NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Natalia Slioussar

# EXPERIMENTAL STUDY OF SEVERAL CORE CONCEPTS OF THEORETICAL MORPHOLOGY (ON THE MATERIAL OF RUSSIAN): REGULARITY, SYNCRETISM, MARKEDNESS

The dissertation was prepared at the National Research University "Higher School of Economics."

**Publications**

Seven publications were selected for the defense:

1. Alexeeva, S., Slioussar, N., Chernova, D. StimulStat: a lexical database for Russian // Behavior Research Methods, online first. DOI 10.3758/s13428-017-0994-3.

2. Kireev, M.V., Slioussar, N., Korotkov, A.D., Chernigovskaya, T.V., Medvedev, S.V. Changes in functional connectivity within the fronto-temporal brain network induced by regular and irregular Russian verb production // Frontiers in Human Neuroscience. Vol. 9. Art. 36. 2015. P. 120–129. DOI 10.3389/fnhum.2015.00036.

3. Slioussar, N. Forms and features: the role of syncretism in number agreement attraction // Journal of Memory and Language. Vol. 101. 2018. P. 51–63. DOI 10.1016/j.jml.2018.03.006.

4. Slioussar, N. Gender, declension and stem-final consonants: an experimental study of gender agreement in Russian // *Computational Linguistics and Intellectual Technologies*. Vol. 17. 2018. P. 688-700.

5. Slioussar, N. Processing of a free word order language: The role of syntax and discourse context // Journal of Psycholinguistic Research. Vol. 40. 2011. P. 291–306. DOI 10.1007/s10936-011-9171-5.

6. Slioussar, N., Kireev, M.V., Chernigovskaya, T.V., Kataeva, G.V., Korotkov, A.D., Medvedev, S.V. An ER-fMRI study of Russian inflectional morphology // Brain and Language. Vol. 130. 2014. P. 33–41. DOI 10.1016/j.bandl.2014.01.006.

7. Slioussar, N., Malko, A. Gender agreement attraction in Russian: production and comprehension evidence // Frontiers in Psychology. Vol. 7. Art. 1651. 2016. P. 1–20. DOI 10.3389/fpsyg.2016.01651.

The results of the present study have also been presented in the following papers:

8. Alexeeva, S., Frolova, A., Slioussar, N. Data from Russian help to determine in which languages the Possible Word Constraint applies. Journal of Psycholinguistic Research. Vol. 46. 2017. P. 629–640.

9. Alexeeva, S., Slioussar, N. Orfografičeskie sosedi v russkom jazyke: baza dannyx i eksperiment, napravlennyj na izučenie morfologičeskoj dekompozicii (in Russian,

'Orthographical neighbors in Russian: a database and an experiment addressing the problem of morphological decomposition') // Voprosy Psikholingvistiki. Vol. 32. 2017. P. 12–27.

10. Magomedova, V., Slioussar, N. Dannye interneta v issledovanii jazykovyx izmenenij: analiz čeredovanij v russkix komparativax i programma dlja raboty s takimi dannymi (in Russian, 'Internet data in the study of language change: A case study of alternations in Russian comparatives and a program to work with such data') // *Computational Linguistics and Intellectual Technologies*. Vol. 13. 2014. P. 379–390.

11. Magomedova, V., Slioussar, N. Paradigm leveling: The decay of consonant alternations in Russian // Perspectives on Morphological Organization: Data and Analyses / ed. by F. Kiefer et al. Leiden: Brill, 2017. P. 123–137.

12. Malko, A., Slioussar, N. Ošibki pri soglasovanii po rodu: eksperimental'noe issledovanie na materiale russkogo jazyka (in Russian, 'Attraction errors in gender agreement: An experimental study on Russian'). St. Petersburg State University Bulletin (Series 9). Vol. 1. 2013. P. 146–154.

13. Slioussar, N. (2008). Russkij porjadok slov v tipologičeskoj perspektive: pozicii podležaščego v imenitel'nom padeže i nekotoryx glagol'nyx form (in Russian, 'Russian word order in the typological perspective: Positions of Nominative subjects and certain verb forms'). St. Petersburg State University Bulletin (Series 9). Vol. 1. 2008. P. 217–225.

14. Slioussar, N. Russian and the EPP requirement in the Tense domain // Lingua. 2011. Vol. 121. P. 2048–2068.

15. Slioussar, N., Alexeeva, S. Orfograficheskie sosedi s zamenoj bukvy pri izuchenii mexanizmov leksicheskogo dostupa (in Russian: "Substitution orthographic neighbors in the study of lexical access") // Computational Linguistics and Intellectual Technologies. Vol. 16. 2017. P. 407–418.

16. Valova, E., Slioussar, N. Sravnenie korpusnogo i eksperimental'nogo metoda na primere issledovanija sintaksičeskix svojstv enklitiki 'že' (in Russian, 'Comparing corpus-based and experimental research methods: a study of the syntactic properties of the Russian enclitic 'zhe'') // Computational Linguistics and Intellectual Technologies. Vol. 15. 2016. P. 792–802.

**Conference presentations and grants**

The main results and conclusions of the present study have been presented in 2011–2018 in 65 oral and poster presentations at 46 international conferences, including:

- Architectures and Mechanisms of Language Processing (AMLaP) Conference (2013, 2015, 2016, 2017, 2018);
- Annual CUNY Conference on Human Sentence Processing (2014, 2015, 2016, 2017);
- International Morphological Processing Conference (2013, 2015, 2017);
- International Conference on the Mental Lexicon (2014, 2016);
- Formal Approaches to Slavic Linguistics (FASL) Conference (2012, 2013, 2015, 2016, 2017);
- Annual Meeting of North East Linguistic Society (NELS) (2016);
- International Morphology Meeting (IMM) (2014, 2016, 2018);
- American International Morphology Meeting (AIMM) (2012);
- "Night Whites" Workshop on Experimental Studies of Speech and Language (2011, 2014, 2018);
- International Organization of Psychophysiology (IOP) World Congress (2016, 2018);
- International Computer Linguistics Conference "Dialogue" (2013, 2015, 2018);
- International Conference on Cognitive Science (2014, 2016).

# 1. Introduction

The papers collected in this dissertation study are dedicated to the experimental study of inflectional morphology in Russian. In six papers, we study the relevance of several core morphological concepts for production and processing, including morphological regularity, productivity, feature markedness, different types of syncretism. As a result, we also learn something new about these concepts, so the main idea of the dissertation is to show how theory and experiment can interact fruitfully. The seventh paper presents a lexical database that contains more than 50 thousand Russian words with more than 1.7 million word forms characterized for more than 70 different parameters. It can be used to facilitate stimulus selection for experimental research. The papers study production and processing of isolated word forms and word forms in a sentence (primarily on the example of number and gender agreement) and in a wider discourse context, using behavioral and neuroimaging methods.

# 2. Production and processing of isolated word forms

Papers selected for the defense: (Slioussar et al. 2014; Kireev et al. 2015).

Inflectional morphology is at the center of an important debate in cognitive science, concerning the general principles according to which the mental lexicon is organized and the foundational properties of our cognitive architecture. This debate focuses on the distinction between morphologically regular and irregular word forms. The so-called "dual route" approach assumes that the former are generated and processed by symbolic rules, while the latter stored in the lexicon, from where they can be retrieved through associative memory mechanisms (e.g. Clahsen 1999; Marslen-Wilson and Tyler 1997; Orsolini and Marslen-Wilson 1997; Pinker 1991, 1999; Pinker and Prince 1988; Ullman 2004). According to the "single route" approach, our cognitive architecture does not rely on symbolic rules and all word forms are computed by a single integrated system (e.g. MacWhinney and Leinbach 1991; McClelland and Patterson 2002; Plunkett and Marchman 1993; Ragnasdóttir et al. 1999; Rumelhart and McClelland 1986).

Behavioral studies testing DR and SR approaches analyze a variety of languages, but neuroimaging studies rely primarily on English and German data (e.g. Beretta et al. 2003; de Diego-Balaguer et al. 2006; Desai et al. 2006; Indefrey et al. 1997; Jaeger et al. 1996; Joanisse and Seidenberg 2005; Marslen-Wilson and Tyler 1998; Münte et al. 1999; Newman et al. 2007; Oh et al. 2011; Sach et al. 2004; Sahin et al. 2006; Ullman et al. 1997). Inflectional morphology

in morphologically richer languages like Finnish, Polish and Arabic was examined in a number of neuroimaging studies, but these studies did not focus on the regular vs. irregular distinction. This is why we decided to address the problem of regularity in an fMRI study of Russian, a language with rich and diverse morphology.

The Russian verb system is very complex, and there are several approaches to dividing verbs into inflectional classes. These classes differ in type frequency, and four or five of them (depending on the chosen classification) are productive. The most frequent productive AJ class was observed to behave as the default class in several previous behavioral studies (Gor and Chernigovskaya 2001, 2003). Thus, there is no obvious division into regular and irregular verbs, and, if any distinctions between different classes are found, we may be able to explore the nature of regularity and to find out which properties (type frequency, productivity or defaultness) are relevant for production and processing.

In our first experiment, we decided to look at the two poles of the verb class system, comparing verbs from the most frequent and productive AJ class to verbs from small unproductive classes (we reasoned that if any differences between these two groups were found, we could compare them to other verbs in subsequent studies). Participants were asked to generate present tense forms from different visually presented real and nonce verbs and to pluralize real and nonce nouns. First we performed a subtractive analysis of the data (Slioussar et al. 2014), then a ROI – whole brain voxel-wise analysis of context dependent changes in functional connectivity (PPI analysis) (Kireev et al. 2015).

We demonstrated that morphological regularity and processing difficulty effects can be teased apart. Activity of the left inferior frontal gyrus (LIFG) was greater for production of irregular verbs (compared to regular ones) and real verbs (compared to nonce ones), so this pattern was explained by processing difficulty (Slioussar et al. 2014). But the functional connectivity of the LIFG with temporal lobe was relatively increased when regular verbs were produced (compared to irregular ones). Nothing similar was found for the real/nonce distinction, so this was concluded to be a genuine regularity effect (Kireev et al. 2015).

Let us discuss the PPI results in more detail. Firstly, we found that functional connectivity between the LIFG and bilaterally distributed clusters in the superior temporal gyri was significantly greater in regular real verb trials than in irregular ones. No other comparisons gave significant results. Secondly, we observed a significant positive covariance between the number of mistakes in irregular real verb trials and the increase in functional connectivity between LIFG

and the right anterior cingulate cortex in these trails as compared to regular ones. Thus, regularity and processing difficulty effects could be dissociated not only using different methods of analysis, but also by the PPI analysis alone.

We found only one previous published PPI study of inflectional morphology (Stamatakis et al. 2005). In this study, functional connectivity between functionally predefined ROIs was assessed during the same/different judgment task. Stimuli were orally presented pairs of English words and nonce words (e.g. *jumped – jump, thought – think, jade – jay*). Thus, the method and materials were very different from ours. Our first finding was similar to what Stamatakis et al. reported, which shows that the observed regularity effect is very robust, being valid crosslinguistically both for production and comprehension.

As for the second finding, Stamatakis et al. have similar results going in the opposite direction. This is also true for the subtractive analysis of their data reported in (Tyler et al. 2005). We hypothesized that the processing difficulty went in the opposite directions in the two studies. Tyler et al. and Stamatakis et al. looked at stimulus pairs like *stayed – stay* vs. *taught – teach.* In regular pairs, the first stimulus was morphologically complex and the second was not, while in irregular pairs, both stimuli were morphologically simple. Thus, regular verb trials induced more processing load. In our study, all verb forms participants read or produced were morphologically complex. But irregular verbs involved various alternations in the stems etc., so irregular verb trials induced more processing load.

The two verb groups used in the first experiment differed by many properties (type frequency, productivity, defaultness), so it was impossible to determine which one was responsible for the observed regularity effect. Therefore we conducted a second fMRI study adding a third verb class — the I class (very frequent, but less frequent than the AJ class, productive, but not default). The results have already been presented at several conferences, but have not been published yet.

In this experiment, we also wanted to find out whether the observed effect would be replicated in comprehension. However, we were worried that in case of isolated word forms, processing of inflectional morphology would be shallow because it cannot be meaningfully interpreted or used for syntactic parsing. The only published fMRI study where regular vs. irregular verbs were compared in comprehension was discussed above (Stamatakis et al. 2005; Tyler et al. 2005), and it is notable that an ingenious design was used to draw participants' attention to the morphological features of stimuli.

This design could not be replicated in Russian, so we came up with the following one. In every trial, participants first saw a real or nonce verb in the infinitive form and a pronoun *ja* 'I' or *on* 'he' below it (600 ms). After an interval, two present tense forms of the previously shown verb appeared on the left and on the right of the screen (1500 ms). One of them agreed with the pronoun, the other did not. Participants were asked to select the correct form. We preferred this design to showing only one form (agreeing or not) because this task would involve agreement violations and would focus participants' attention on error detection.

We analyzed BOLD signal changes associated both with the 1st stimulus (an infinitive and a pronoun) and the 2nd stimulus (two present tense forms). Based on these analyses we selected ROIs for the PPI analysis. All non-trivial results were associated with the 2nd stimulus. We showed that the effects of interest are the same in production and in comprehension. The subtractive analysis showed that the activity of the LIFG gradually increased from the AJ class to I class and then to irregular verbs. The effect was analogous to the processing difficulty effect from (Slioussar et al. 2014).

The PPI analysis revealed a connectivity pattern that was very similar to the ones reported in (Kireev et al. 2015) and (Stamatakis et al. 2005). The fact that it was found in the two languages with relatively poor and relatively rich inflectional morphology and in the studies using three different tasks proves that this effect is reliable. Moreover, we found out whether it can be associated with type frequency, productivity or defaultness. The latter was true: an increase in functional connectivity of the LIFG was observed for the AJ class (as opposed to the I class and irregular verbs). Notably, this can be explained only in the dual route approach to inflectional morphology postulating a categorical distinction between the default class and the other classes. In the other approaches, regularity effects, if present at all, are expected to correlate with type frequency and productivity.

## 3. Production and processing of word forms in a sentence

### 3.1. Number agreement attraction

Paper selected for the defense: (Slioussar 2018a).

Three papers presented in this section are dedicated to production and processing of predicative agreement. For this domain, the phenomenon of agreement attraction plays an important role. An example of an attraction error is given (1a): the verb agrees not with the head of the subject NP but with a dependent NP (an attractor).

(1)  a.  *The key to the cabinets were rusty.*

   b.  *The key to the cabinet were rusty.*

Across languages, such errors in number agreement have been shown to arise more frequently than errors of the type exhibited in (1b), where no attraction is possible (e.g. Bock and Miller 1991; Eberhard et al. 2005; Franck et al. 2002, 2006; Hartsuiker et al. 2003; Solomon and Pearlmutter 2004; Staub 2009, 2010; Vigliocco et al. 1995, 1996). In comprehension experiments, attraction errors have been demonstrated to trigger more grammaticality judgment mistakes and to provoke less pronounced effects in reading time and EEG studies than other agreement errors (e.g. Clifton et al. 1999; Dillon et al. 2013; Pearlmutter et al. 1999; Tanner et al. 2014; Wagers et al. 2009).

Two major approaches to agreement attraction can be identified in the literature: representational and retrieval approaches. According to the representational approach (e.g. Brehm and Bock 2013; Eberhard et al. 2005; Franck et al. 2002; Nicol et al. 1997; Staub 2010), agreement attraction takes place because the mental representation of the subject NP's number feature is faulty or ambiguous. Some authors assume that the number feature can "percolate" from the embedded NP to the subject NP. Others, relying primarily on the Marking and Morphing model (Eberhard et al. 2005), argue that the number value of the subject NP is a continuum. The more plural the subject NP, the higher the possibility of choosing a plural verb. This plurality depends on properties of the subject NP as a whole and of its head, such as collectivity, distributivity etc. The retrieval approach (e.g. Badecker and Kuminiak 2007; Dillon et al. 2013; Solomon and Pearlmutter 2004; Wagers et al. 2009) claims that the number feature on the subject NP is always represented unambiguously and correctly and attraction errors arise when the subject NP is accessed via cue-based retrieval to determine the number of the agreeing verb because several nouns are simultaneously active.

Experiments on agreement attraction have been useful to study many important questions: how agreement is produced and processed, whether this morphosyntactic operation can be influenced by semantics and phonology etc. Here, we will focus on the ones directly relevant for our studies. Firstly, in all examined languages, significant attraction effects were found only in the sentences with singular heads and plural attractors, but not in the opposite configuration. We will come back to this asymmetry in section 2.2.

Secondly, data from languages with morphologically marked cases show that syncretism is an important factor. For example, in an experiment on German (Hartsuiker et al. 2003), subjects

like (2a), in which the form of the attractor is ambiguous between accusative and nominative, provoked significantly more errors than subjects like (2b), in which the attractor is unambiguously dative.

(2) a. *die        Stellungnahme gegen   die        Demonstrationen*
       the_{F.NOM.SG} position      against the_{ACC.PL} demonstrations

    b. *die        Stellungnahme zu  den        Demonstrationen*
       the_{F.NOM.SG} position      on  the_{DAT.PL} demonstrations

Notably, the syncretism factor has been examined only in production and never been examined separately from number: in (2a), the attractor is plural and morphologically ambiguous with nominative plural, and in (2b), it is plural and not ambiguous. This is possible to do in Russian, which has forms like *večerinki* 'party_{GEN.SG/NOM.PL/ACC.PL}'. In one production and two comprehension experiments (a speeded grammaticality judgement task and a self-paced reading task), we compared number agreement attraction effects with eight types of subject NPs exemplified in (3)–(4).

(3) a. *trassa / trassy       čerez  pole*
       highway_{NOM.SG / NOM.PL} across field_{ACC.SG(=NOM.SG)} — so-called systematic syncretism

    b. *trassa / trassy       čerez  polja*
       highway_{NOM.SG / NOM.PL} across field_{ACC.PL(=NOM.PL)} — so-called systematic syncretism

(4) a. *komnata / komnaty dlja večerinki*
       room_{NOM.SG / NOM.PL} for  party_{GEN.SG(=NOM.PL)} — so-called accidental syncretism

    b. *komnata / komnaty dlja večerinok*
       room_{NOM.SG / NOM.PL} for  party_{GEN.PL}

No effects were found with plural heads, as it was expected based on the previous findings. With singular heads, both in production and in comprehension syncretic genitive singular forms were found to trigger larger attraction effects than morphologically unambiguous genitive plural forms. Accusative plural forms coinciding with the nominative plural were shown to be the most effective attractors.

These results have implications for different models of attraction and for other discussions in morphology concerning ambiguity processing, different approaches to syncretism and the problem of lexical insertion. In particular, they show that syncretism does not simply boost the effect of the plural feature — it is an independent factor. The paper argues that the results cannot be explained as local coherence phenomena and suggests an explanation in terms of the cue-based retrieval approach to agreement attraction, with two notable modifications. First, the study shows that retrieval cues can tap into alternative feature sets of syncretic forms. Notably,

unlike in the previously discussed cases, this is possible although no parallel syntactic parses are considered and no reanalysis is attempted. Second, it demonstrates that retrieval should rely on compound cues such as {nominative+plural} instead of a set of individual cues.

## 3.2. Gender agreement attraction

<u>Paper selected for the defense</u>: (Slioussar, Malko 2016).

As we noted in the previous section, in all studied languages, attraction effects were found to be asymmetric. They can be observed when the head is singular, and the attractor is plural, but are much weaker or virtually non-existent in the opposite configuration. In the majority of agreement attraction studies, this asymmetry is explained in terms of feature markedness. Plural is assumed to be the marked value of number feature, and the asymmetry is attributed to the to the fact that attractors with a marked feature are more disruptive.

However, many approaches to markedness have been proposed in the literature: the marked value of a certain feature can be determined based on frequency, presence of a non-zero affix, default use of a form (e.g., in impersonal sentences), various semantic tests etc. An overview can be found in (Haspelmath 2006). It is impossible to evaluate which approach is psycholinguistically relevant looking only at singular vs. plural — other features should be studied. We conducted one production and three self-paced reading experiments on gender agreement attraction in Russian (Slioussar, Malko 2016). There were very few production experiments (Badecker and Kuminiak 2007; Franck et al. 2008; Martin et al. 2014; Vigliocco, Franck, 1999) and no previous comprehension studies looking at gender.

We compared sentences with subject NPs with different gender combinations, as in (5a-b). Badecker and Kuminiak's (2007) study on Slovak was especially relevant for us. Both Slovak and Russian have three genders. Masculine is the most frequent, while neuter is used as grammatical default (in impersonal sentences). Badecker and Kuminiak found that in production, feminine behaves as more marked than masculine and neuter, and masculine as more marked than neuter. Our production experiments observed the same pattern in Russian.

(5) a. *okno      v  pole / vo dvor*
       window$_N$ to field$_N$  to  yard$_M$

    b. *vyxod  v  les  /  v  pole*
       exit$_M$  to forest$_M$ to  field$_N$

Number agreement attraction studies found parallel patterns in production and comprehension, but our comprehension experiments revealed a different pattern: masculine, the most frequent

gender, triggered weakest effects. Thus, our study shows that different approaches to markedness are psycholinguistically relevant. There was another important finding: in the comprehension experiments attraction was observed for all dependent noun genders, but only for a subset of head noun genders. This goes against the traditional assumption that the features of the dependent noun are crucial for attraction, showing the features of the head are more important. We demonstrated that this approach could be extended to previous findings on attraction and that there existed other evidence for it.

**3.3. Prototypicality of inflections in agreement processing**

Paper selected for the defense: (Slioussar 2018b).

The experiment presented in this section (Slioussar 2018b) is also dedicated to subject-predicate gender agreement processing, but studies it without attraction, focusing on two factors: the gender of the noun and its inflectional class, or declension. Russian nouns are inflected for case and number, and, depending on the set of their inflections, are divided into several declensions. We will rely on the most widely accepted system presented in Table 1. As Table 1 shows, most consonant-final nominative singular forms (having zero inflection) are masculine, and most feminine nominative singular forms end in -*a/ja*, with 3rd declension consonant-final feminine nouns (3D-F) being less usual, or non-typical.

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

Table 1. Declension and gender in Russian nouns (excluding substantivized adjectives).

Many experimental studies found differences between nouns with more and less typical inflections in a variety of languages (e.g. Andonova et al. 2004; Bates et al. 1995; Gollan and Frost 2001; Spalek et al. 2008). However, these studies usually looked at the processing of isolated nouns. Among the few sentence-processing studies, 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

---

[1] Slioussar and Samoilova (2015) identified the frequency of nouns with different grammatical characteristics in the grammatically disambiguated subcorpus of the Russian National Corpus (http://www.ruscorpora.ru).

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 gender agreement attraction.

Only one paper (Taraban and Kempe, 1999) addressed this problem in Russian, and no differences between 2$^{nd}$ declension masculine and 3$^{rd}$ declension feminine nouns were found for native speakers. The task involved selecting a correct verb form for different subject nouns. We decided to come back to this question using the self-paced reading method, which could allow us tracing earlier processing stages. Materials were sentence sets in six experimental conditions like in (6).

(6)   *Xalat / kurtka / šinel' + byl / byla potrepannym / potrepannoj + ot mnogoletnej noski.*
      robe$_{2D-M}$ / jacket$_{1D-F}$ / overcoat$_{3D-F}$ + was$_{M/F}$ shabby$_{M/F}$ + from years-long wear

Firstly, the influence of the two factors of interest could 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). Secondly, declension played a role at a very early stage and its effect was very short-lived, while the role of gender became visible later and its effect was more pronounced. Namely, agreement errors were noticed significantly later with non-typical 3D-F nouns than with typical 1D-F or 2D-M ones. But error-related delay on subsequent words was significantly more pronounced for M subjects than for F ones (both 1D and 3D).

The latter result suggests that predictions we make about predicate gender are stronger for M subjects. This finding agrees with Slioussar and Malko's (2016) study presented in the previous section, as well as with experimental studies of adjectival agreement in Russian (Akhutina et al. 1999, 2001; Romanova and Gor 2017). The former result can be explained by the fact that 3D-F nouns essentially look like masculine. But an alternative explanation is also possible: all agreement errors (in M or in N) may be harder to detect after 3D-F nouns, i.e. their gender can be in general harder to retrieve. A follow-up experiment is conducted to tease apart these two scenarios. Another current experiment looks at non-typical masculine nouns like *papa* 'dad'.

## 3.4. Processing different word orders

Paper selected for the defense: (Slioussar 2011).

The paper presented in this section (Slioussar 2011) studies how different word forms are processed depending on the word order in the sentence and on its information structure (IS),

i.e. what information is given or new. Russian has flexible constituent order. Available orders are used to encode IS, which means that they have different context requirements: for example, an OVS sentence is felicitous when the object is given and the subject is new, but not in zero context.

However, most experimental studies of flexible constituent order in different languages have focused on its syntactic properties, ignoring the context factor. Canonical and noncanonical orders were compared in zero context, and the latter were found to be more difficult to process (e.g., Bader and Meng, 1999; Erdocia, Laka, Mestres-Misse, and Rodriguez-Fornells, 2009; Frazier and Flores d'Arcais, 1989; Hyönä and Hujanen, 1997; Miyamoto and Takahashi, 2002, 2004; Stojanović, 1999; Vasishth, 2002). This was explained by the increased syntactic complexity of noncanonical orders. But an alternative explanation is also possible: noncanonical orders are infelicitous in zero context, which violates their IS requirements.

Context-sensitive studies of scrambling are very rare (Bornkessel and Schlesewsky 2006; Bornkessel et al. 2003; Kaiser and Trueswell 2004).[2] Kaiser and Trueswell (2004) compared Finnish SVO and OVS structures in appropriate and inappropriate contexts. In Finnish, as in most other free word order languages, the vast majority of narrative sentences have a 'given-new' constituent order. So appropriate contexts presupposed such an order in target constructions, while inappropriate ones presupposed a 'new-given' order — i.e., essentially violating the IS requirements of target sentences. Kaiser and Trueswell showed that noncanonical constructions in appropriate contexts were processed faster than those in inappropriate contexts, but were still slower than the canonical order.

In our study on Russian (Slioussar 2011), we used the same methodology, but decided to take longer constructions and more extensive context sentences (making two out of three constituents in the target sentence given). In the first self-paced reading experiment, we compared S V IO DO (canonical), DO S V IO and DO IO V S constructions in appropriate vs. inappropriate contexts. Examples of context sentences are given in (7a-b), examples of target sentences are given in (8a-b). (7a) was used as appropriate context for (8a) and as inappropriate context for (8b) and for another target sentence with the DO S V IO order.

(7) a. *Na 8 marta Danja Kašin pozdravil      Mašu Smolinu.*
        on March 8 [D.K.].NOM   congratulated [M.S.].ACC

---

[2] Sekerina (2003) also compared Russian PP S V IO DO PP and PP DO S V IO PP constructions in isolation and in the *same* context.

b. *Na 8 marta Maša Smolina polučila pljuševogo slona    i    konfety.*
      on March 8 [M.S.].NOM    received [plush elephant].ACC  and  [candies].ACC

(8)  a. *Danja Kašin  podaril  Maše Smolinoj  pljuševogo slona.*    S V IO DO
        [D.K.].NOM    gave     [M.S.].DAT      [plush elephant].ACC

   b. *Pljuševogo slona      Maše Smolinoj  podaril  Danja Kašin.*  DO IO V S
      [plush elephant].ACC [M.S.].DAT       gave     [D.K.].NOM

The context factor was significant, while the syntax factor was not. The less pronounced context effect evidenced in (Kaiser and Trueswell 2004) might be due to the use of shorter target and context sentences. We also demonstrated that the slow-down starts at the first contextually inappropriate constituent, which shows that the information about context requirements is taken into account immediately, but that it develops faster on preverbal subjects and postverbal indirect objects (occupying their canonical symtactic positions) than on preverbal indirect objects (occupying a noncanonical position, or scrambled). In the seconf self-paced reading experiment, these findings were replicated for IO S V DO and IO DO V S orders. S V IO DO orders with a continuation were used to show that there is no additional effect of inappropriate context at the end of the sentence. Thus, processing of word forms in a sentence depends both on their IS status and syntactic position.

## 4. Creating resources for experimental research

Paper selected for the defense: (Alexeeva, Slioussar, Chernova 2017).

Experimental studies identified a large list of word properties that play a role for speech production and comprehension. They include lemma and form frequency, word length, the number of syllables, stress pattern, whether the word has homonyms, homographs or orthographical neighbors, whether it has multiple senses etc. Of course, various grammatical characteristics (part of speech, inflectional paradigm etc.) are also important. For several languages, databases with search tools were designed to take these properties into account during stimulus selection (e.g. Balota et al. 2007; Colheart 1981; Davis 2005; Davis, Perea 2005; Duchon et al. 2013; Heister et al. 2011; New et al. 2004).

As for Russian, some required characteristics were not represented in electronic dictionaries and databases at all, while the others were scattered across frequency lists, grammatical or explanatory dictionaries or others resources. Therefore, we created StimulStat – a lexical database for Russian in the form of a web application (http://stimul.cognitivestudies.ru). The database contains more than 50,000 most frequent Russian words (1.7 million word forms). These words and forms are characterized according to more than 70 properties, including

frequency, length, phonological and grammatical properties, orthographic neighbourhood frequency and size, homonymy and polysemy. Some properties were retrieved from various dictionaries and presented collectively in a searchable form for the first time, the others were computed specifically for the database. In the paper describing the database (Alexeeva, Slioussar, Chernova 2017) we also present some interesting crosslinguistic differences that can be identified using such databases.

## 5. Conclusions

Experiments collected in this dissertation study how word forms are produced and processed in isolation and inside sentences on the material of Russian. Many experiments were the first or among the first to look at a certain phenomenon (e.g. at the factors of syncretism and declension in agreement processing studies, at gender agreement attraction in comprehension) or to apply a particular method (e.g. the PPI method of fMRI data analysis in a linguistic study). The study shows that several notions discussed in theoretical morphology (including morphological regularity, markedness, syncretism) are psycholinguistically relevant and that experimental data can shed new light on them, demonstrating how theory and experiment can interact fruitfully.

# References

Akhutina, T., Kurgansky, A., Kurganskaya, M., Polinsky, M., Polonskaya, N., Larina, O., Bates, E., Appelbaum, M. Processing of grammatical gender in normal and aphasic speakers of Russian // Cortex. Vol. 37. 2001. P. 295–326.

Akhutina, T., Kurgansky, A., Polinsky, M., Bates, E. Processing of grammatical gender in a three-gender system: Experimental evidence from Russian // Journal of Psycholinguistic Research. Vol. 28. 1999. P. 695–713.

Alexeeva, S., Slioussar, N., Chernova, D. StimulStat: a lexical database for Russian // Behavior Research Methods, online first. DOI 10.3758/s13428-017-0994-3.

Andonova, E., D'Amico, S., Devescovi, A., Bates, E. Gender and lexical access in Bulgarian // Perception and Psychophysics. Vol. 66. 2004. P. 496–507.

Badecker, W., Kuminiak, F. Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak // Journal of Memory and Language. Vol. 56. 2007. P. 65–85.

Bader, M., Meng, M. Subject-object ambiguities in German embedded clauses: an across-the-board comparison // Journal of Psycholinguistic Research. Vol. 28. 1999. P. 121–144.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Treiman, R. The English Lexicon Project // Behavior Research Methods. Vol. 39. 2007. P. 445–459.

Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., Hernandez, A. Gender and lexical access in Italian // Perception and Psychophysics. Vol. 57. 1995. P. 847–862.

Beretta, A., Campbell, C., Carr, T. H., Huang, J., Schmitt, L. M., Christianson, K., Cao, Y. An ER-fMRI investigation of morphological inflection in German reveals that the brain makes a distinction between regular and irregular forms // Brain and Language. Vol. 85. 2003. P. 67–92.

Bock, K., Miller, C. A. Broken agreement // Cognitive Psychology. Vol. 23. 1991. P. 45–93.

Bornkessel, I., Schlesewsky, M. The role of contrast in the local licensing of scrambling in German: Evidence from online comprehension // Journal of Germanic Linguistics. Vol. 18. 2006. P. 1-43.

Bornkessel, I., Schlesewsky, M., Friederici, A. Contextual information modulates initial processes of syntactic integration: The role of inter-versus intrasentential predictions // Journal of Experimental Psychology: Learning, Memory and Cognition. Vol. 29. 2003. P. 871–882.

Brehm, L., Bock, K. What counts in grammatical number agreement? // Cognition. Vol. 128. 2013. P. 149–169.

Caffarra, S., A. Siyanova-Chanturia, F. Pesciarelli, F. Vespignani, Cacciari, C. Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian // Psychophysiology. Vol. 52. 2015. P. 1019–1030.

Clahsen, H. Lexical entries and rules of language: a multidisciplinary study of German inflection // Behavioral and brain sciences. Vol. 22. 1999. P. 991–1060.

Clifton, Ch., Jr., Frazier, L., Deevy, P. Feature manipulation in sentence comprehension // Rivista di Linguistica. Vol. 11. 1999. P. 11–39.

Coltheart, M. The MRC psycholinguistic database // The Quarterly Journal of Experimental Psychology, Section A. Vol. 33. 1981. P. 497–505.

Davis, C. J. N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics // Behavior Research Methods. Vol. 37. 2005. P. 65–70.

Davis, C. J., Perea, M. BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish // Behavior Research Methods. Vol. 37. 2005. P. 665–671.

de Diego-Balaguer, R., Rodriguez-Fornells, A., Rotte, M., Bahlmann, J., Heinze, H. J., Münte, T. F. Neural circuits subserving the retrieval of stems and grammatical features in regular and irregular verbs // Human Brain Mapping. Vol. 27. 2006. P. 874–888.

Desai, R., Conant, L. L., Waldron, E., Binder, J. R. fMRI of past tense processing: the effects of phonological complexity and task difficulty // Journal of Cognitive Neuroscience. Vol. 18. 2006. P. 278–297.

Dillon, B., Mishler, A., Sloggett, S., Phillips, C. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence // Journal of Memory and Language. Vol. 69. 2013. P. 85–103.

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., Carreiras, M. EsPal: One-stop shopping for Spanish word properties // Behavior Research Methods. Vol. 45. 2013. P. 1246–1258.

Eberhard, K. M., Cutting, J. C., Bock, K. Making syntax of sense: Number agreement in sentence production // Psychological Review. Vol. 112. 2005. P. 531–559.

English: Evidence from event-related brain potentials // Cognitive Brain Research. Vol. 7. 1999. P. 241–253.

Erdocia, K., Laka, I., Mestres-Misse, A., Rodriguez-Fornells, A. Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque // Brain and Language. Vol. 109. 2009. P. 1–17.

Franck, J., Lassi, G., Frauenfelder, U. H., Rizzi, L. Agreement and movement: A syntactic analysis of attraction // Cognition. Vol. 101. 2006. P. 173–216.

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., Frauenfelder, U. H. The interplay of syntax and form in language production: A cross-linguistic study of form effects on agreement. Language and Cognitive Processes. Vol. 23. 2008. P. 329–374.

Franck, J., Vigliocco, G., Nicol, J. Subject-verb agreement errors in French and English: The role of syntactic hierarchy // Language and Cognitive Processes. Vol. 17. 2002. P. 371–404.

Frazier, L., Flores d'Arcais, G. B. Filler-driven parsing: a study of gap-filling in Dutch // Journal of Memory and Language. Vol. 28. 1989. P. 331–344.

Gollan, T.H., Frost, R. Two routes to grammatical gender: Evidence from Hebrew // Journal of Psycholinguistic Research. Vol. 30. 2001. P. 627–651.

Gor, K., Chernigovskaya, T. Mental Lexicon Structure in L1 and L2 Acquisition: Russian Evidence // Glossos. Vol. 4. 2003 (http://slaviccenters.duke.edu/projects/glossos-journal/issues/issue-4).

Gor, K., Chernigovskaya, T. Rules in the processing of Russian verbal morphology // Current Issues in Formal Slavic Linguistics / ed. G. Zybatow, U. Junghanns, G. Mehlhorn, L. Szucsich. Frankfurt am Main: Peter Lang, 2001. P. 528–536.

Hartsuiker, R. J., Schriefers, H. J., Bock, K., Kikstra, G. M. Morphophonological influences on the construction of subject-verb agreement // Memory and Cognition. Vol. 31. 2003. P. 1316–1326.

Haspelmath, M. Against markedness (and what to replace it with) // Journal of Linguistics. Vol. 42. 2006. P. 25–70.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., Kliegl, R. dlexDB — eine lexikalische Datenbank für die psychologische und linguistische Forschung // Psychologische Rundschau. Vol. 62. 2011. P. 10–20.

Hyönä, J., Hujanen, H. Effects of case marking and word order on sentence parsing in Finnish: an eye fixation analysis // The Quarterly Journal of Experimental Psychology: A Human Experimental Psychology. Vol. 50A. 1997. P. 841–858.

Indefrey, P., Brown, C., Hagoort, P., Herzog, H., Sach, M., Seitz, R. A PET study of cerebral activation patterns induced by verb inflection // Neuroimage. Vol. 5. 1997. P. S548.

Jaeger, J. J., Lockwood, A. H., Kemmerer, D. L., Van Valin, R. D., Jr., Murphy, B. W., Khalak, H. G. A positron emission tomographic study of regular and irregular verb morphology in English // Language. Vol. 72. 1996. P. 451–497.

Joanisse, M. F., Seidenberg, M. S. Imaging the past: Neural activation in frontal and temporal regions during regular and irregular past-tense processing // Cognitive, Affective, and Behavioral Neuroscience. Vol. 5. 2005. P. 282–296.

Kaiser, E., Trueswell, J. C. The role of discourse context in the processing of a flexible word-order language // Cognition. Vol. 94. 2004. P. 113–147.

Kireev, M.V., Slioussar, N., Korotkov, A.D., Chernigovskaya, T.V., Medvedev, S.V. Changes in functional connectivity within the fronto-temporal brain network induced by regular and irregular Russian verb production // Frontiers in Human Neuroscience. Vol. 9. Art. 36. 2015. P. 1–10. DOI 10.3389/fnhum.2015.00036.

MacWhinney, B., Leinbach, J. Implementations are not conceptualizations: revising the verb learning model // Cognition. Vol. 40. 1991. P. 121–157.

Marslen-Wilson, W. D., Tyler, L. K. Dissociating types of mental computation // Nature. Vol. 387. 1997. P. 592–594.

Marslen-Wilson, W. D., Tyler, L. K. Rules, representations, and the English past tense // Trends in Cognitive Sciences. Vol. 2. 1998. P. 428–435.

Martin, A. E., Nieuwland, M. S., Carreiras, M. Agreement attraction during comprehension of grammatical sentences: ERP evidence from ellipsis. Brain and Language. Vol. 135. 2014. P. 42–51.

McClelland, J. L., Patterson, K. Rules or connections in past-tense inflections: what does the evidence rule out? // Trends in Cognitive Sciences. Vol. 6. 2002. P. 465–472.

Miyamoto, E., Takahashi, S. Sources of difficulty in processing scrambling in Japanese // Sentence processing in East-Asian languages / Ed. by M. Nakayama. Stanford, CA: CSLI Publications, 2002. P. 167–188.

New, B., Pallier, C., Brysbaert, M., Ferrand, L. Lexique 2: A new French lexical database. Behavior Research Methods, Instruments, and Computers. Vol. 36. 2004. P. 516–524.

Newman, A. J., Ullman, M. T., Pancheva, R., Waligura, D. L., Neville, H. J. An ERP study of regular and irregular English past tense inflection // NeuroImage. Vol. 34. 2007. P. 435–445.

Nicol, J., Forster, K., Veres, C. Subject–verb agreement processes in comprehension // Journal of Memory and Language. Vol. 36. 1997. P. 569–587.

Oh, T. M., Tan, K. L., Ng, P., Berne, Y. I., Graham, S. The past tense debate: is phonological complexity the key to the puzzle? // Neuroimage. Vol. 57. 2011. P. 271–280.

Orsolini, M., Marslen-Wilson, W. D. Universals in morphological representation: evidence from Italian // Language and Cognitive Processes. Vol. 12. 1997. P. 1–47.

Pearlmutter, N. J., Garnsey, S. M., Bock, K. Agreement processes in sentence comprehension // Journal of Memory and Language. Vol. 41. 1999. P. 427–456.

Pinker, S. Rules of language // Science. Vol. 253. 1991. P. 530–535.

Pinker, S. Words and rules: the ingredients of language. New York: Harper Collins, 1999.

Pinker, S., Prince, A. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition // Cognition. Vol. 28. 1988. P. 73–193.

Plunkett, K., Marchman, V. From rote learning to system building: acquiring verb morphology in children and connectionist nets // Cognition. Vol. 48. 1993. P. 21–69.

Ragnasdóttir, H., Simonsen, H. G., Plunkett, K. The acquisition of past tense morphology in Icelandic and Norwegian children: an experimental study // Journal of Child Language. Vol. 26. 1999. P. 577–618.

Romanova, N., Gor, K. Processing of gender and number agreement in Russian as a second language // Studies in Second Language Acquisition. Vol. 39. 2017. P. 97–128.

Rumelhart, D., McClelland, J. On learning the past tenses of English verbs // Parallel distributed processing: explorations in the microstructure of cognition. Vol. 2. Psychological and biological models / ed. D. Rumelhart, J. McClelland. Cambridge, MA: MIT Press, 1986. P. 216–271.

Sach, M., Seitz, R., Indefrey P. Unified inflectional processing of regular and irregular verbs: a PET study // NeuroReport. Vol. 15. 2004. P. 533–537.

Sahin, N., Pinker, S., Halgren, E. Abstract grammatical processing of nouns and verbs in Broca's area: Evidence from fMRI // Cortex. Vol. 42. 2006. P. 540–562.

Sekerina, I. A. Scrambling and processing: dependencies, complexity, and constraints // Word order and scrambling / Ed. by S. Karimi. Oxford: Blackwell, 2003. P. 301–324.

Slioussar, N. Forms and features: the role of syncretism in number agreement attraction // Journal of Memory and Language. Vol. 101. 2018. P. 51–63. DOI 10.1016/j.jml.2018.03.006.

Slioussar, N. Gender, declension and stem-final consonants: an experimental study of gender agreement in Russian // Computational Linguistics and Intellectual Technologies. Vol. 17. 2018. P. 688-700.

Slioussar, N. Processing of a free word order language: The role of syntax and discourse context // Journal of Psycholinguistic Research. 2011. Vol. 40. P. 291–306. DOI 10.1007/s10936-011-9171-5.

Slioussar, N., Kireev, M.V., Chernigovskaya, T.V., Kataeva, G.V., Korotkov, A.D., Medvedev, S.V. An ER-fMRI study of Russian inflectional morphology // Brain and Language. Vol. 130. 2014. P. 33–41. DOI 10.1016/j.bandl.2014.01.006.

Slioussar, N., Malko, A. Gender agreement attraction in Russian: production and comprehension evidence // Frontiers in Psychology. Vol. 7. Art. 1651. P. 1–20. DOI 10.3389/fpsyg.2016.01651.

Solomon, E. S., Pearlmutter, N. J. Semantic integration and syntactic planning in language production // Cognitive Psychology. Vol. 49. 2004. P. 1–46.

Spalek, K., Franck, J., Schriefers, H., Frauenfelder, U.H. Phonological regularities and grammatical gender retrieval in spoken word recognition and word production // Journal of Psycholinguistic Research. Vol. 37. 2008. P. 419–442.

Stamatakis, E. A., Marslen-Wilson, W. D., Tyler, L. K., Fletcher, P. C. Cingulate control of fronto-temporal integration reflects linguistic demands: a three-way interaction in functional connectivity // Neuroimage. Vol. 28. 2005. P. 115–121.

Staub, A. Response time distributional evidence for distinct varieties of number attraction // Cognition. Vol. 114. 2010. P. 447–454.

Stojanović, D. Parsing and acquisition: Evidence from Serbo-Croatian. Doctoral dissertation. University of Ottawa. 1999.

Tanner, D., Nicol, J., Brehm, L. The time-course of feature interference in agreement comprehension: Multiple mechanisms and asymmetrical attraction // Journal of Memory and Language. Vol. 76. 2014. P. 195–215.

Taraban, R., Kempe, V. Gender processing in native and nonnative Russian speakers // Applied Psycholinguistics. Vol. 20. 1999. P. 119–148.

Tyler, L. K., Stamatakis, E. A., Post, B., Randall, B., Marslen-Wilson, W. D. Temporal and frontal systems in speech comprehension: An fMRI study of past tense processing // Neuropsychologia. Vol. 43. 2005. P. 1963–1974.

Ullman, M. T. Contributions of memory circuits to language: the declarative/procedural model // Cognition. Vol. 92. 2004. P. 231–270.

Ullman, M. T., Bergida, R., O'Craven, K. Distinct fMRI activation patterns for regular and irregular past tense // NeuroImage. Vol. 5. 1997. S549.

Vasishth, S. Working memory in sentence comprehension: processing Hindi center embeddings. Doctoral dissertation. Ohio State University, Columbus, OH. 2002.

Vigliocco G., Butterworth, B., Garrett, M. F. Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints // Cognition. Vol. 61. 1996. P. 261–298.

Vigliocco, G., Butterworth, B., Semenza, C. Constructing subject-verb agreement in speech: The role of semantic and morphological factors // Journal of Memory and Language. Vol. 34. 1995. P. 186–215.

Vigliocco, G., Franck, J. When sex and syntax go hand in hand: Gender agreement in language production // Journal of Memory and Language. Vol. 40. 1999. P. 455–478.

Vigliocco, G., Zilli, T. Syntactic accuracy in sentence production: the case of gender disagreement in Italian language-impaired and unimpaired speakers // Journal of Psycholinguistic Research. Vol. 28. 1999. P. 623–648.

Wagers, M. W., Lau, E. F., Phillips, C. Agreement attraction in comprehension: Representations and processes // Journal of Memory and Language. Vol. 61. 2009. P. 206–223.