestingly, it roughly resembles Caha's (2008) nanosyntactic approach to Russian case, see (4b).

(4)   a.   -*y* < Loc < Dat < Ins, Loc ≈ Gen ≈ Dat, Gen < Ins   (our data: exp. 2)

b. [Ins [Dat [Loc [Gen [Acc [Nom]]]]]]

Here, the only diverging case is genitive. Unlike all other cases in plural, it is spelled out differently for our two target inflectional classes. Hence, at the checking or licensing stage (see, e.g., Bertram et al. 2000), which follows the decomposition of the wordform into morphemes, it is verified whether the inflectional class of the lexeme matches the inflectional class of the ending. For other oblique case forms, such a procedure is not needed, as they are uniform for both classes. As a consequence, we observe longer reaction times than those that could be expected if genitive plural meaning was expressed in only one way.

Zero oblique inflection   . In line with Gor et al. (2017), response latencies for the zero oblique ∅ 'GEN.PL' did not differ significantly from the overt oblique *-ov* 'GEN.PL'. Yet, the zero genitive yielded higher error rates than the overt genitive. We doubt that low accuracy stems from a greater processing cost associated with zero inflection compared to overt inflection, especially as this is not attested in the RT analysis. A more plausible source for the high error rate is homonymy. Feminine genitive plural wordforms having a phonologically null ending are ambiguous with the stem itself. This homonymy might lead to a competition in the recognition process: if the wordform reading wins, the correct answer is produced in the lexical decision task; if the stem reading wins, non-word answer is selected, as Russian does not allow for bare stems.

Nominative ambiguity   . We failed to find evidence supporting the claim that the ambiguous feminine *-y* 'NOM.PL'/'GEN.SG' is easier to be recognized than the non-ambiguous masculine *-y* 'NOM.PL' due to its wider distribution.

Singular.   In singular, the following generalization holds for both nouns:

(5)   Nom < Ins ≈ *-u* < *-e*, , where *-u* corresponds to 'ACC.FEM'/'DAT.MASC' and *-e* corresponds to 'DAT/LOC.FEM'/'LOC.MASC'

Instrumental.   Instrumental singular wordforms, despite their relatively low frequency (Samojlova & Slioussar 2014), are one of the easiest obliques to be recognized due to their unambiguity. In §2, we hypothesized that feminine instrumental *-oj* could be processed faster than masculine instrumental *-om* due to their homonymy with adjectival endings, and this prediction was borne out.

Masculine -*om* marks different cases in nouns and adjectives ('ɪɴꜱ' vs. 'ʟᴏᴄ', respectively), and this feature mismatch might negatively affect their processing. Feminine adjectival -*oj*, on the other hand, includes 'ɪɴꜱ' as one of its possible interpretations; consequently, no conflict arises.

-U forms   The -*u* forms ('ᴀᴄᴄ.ꜰᴇᴍ' and 'ᴅᴀᴛ.ᴍᴀꜱᴄ') behave similarly to the unambiguous instrumentals, but this does not signify that their homonymy is accidental. The lack of significant difference in the processing speed of the two forms is compatible with the hypothesis of a shared underspecified representation, as suggested in Müller (2004); Wunderlich (2004). However, this evidence is not enough to reject the accidental homonymy hypothesis. A better insight into this problem might be gained if we compare the processing of dative -*u* in the accusative environment and vice versa. If there is one shared representation for the two -*u*-s in the mental lexicon, such sentences, following Penke et al. (2004); Opitz et al. (2013), should show reduced ungrammaticality effects, if any.

-E forms.   The -*e* forms ('ᴅᴀᴛ/ʟᴏᴄ.ꜰᴇᴍ' vs. 'ʟᴏᴄ.ᴍᴀꜱᴄ') are most difficult to process in both noun groups, triggering longer RTs and lower accuracy, masculine -*e* being even more difficult with the slowest reaction times and the highest error rates. Feminine -*e* is largely believed to have a shared semantic representation for its two interpretations (Müller 2004; Wiese 2004; Wunderlich 2004). But a shared representation on its own is not a plausible source for such a processing cost. Masculine -*e* wordforms, on the contrary, are not ambiguous, but they are always governed by a preposition, and in the present study locatives were presented preposition-less in the experimental conditions. In experiment 2, locative plural, which is also preposition-dependent, actually, turned out to be one of the easiest oblique cases. Thus, absence of the preposition is not the main reason for poor participants' performance on singular masculine locatives.

We suggest that this finding could be accounted for in a model of Russian case where all -*e*-s have one shared representation. The features distinguishing between two cases compete with each other during wordform processing. Locative, as the more frequent reading (Samojlova & Slioussar 2014), has by default more weight, while dative gets more weight in the appropriate context, i.e. in the preposition-less environment. This competition slows down the recognition process. If we assume that the context cue prevails over the frequency cue, then for feminine nouns the dative reading succeeds. With masculine nouns, the context cue will lead to the incorrect selection of the dative features and cause non-word answers. The reanalysis of -*e* as 'ʟᴏᴄ' is, thus, warranted. As any reanalysis, it

requires additional time cost, which explains the superiority of feminine *-e* forms over masculine *-e* forms in the processing speed.

Genitive. Genitive wordforms behave differently in the two inflectional classes. Masculine genitive pattern together with the difficult *-e* in the RT analysis. Feminine genitive falls in the "easier" oblique group. Both genitive endings are homonymous: feminine genitive *-y* coincides with nominative plural, masculine genitive *-a* – with the feminine nominative *-a*.

As far as the masculine genitive *-a* is concerned, an analysis similar to the analysis of *-e* forms is plausible. Two *-a*-s ('GEN.MASC' and 'NOM.FEM') have a shared representation in the mental lexicon. In a context-less condition, the nominative reading is preferred . Shared representation for these morphemes was previously proposed by Müller (2004); Wunderlich (2004). However, this analysis does not capture the fact that masculine genitive is processed more accurately than masculine locative.

As for feminine genitive in *-y*, the unanimous position (Müller 2004; Wunderlich 2004; Wiese 2004) stands for accidental homonymy. Genitive singular *-y* is more frequent than nominative plural (Samojlova & Slioussar 2014). Thus, fullform storage is more likely for nominative plural, following the suggestion by Bertram et al. (2000). Fullform access is assumed to be faster than the decomposition route (i.e., Bertram et al. 2000), yet we do not have enough evidence to claim that our *-y* forms were always processed as nominative plurals. In the singular environment of experiment 1, the singular reading might be chosen due to interstimulus priming. Nevertheless, whichever interpretation is chosen, it is easier to process than the ambiguous *-e*.

Zero nominative inflection . The visual lexical decision task hints at a processing advantage for the phonologically non-overt inflection (-∅ 'NOM/ACC.MASC') over the phonologically overt inflection (-*a* 'NOM.FEM'). This contrasts with the null effect obtained previously in the auditory modality (Gor et al. 2017); note that non-significant effects are actually misleading, as they do not allow to conclude anything. Strictly speaking, these two forms differ not only in phonological overtness, but also in ambiguity: the -*a* wordform is unambiguous, while the -∅ wordform also marks accusative in the discussed set of nouns. So this finding should be treated with caution.

6 Conclusion

The results of our two experiments disagree with previous findings in Finnish, Uyghur, and Serbian, suggesting that differences in oblique case processing exist. Moreover, these differences arise both in transparent systems of case marking (Russian plural) and opaque or highly syncretic systems of case marking (Russian singular).

Data from the experiment with plural nouns suggests that case processing might be guided by Caha's (2008) functional case sequence. Results for singular nouns imply that different types of ambiguity are present in Russian declension. When the ambiguity is not accidental, context plays a major role in the selection of the interpretation.

Abbreviations

ACC	accusative	LOC	locative
DAT	dative	MASC	masculine
GEN	genitive	NOM	nominative
FEM	feminine	PL	plural
INS	instrumental	SG	singular

Acknowledgements

References

Abulizi, Abudoukelimu, Minghu Jiang, Dengfeng Yao & Halidanmu Abudouke-limu. 2016. Cognitive neural dynamic Effects for processing of Uyghur inflectional and derivational words: Evidence from behavioral results. In Rubin Wang & Xiaochuan Pan (eds.), *Advances in cognitive neurodynamics*, vol. V, 553–559. Singapore: Springer.   DOI:10.1007/978-981-10-0207-6_75

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Bates, Douglas, Martin Maechler, Ben Bolker, Steven Walker, et al. 2014. Lme4: Linear mixed-effects models using Eigen and S4. *R package version* 1(7). 1–23. http://CRAN.R-project.org/package=lme4.

Benjamini, Yoav & Yosef Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)* 57(51). 289–300. http://www.jstor.org/stable/2346101.

Bertram, Raymond, Jukka Hyönä & Matti Laine. 2000. The role of context in morphological processing: Evidence from Finnish. *Language and Cognitive Processes* 15(4-5). 367–388.   DOI:10.1080/01690960050119634

Bosch, Sina & Harald Clahsen. 2016. Accessing morphosyntax in L1 and L2 word recognition. *The Mental Lexicon* 11(1). 26–54.   DOI:10.1075/ml.11.1.02bos

Bosch, Sina, Helena Krause & Alina Leminen. 2017. The time-course of morphosyntactic and semantic priming in late bilinguals: A study of German adjectives. *Bilingualism: Language and Cognition* 20(3). 435–456. DOI:10.1017/S1366728916000055

Brysbaert, Marc & Kevin Diependaele. 2013. Dealing with zero word frequencies: A review of the existing rules of thumb and a suggestion for an evidence-based choice. *Behavior Research Methods* 45(2). 422–430.   DOI:10.3758/s13428-012-0270-5

Caha, Pavel. 2008. The case hierarchy as functional sequence. In Marc Richards & Andrej L. Malchukov (eds.), *Linguistische Arbeitsberichte 86: Scales*, 247–276. Leipzig: University of Leipzig.

Clahsen, Harald, Sonja Eisenbeiss, Meike Hadler & Ingrid Sonnenstuhl. 2001. The mental representation of inflected words: An experimental study of adjectives and verbs in German. *Language* 77(3). 510–543. http://www.jstor.org/stable/3086943.

Draft of January 21, 2019, 13:51

Feldman, Laurie B. & Carol A. Fowler. 1987. The inflected noun system in Serbo-Croatian: Lexical representation of morphological structure. *Memory & Cognition* 15(1). 1–12. DOI:10.3758/BF03197707

Filipović Đurđević, Dušica & Aleksandar Kostić. 2003. Processing of inflected adjectives. *Psihologija* 36(3). 353–378. DOI:10.2298/PSI0303353F

Gor, Kira, Anna Chrabaszcz & Svetlana Cook. 2017. Processing of native and non-native inflected words: Beyond affix stripping. *Journal of Memory and Language* 93. 315–332. DOI:10.1016/j.jml.2016.06.014

Hyönä, Jukka, Matti Laine & Jussi Niemi. 1995. Effects of a word's morphological complexity on readers' eye fixation patterns. In John M. Findlay, Robin Walker & Robert W. Kentridge (eds.), *Eye movement research: Mechanisms, processes and applications*, 445–452. Amsterdam: Elsevier. DOI:10.1016/S0926-907X(05)80037-6

Hyönä, Jukka, Seppo Vainio & Matti Laine. 2002. A morphological effect obtains for isolated words but not for words in sentence context. *European Journal of Cognitive Psychology* 14(4). 417–433. DOI:10.1080/09541440143000131

Katz, Leonard, Suzanne Boyce, Louis M. Goldstein & Georgije Lukatela. 1987. Grammatical information effects in auditory word recognition. *Cognition* 25(3). 235–263. DOI:10.1016/S0010-0277(87)80005-7

Keuleers, Emmanuel. 2013. VWR: Useful functions for visual word recognition research. *R package version* 0.3.0. http://CRAN.R-project.org/package=vwr.

Kostić, Aleksandar. 1991. Informational approach to the processing of inflected morphology: Standard data reconsidered. *Psychological research* 53(1). 62–70. DOI:10.1007/BF00867333

Kostić, Aleksandar. 1995. Information load constraints on processing inflected morphology. In Laurie Beth Feldman (ed.), *Morphological aspects of language processing*, 317–343. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kostić, Aleksandar & Leonard Katz. 1987. Processing differences between nouns, adjectives, and verbs. *Psychological Research* 49(4). 229–236. DOI:10.1007/BF00309031

Kostić, Aleksandar, Tanja Markovic & Aleksandar Baucal. 2003. Inflectional morphology and word meaning: Orthogonal or co-implicative cognitive domains? In R. Harald Baayen & Robert Shreuder (eds.), *Morphological structure in language processing*, 1–43. Berlin: de Gruyter.

Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2017. lmerTest package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82(13). DOI:10.18637/jss.v082.i13

Laine, Matti & Mika Koivisto. 1998. Lexical access to inflected words as measured by lateralized visual lexical decision. *Psychological Research* 61(3). 220–229. DOI:10.1007/s004260050027

Laine, Matti, Seppo Vainio & Jukka Hyönä. 1999. Lexical access routes to nouns in a morphologically rich language. *Journal of memory and language* 40(1). 109–135. DOI:10.1006/jmla.1998.2615

Lenth, Russell V. 2016. Least-squares means: The R package lsmeans. *Journal of statistical software* 69(1). 1–33. DOI:10.18637/jss.v069.i01

Lukatela, Georgije, Claudia Carello & Michael T. Turvey. 1987. Lexical representation of regular and irregular inflected nouns. *Language and Cognitive Processes* 2(1). 1–17. DOI:10.1080/01690968708406349

Lukatela, Georgije, Branislaw Gligorijević, Aleksandar Kostić & Michael T. Turvey. 1980. Representation of inflected nouns in the internal lexicon. *Memory & Cognition* 8(5). 415–423. DOI:10.3758/BF03211138

Lukatela, Georgije, Z. Mandić, Branislav Gligorijević, Aleksandar Kostić, M. Savić & Michael T. Turvey. 1978. Lexical decision for inflected nouns. *Language and Speech* 21(2). 166–173. DOI:10.1177/002383097802100203

Lyashevskaya, Olga N. & Sergej Sharoff. 2009. *Frequency dictionary of the Modern Russian language (on the base of the National Corpus of the Russian Language)*. http://dict.ruslang.ru/freq.php.

Milin, Petar, Dušica Filipović Đurđević & Fermín Moscoso del Prado Martín. 2009. The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language* 60(1). 50–64. DOI:10.1016/j.jml.2008.08.007

Moscoso del Prado Martín, Fermín, Aleksandar Kostić & R. Harald Baayen. 2004. Putting the bits together: An information theoretical perspective on morphological processing. *Cognition* 94(1). 1–18. DOI:10.1016/j.cognition.2003.10.015

Müller, Gereon. 2004. On decomposing inflection class features: Syncretism in Russian noun inflection. In Gereon Müller, Lutz Gunkel & Gisela Zifonun (eds.), *Explorations in nominal inflection*, 189–228. Berlin: Walter de Gruyter.

Niemi, Jussi, Matti Laine & Päivi Koivuselkä-Sallinen. 1991. Recognition of Finnish polymorphemic words: Effects of morphological complexity and inflection vs. Derivation. In Jussi Niemi (ed.), *Papers from the 18th Finnish Conference of Linguistics*, 114–132. University of Joensuu.

Niemi, Jussi, Matti Laine & Juhani Tuominen. 1994. Cognitive morphology in Finnish: Foundations of a new model. *Language and Cognitive Processes* 9(3). 423–446. DOI:10.1080/01690969408402126

Opitz, Andreas, Stefanie Regel, Gereon Müller & Angela D. Friederici. 2013. Neurophysiological evidence for morphological underspecification in German strong adjective inflection. *Language* 89(2). 231–264. DOI:10.1353/lan.2013.0033

Penke, Martina, Ulrike Janssen & Sonja Eisenbeiss. 2004. Psycholinguistic evidence for the underspecification of morphosyntactic features. *Brain and Language* 90(1–3). 423–433. DOI:10.1016/S0093-934X(03)00453-X

R Core Team. 2014. *A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. https://www.r-project.org.

Samojlova, Maria & Natalia Slioussar. 2014. *Frequencies of different grammatical features and inflectional affixes of Russian nouns: A database.* http://www.slioussar.ru/freqdatabase.html.

Ševa, Nada & Aleksandar Kostić. 2003. Annotated corpus and the empirical evaluation of probability estimates of grammatical forms. *Psihologija* 36(3). 255–270. DOI:10.2298/PSI0303255S

Todorović, Dejan. 1988. Hemispheric differences in case processing. *Brain and Language* 33(2). 365–389. DOI:10.1016/0093-934X(88)90073-9

Vasilyeva, Maria D., Maria V. Falikman & Olga V. Fedorova. 2014. *Slovoizmenenie i slovoobrazovanie v mental'nom leksikone: Eksperimental'noe issledovanie na materiale russkogo yazyka [in Russian, 'Russian nominal inflection and derivation: Evidence from Russian'].* Published in the book of abstracts of *The Sixth International Conference on Cognitive Science,* 197–199, Kaliningrad.

Wiese, Bernd. 2004. Categories and paradigms. On underspecification in Russian declension. In Gereon Müller, Lutz Gunkel & Gisela Zifonun (eds.), *Explorations in nominal inflection*, 321–372. Berlin: Mouton de Gruyter, Berlin.

Wunderlich, Dieter. 1996. Minimalist morphology: The role of paradigms. In Jaap van Marle Geert Booij (ed.), *Yearbook of morphology 1995*, 93–114. Dordrecht: Kluwer.

Wunderlich, Dieter. 2004. Is there any need for the concept of directional syncretism. In Gereon Müller, Lutz Gunkel & Gisela Zifonun (eds.), *Explorations in nominal inflection*, 373–395. Berlin: Mouton de Gruyter, Berlin.

Zaliznyak, Andrey. 1977. *Grammatical dictionary of Russian.* Russky yazyk, Moscow.

Appendix A: Experimental items

Lemma frequency counts are given in brackets.

Feminine nouns (47.39)

anketa 'questionnaire' (14.4), *arfa* 'harp' (2.6), *astra* 'aster' (3.8), *aura* 'aura' (5.1), *beseda* 'conversation' (87.5), *bukva* 'letter. character' (63.5), *data* 'date' (49.5), *doza* 'dose' (22.4), *dyuna* 'dune' (2), *fleita* 'flute' (5.8), *gazeta* 'newspaper' (237.5), *gitara* 'guitar' (22.2), *kareta* 'carriage' (9.4), *karta* 'map' (103), *kassa* 'cashier's desk' (20.9), *klumba* 'flower-bed' (8.7), *klyaksa* 'blot' (4.5), *kofta* 'jacket' (7.7), *lampa* 'lamp' (34), *lapa* 'paw' (39.7), *lenta* 'ribbon' (35.9), *lira* 'lyre' (8), *lyustra* 'lustre' (9.9), *mera* 'measure' (284.3), *minuta* 'minute' (344.2), *moneta* 'coin' (17.5), *norma* 'norm' (111.3), *orbita* 'orbite' (15), *pal'ma* 'palm tree' (14.3), *pasta* 'paste' (6.3), *pochva* 'soil' (56.2), *poza* 'pose' (29.8), *raketa* 'rocket' (62.9), *rama* 'frame' (21.2), *rana* 'wound' (29.4), *rasa* 'race' (5.9), *rifma* 'rhyme' (8.5), *roza* 'rose' (42.7), *shakhta* 'pit' (20.7), *shina* 'tire' (15.3), *shirma* 'folding-screen' (5.3), *shkola* 'school' (316), *shlyapa* 'hat' (34.2), *shuba* 'furcoat' (18.7), *shvabra* 'mop' (3.4), *summa* 'sum' (130.6), *trassa* 'route' (32.5), *travma* 'trauma' (19.6), *tsifra* 'numeric' (62.2), *tsitata* 'citation' (21.5), *tykva* 'pumpkin' (5), *vaza* 'vase' (14.3), *yakhta* 'yacht' (9.5), *yurta* 'yurt' (2.7)

Masculine nouns (47.37)

al'bom 'album' (23.7), *ananas* 'pineapple' (3.6), *aromat* 'aroma' (22.9), *aspekt* 'aspect' (35.6), *atom* 'atom' (20.5), *banan* 'banana' (7.3), *baton* 'loaf (of bread)' (5.3), *bufet* 'buffet' (20), *buton* 'bud' (4.6), *desert* 'dessert' (4), *diplom* 'diploma' (25.8), *divan* 'sofa' (60.1), *dzhip* 'jeep' (14.7), *fontan* 'fountain' (18.4), *frukt* 'fruit' (21.6), *gimn* 'hymn' (14.8), *ideal* 'ideal' (36), *kanat* 'rope' (9), *kapriz* 'caprice' (7.1), *kedr* 'cedar' (6.1), *khalat* 'bathrobe' (36.1), *klad* 'treasure' (7.5), *komod* 'dresser' (5.2), *kontur* 'contour' (15.3), *kostyum* 'costume. suit' (81.3), *kurort* 'resort' (12.8), *metall* 'metal' (57.5), *moment* 'moment' (306.8), *nrav* 'temper' (17.8), *ofis* 'office' (34.1), *period* 'period' (204.2), *plan* 'plan' (235.3), *pled* 'plaid' (4.8), *reis* 'flight. voyage' (22), *remont* 'reparation' (64.2), *ritm* 'rhythm' (30.6), *romb* 'rhombus' (1.8), *rulon* 'roll' (4.3), *servis* 'service' (14.6), *sezon* 'season' (69.2), *shram* 'scar' (10.7), *shtraf* 'forfeit' (32.3), *simvol* 'symbol' (46.4), *sous* 'sauce' (10.8), *syuzhet* 'storyline' (56.6), *teatr* 'theater' (305.3), *tekst* 'text' (146.2), *temp* 'tempo' (49), *tovar* 'item of goods' (115.5), *tsikl* 'cycle' (43.6), *virus* 'virus' (106.5), *vulkan* 'volcano' (6), *yarus* 'tier. layer' (6.5), *zhanr* 'genre' (36)

Appendix B: Results of statistical analyses

Table 5: Experiment 1 (singular nouns): final models for RTs and accuracy

Random effects

Groups	RTs			Accuracy		
	Name	Variance	SD	Name	Variance	SD
Subject	(Intercept)	0.001	0.031	(Intercept)	0.143	0.378
Lexeme	(Intercept)	0.027	0.163	(Intercept)	0.96	0.98
Residual		0.032	0.179			

Fixed effects

	RTs					Accuracy			
	B	SE	df	t-value	Pr(>|t|)	B	SE	z-value	Pr(>|z|)
(Intercept)	6.58	0.026	64	257.46	<0.001	4.71	0.414	11.4	<0.001
Trial	−0.012	0.003	4541	−4.45	<0.001				
PC2	0.018	0.003	121	6.03	<0.001	−0.364	0.062	−5.86	<0.001
PC3	−0.022	0.006	103	−3.68	<0.001	0.509	0.104	4.89	<0.001
Case[Acc]	0.067	0.013	4538	5.22	<0.001	−1.282	0.442	−2.9	0.004
Case[Dat]	0.1	0.013	4490	7.7	<0.001	−1.906	0.417	−4.57	<0.001
Case[Gen]	0.048	0.013	4565	3.72	<0.001	−1.356	0.445	−3.05	0.002
Case[Ins]	0.042	0.013	4472	3.32	0.001	−0.952	0.454	−2.1	0.036
Case[Loc]	0.089	0.013	4488	6.82	<0.001	−2.07	0.413	−5.02	<0.001
Gender[m]	−0.023	0.014	816	−1.63	0.104	0.001	0.536	0.003	0.998
Case[Acc] : Gender[m]	−0.077	0.018	4510	−4.32	<0.001	0.655	0.643	1.02	0.308
Case[Dat] : Gender[m]	−0.0005	0.018	4521	−0.027	0.979	0.503	0.594	0.847	0.397
Case[Gen] : Gender[m]	0.100	0.018	4514	5.56	<0.001	−0.545	0.607	−0.898	0.369
Case[Ins] : Gender[m]	0.056	0.018	4470	3.13	0.002	0.539	0.667	0.808	0.419
Case[Loc] : Gender[m]	0.082	0.019	4493	4.38	<0.001	−0.915	0.571	−1.6	0.109
PC3 : Gender[m]	0.024	0.008	103	3.2	0.002	−0.472	0.13	−3.63	<0.001

Table 6: Experiment 2 (plural nouns): final models for RTs and accuracy

Random effects

Groups		RTs				Accuracy		
	Name	Variance	SD			Name	Variance	SD
Subject	(Intercept)	0.002	0.041			(Intercept)	0.323	0.569
Lexeme	(Intercept)	0.021	0.145			(Intercept)	0.612	0.782
Residual		0.035	0.188					

Fixed effects

	B	SE	df	t-value	Pr(>|t|)	B	SE	z-value	Pr(>|z|)
(Intercept)	6.68	0.024	76	279.9	<0.001	3.265	0.271	12.03	<0.001
PC2	0.019	0.004	126	4.98	<0.001	−0.285	0.068	−4.21	<0.001
PC3	−0.022	0.007	111	−2.91	0.004	0.442	0.12	3.7	<0.001
Case[Acc]	0.034	0.013	4342	2.52	0.012	0.205	0.332	0.618	0.537
Case[Dat]	0.094	0.014	3769	6.49	<0.001	−0.109	0.3	−0.364	0.716
Case[Gen]	0.087	0.015	3645	5.76	<0.001	−1.505	0.268	−5.61	<0.001
Case[Ins]	0.103	0.014	4419	7.42	<0.001	−0.385	0.294	−1.31	0.191
Case[Loc]	0.067	0.014	4311	4.77	<0.001	−0.345	0.292	−1.18	0.237
Gender[m]	0.017	0.016	589	1.06	0.288	0.179	0.345	0.52	0.603
Case[Acc] : Gender[m]	−0.046	0.019	4342	−2.41	0.016	−0.481	0.455	−1.06	0.29
Case[Dat] : Gender[m]	−0.011	0.019	4437	−0.58	0.565	−0.357	0.42	−0.85	0.395
Case[Gen] : Gender[m]	−0.026	0.020	4269	−1.29	0.197	1.534	0.428	3.58	<0.001
Case[Ins] : Gender[m]	0.015	0.019	4391	0.76	0.448	−0.183	0.42	−0.44	0.662
Case[Loc] : Gender[m]	−0.010	0.019	4426	−0.53	0.594	−0.143	0.415	−0.34	0.731
PC3 : Gender[m]	0.021	0.009	107	2.28	0.025	−0.423	0.155	−2.73	0.006

Table 7: Experiment 1 (singular nouns): paired contrasts for feminine and masculine nouns in different cases (analyses with $p \leq .05$ are given in bold)

|				RTs				accuracy		
				Δ	df	t	p	Δ	z	p
f	Nom	vs.	Gen	**−0.048**	**4565**	**−3.72**	**<0.001**	**1.356**	**3.05**	**0.006**
			Acc	**−0.067**	**4538**	**−5.22**	**<0.001**	**1.282**	**2.9**	**0.008**
			Ins	−0.042	4472	−3.32	0.002	0.952	2.1	0.055
			-e	**−0.094**	**4495**	**−8.44**	**<0.001**	**1.99**	**5.02**	**<0.001**
	Gen	vs.	Acc	−0.019	4497	−1.48	0.174	−0.074	−0.22	0.836
			Ins	0.006	4565	0.47	0.666	−0.404	−1.16	0.321
			-e	**−0.046**	**4561**	**−4.06**	**<0.001**	**0.633**	**2.36**	**0.032**
	Acc	vs.	Ins	0.025	4516	1.95	0.07	−0.33	−0.96	0.421
			-e	**−0.027**	**4503**	**−2.38**	**0.027**	**0.708**	**2.68**	**0.015**
	Ins	vs.	-e	**−0.052**	**4476**	**−4.65**	**<0.001**	**1.037**	**3.65**	**0.001**
m	-∅	vs.	Gen	**−0.154**	**4505**	**−13.8**	**<0.001**	**1.587**	**5.39**	**<0.001**
			Dat	**−0.104**	**4524**	**−9.42**	**<0.001**	**1.089**	**3.54**	**<0.001**
			Ins	**−0.103**	**4472**	**−9.53**	**<0.001**	0.1	0.25	0.836
			Loc	**−0.176**	**4491**	**−14.69**	**<0.001**	**2.675**	**9.98**	**<0.001**
	Gen	vs.	Dat	**0.049**	**4557**	**3.77**	**<0.001**	−0.499	−1.85	0.092
			Ins	**0.051**	**4494**	**3.98**	**<0.001**	**−1.488**	**−4.1**	**<0.001**
			Loc	−0.022	4514	−1.6	0.141	**1.088**	**4.95**	**<0.001**
	Dat	vs.	Ins	0.002	4506	0.13	0.899	**−0.989**	**−2.64**	**0.015**
			Loc	**−0.071**	**4498**	**−5.2**	**<0.001**	**1.586**	**6.65**	**<0.001**
	Ins	vs.	Loc	**−0.073**	**4483**	**−5.41**	**<0.001**	**2.575**	**7.54**	**<0.001**
f	vs.		m	−0.005	100	−0.59	0.608	−0.035	−0.21	0.836
Ins	f vs. m			**−0.034**	**830**	**−2.4**	**0.027**	−0.534	−1.29	0.27
$-u_f$ vs. $-u_m$				−0.01	836	−0.7	0.548	0.125	0.39	0.786
$-e_f$ vs. $-e_m$				**−0.054**	**739**	**−3.99**	**<0.001**	**1.004**	**5.14**	**<0.001**
$-e_f$ vs. $-u_m$				0.017	589	1.31	0.226	**−0.583**	**−2.34**	**0.032**
$-a_f$ vs. $-∅_m$				**0.027**	**551**	**2.16**	**0.045**	0.318	0.71	0.566

Table 8: Experiment 2 (plural nouns): paired contrasts for feminine and masculine nouns in different cases (analyses with $p ≤ .05$ are given in bold)

			RTs				accuracy		
			Δ	df	*t*	*p*	Δ	*z*	*p*
-*y*	vs.	Gen	**−0.068**	**4096**	**−7.7**	**<0.001**	**0.72**	**4**	**<0.001**
		Dat	**−0.082**	**3367**	**−9.26**	**<0.001**	0.27	1.51	0.245
		Ins	**−0.105**	**4400**	**−12.39**	**<0.001**	**0.459**	**2.62**	**0.038**
		Loc	**−0.056**	**4134**	**−6.48**	**<0.001**	0.399	2.29	0.058
Gen	vs.	Dat	−0.014	4409	−1.41	0.229	−0.45	−2.38	0.057
		Ins	**−0.036**	**4439**	**−3.58**	**0.001**	−0.261	−1.36	0.251
		Loc	0.012	4380	1.25	0.276	−0.321	−1.72	0.186
Dat	vs.	Ins	**−0.022**	**4290**	**−2.2**	**0.046**	0.189	0.98	0.424
		Loc	**0.026**	**4390**	**2.71**	**0.013**	0.129	0.7	0.575
Ins	vs.	Loc	**0.049**	**4440**	**4.9**	**<0.001**	−0.06	−0.32	0.785
f	vs.	m	−0.005	102	−0.48	0.686	−0.232	−1.4	0.251
-*y*	f vs. m		−0.005	255	−0.4	0.692	0.07	0.27	0.785
Gen	f vs. m		0.008	644	0.49	0.686	**−1.704**	**−5.68**	**<0.001**

Did you like this book?

This book was brought to you for free
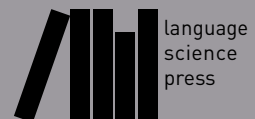
Advances in formal Slavic linguistics 2016

Advances in Formal Slavic Linguistics 2016 initiates a new series of collective volumes on formal Slavic linguistics. It presents a selection of high quality papers authored by young and senior linguists from around the world and contains both empirically oriented work, underpinned by up-to-date experimental methods, as well as more theoretically grounded contributions. The volume covers all major linguistic areas, including morphosyntax, semantics, pragmatics, phonology, and their mutual interfaces. The particular topics discussed include argument structure, word order, case, agreement, tense, aspect, clausal left periphery, or segmental phonology. The topical breadth and analytical depth of the contributions reflect the vitality of the field of formal Slavic linguistics and prove its relevance to the global linguistic endeavour. Early versions of the papers included in this volume were presented at the conference on Formal Description of Slavic Languages 12 or at the satellite Workshop on Formal and Experimental Semantics and Pragmatics, which were held on December 7–10, 2016 in Berlin.

9 783961 101276