

УДК 8:372.8
ISSN 2500-2953

Rhema. Рема

3.2024

Издается с 2002 г.

**Учредитель
и издатель:**
Московский
педагогический
государственный
университет

Выходит 4 раза в год

Журнал входит в Перечень ведущих рецензируемых научных журналов и изданий ВАК РФ:

Педагогика

5.8.2. Теория и методика обучения и воспитания

Филология

5.9.1. Русская литература и литературы народов Российской Федерации

5.9.3. Теория литературы

5.9.5. Русский язык. Языки народов России

5.9.6. Языки народов зарубежных стран

5.9.8. Теоретическая, прикладная и сравнительно-сопоставительная лингвистика

Когнитивные науки

(филологические науки)

5.12.3. Междисциплинарные исследования языка

Свидетельство
о регистрации СМИ:
ПИ № ФС 77–67769
от 17.11.2016

Адрес редакции:
109240, Москва,
ул. В. Радищевская,
д. 16–18, комн. 223

Сайт: rhema-journal.com

E-mail: rhema.pema@gmail.com

Подписной индекс журнала по Объединенному каталогу
«Пресса России» – **85006**

ISSN 2500-2953

Rhema. Pema

3.2024

**The Founder
and Publisher:**

Moscow Pedagogical
State University

Mass media
registration
certificate

ПИ № ФС 77-67769
as of 17.11.2016

Editorial office:

Moscow, Russia, Verh-
nyaya
Radishchevskaya str.,
16–18, room 223,
109240

The journal is included in the list of the leading peer-reviewed scholarly journals the Higher Attestation Commission of the Ministry of Science and Higher Education of the Russian Federation recommended to PhD candidates and those working for their habilitation who wish to publish the results of their research.

The journal has been published since 2002

The journal is published 4 times a year

E-mail: rhema.pema@gmail.com

Information on journal can be accessed via: rhema-journal.com

Редакционная коллегия

Главный редактор

Антон Владимирович Циммерлинг – доктор филологических наук; профессор кафедры общего языкознания и русского языка Государственного института русского языка имени А.С. Пушкина; ведущий научный сотрудник сектора типологии Института языкознания РАН, г. Москва.

Заместитель главного редактора

Екатерина Анатольевна Лютикова – доктор филологических наук, доцент; профессор кафедры теоретической и прикладной лингвистики филологического факультета Московского государственного университета имени М.В. Ломоносова.

Ответственный секретарь

Анастасия Алексеевна Герасимова – кандидат филологических наук; научный сотрудник лаборатории автоматизированных лексикографических систем Научно-исследовательского вычислительного центра Московского государственного университета имени М.В. Ломоносова.

Татьяна Михайловна Воителева – доктор педагогических наук, профессор; профессор кафедры методики преподавания русского языка и литературы Московского государственного областного университета.

Наталья Дмитриевна Гальскова – доктор педагогических наук, профессор; профессор кафедры лингводидактики Московского государственного областного педагогического университета.

Елена Валентиновна Гетманская – доктор педагогических наук; профессор кафедры методики преподавания литературы Института филологии Московского педагогического государственного университета.

Павел Валерьевич Гращенков – доктор филологических наук; доцент кафедры теоретической и прикладной лингвистики филологического факультета Московского государственного университета имени М.В. Ломоносова; старший научный сотрудник отдела языков народов Азии и Африки Института востоковедения РАН, г. Москва.

Атле Грённ – профессор кафедры литературы, страноведения и европейских языков Университета Осло, Норвегия.

Сурен Тигранович Золян – доктор филологических наук, профессор; ведущий научный сотрудник отдела теоретической философии Института философии, социологии и права Национальной Академии наук Армении, г. Ереван, Армения; профессор Института гуманитарных наук Балтийского федерального университета имени Иммануила Канта, г. Калининград.

Алексей Александрович Корнев – кандидат педагогических наук; доцент кафедры теории преподавания иностранных языков факультета иностранных языков и регионоведения Московского государственного университета имени М.В. Ломоносова.

Анатолий Симонович Либерман – доктор филологических наук; профессор кафедры немецкого, нидерландского и скандинавских языков Университета Миннесоты, г. Миннеаполис, США.

Сильвия Лураги – PhD (филология); профессор факультета гуманитарных наук, Университет Павии, Италия.

Михаил Николаевич Михайлов – PhD (филология); профессор переводоведения (русский и финский языки) Института языков, перевода и литературы Университета Тампере, Финляндия.

Нерея Мадарьяга Писано – PhD (филология); доцент кафедры классических языков (секция славянской филологии) Университета Страны Басков, г. Витория, Испания.

Владимир Александрович Плунгян – доктор филологических наук, профессор, академик РАН; заместитель директора Института русского языка им. В.В. Виноградова РАН; заведующий сектором типологии Института языкознания РАН; профессор кафедры теоретической и прикладной лингвистики Московского государственного университета имени М.В. Ломоносова.

Велка Александрова Попова – кандидат филологических наук; доцент кафедры болгарского языка, сотрудник лаборатории прикладной лингвистики факультета гуманитарных наук Шуменского университета имени Епископа Константина Преславского, Болгария.

Наталья Вадимовна Сердобольская – кандидат филологических наук; доцент Учебно-научного центра лингвистической типологии Института лингвистики Российского государственного гуманитарного университета.

Андрей Стоянович – доктор филологических наук, профессор; заведующий кафедрой иностранных языков Белградского университета, Сербия.

Младен Ухлик – доктор филологических наук, профессор; заведующий кафедрой русского языка Отделения славянских языков философского факультета Университета Любляны, Словения.

Александр Иосифович Федута – доктор филологических наук; редактор биографического альманаха «Асоба і час» («Личность и время»), г. Минск, Республика Беларусь.

Любовь Георгиевна Чапаева – доктор филологических наук; профессор кафедры общего и прикладного языкознания Московского педагогического государственного университета.

Editorial Board

Editor-in-Chief

Anton V. Zimmerling – Dr. Phil. Hab.; principal research fellow, Institute of Linguistics, Russian Academy of Sciences; professor at the Department of General Linguistics and Russian Language, Pushkin State Russian Language Institute, Moscow, Russian Federation.

Deputy chief editor

Ekaterina A. Lyutikova – Dr. Phil. Hab.; professor at the Department of Theoretical and Applied Linguistics, Lomonosov Moscow State University, Russian Federation.

Executive secretary

Anastasia A. Gerasimova – PhD in Linguistics; researcher at the Research Computing Center, Lomonosov Moscow State University, Moscow, Russian Federation.

Lyubov G. Chapaeva – Dr. Phil. Hab.; professor at the Department of General and Applied Linguistics, Institute of Philology and Foreign languages, Moscow Pedagogical State University, Russian Federation.

Aleksandr Feduta – Dr. Phil. Hab.; editor-in-chief of the biographical almanac «Personality and Time», Minsk, Belarus.

Natalia D. Galskova – Dr. Ped. Hab.; professor at the Department of Linguodidactics, Moscow Region State Pedagogical University, Russian Federation.

Elena V. Getmanskaya – Dr. Ped. Hab.; professor at the Department of Methods of Teaching Literature, Moscow Pedagogical State University, Russian Federation.

Pavel V. Grashchenkov – Dr. Phil. Hab.; associate professor at the Department of Theoretical and Applied Linguistics, Faculty of Philology, Lomonosov Moscow State University; research fellow at the Institute of Oriental Studies, Russian Academy of Sciences, Moscow, Russian Federation.

Atle Grønn – Dr. Phil. Hab.; professor at ILOS – Department of Literature, Area Studies and European Languages, University of Oslo, Norway.

Alexey A. Korenev – PhD in Pedagogy; associate professor at the Department of Foreign Language Teaching Theory, Faculty of Foreign Languages and Area Studies, Lomonosov Moscow State University, Russian Federation.

Anatoly Liberman – Dr. Phil. Hab.; professor at the Department of German, Dutch and Scandinavian, University of Minnesota, USA.

Silvia Luraghi – PhD in Philology; associate professor at the Department of Humanities, Section of General and Applied Linguistics, University of Pavia, Italy.

Mikhail Mikhailov – PhD in Philology; professor at the School of Language, Translation and Literary Studies, University of Tampere, Finland.

Nerea Madariaga Pisano – PhD in Philology; associate professor at the Department of Classical Studies (Section of Slavic Philology), University of the Basque Country, Vitoria, Spain.

Vladimir A. Plungian – Dr. Phil. Hab., full member of the Russian Academy of Sciences; deputy director of the Vinogradov Institute of Russian Language, Russian Academy of Sciences; head of the Sector of Typology, Institute of linguistics, Russian Academy of Sciences; professor at the Department of Theoretical and Applied Linguistics, Lomonosov Moscow State University, Russian Federation.

Velka A. Popova – PhD in Philology; associate professor at the Department of Bulgarian, research fellow at the Laboratory of Applied Linguistics, Faculty of Humanities, the Constantin of Preslav Bishop Shumen University, Bulgaria.

Natalia V. Serdobolskaya – PhD in Philology; associate professor at the Training and Research Center for Linguistic Typology, Institute of Linguistics, Russian State University for the Humanities, Moscow, Russian Federation.

Andrej Stojanović – Dr. Phil. Hab., full professor; head of the Department of Foreign Languages, University of Belgrade, Serbia.

Mladen Uhlik – Dr. Phil. Hab., professor; head of the Department of Russian, Institute of Slavic Languages, Faculty of Philosophy, University of Ljubljana, Slovenia.

Tatiana M. Voiteleva – Dr. Ped. Hab.; professor at the Department of the Methods of Teaching Russian Language and Literature, Moscow Region State University, Russian Federation.

Suren T. Zolyan – Dr. Phil. Hab., professor; leading research fellow at the Department of Theoretical Philosophy, Institute of Philosophy, Sociology and Law Studies, National Academy of Sciences, Yerevan, Armenia; professor at the Institute for the Humanities, Immanuel Kant Baltic Federal University, Kaliningrad, Russian Federation.

Содержание

ЛИНГВИСТИКА

А.А. Герасимова, Е.А. Лютикова

От теоретических проблем
к экспериментальным исследованиям и обратно 9

Э.Е. Буров

О связи между корпусными данными и интуицией
носителей языка (на материале сочетаний типа
«что-то/кто-то + полное прилагательное») 20

Т.И. Давидюк

Лично-числовое согласование
с сочинительными конструкциями с союзом *ни ... ни*:
корпусное и экспериментальное исследование 33

А.В. Прокопьева

Согласование предиката по числу
с сочинительными конструкциями
с союзами *как ... так и* и *не только ... но и*:
экспериментальное исследование 54

А.Г. Анисимова

Предикативное согласование по роду
с сочинительными конструкциями
с двойным заместительным союзом («X, а не Y») 74

А.Ю. Хоменко, Д.П. Бальба, Д.А. Исаков

Кластеризация лексики для решения задачи
диагностики афазии 87

П.А. Пилипец, В.Н. Староверова

Понимание сложных синтаксических конструкций
школьниками младших классов 112

Contents

LINGUISTICS

A. Gerasimova, E. Lyutikova

From theoretical problems
to experimental data and back 9

E. Burov

On the relationship between corpus data
and native speakers' intuition
(on the material of combinations
of the type “*čto-to* ‘something’ /
kto-to ‘somebody’ + full adjective”) 20

T. Davidyuk

Person-number agreement with constructions conjoined
by *ni ... ni* ‘neither ... nor’:
A corpus and experimental study 33

A. Prokopeva

Predicative number agreement
with coordinated subjects containing *kak ... tak i* ‘both ... and’
and *ne tol’ko ... no i* ‘not only ... but also’:
An experimental study 54

A. Anisimova

Predicate gender agreement with the coordinate constructions
with a double substitution conjunction
(“*X, a ne Y*”) 74

A. Khomenko, D. Balba, D. Isakov

Vocabulary clustering for solving problems
of aphasia diagnostics 87

P. Pilipets, V. Staroverova

Understanding complex syntactic constructions
by primary school students 112

DOI: 10.31862/2500-2953-2024-3-9-19

УДК 81-114.4

А.А. Герасимова, Е.А. Лютикова

Московский государственный университет имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация

От теоретических проблем к экспериментальным исследованиям и обратно

Статья предваряет два специальных выпуска журнала, составленные по материалам Всероссийской конференции для молодых исследователей «Экспериментальные исследования языка 2024». В статье мы рассматриваем соотношение эмпирических данных и теории в языкознании и показываем циклическую связь между экспериментальными данными и их теоретической интерпретацией. Далее мы обсуждаем необходимые составляющие развития культуры проведения лингвистических экспериментов. Статью завершает краткий обзор статей специальных выпусков.

Ключевые слова: эксперимент, экспериментальный синтаксис, психолингвистика, теория языка

Благодарности. Исследование выполнено в рамках проекта РНФ № 22-18-00037 «Параметрическая модель согласования в свете экспериментальных данных», реализуемого в Московском государственном университете имени М.В. Ломоносова, <https://rscf.ru/project/22-18-00037/>.

Для ЦИТИРОВАНИЯ: Герасимова А.А., Лютикова Е.А. От теоретических проблем к экспериментальным исследованиям и обратно // Рема. Rhema. 2024. № 3. С. 9–19. DOI: 10.31862/2500-2953-2024-3-9-19

© Герасимова А.А., Лютикова Е.А., 2024



Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.31862/2500-2953-2024-3-9-19

A. Gerasimova, E. Lyutikova

Lomonosov Moscow State University,
Moscow, 119991, Russian Federation

From theoretical problems to experimental data and back

This paper presents the two special issues of the journal, compiling the experimental studies presented at the All-Russian Conference for Young Researchers “Experimental Studies of Language 2024”. In the paper, we examine the relationship between empirical data and theory in linguistics and demonstrate the cyclical connection between experimental data and their theoretical interpretation. Furthermore, we discuss the components necessary for the development of the culture of conducting linguistic experiments. The paper concludes with a brief overview of the special issues’ content.

Key words: experiment, experimental syntax, psycholinguistics

Acknowledgments. This research is supported by Russian Science Foundation, RSF project 22-18-00037 realized at Lomonosov Moscow State University, <https://rscf.ru/en/project/22-18-00037/>

FOR CITATION: Gerasimova A., Lyutikova E. From theoretical problems to experimental data and back. *Rhema*. 2024. No. 3. Pp. 9–19. (In Rus.). (In Rus.). DOI: 10.31862/2500-2953-2024-3-9-19

1. Введение

Экспериментальный метод занял прочные позиции в современном языкознании. Возможность контролировать условия, в которых измеряется реакция носителей на языковые выражения, позволяет получать более объективные данные о языковой действительности и проверять сложные гипотезы об устройстве грамматики и процессов обработки языка. Метод эксперимента обеспечивает объективность в процессе получения данных и, как следствие, делает этот процесс более предсказуемым и воспроизводимым. Этим объясняется широкая востребованность

экспериментальных методов в самых разных областях лингвистики, охватывающих уровни языка от фонетики до дискурса.

Объективность и непредвзятость экспериментального метода – те характеристики, которые, пусть и в неявном виде, отмечают многие исследователи. Действительно, больший контроль за языковым материалом, репрезентативность выборки выгодно отличают эксперимент от интроспекции. Тем не менее, в данном методе, как и во всех методах, применяемых в лингвистике, сохраняется проблема черного ящика – отсутствия непосредственного доступа к механизмам, обеспечивающим языковую способность человека. Тем самым ключевой характеристикой эксперимента становится на самом деле бóльшая прозрачность гипотез, которые связывают эмпирический материал с его теоретической интерпретацией.

Переход к новой методологической парадигме предполагает развитие соответствующей инфраструктуры, с одной стороны, и научной культуры – с другой. В этой заметке мы хотим обсудить роль интерпретации эмпирических данных в развитии экспериментального направления лингвистики и в этом свете представить специальные выпуски по материалам конференции для молодых исследователей «Экспериментальные исследования языка 2024», которая задумывалась как площадка для реализации указанных двух составляющих развития экспериментального метода. В первую очередь мы обратимся к вопросу о соотношении эмпирических данных и теории в языкознании, а затем представим наше видение того, как развивается культура проведения лингвистических экспериментов. В завершении будет дан обзор содержания двенадцати статей, посвященных экспериментальным исследованиям самых разных феноменов русского языка.

2. Насколько объективным может быть экспериментальный метод?

Привлекательность экспериментального метода состоит в том, что эксперимент, как всякий количественный подход, позволяет обобщить выявляемые закономерности в некотором стандартизированном формате. Достоинства экспериментального метода могут не только создать впечатление единственно верного инструмента для получения данных, но и поставить под сомнение обобщения, полученные другими способами. Действительно, если в случае психолингвистических исследований онлайн-методы оказываются основным инструментом проверки гипотез, в тех ситуациях, когда используются экспериментальные офлайн-методы, часто есть более доступные альтернативы, например,

наблюдение по лингвистическому корпусу или полуструктурированная эlicitация в полевых условиях. Должен ли современный исследователь полностью отказаться от этого инструментария и предпочесть всему эксперимент?

Мы считаем, что эксперимент может быть вызван к жизни исключительно необходимостью разрешения теоретической проблемы и невозможностью разрешить эту проблему каким-либо иным способом. Методика эксперимента не нужна для того, чтобы подтверждать наблюдения, которые не вызывают никаких сомнений у исследователя. Несмотря на то, что порог входа в экспериментальную лингвистику остается достаточно невысоким, проведение эксперимента на качественно приемлемом уровне оказывается затратным с точки зрения как временных, так и материальных ресурсов. Самой крупной «статьей расходов» в организации экспериментального исследования остается набор респондентов: приглашение к участию репрезентативной выборки носителей языка либо занимает много времени, либо, как в случае краудсорсинга, требует денежных затрат, при этом зачастую без гарантии качественного и вовлеченного участия. Во многом поэтому бездумно собирать экспериментальные данные обо всех языковых феноменах не имеет смысла.

Когда же метод эксперимента оказывается необходим? Во-первых, тогда, когда эксперимент позволяет заполнить лакуны в эмпирических данных, притом, что любые другие способы получения данных недоступны. Во-вторых, в тех случаях, когда эксперимент позволяет решать значительно более глубокие исследовательские вопросы, которые подразумевают корреляцию или причинно-следственную связь между отдельными языковыми параметрами. Примечательно то, что обе эти ситуации возникают тогда, когда исследователь уже пытается анализировать данные в рамках определенной теории: в результате он приходит к пониманию недостатка данных и/или формулирует гипотезы о взаимосвязях в данных.

Таким образом возникает цикличность в переходе от теории языка к экспериментальным данным. Постановка экспериментальной задачи невозможна без формулирования предварительных гипотез, вытекающих по большей части не просто из эмпирического материала, собранного на предыдущих этапах исследования, но из теоретического осмысления этого самого материала. Однако и дальнейшие шаги, связанные с интерпретацией вновь полученных данных, не могут обойтись без опоры на теоретические представления о предполагаемых языковых закономерностях. Так, например, замедление скорости реакции или оценки приемлемости в нижней части шкалы Ликерта интерпретируется

как наличие трудностей в когнитивной обработке языкового материала, но то, как исследователь будет интерпретировать сами «трудности», зависит уже от того, в какой научной парадигме он работает и на какие представления об устройстве грамматики он опирается.

3. Как сделать экспериментальный метод более объективным?

Несмотря на высокую зависимость значимости экспериментальных данных от их интерпретации в рамках конкретной языковой модели (см. [Almeida et al., 2023]), мы считаем, что возможно добиться большей прозрачности метода и приблизиться к нивелированию проблемы черного ящика. В качестве возможного пути мы видим распространение экспериментального метода как исследовательской процедуры на решение все большего количества научных задач.

Именно массовость в применении экспериментального метода оказывается залогом успешной интерпретации экспериментальных результатов. Во-первых, множественное применение технологии эксперимента позволяет убедиться в том, что получаемые результаты неслучайны и воспроизводимы. Если экспериментальная проверка одного и того же фактора повторяется многократно и при этом результаты остаются стабильными, мы получаем подтверждение, что перед нами – языковая закономерность. Весьма продуктивным в этом отношении становится обращение к материалу разных языков, когда в рамках одной экспериментальной парадигмы получается объединить языковые параметры различного порядка. Так, например, не прекращаются исследования раннего и позднего закрытия ([Cuetos, Mitchell, 1988], см. недавние работы для русского языка [Алексеев, Федорова, 2023] и для китайского [Liu, Lin, 2024]) или эффектов аттракции ([Bock, Miller, 1991; Bock, Eberhard, 1993], см. недавние работы для турецкого [Türk, Logačev, 2024] и для чешского [Lacina, Dotlačil, 2024]): привлечение все новых языков приводит ко все большей детализации исследовательских гипотез и, как следствие, к существенному продвижению научного знания.

Во-вторых, по мере накопления экспериментальных результатов все чаще исследователи обращаются к вопросам, связанным с совершенствованием самого метода. В этом отношении работа может вестись в направлении уточнения аспектов экспериментальной процедуры (ср. [Marty et al., 2020; Arehalli, Wittenberg, 2021; Schoenmakers, 2023]), деталей статистической обработки (ср. переход от фреквентистской статистической парадигмы к байесовской [Vasishth, Gelman, 2021; Vasishth, 2023]). Эффективность применения метода определяется в том числе

нашим понимаем того, какие ограничения он в себе несет. Соответственно, из набора возможных интерпретаций получается исключить те, которые связаны с особенностями процедуры проведения эксперимента.

Наконец, массовый подход позволяет собрать обширные данные, что способствует более глубокому анализу эмпирического материала и более точному выявлению закономерностей. По мере накопления материала появляется возможность увидеть за действием конкретного фактора в определенном контексте более общую картину, а именно, понять, как в принципе следует интерпретировать данные, полученные для нового феномена. Например, установлено, что наличие отрицания обычно снижает оценки приемлемости для предложения, что необходимо учитывать при сравнении любых условий, которые включают противопоставление по наличию/отсутствию отрицания.

4. К распространению метода эксперимента в российской лингвистике

Специализированная конференция «Экспериментальные исследования языка», равно как и представляемые вниманию читателей специальные выпуски журнала «Рема. Rhema», ставят своей целью укрепление культуры проведения экспериментальных исследований в отечественном языкознании и воспитание нового поколения исследователей, для которых массовость экспериментального метода становится реальностью научной работы. В качестве идеальной мы видим картину, когда каждое экспериментальное исследование важно и значимо не только само по себе, в рамках конкретного исследовательского вопроса, который решает экспериментатор, но вносит вклад в общее понимание того, как экспериментальные данные могут быть использованы для моделирования языковой способности и языковой деятельности человека.

Необходимым условием для реализации этой программы становится наработка культуры интерпретации экспериментальных результатов. Постановка экспериментальной задачи, планирование эксперимента и разработка его дизайна невозможны без оглядки на всю совокупность ранее полученных данных. Мы выражаем надежду, что работы, представленные в двух специальных выпусках журнала, будут способствовать укреплению позиций отечественной экспериментальной лингвистики и повышению общего научного уровня интерпретации экспериментальных результатов.

* * *

Первый специальный выпуск открывает работа Э.Е. Бурава, посвященная особенностям функционирования местоимений *кто-то* и *что-то*,

в частности, их сочетаемости с постпозитивными прилагательными. Примечательно предпринятое в рамках статьи сопоставление результатов исследований на базе корпуса и эксперимента. Хотя экспериментальные переменные задаются контекстом употребления и, соответственно, оценки во многом связаны с лексическим наполнением конкретных стимулов, фиксируется прямая связь между частотностью сочетаний и оценками приемлемости. Исследование вносит вклад в уточнение наших представлений о связи между эмпирическими данными разного происхождения и демонстрирует пример вариативного феномена, для которого узус прямо коррелирует с языковой интуицией.

Вопрос связи корпусных и экспериментальных данных также затрагивает статья Т.И. Давидюк, в которой анализируется предикативное согласования в конструкциях с сочиненным подлежащим с союзом *ни ... ни* в сравнении с союзом *и ... и*. Хотя корпусные данные показали асимметрию в распределении частот согласовательных стратегий для сочинения с различными союзами, экспериментальное исследование не выявляет подобных различий. Примечательно, что на распределение оценок для предложений с союзом *и ... и* повлияло наличие отрицания в стимульном материале, что было бы невозможно установить без опоры на предыдущие сходные экспериментальные исследования.

Исследование роли различных типов союзов при вариативном предикативном согласовании в русском языке продолжает работа А.В. Прокöpfeвой. В частности, проверяется гипотеза, ранее сформулированная на материале немецкого языка, о том, что при союзах вида *как ... так и* и *не только ... но и* предпочтение должно всегда отдаваться согласованию по множественному числу. С привлечением русского языкового материала автор показывает, что вариативность в согласовании при подлежащих данного типа определяется фактором порядка слов и семантическими свойствами собственно союзов. Тем самым, экспериментальный метод позволяет значительно уточнить картину языкового явления по сравнению с методами интроспекции.

В статье А.Г. Анисимовой ставится вопрос о том, как могут взаимодействовать факторы, для которых в ситуации вариативности ожидается противоположное действие. В частности, рассматриваются согласования с сочиненными подлежащими с двойным заместительным союзом *а не* с позиции двух тенденций к согласованию: формального с конъюнктом, ближайшим к предикату, и семантического с конъюнктом, который не находится под отрицанием. В дополнение в эксперимент включаются факторы порядка слов и наличия/отсутствия материала между подлежащим и сказуемым. Результаты показывают однозначный выбор семантической стратегии согласования, что существенно дополняет

представления о том, какие ограничения на варьирование может накладывать семантический тип союза.

Важным направлением в экспериментальных исследованиях становится определение эталонных ответов в некоторых стандартизированных условиях. В статье А.Ю. Хоменко, Д.А. Исакова и Д.П. Бальба с использованием теста на семантическую вербальную беглость и алгоритмов автоматической кластеризации разрабатывается функциональная модель для определения афазии у носителей русского языка. Суть теста состоит в назывании максимального количества слов, соответствующих некоторой заданной категории, за определенный промежуток времени. Автоматическая кластеризация лексики в ответах носителей на вопросы теста позволила оценить ассоциативные связи внутри лексических подкатегорий и установить различия для контрольной группы и группы респондентов с афазией.

Вопросы стандартизации тестов синтаксических компетенций также освещены в работе П.А. Пилипец и В.Н. Старовой. Экспериментальное исследование, участниками которого стали школьники 1–4 классов, позволило определить особенности понимания сложных синтаксических конструкций младшими когортами респондентов. Авторы показывают, что с течением времени у детей младшего школьного возраста происходит переход к взрослой стратегии восприятия конструкций с ранним и поздним закрытием и сравнительных конструкций. Исследование оказывается первой работой, в которой обработка сложных конструкций разных типов представлена в сопоставительном аспекте, что обеспечивает основу для разработки стандартизированного теста для оценки синтаксических компетенций у детей.

Первой во втором специальном выпуске журнала становится статья А.И. Добрыниной, которая посвящена связи жестов рук и интерпретации неоднозначных предложений с квантором *весь/все* и отрицанием. Сочетание жестикуляции с кванторными словами исследуется экспериментально. Автор проверяет, какую интерпретацию для слушающего, с узкой или широкой сферой действия квантора, вызывает жестикуляция говорящего. Работа А.И. Добрыниной – образец исследования, в котором экспериментальный метод дает совершенно новую перспективу уже знакомой и хорошо исследованной проблематике. Получен важный результат: влияние жестикуляции на восприятие интерпретации проявляется только для тех интерпретаций, которые менее ожидаемы для слушающего.

Следующие две статьи посвящены морфологическим сюжетам: к сожалению, до сих пор это нечастое направление исследований в экспериментальной лингвистике, но весьма многообещающее. В статье

Г.Н. Бубнова тестируется гипотеза о связи выбора чистовидовой приставки с семантикой глагола. Гипотеза рассматривается на материале глаголов-неологизмов, связанных с тематикой компьютерных игр. Работа представляет собой пример экспериментального исследования, которое проверяет теоретические заключения, изначально сформулированные на материале стандартного узуса, обращаясь к материалу заимствований.

В статье В.А. Борисовой используется гипотеза о связи времени обработки языкового выражения с его сложностью в вопросе моделирования русских глаголов с лексическими и супралексическими префиксами. Результаты не подтверждают предположения о том, что глаголы с супралексическими префиксами вызывают более длительное время реакции ввиду своей сложности. Эксперименты для сходных феноменов позволяют выдвинуть гипотезу о природе асимметрии между классами глагольных префиксов, предположив дополнительную некомпозициональность в случае лексических префиксов и, как следствие, сходные затраты на ее обработку.

Влияние семантических факторов традиционно изучается и для вариативных синтаксических конструкций. И.О. Гриценко обращается к семантическим факторам вариативного согласования. Рассматриваются предложения с сочиненным подлежащим, в котором конъюнкты синтаксически и семантически разнородны. Помимо того, что работа достоверным образом подтверждает гипотезы, сформулированные с помощью методов интроспекции, она обнаруживает новый результат: существенным параметром вариативности становится лицо местоименного конъюнкта. Этот аспект исследования существенно дополняет общую картину устройства признаковых оппозиций в русском языке.

В работе А.В. Крайновой рассматриваются конструкции с количественными существительными в русском языке. Предыдущие исследования данных конструкций, как корпусные, так и экспериментальные, были посвящены проблеме предикативного согласования. Согласно результатам, на предпочтительность той или иной стратегии влияет степень абстрактности семантики количественного существительного: чем более неопределенное количество существительное выражает, тем выше вероятность употребления единственного числа сказуемого, и наоборот, чем оно более конкретно, тем выше вероятность выбора множественного числа. В исследовании автор рассматривает семантическое атрибутивное согласование в конструкциях с количественными существительными и проверяет, действуют ли в данном случае аналогичные закономерности.

Наконец, завершает номер статья, представляющая собой пример сложного многофакторного исследования. В работе Е.П. Патраковой экспериментальными методами проверяется диагностическая модель поиска недостоверной информации в речи. В исследовании проводится одновременный анализ вербальных и невербальных факторов в различных модальностях вместе с учетом экстралингвистических характеристик носителей языка, в том числе возраста. В результате обнаруживается, что больший вес имеют невербальные признаки недостоверной информации.

* * *

Как видим, статьи специальных выпусков охватили самые разные направления современной экспериментальной лингвистики и привлекли в качестве материала языковые единицы всего спектра уровней языка. Верим, что заинтересованный читатель найдет в этом многообразии те идеи и приемы, которые не только будут способствовать углубленному пониманию отдельных лингвистических сюжетов, но и помогут продвинуться в собственных экспериментальных исследованиях.

Библиографический список / References

Алексеев, Федорова, 2023 – Алексеев Д.А., Федорова О.В. Откуда на Руси есть пошел запрет на «раннее закрытие» // Вестник Московского университета. Серия 9: Филология. 2023. № 5. С. 36–45. [Alekseev D.A., Fedorova O.V. The origins of the early closure ban in the Russian literary standard. *Lomonosov Philology Journal. Series 9: Philology*. 2023. No. 5. Pp. 36–45. (In Rus.)]

Almeida et al., 2023 – Almeida D., Breen M., Carlson K. et al. The future of experimental syntax. *The Oxford Handbook of Experimental Syntax*. J. Sprouse (ed.). Oxford University Press, 2023. Pp. 643–664.

Arehalli, Wittenberg, 2021 – Arehalli S., Wittenberg E. Experimental filler design influences error correction rates in a word restoration paradigm. *Linguistics Vanguard*. 2021. Vol. 7. No. 1. Art. 20200052.

Bock, Miller, 1991 – Bock K., Miller C. Broken agreement. *Cognitive Psychology*. 1991. No. 23. Pp. 45–93.

Bock, Eberhard, 1993 – Bock K., Eberhard K.M. Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*. 1993. Vol. 8. No. 1. Pp. 57–99.

Cuetos, Mitchell, 1988 – Cuetos F., Mitchell D.C. Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in Spanish. *Cognition*. 1988. No. 30. Pp. 73–105.

Lacina, Dotlačil, 2024 – Lacina R., Dotlačil J. Grammaticality illusions in Czech: A speeded acceptability study of agreement attraction. *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 46. L.K. Samuelson, S.L. Frank, M. Toneva et al. (eds.). URL: <https://escholarship.org/uc/item/86945505> (date accessed: 01.08.2024).

Liu, Lin, 2024 – Liu F., Lin C.-J.C. Relative clause attachment in Mandarin Chinese: Insights from classifier-noun agreement. *Frontiers in Language Sciences*. 2024. Vol. 3. Art. 1438323.

Marty et al., 2020 – Marty P., Chemla E., Sprouse J. The effect of three basic task features on the sensitivity of acceptability judgment tasks. *Glossa: A Journal of General Linguistics*. 2020. No. 5. Pp. 1–23.

Schoenmakers, 2023 – Schoenmakers G.J. Linguistic judgments in 3D: The aesthetic quality, linguistic acceptability, and surface probability of stigmatized and non-stigmatized variation. *Linguistics*. 2023. Vol. 61. No. 3. Pp. 779–824.

Türk, Logačev, 2024 – Türk U., Logačev P. Agreement attraction in Turkish: The case of genitive attractors. *Language, Cognition and Neuroscience*. 2024. Vol. 39. No. 4. Pp. 448–454. DOI: 10.1080/23273798.2024.2324766

Vasishth, 2023 – Vasishth S. New directions in statistical analysis for experimental linguistics. *The Routledge Handbook of Experimental Linguistics*. S. Zufferey, P. Gygas (eds.). Routledge, 2023. Pp. 313–329.

Vasishth, Gelman, 2021 – Vasishth S., Gelman A. How to embrace variation and accept uncertainty in linguistic and psycholinguistic data analysis. *Linguistics*. Vol. 59. No. 5. Pp. 1311–1342. DOI: 10.1515/ling-2019-0051

Статья поступила в редакцию 30.08.2024

The article was received on .30.08.2024

Сведения об авторах / About the authors

Герасимова Анастасия Алексеевна – кандидат филологических наук; научный сотрудник Научно-исследовательского вычислительного центра, Московский государственный университет имени М.В. Ломоносова

Anastasia A. Gerasimova – PhD in Linguistics; researcher at the Research Computing Center, Lomonosov Moscow State University, Moscow, Russian Federation

ORCID: <https://orcid.org/0000-0003-4686-5598>

E-mail: anastasiagerasimova432@gmail.com

Лютикова Екатерина Анатольевна – доктор филологических наук; профессор кафедры теоретической и прикладной лингвистики филологического факультета, ведущий научный сотрудник Научно-исследовательского вычислительного центра, Московский государственный университет имени М.В. Ломоносова

Ekaterina A. Lyutikova – Dr. Phil. Hab.; Professor at the Department of Theoretical and Applied Linguistics of the Philological Faculty, leading researcher at the Research Computing Center, Lomonosov Moscow State University, Moscow, Russian Federation

ORCID: <https://orcid.org/0000-0003-4439-0613>

E-mail: lyutikova2008@gmail.com

Все авторы прочитали и одобрили окончательный вариант рукописи

All authors have read and approved the final manuscript

DOI: 10.31862/2500-2953-2024-3-20-32

УДК 81-114.4

Э.Е. Буров

Иркутский государственный университет,
664003 г. Иркутск, Российская Федерация

О связи между корпусными данными и интуицией носителей языка (на материале сочетаний типа «что-то/кто-то + полное прилагательное»)

Статья посвящена экспериментальному исследованию связи между корпусными данными и интуицией носителей языка. Указанная проблема рассматривается на материале сочетаний типа «что-то/кто-то + полное прилагательное». Гипотеза исследования состоит в том, что на данном участке между корпусными данными и интуицией носителей языка существует прямая связь, и предложения, содержащие высокоустойчивые со статистической точки зрения сочетания, значимо более приемлемы, чем предложения с сочетаниями, не обладающими статистической устойчивостью. Это утверждение проверяется в ходе двух экспериментов на оценку приемлемости. Используется шкала Ликерта от 1 до 7. Первый эксперимент нацелен на исследование сочетаний типа «что-то + полное прилагательное», второй – на исследование сочетаний типа «кто-то + полное прилагательное». Делается вывод о том, что данные обоих экспериментов полностью подтверждают выдвинутую гипотезу. Ставится вопрос о возможных причинах наличия прямой связи между корпусными и экспериментальными данными.

Ключевые слова: эксперимент на оценку приемлемости, мера logDice, языковая интуиция, корпусные данные, неопределенные местоимения

ДЛЯ ЦИТИРОВАНИЯ: Буров Э.Е. О связи между корпусными данными и интуицией носителей языка (на материале сочетаний типа «что-то/кто-то + полное прилагательное») // Рема. Rhema. 2024. № 3. С. 20–32. DOI: 10.31862/2500-2953-2024-3-20-32

© Буров Э.Е., 2024

Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License



E. Burov

Irkutsk State University,
Irkutsk, 664003, Russian Federation

On the relationship between corpus data and native speakers' intuition (on the material of combinations of the type “*čto-to* ‘something’ / *kto-to* ‘somebody’ + full adjective”)

This paper presents an experimental study of the relationship between the corpus data and the intuition of native speakers. This issue is considered using the material of Russian combinations of the type “*čto-to* ‘something’ / *kto-to* ‘somebody’ + full adjective”. The hypothesis of the study is that for this phenomenon there is a direct connection between corpus data and the intuition of native speakers, and sentences containing combinations that are statistically highly stable are significantly more acceptable than sentences with combinations that do not have statistical stability. This statement is verified with the help of two acceptability judgment experiments using 1–7 Likert scale. The first experiment is aimed at studying combinations of the type “*čto-to* ‘something’ + full adjective”, and the second one is aimed at studying combinations of the type “*kto-to* ‘somebody’ + full adjective”. The author concludes that the data from both experiments fully confirm the hypothesis and raise the question about the possible reasons for the existence of a direct connection between the corpus and experimental data.

Key words: acceptability judgment task, logDice measure, patterns of knowledge, patterns of use, indefinite pronouns.

FOR CITATION: Burov E. On the relationship between corpus data and native speakers' intuition (on the material of combinations of the type “*čto-to* ‘something’ / *kto-to* ‘somebody’ + full adjective”). *Rhema*. 2024. No. 3. Pp. 20–32. (In Rus.). DOI: 10.31862/2500-2953-2024-3-20-32

1. Введение

Одним из нетривиальных вопросов современной лингвистики является вопрос о связи между корпусными данными и интуицией носителей языка (см. [Bresnan, 2007; Divjak, 2017; Lau et al., 2017], обзор представлен в [Gerasimova, Lyutikova, 2020]). И.Б. Левонтина отмечает: «Априорно нельзя утверждать, что результаты анализа корпуса (даже если он достаточно большой и претендует на репрезентативность) отражают интуицию носителей языка. Иными словами, языковая интуиция (“patterns of knowledge”) не может быть приравнена к “patterns of use” (как они, в частности, фиксируются в корпусах). <...> Таким образом, может оказаться, что корреляции между корпусными данными и языковой интуицией носителя довольно сложны» [Левонтина, 2022, с. 16].

Представляется значимым провести экспериментальное исследование, направленное на изучение связи между корпусными данными и интуицией носителей языка. Хорошую возможность сделать это дают сочетаемостные свойства неопределенных местоимений-существительных *что-то* и *кто-то*. Как указывает Е.В. Падучева, «*что-то* часто употребляется в контексте прилагательного» [Падучева, 2016]. То же, хотя и в меньшей степени, относится к местоимению *кто-то*. Удобно изучать, с одной стороны, статистическую устойчивость сочетаний типа «*что-то/кто-то* + полное прилагательное» в корпусе, а с другой – приемлемость подобных сочетаний для носителей языка в эксперименте, а затем сопоставлять корпусные и экспериментальные данные.

Корпусный анализ¹ показывает, что одни сочетания типа «*что-то/кто-то* + полное прилагательное» являются высокоустойчивыми со статистической точки зрения, а другие, напротив, неустойчивы. Высокоустойчивые сочетания – это такие сочетания, которые встречаются в текстах намного чаще, чем они встречались бы при случайном соединении их компонентов; неустойчивые, наоборот, встречаются реже, чем при случайном соединении их компонентов. Одним из способов выявления высокоустойчивых и неустойчивых сочетаний является применение метрики \logDice , указывающей на силу синтагматической связи между компонентами словосочетания (см. подробнее [Буров, 2023]), а также корпусные исследования [Захаров, Хохлова, 2014; Хохлова, 2017; Ольховская, Зельникова, 2023], где указанная метрика применяется для анализа сочетаемости русских слов.

Если значение метрики \logDice отрицательное, соответствующее сочетание считается статистически неустойчивым, если положительное –

¹ В настоящем исследовании привлекаются данные основного корпуса Национального корпуса русского языка (НКРЯ).

статистически устойчивым (в той или иной степени). Высокоустойчивым будем считать сочетание с показателем \logDice , превышающим 2.

Например, высокоустойчивы следующие сочетания: *что-то знакомое* ($\logDice = 6,92$), *что-то белое* ($\logDice = 5,40$), *что-то грустное* ($\logDice = 5,17$), *что-то вкусное* ($\logDice = 4,92$), *кто-то чужой* ($\logDice = 5,57$), *кто-то огромный* ($\logDice = 3,63$), *кто-то страшный* ($\logDice = 2,99$), *кто-то умный* ($\logDice = 3,33$). Не обладают статистической устойчивостью следующие сочетания: *что-то популярное* ($\logDice = -0,19$), *что-то бесцветное* ($\logDice = -0,07$), *что-то невероятное* ($\logDice = -0,02$), *что-то сытное* ($\logDice = -0,04$), *кто-то свой* ($\logDice = -0,56$), *кто-то низкий* ($\logDice = -0,43$), *кто-то красивый* ($\logDice = -1,03$), *кто-то глупый* ($\logDice = -0,16$).

Эксперимент позволяет проверить, как соотносятся между собой фиксируемая в корпусе степень статистической устойчивости сочетания типа «*что-то/кто-то* + полное прилагательное» и приемлемость этого сочетания для носителей языка.

Гипотеза состоит в том, что между корпусными данными и интуицией носителей языка существует прямая связь, и предложения с высокоустойчивыми сочетаниями значимо более приемлемы, чем предложения с сочетаниями без статистической устойчивости, несмотря на то что в последних не нарушены какие бы то ни было сочетаемостные ограничения (семантические, лексические, морфо-синтаксические) в смысле [Апресян, 1974, с. 60–62]. Подтверждение этой гипотезы означало бы, что носители языка способны не только отличать противоречащее языковым правилам от не противоречащего им (например, осознавать грамматичность сочетания *что-то новое* и неграмматичность сочетания *что-то новая*), но и определять, что из соответствующего языковым правилам встречается в русской речи нетривиально часто, а что – нетривиально редко.

Было проведено два эксперимента: в первом исследовались сочетания типа «*что-то* + полное прилагательное», во втором – типа «*кто-то* + полное прилагательное».

2. Описание первого эксперимента (сочетание «*что-то* + полное прилагательное»)

В ходе первого эксперимента носители языка оценивали предложения, в которых были представлены либо высокоустойчивые сочетания типа «*что-то* + полное прилагательное», либо неустойчивые сочетания того же типа. В эксперименте, таким образом, исследовался фактор «статистическая устойчивость сочетания», который имел 2 уровня:

«высокая степень статистической устойчивости» и «отсутствие статистической устойчивости».

Сочетания для эксперимента отбирались из списка, автоматически сформированного в основном корпусе НКРЯ по запросу «ключ *что-то*, *SPRO*, коллокат *A & plen*, на расстоянии от 1 до 1 от ключа» (поле «Поиск коллокаций»).

Отбор производился с установкой на формирование пар, в которых постпозитивные прилагательные сходны по семантике (*знакомый – популярный, белый – бесцветный, грустный – нерадостный, вкусный – сытный*). Подобных пар, в которых одно прилагательное образует вместе с местоимением *что-то* высокоустойчивое сочетание, а другое прилагательное – сочетание, не обладающее статистической устойчивостью, обнаружилось очень мало. В связи с этим пришлось ограничиться небольшим количеством лексикализаций на условие – четырема: **Он напевал *что-то* знакомое / *что-то* популярное. У него в руке было *что-то* белое / *что-то* бесцветное. Произошло *что-то* грустное / *что-то* нерадостное. Мне дали съесть *что-то* вкусное / *что-то* сытное.**

Предложения распределялись по экспериментальным листам в соответствии с правилом латинского квадрата: 2 листа, т.к. имеется два экспериментальных условия, в каждом листе по 4 собственно экспериментальных, или стимульных, предложения, всего 8 экспериментальных предложений.

В экспериментальных листах содержались как собственно экспериментальные предложения, так и филлеры. Филлеры были грамматически приемлемыми, в них местоимение *что-то* употреблялось без зависимых прилагательных. Филлеры и стимульные предложения находились в соотношении 1 : 1. Предложения 1-го экспериментального листа с комментариями см. в табл. 1.

Участниками эксперимента стали 23 человека – студенты факультета иностранных языков Института филологии, иностранных языков и медиакоммуникации Иркутского государственного университета, а также выпускники факультета теоретической и прикладной филологии того же института. Респонденты распределились по экспериментальным листам следующим образом: 1-й лист – 11 участников, 2-й лист – 12 участников.

Эксперимент распространялся через социальные сети и электронную почту. Опрос проводился на платформе PCIBexFarm (<https://farm.pcibex.net>) с применением методики оценки приемлемости по шкале Ликерта от 1 до 7. Анализ данных был осуществлен в среде разработки RStudio с использованием языка программирования R [R Core Team 2023].

Результаты эксперимента представлены в табл. 2.

**Предложения 1-го опросного листа эксперимента,
направленного на исследование сочетаний типа
«что-то + полное прилагательное», с комментариями
[Sentences of the 1st questionnaire of the experiment
aimed at studying combinations of the type
“*čto-to* ‘something’ + full adjective”, with comments]**

Предложение [Sentence]	Комментарий [Comment]
<i>Он напевал что-то знакомое.</i>	Высокая устойчивость (1-я лексикализация) [High stability (1st lexicalization)]
<i>На столе лежало что-то вроде фотоаппарата.</i>	Филлер [Filler]
<i>У него в руке было что-то бесцветное.</i>	Неустойчивость (2-я лексикализация) [Instability (2nd lexicalization)]
<i>Он что-то достал из кармана.</i>	Филлер [Filler]
<i>Произошло что-то грустное.</i>	Высокая устойчивость (3-я лексикализация) [High stability (3st lexicalization)]
<i>Вдруг я услышал что-то из Бетховена.</i>	Филлер [Филлер]
<i>Мне дали съесть что-то сытное.</i>	Неустойчивость (4-я лексикализация) [Instability (4nd lexicalization)]
<i>С ним что-то случилось.</i>	Филлер [Filler]

Средняя нормализованная оценка предложений с высокоустойчивыми сочетаниями типа «что-то + полное прилагательное» составляет 0,19, а средняя нормализованная оценка предложений с неустойчивыми сочетаниями того же типа – минус 0,22. При этом во всех парах сочетаний с семантически сходными прилагательными, кроме пары *что-то грустное – что-то нерадостное*, неустойчивое сочетание имеет более низкую оценку, чем высокоустойчивое.

Существенно, что разница между оценками, данными предложениям с высокоустойчивыми сочетаниями, и оценками, данными предложениям с неустойчивыми сочетаниями, статистически значима: $df = 1$; $F = 4,413$; $p = 0,0385$ (однофакторный дисперсионный анализ; зависимость оценки от того, что употребляется в предложении – высокоустойчивое сочетание или сочетание, не обладающее статистической устойчивостью).

Таблица 2

**Средние нормализованные оценки предложений с сочетаниями
типа «что-то + полное прилагательное»
[Average normalized ratings of sentences with combinations
of the type “*что-то* ‘something’ + full adjective”]**

Высокоустойчивое сочетание [Highly stable combination]	Средняя нормализованная оценка предложения, содержащего данное высокоустойчивое сочетание [Average normalized score of a sentence containing a given highly stable combination]	Статистически неустойчивое сочетание [Statistically unstable combination]	Средняя нормализованная оценка предложения, содержащего данное статистически неустойчивое сочетание [The average normalized score of a sentence containing a given statistically unstable combination]
<i>Что-то знакомое</i>	0,53	<i>Что-то популярное</i>	-0,22
<i>Что-то белое</i>	0,94	<i>Что-то бесцветное</i>	0,12
<i>Что-то грустное</i>	-0,35	<i>Что-то нерадостное</i>	-0,19
<i>Что-то вкусное</i>	-0,38	<i>Что-то сытное</i>	-0,60
В среднем [On average]	0,19	В среднем [On average]	-0,22

3. Описание 2-го эксперимента (сочетание «кто-то + полное прилагательное»)

Второй эксперимент был устроен аналогично первому.

На каждое из двух экспериментальных условий было предложено 4 лексикализации: *Что делать, если я встречу кого-то чужого / кого-то своего? На пороге показался кто-то огромный / кто-то низкий. Около двери сидел кто-то страшный / кто-то красивый. Это сказал кто-то умный / кто-то глупый.*

Как и в случае первого эксперимента, в экспериментальных листах содержались, помимо собственно экспериментальных предложений, грамматически приемлемые филлеры. В них местоимение *кто-то* употреблялось без зависимых прилагательных. Филлеры и стимульные предложения находились в соотношении 1 : 1. Предложения 1-го экспериментального листа с комментариями см. в табл. 3.

Таблица 3

**Предложения 1-го опросного листа эксперимента,
направленного на исследование сочетаний типа
«кто-то + полное прилагательное», с комментариями
[Sentences of the 1st questionnaire of the experiment
aimed at studying combinations of the type
“kto-to ‘somebody’ + full adjective”, with comments]**

Предложение [Sentence]	Комментарий [Comment]
<i>Что делать, если я встречу кого-то чужого?</i>	Высокая устойчивость (1-я лексикализация) [High stability (1st lexicalization)]
<i>Ко мне подошел кто-то в сером.</i>	Филлер [Filler]
<i>На пороге показался кто-то низкий.</i>	Неустойчивость (2-я лексикализация) [Instability (2nd lexicalization)]
<i>Кто-то заходил в комнату пять минут назад.</i>	Филлер [Филлер]
<i>Около двери сидел кто-то страшный.</i>	Высокая устойчивость (3-я лексикализация) [High stability (3rd lexicalization)]
<i>Он разговаривал с кем-то из съемочной группы.</i>	Филлер [Филлер]
<i>Это сказал кто-то глупый.</i>	Неустойчивость (4-я лексикализация) [Instability (4nd lexicalization)]
<i>Ты, наверное, разговорился с кем-то.</i>	Филлер [Filler]

Таблица 4

**Средние нормализованные оценки предложений с сочетаниями
типа «кто-то + полное прилагательное»
[Average normalized ratings of sentences with combinations
of the type “*kto-to* + full adjective”]**

Высокоустойчивое сочетание [Highly stable combination]	Средняя нормализованная оценка предложения, содержащего данное высокоустойчивое сочетание [Average normalized score of a sentence containing a given highly stable combination]	Статистически неустойчивое сочетание [Statistically unstable combination]	Средняя нормализованная оценка предложения, содержащего данное статистически неустойчивое сочетание [The average normalized score of a sentence containing a given statistically unstable combination]
<i>Кто-то чужой</i>	-0,07	<i>Кто-то свой</i>	-0,14
<i>Кто-то огромный</i>	-0,18	<i>Кто-то низкий</i>	-0,74
<i>Кто-то страшный</i>	0,09	<i>Кто-то красивый</i>	-0,88
<i>Кто-то умный</i>	-0,40	<i>Кто-то глупый</i>	-0,52
В среднем [On average]	-0,15	В среднем [On average]	-0,57

В эксперименте принял участие 21 человек. Респонденты распределились по экспериментальным листам следующим образом: 1-й лист – 10 участников, 2-й лист – 11 участников.

Результаты второго эксперимента представлены в табл. 4.

Предложения с высокоустойчивыми сочетаниями типа «кто-то + полное прилагательное» получили среднюю нормализованную оценку $-0,15$, а предложения с неустойчивыми сочетаниями того же типа – минус $0,57$. В каждой из 4 пар сочетаний с семантически сходными прилагательными неустойчивое сочетание имеет более низкую оценку, чем высокоустойчивое.

Разница между оценками предложений с высокоустойчивыми сочетаниями и оценками предложений с неустойчивыми сочетаниями, как и в случае первого эксперимента, статистически значима: $df = 1$; $F = 4,268$; $p = 0,042$ (однофакторный дисперсионный анализ; зависимость оценки от того, что употребляется в предложении – высокоустойчивое сочетание или сочетание, не обладающее статистической устойчивостью).

4. Обсуждение

Экспериментальные данные полностью подтверждают исходную гипотезу. Носители языка действительно воспринимают как менее приемлемые те сочетания, которые менее характерны для русской речи, несмотря на то что никакие сочетаемостные ограничения в этих сочетаниях не нарушены. Наблюдается прямая связь между корпусными данными и языковой интуицией носителей.

В завершение следует сказать об одном дополнительном наблюдении.

Обращает на себя внимание, что предложения первого эксперимента получили в среднем более высокие оценки, чем предложения второго. Средняя нормализованная оценка предложений с сочетанием «что-то + полное прилагательное» составляет $-0,02$, средняя нормализованная оценка предложений с сочетанием «кто-то + полное прилагательное» – минус $0,36$.

Если объединить результаты первого и второго эксперимента и принять за зависимую переменную оценку предложения, а за независимую – тип сочетания с точки зрения употребленного в нем местоимения («что-то + полное прилагательное» или «кто-то + полное прилагательное»), то однофакторный дисперсионный анализ покажет, что разница между оценками предложений с сочетанием «что-то + полное прилагательное» и оценками предложений с сочетанием «кто-то + полное прилагательное» статистически значима: $df = 1$; $F = 5,621$; $p = 0,0188$ (зависимость оценки от типа сочетания – «что-то + полное прилагательное» / «кто-то + полное прилагательное»).

Показательно также, что предложения с сочетанием «*кто-то* + полное прилагательное» имеют в среднем значимо более низкие оценки, чем грамматичные филлеры с местоимением *кто-то* ($df = 1$; $F = 24,35$; $p = 1,94 \cdot 10^{-6}$), в то время как предложения с сочетанием «*что-то* + полное прилагательное» значимо не отличаются по оценкам от грамматичных филлеров с местоимением *что-то* (хотя и оцениваются в среднем ниже, чем грамматичные филлеры).

Это коррелирует с тем фактом, что в текстах существенно большей склонностью употребляться с постпозитивными полными прилагательными обладает местоимение *что-то*: согласно данным основного корпуса НКРЯ, показатель \logDice для сочетания «*что-то* + полное прилагательное» составляет 5,31, а для сочетания «*кто-то* + полное прилагательное» – 1,25.

Можно, таким образом, констатировать, что в данном случае, как и в случае основных двух экспериментов, обнаруживается прямая связь между экспериментальными и корпусными данными.

Однако характер наблюдаемой связи, как кажется, здесь другой.

В случае основных двух экспериментов причину корреляции между корпусными и экспериментальными данными можно усматривать в том, что узус влияет на языковую интуицию: носители языка, нетривиально часто встречаясь в текстах с сочетаниями типа *что-то знакомое*, *что-то белое*, *кто-то страшный*, начинают воспринимать их как более приемлемые, чем сочетания типа *что-то популярное*, *что-то бесцветное*, *кто-то красивый*, которые, наоборот, встречаются в текстах нетривиально редко. При этом объяснить, почему сочетание *что-то знакомое* высокоустойчиво, а сочетание *что-то популярное* неустойчиво, затруднительно. По-видимому, это связано с частотностью прилагательных *знакомый* и *популярный*: *знакомый* частотнее, чем *популярный* (см. [Ляшевская, Шаров, 2009]). Но сама эта частотность представляет собой речевую данность, которую сложно каким-либо образом объяснить.

В случае же с дополнительным наблюдением связь между корпусными и экспериментальными данными имеет другой характер, т.к. и те, и другие данные указывают на одну и ту же особенность семантики местоимений *что-то* и *кто-то*: «понятие “предмета”, которое стоит за *что*-местоимениями..., гораздо абстрактнее понятия “лица”, стоящего за *кто*-местоимениями» [Кузьмина, 1989, с. 202].

Итак, местоимение *что-то* обладает более неопределенной семантикой, чем *кто-то*, а значит, ему более естественно сочетаться с постпозитивными полными прилагательными, которые в некоторой мере компенсируют неопределенность. Большая естественность конструкции «*что-то* + полное прилагательное» находит отражение, с одной

стороны, в большей устойчивости этой конструкции в корпусе, а с другой – в большей приемлемости этой конструкции для носителей языка. В результате корпусные и экспериментальные данные и оказываются напрямую связаны.

Можно заключить, что экспериментальный анализ сочетаний типа «что-то / кто-то + полное прилагательное» позволяет не только констатировать наличие на этом участке прямой связи между корпусными данными и интуицией носителей языка, но и выявить, что природа прямой корреляции между корпусными и экспериментальными данными может быть различной.

Библиографический список / References

Апресян, 1974 – Апресян Ю.Д. Лексическая семантика. Синонимические средства языка. М., 1974. [Apresyan Yu.D. Leksicheskaya semantika. Sinonimicheskie sredstva yazyka [Lexical semantics. Synonymous means of language]. Moscow, 1974.]

Буров, 2023 – Буров Э.Е. Выявление сочетаемостных свойств слова с помощью статистических мер ассоциации // «Цифра» в социально-гуманитарных исследованиях: метод, поле, реальность: материалы конференции молодых ученых. Иркутск, 14–16 ноября 2022 г. Иркутск, 2023. С. 8–11. [Burov E.E. Identification of words' syntagmatic properties by means of statistical measures of association. "Tsifra" v sotsialno-gumanitarnykh issledovaniyakh: metod, pole, realnost. Irkutsk, 2023. Pp. 8–11. (In Rus.)]

Захаров, Хохлова, 2014 – Захаров В.П., Хохлова М.В. Выделение терминологических словосочетаний из специальных текстов на основе различных мер ассоциации // Технологии информационного общества в науке, образовании и культуре: сборник научных статей. Труды XVII Всероссийской объединенной конференции «Интернет и современное общество» (IMS–2014). СПб., 2014. С. 290–293. [Zakharov V.P., Khokhlova M.V. Extracting terminological phrases by different association measures. Tekhnologii informatsionnogo obshchestva v nauke, obrazovanii i kulture. Trudy XVII Vserossiiskoi obedinennoi konferentsii «Internet i sovremennoe obshchestvo» (IMS–2014). St. Petersburg, 2014. Pp. 290–293. (In Rus.)]

Кузьмина, 1989 – Кузьмина С.М. Семантика и стилистика неопределенных местоимений // Грамматические исследования. Функционально-стилистический аспект / отв. ред. Д.Н. Шмелев. М., 1989. С. 158–231. [Kuzmina S.M. Semantics and stylistics of indefinite pronouns. Grammaticheskie issledovaniya. Funktsionalno-stilisticheskii aspekt. D.N. Shmelev (ed.). Moscow, 1989. Pp. 158–231. (In Rus.)]

Левонтина, 2022 – Левонтина И.Б. Частицы речи. М., 2022. [Levontina I.B. Chastitsy rechi [Particles of speech]. Moscow, 2022.]

Ляшевская, Шаров 2009 – Ляшевская О.Н., Шаров С.А. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М., 2009. [Lyashevskaya O.N., Sharov S.A. Chastotnyi slovar sovremennogo russkogo yazyka (na materialakh Natsionalnogo korpusa russkogo yazyka) [Frequency dictionary of modern Russian based on the Russian National Corpus]. Moscow, 2009.]

Ольховская, Зельникова, 2023 – Ольховская А.И., Зельникова А.А. Математические инструменты измерения коллокаций в аспекте изучения фразеологии // III Костомаровский форум. Славянская культура: истоки, традиции,

взаимодействие: материалы Международной научно-практической конференции. М., 2023. С. 418–423. [Olkhovskaya A.I., Zelnikova A.A. Mathematical tools for measuring collocations in the aspect of phraseology research. III Kostomarovskii forum. Slavyanskaya kultura: istoki, traditsii, vzaimodeistvie. Moscow, 2023. Pp. 418–423. (In Rus.)]

Падучева, 2016 – Падучева Е.В. Местоимения неизвестности. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru>). На правах рукописи. М., 2016. [Paducheva E.V. Mestoimeniya neizvestnosti. Materialy dlya proekta korpusnogo opisaniya russkoi grammatiki (<http://rusgram.ru>) (Pronouns of the Unknown. Materials for the project of corpus description of Russian grammar (<http://rusgram.ru>)). On the praach of the manuscript. Moscow, 2016.]

Хохлова, 2017 – Хохлова М.В. Сопоставительный анализ статистических мер на примере частеречных предпочтений сочетаемости существительных // Компьютерная лингвистика и вычислительные онтологии. 2017. Вып. 1. С. 166–171. [Khokhlova M.V. A comparative analysis of the association measures based on the analysis of part-of-speech collocational preferences of the Russian nouns. *Kompyuternaya lingvistika i vychislitelnye ontologii*. 2017. Iss. 1. Pp. 166–171. (In Rus.)]

Bresnan, 2007 – Bresnan J. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. *Roots: Linguistics in Search of its Evidential base*. Vol. 96. S. Featherston, W. Sternefeld (eds.). Berlin, 2007. Pp. 77–96.

Divjak, 2017 – Divjak D. The role of lexical frequency in the acceptability of syntactic variants: Evidence from *that* clauses in Polish. *Cognitive Science*. 2017. No. 4. Pp. 354–382. DOI: 10.1111/cogs.12335

Gerasimova, Lyutikova, 2020 – Gerasimova A., Lyutikova E. Intralingual variation in acceptability judgments and production: Three case studies in Russian grammar. *Frontiers in Psychology*. 2020. No. 11. 348. DOI: 10.3389/fpsyg.2020.00348

Lau et al., 2017 – Lau J.H., Clark A., Lappin S. Grammaticality, acceptability, and probability: A probabilistic view of linguistic knowledge. *Cognitive Science*. 2017. No. 41. Pp. 1201–1241. DOI: 10.1111/cogs.12414

R Core Team, 2023 – R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2023.

Статья поступила в редакцию 12.08.2024

The article was received on 12.08.2024

Сведения об авторе / About the author

Буров Эдуард Евгеньевич – аспирант кафедры русского языка и общего языкознания факультета теоретической и прикладной филологии, Иркутский государственный университет; преподаватель кафедры русского языка и общего языкознания факультета теоретической и прикладной филологии, Иркутский государственный университет

Eduard E. Burov – postgraduate at the Department of Russian Language and General Linguistics of the Faculty of Theoretical and Applied Philology, Irkutsk State University; lecturer at the Department of Russian Language and General Linguistics of the Faculty of Theoretical and Applied Philology, Irkutsk State University, Irkutsk, Russian Federation

E-mail: zhenyabur@mail.ru

Т.И. Давидюк

Московский государственный университет имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация;

Институт языкознания Российской академии наук,
125009 г. Москва, Российская Федерация

Лично-числовое согласование с сочинительными конструкциями с союзом *ни ... ни*: корпусное и экспериментальное исследование

В статье рассматривается согласование в конструкциях с сочиненным подлежащим с союзом *ни ... ни*, где первым конъюнктом выступает личное местоимение *я*. Для сравнения стратегий согласования с конструкциями с союзом *ни ... ни* и союзом *и ... и* (который считается положительным эквивалентом для первого) в исследовании используются корпусные и экспериментальные данные. Корпусные данные выявили возможность согласования с ближайшим конъюнктом при порядке слов, когда предикат следует за подлежащим, для конструкций с *ни ... ни*, тогда как для конструкций с *и ... и* такая стратегия согласования оказалась более редкой; однако эксперимент не подтвердил значимых различий между конструкциями с *ни ... ни* и *и ... и*. Приемлемыми стратегиями согласования оказались только согласование по правилам разрешения и согласование по 3 лицу множественного числа. В статье также обсуждается, почему конструкции с союзом *и ... и* получили более низкие оценки, и предполагается, что это связано с наличием отрицания в экспериментальных предложениях, не связанного с требованием отрицательного согласования, которое имеет место для предложений с *ни ... ни*. Отдельное внимание уделяется роли нереферентных конъюнктов в выборе стратегии согласования.

Ключевые слова: лично-числовое согласование, сочиненное подлежащее, русский язык, экспериментальный синтаксис, корпусное исследование

Благодарности. Исследование выполнено в рамках проекта Российского научного фонда № 22-18-00037, реализуемого в Московском государственном университете имени М.В. Ломоносова, <https://rscf.ru/project/22-18-00037/>

Благодарю рецензента и редколлегию журнала за высказанные замечания и внимательное отношение к статье.

ДЛЯ ЦИТИРОВАНИЯ: Давидюк Т.И. Лично-числовое согласование с сочинительными конструкциями с союзом *ни ... ни*: корпусное и экспериментальное исследование // Рема. Rhema. 2024. № 3. С. 33–53. DOI: 10.31862/2500-2953-2024-3-33-53

DOI: 10.31862/2500-2953-2024-3-33-53

T. Davidyuk

Lomonosov Moscow State University,
Moscow, 119991, Russian Federation;
Institute of Linguistics, Russian Academy of Sciences,
Moscow, 125009, Russian Federation

Person-number agreement with constructions conjoined by *ni ... ni* ‘neither ... nor’: A corpus and experimental study

The article examines agreement in constructions with conjoined subjects linked by the conjunction *ni ... ni* ‘neither ... nor’, where the first conjunct is the personal pronoun *ya* ‘I’. The study is based on corpus and experimental data, comparing agreement strategies for constructions with *ni ... ni* ‘neither ... nor’ and the conjunction *i ... i* ‘both ... and’, which is considered to be its positive counterpart. The corpus data revealed the possibility of the closest conjunct agreement in SV word order for constructions with *ni ... ni* ‘neither ... nor’, whereas this strategy was less common for constructions with *i ... i* ‘both ... and’; however, the experiment did not confirm significant differences between the *ni ... ni* ‘neither ... nor’ and *i ... i* ‘both ... and’ constructions. The only acceptable agreement strategies were resolved agreement and the 3rd person plural agreement. The article also discusses

why constructions with *i ... i* 'both ... and' received lower ratings, suggesting that this may be due to the presence of negation in the experimental sentences. Special attention is given to the role of non-referential conjuncts in determining the agreement strategy.

Key words: person-number agreement, conjoined subject, Russian language, experimental syntax, corpus study

Acknowledgments. The study was carried out within the framework of the Russian Science Foundation project No. 22-18-00037, implemented at Lomonosov Moscow State University, <https://rscf.ru/project/22-18-00037/>

I thank the reviewer and the editorial board of the journal for their comments and attentive attitude to the article.

FOR CITATION: Davidyuk T. Person-number agreement with constructions conjoined by *ni ... ni* 'neither ... nor': A corpus and experimental study. *Rhema*. 2024. No. 3. Pp. 33–53. (In Rus.). DOI: 10.31862/2500-2953-2024-3-33-53

1. Введение

Данная статья посвящена исследованию лично-числового согласования с подлежащими, сочиненными союзом *ни ... ни*. Материал согласования с сочинительными конструкциями интересен в нескольких аспектах.

Во-первых, внутри сочинительной конструкции достаточно свободно можно сочетать именные группы с различными признаками, среди которых признак лица имеет некоторую специфику. Так, в языках мира обнаруживается, что согласование по лицу оказывается более ограниченным по сравнению с согласованием по другим именным категориям. Это проявляется в эффектах лично-падежных ограничений (англ. *Person Case Constraint*) [Haspelmath, 2004], при которых в некоторых языках оказывается невозможным определенное сочетание лиц у прямого и непрямого объектов, но подобных ограничений для других признаков не наблюдается. Кроме того, существуют конфигурации, в которых согласование по лицу оказывается невозможным, в отличие от согласования по другим именным категориям, в частности, при дистантном согласовании (англ. *long-distance agreement*) [Baker, 2011]. Таким образом, возможность комбинирования различных признаков и их значений представляется важным свойством сочинительных конструкций в контексте исследования признаковых асимметрий.

Во-вторых, одним из основных механизмов согласования с сочинительными конструкциями является разрешение признаков, однако оно может осуществляться по-разному для разных признаков, что также

подчеркивает их противопоставление. В языках мира наблюдаются определенные закономерности при разрешении признаков в рамках сочиненной конструкции [Corbett, 1983] – так называемые правила разрешения (англ. *resolution rules*), которые позволяют осуществлять согласование на основе «подсчета» признаков всех конъюнктов (такое согласование мы будем называть согласованием по правилам разрешения).

Разрешение признака числа в языках с двухчастным противопоставлением (ед. ч. / мн. ч.), к которым относится и современный русский язык, устроено таким образом, что сочетание нескольких признаков единственного числа (а также сочетание нескольких признаков единственного и множественного числа) дает значение множественного числа. Правила разрешения для признака лица в языках мира строятся на иерархии лиц $1 > 2 > 3$ [Zwicky, 1977]. Если в состав сочинительной конструкции входит местоимение 1 лица, то согласование будет осуществляться по 1 лицу вне зависимости от лица других конъюнктов; если в состав сочинительной конструкции входит местоимение 2 лица (но при этом не входит местоимение 1 лица), то выбор результирующего значения лица будет осуществлен в пользу 2 лица, а не 3 лица. Иллюстрация работы разрешения признака лица на основе иерархии лиц приведена в (1а–с) на примере чешского языка. Что касается признака рода, то он является более вариативным в плане правил разрешения и в некоторых языках оказывается зависим от одушевленности конъюнктов [Adamson, Anagnostopoulou, in press].

- (1) a. *já a ty zůstaneme doma.*
я и ты оставаться.FUT.1PL¹ дома
'Я и ты останемся дома'.
- b. *bratr a já se učíme hrát na klavír.*
брат и я REFL учиться.PRS.1PL играть на фортепиано
'Брат и я учимся играть на фортепиано'.
- c. *tvůj otec a ty jste si podobní.*
твой отец и ты быть.PRS.2PL друг на друга похожи
'Твой отец и ты похожи друг на друга'. [Corbett, 1983, p. 176]

В-третьих, на согласование с сочинительными конструкциями могут влиять постсинтаксические факторы, однако результат взаимодействия этих факторов также может зависеть от конкретных признаков,

¹ Условные сокращения: 1–3 – 1–3 лицо; FUT – будущее время; PL – множественное число; PRS – настоящее время; REFL – рефлексив; SV – порядок слов, при котором предикат следует за подлежащим; VS – порядок слов, при котором предикат предшествует подлежащему.

представленных в сочинительной конструкции. Например, постсинтаксические механизмы согласования приобретают большее значение, если в языке отсутствует достаточно устойчивое правило разрешения для определенного признака. В частности, как отмечено в [Nevins, Weisser, 2019, p. 225], отсутствие типологически «общепринятых» правил для разрешения признака рода приводит к тому, что при конфликте признака рода между конъюнктами в языках мира чаще выбирается согласование только с одним из конъюнктов.

Согласование с сочинительными конструкциями демонстрирует значительную вариативность, и, учитывая гибкость сочетаний различных признаков внутри таких конструкций, а также возможное влияние внешнего синтаксиса, экспериментальные исследования представляются эффективным методом его изучения. Экспериментальных исследований согласования с сочинительными конструкциями при конфликте значений признака лица у конъюнктов сравнительно немного, что, вероятно, связано с тем, что действие иерархии лиц, лежащей в основе правил разрешения, считается практически универсальным механизмом [Corbett, 2006]. Однако имеющиеся экспериментальные исследования на материале германских языков [Timmermans et al., 2004; Himmelreich, Hartmann, 2023] и русского языка [Белова, Давидюк, 2023; Гриценко, 2024; Давидюк, 2024] показывают нестрогость действия иерархии лиц при согласовании с сочинительными конструкциями, проявляющуюся в возможности выбрать в качестве дефолтного значения 3 лицо. В перечисленных работах содержатся экспериментальные исследования, выполненные на материале конструкций не только с конъюнктивными, но и с дизъюнктивными союзами, однако конструкции с союзом *ни ... ни* остаются вне рассмотрения. Данная статья направлена на восполнение этой лакуны.

Союз *ни ... ни* в русском языке считается отрицательным «аналогом» союза *и ... и* [РГ, 1980, с. 618; Санников, 2008, с. 300]. Однако в корпусном исследовании О.Е. Пекелис между конструкциями с этими двумя союзами обнаруживаются различия в согласовании: при обычном порядке слов с конструкциями с *ни ... ни* свободно допускается согласование по единственному числу даже при референтных одушевлённых конъюнктах (2). В работе подчеркивается, что эта особенность конструкций с союзом *ни ... ни* отмечается также и в [Граудина и др., 2008, с. 39], и в ней же, вопреки традиции, данный союз относят к классу разделительных. О.Е. Пекелис предлагает семантическое объяснение этой согласовательной особенности конструкций с союзом *ни ... ни*, предполагая, что в значении союза действительно присутствует разделительный компонент [Пекелис, 2013а].

- (2) Ни отец, ни мать не **помнила**, которого числа я родился.
[Пекелис, 2013а, с. 81]

Конструкции, подобные русским конструкциям с союзом *ни ... ни*, становились предметом отдельных экспериментальных исследований на материале других языков. В ранней работе [Peterson, 1986], использующей методику выбора между альтернативами (англ. *forced choice*), испытуемым предлагалось выбрать один из предложенных вариантов согласования (с возможностью также не выбирать ничего или выбрать все предложенные варианты) при конъюнктах, совпадающих и не совпадающих по числу. Исследование проводилось на материале английского языка и включало конструкции с дизъюнктивными союзами (*or, either ... or*) и с союзом *neither ... nor* 'ни ... ни'. Было выявлено, что при конструкциях с союзом *neither ... nor* 'ни ... ни' преобладает согласование по множественному числу, даже в случае конъюнктов единственного числа. Этот результат отличает данные конструкции от конструкций с союзами *or* 'или' и *either ... or* 'или ... или', где согласование по множественному числу с конъюнктами в единственном числе наблюдалось значительно реже.

В другом исследовании, выполненном на материале немецкого языка [Felser, Jessen, 2021], сравнивались стратегии согласования для конструкций с союзами *sowohl ... als auch* 'как ... так и' и *weder ... noch* 'ни ... ни'. В рамках исследования было проведено два эксперимента: в первом использовалась методика оценки приемлемости, а во втором – методика чтения с саморегуляцией скорости, при которой участники должны были выбрать подходящий вариант согласования. В обоих экспериментах было выявлено превалирование согласования по множественному числу. Значимое различие между этими двумя союзами было выявлено только в первом эксперименте, в котором оценки согласования по множественному числу для конструкций с союзом *weder ... noch* 'ни ... ни' оказались несколько снижены по сравнению с аналогичным условием для *sowohl ... als auch*.

Таким образом, исследование П. Петерсона (1986) демонстрирует, что конструкции с союзом 'ни ... ни' в меньшей степени обладают согласовательными свойствами, характерными для разделительных дизъюнктивных союзов. Исследование К. Фелсер и А. Йессен (2021) также выявляет сходство между конструкциями с 'ни ... ни' и конструкциями с конъюнкцией, но при этом обнаруживается некоторое снижение оценок приемлемости для согласования по правилам разрешения при конструкциях с 'ни ... ни'. Однако в перечисленных исследованиях не рассматривались случаи конфликта признака лица у конъюнктов.

Наше исследование лично-числового согласования с конструкциями с союзом *ни ... ни* и его сравнение со стратегиями согласования для конструкций с союзом *и ... и* структурировано следующим образом. В разделе 2 содержатся корпусные данные по исследуемым конструкциям. Раздел 3 описывает проведенное нами экспериментальное исследование. В разделе 4 обсуждаются полученные результаты. В разделе 5 приводится заключение.

2. Корпусное исследование

Хотя основное внимание в данной работе сосредоточено на экспериментальных исследованиях, обращение к корпусным данным² помогает наметить возможные гипотезы относительно исследуемого явления. Так как в экспериментальном исследовании (разд. 3) в качестве стимульного материала используются предложения с сочиненным подлежащим, в котором первым конъюнктом выступает личное местоимение *я*, мы рассмотрели подобные конструкции в основном подкорпусе Национального корпуса русского языка (НКРЯ) [Савчук и др., 2024].

Результаты поиска в основном подкорпусе НКРЯ для конструкции «*и я, и X*» представлены в табл. 1. Структура запроса в корпусе: *и + я*, на расстоянии от 1 до 1 от Слова 1 + *и*, на расстоянии от 1 до 1 от Слова 2 + (S | SPRO) & *пом*, на расстоянии от 1 до 5 от Слова 3, где S – существительное, SPRO – местоименное существительное, *пом* – именительный падеж. В итоговую таблицу не включались примеры с более чем двумя конъюнктами и примеры с отсутствием запятой между конъюнктами.

Для подлежащих вида «*и я, и X*» при обычном порядке слов – SV – преобладает согласование по правилам разрешения (3). Также встретилось пять примеров с согласованием по 3 лицу множественного числа. Три из них можно расценивать как согласование с ближайшим конъюнктом, т.к. последним конъюнктом является имя существительное множественного числа (4); остальные два случая нельзя считать согласованием с ближайшим конъюнктом (5).

(3) И я, и Анна очень **рассчитываем** на это.
(НКРЯ. Виктор Пелевин. Чапаев и Пустота (1996))

(4) Но и я, и все мои знакомые девушки **обращают** особое внимание на парней, которые пахнут туалетной водичкой. (НКРЯ. Форум: Фэндомная Правда (2011))

² Исследование проведено на материале текстов Национального корпуса русского языка (ruscorpora.com).

- (5) Но помимо всего этого и я, и группа **имеют** собственные особенности. (НКРЯ. Иэн Гиллан, Макс Хаген.
«Ты не можешь быть классиком, пока не умер» (24.10.2008) // <http://www.rbcdaily.ru/2008/10/24/lifestyle/387361.shtml>)

Таблица 1

Предикативное согласование с сочиненным подлежащим вида «и я, и X»: данные Национального корпуса русского языка
[Predicate agreement with a coordinated subject of the form “i ja, i X”
‘both I, ... and X’: data from the Russian National Corpus]

Тип согласования [Type of agreement]	Порядок слов [Word order]	
	SV	VS
Согласование по лицу и числу [Agreement by person and number]		
1 лицо, множественное число [1st person, plural]	48	2
1 лицо, единственное число [1st person, singular]	–	3
3 лицо, множественное число [3rd person, plural]	5	–
Согласование по числу [Agreement by number]		
Множественное число [Plural]	215	18
Единственное число [Singular]	–	11

Предложений с порядком слов VS с предикатом в непрошедшем времени, согласующимся по лицу и числу, мало, и в них встречается как согласование по правилам разрешения, так и согласование с ближайшим конъюнктом (6)–(7). В контекстах только с числовым согласованием встречается и согласование по множественному, и согласование по единственному числу (8)–(9).

- (6) **Живем** давно в печали и я, и мать моя. (НКРЯ. Василий Ян. Финикийский корабль (1931))
- (7) Как они не понимают – **беспокоюсь** и я, и мать.
(НКРЯ. К.А. Коровин. Утенок (1930–1938))
- (8) Характеристику написал Иосиф Шкловский, тоже астрофизик, член-корреспондент АН СССР, которого, независимо, хорошо **знали** и я, и Люся. (НКРЯ. А.Д. Сахаров. Воспоминания (1983–1989))

(9) В этом **убедился** и я, и Владимир Ильич. (НКРЯ. Л.Д. Троцкий. Моя жизнь (1929–1933))

Результаты поиска в основном подкорпусе НКРЯ для конструкции «*ни я, ни X*» представлены в табл. 2. Структура запроса в корпусе: *ни + я*, на расстоянии от 1 до 1 от Слова 1 + *ни*, на расстоянии от 1 до 1 от Слова 2 + (S | SPRO) & *пом*, на расстоянии от 1 до 5 от Слова 3, где S – существительное, SPRO – местоименное существительное, *пом* – именительный падеж. В итоговую таблицу не включались примеры с более чем двумя конъюнктами и примеры с отсутствием запятой между конъюнктами.

Таблица 2

Предикативное согласование с сочиненным подлежащим вида «*ни я, ни X*»: данные Национального корпуса русского языка
[Predicate agreement with a coordinated subject of the form «*ni ja, ni X*» ‘neither I, nor X’: data from the Russian National Corpus]

Тип согласования [Type of agreement]	Порядок слов [Word order]	
	SV	VS
Согласование по лицу и числу [Agreement by person and number]		
1 лицо, множественное число [1st person, plural]	124	5
1 лицо, единственное число [1st person, singular]	–	3
3 лицо, множественное число [3rd person, plural]	18	–
3 лицо, единственное число [3rd person, singular]	18	–
Согласование по числу [Agreement by number]		
Множественное число [Plural]	279	12
Единственное число [Singular]	15	8

В предложениях с конструкцией с *ни ... ни* при обычном порядке слов в контекстах с лично-числовым согласованием также преобладает согласование по правилам разрешения (10). Было обнаружено 18 примеров с согласованием по 3 лицу множественного числа, которое в 11 случаях можно причислить к согласованию с ближайшим конъюнктом, т.к. ближайшим конъюнктом является именная группа 3 лица множественного числа (11); в остальных 7 примерах мы действительно наблюдаем

согласование по 3 лицу множественного числа (12). Среди 18 примеров с согласованием по 3 лицу единственному числу подавляющее большинство примеров – это примеры, в которых вторым конъюнктом выступает неопределенное местоимение (13)–(14) или другие нереферентные выражения (15); лишь единичные примеры с таким согласованием содержат в качестве второго конъюнкта референтную именную группу (16).

- (10) Ни я, ни мама о тебе так не **думаем**. (НКРЯ. А.Н. Бузулукский. Пальчиков // Волга. 2014)
- (11) Содержание документа столь чудовищно, что ни я, ни мои друзья не **верят**, что это действительно официальная бумага, написанная с Вашего ведома. (НКРЯ. К.И. Чуковский. Дневник (1922))
- (12) Ни я, ни Влас Михайлович услуг не **забывают**. (НКРЯ. Дон Аминадо. Поезд на третьем пути (1954))
- (13) А писал я приятелю о том, что ни я, ни кто-либо из моих однокашников писателем, конечно, не **станет**, писатель – это нечто высшее. (НКРЯ. Г.Я. Бакланов. Жизнь, подаренная дважды (1999))
- (14) Ни я, ни кто другой ничего не **записывает**. (НКРЯ. Анатолий Жигулин. Черные камни (1988))
- (15) Этого человека ни я, ни сам черт не **знает**, – сказал он. (НКРЯ. Ф.В. Булгарин. Иван Иванович Выжигин (1829))
- (16) Ни я, ни, тем более, Дарвин этого не **утверждает**, – сказал Протасов. (НКРЯ. В.Я. Шишков. Угрюм-река. Ч. 1–4 (1928–1933))

В контекстах с числовым согласованием при обычном порядке слов согласование по единственному числу также возникает в случаях, когда вторым конъюнктом является неопределенное местоимение (17); только в одном найденном нами примере на подобную стратегию согласования вторым конъюнктом выступает референтная именная группа (однако неодушевленная, в некоторых работах утверждается, что согласованию по единственному числу способствует неодушевленность конъюнктов [Санников, 2008, с. 157; Пекелис, 2013b]) (18).

- (17) Во-первых, ни я, ни кто-либо еще из чиновников не **собирался** выстраивать постоянно действующий механизм контроля за СМИ и влияния на их работу. (НКРЯ. Светлана Бабаева. Михаил Лесин: «Террористы имели четкий медиа-план» // Известия. 31.10.2002)

(18) <...> благо ни я, ни машина не **пострадала** <...> (НКРЯ. Учимся водить (2007–2008))

При обратном порядке слов – VS – возможно как согласование по правилам разрешения (19), так и согласование с ближайшим конъюнктом (20) (в контекстах с числовым согласованием – согласование по единственному числу (21)).

(19) Но почему об этом ничего не **знаем** ни я, ни отец?
(НКРЯ. Евгений Велтистов. Электроник – мальчик из чемодана (1964))

(20) А если это изобильное счастье и достойная демократическая жизнь для простых людей скажется в таком отдаления, что до него не **дотянусь** ни я, ни Маринка, а? (НКРЯ. Сергей Есин. Стоящая в дверях // Наш современник. 1992)

(21) Ни в четверг, ни в пятницу мы Радде Юльевне не позвонили.
Не позвонил ни я, ни Рита. (НКРЯ. Юрий Трифонов. Предварительные итоги (1970))

Сравнение полученных данных для исследуемых конструкций с союзами *и ... и* и *ни ... ни* позволяет заключить, что согласование с ближайшим конъюнктом при обычном порядке слов для конструкций с *ни ... ни* действительно проявляется в большей степени, что соотносится с выводами корпусного исследования [Пекелис, 2013а]. Согласование по правилам разрешения и согласование по 3 лицу множественного числа являются общими стратегиями для обоих типов сочиненных подлежащих.

3. Экспериментальное исследование

Таким образом, корпусное исследование выявило некоторые различия между конструкциями с *и ... и* и *ни ... ни*: в конструкциях с *ни ... ни* чаще встречаются случаи согласования с ближайшим конъюнктом при порядке слов SV по сравнению с конструкциями с *и ... и*. Данное наблюдение предполагает, что конструкции с союзом *ни ... ни* могут быть по своим согласовательным свойствам ближе к конструкциям с разделительными союзами. Для проверки этого предположения было проведено экспериментальное исследование.

В этом разделе представлены гипотезы исследования, сформулированные на основе корпусных данных и результатов предыдущих исследований, дизайн эксперимента, а также информация об участниках эксперимента и полученные экспериментальные данные.

3.1. Гипотезы

Эксперимент, представленный в данной работе, основан на методике оценки приемлемости предложений по шкале Ликерта от 1 (полностью неприемлемое предложение) до 7 (полностью приемлемое предложение). На основе корпусных данных и предыдущих исследований можно выдвинуть следующие гипотезы.

Ожидается, что согласование по правилам разрешения – по 1 лицу множественного числа – будет оцениваться наиболее высоко. В исследованиях по лично-числовому согласованию конструкций с союзами *и; и ... и; или и или ... или* при порядке слов SV эта стратегия получала устойчивые высокие оценки [Белова, Давидюк, 2023; Давидюк, 2024]. Более того, в корпусных данных именно эта стратегия является наиболее частотной. Кроме того, нормативные грамматики и справочники по русскому языку рекомендуют использование стратегии согласования по правилам разрешения в ситуациях конфликта значений признака лица у конъюнктов [РГ, 1980, с. 243–244; Розенталь и др., 1994, с. 272]. Что касается типа союза, то в корпусном материале количество примеров согласования по правилам разрешения для подлежащих с союзом *ни ... ни* более чем в два раза превышает количество аналогичных примеров с союзом *и ... и*. Исходя из этих данных, можно предположить, что согласование по правилам разрешения будет оцениваться выше в конструкциях с союзом *ни ... ни* по сравнению с конструкциями с союзом *и ... и*. С другой стороны, в экспериментальном исследовании согласования с сочинительными конструкциями с союзами *sowohl ... als auch* ‘как ... так и’ и *weder ... noch* ‘ни ... ни’ в немецком языке наблюдалось некоторое падение оценок приемлемости для конструкций с союзом ‘ни ... ни’ [Felser, Jessen, 2021] (см. разд. 1); однако это исследование не фокусировалось на случаях конфликта значений признака лица у конъюнктов.

Помимо согласования по правилам разрешения, ожидается, что будут приемлемы и другие стратегии согласования.

Согласование по 3 лицу множественного числа было обнаружено в корпусном материале. Кроме того, эта стратегия согласования оценивалась как приемлемая в экспериментах с аналогичными конструкциями с союзом *и* [Белова, Давидюк, 2023]. Следовательно, эта стратегия, вероятно, будет обнаружена и в рамках нашего эксперимента.

Как показывают корпусные данные, конструкции с союзом *ни ... ни* при порядке слов SV могут демонстрировать согласование с ближайшим конъюнктом. Таким образом, можно ожидать приемлемость этого варианта согласования в рамках экспериментального исследования.

3.2. Дизайн

Эксперимент имеет дизайн 2×4 . Первая независимая переменная – тип союза – имеет два уровня: *и ... и / ни ... ни*. Сравнение конструкций с *ни ... ни* с конструкциями с *и ... и* мотивировано тем, что, как отмечается в разделе 1, союз *ни ... ни* считается отрицательным «аналогом» союза *и ... и*. Вторая независимая переменная – это тип согласования, имеющая четыре уровня: 1 лицо единственного числа, 1 лицо множественного числа, 3 лицо единственного числа и 3 лицо множественного числа.

Согласование по 1 лицу множественного числа соответствует согласованиям по правилам разрешения, а согласование по 3 лицу единственного числа – стратегии согласования с ближайшим конъюнктом; вероятность приемлемости этих двух стратегий согласования входит в гипотезы эксперимента (см. разд. 3.1). Кроме того, предполагается, что согласование по 3 лицу множественного числа также окажется среди приемлемых стратегий. Стратегия согласования с первым конъюнктом (по 1 лицу единственного числа) также включена в эксперимент, т.к. ее возможность была отмечена в экспериментальных исследованиях на материале других славянских языков, и, следовательно, она не может быть исключена. В стимульных предложениях был зафиксирован порядок слов (SV) и порядок конъюнктов. Первым конъюнктом выступало личное местоимение *я*, вторым – имя собственное мужского рода. Тип второго конъюнкта был зафиксирован, т.к. этот фактор отмечается важным для согласования [Санников, 2008, с. 157; Пекелис, 2013b]. Для обеспечения единообразия стимулов в предложения с сочиненными подлежащими с союзом *и ... и* было добавлено отрицание. Пример одной из лексикализаций, использованных в эксперименте, приведен в (22).

(22) а. И я, и Максим не [хожу/ходим/ходит/ходят] на местный рынок.

б. Ни я, ни Максим не [хожу/ходим/ходит/ходят] на местный рынок.

На каждое из экспериментальных условий приходилось по четыре лексикализации. Соотношение экспериментальных предложений и филлеров составляло 1 : 1. Перед началом эксперимента участникам демонстрировались четыре тренировочных предложения. Таким образом, суммарно эксперимент содержал 68 предложений. В середине эксперимента участникам была предоставлена возможность сделать небольшой перерыв.

Филлеры делились на грамматичные и неграмматичные. Грамматичные филлеры состояли из предложений, в которых с помощью союзов *и* и *и ... и* были сочинены имена собственные мужского рода. При этом половина предложений содержала отрицание (*Кеша и Миша не проживают по данному адресу. И Артём, и Егор не занимаются в тренажерном зале.*), а другая половина – нет (*Вася и Юра гуляют в городском парке. И Саша, и Дима уезжают в другой город.*). Отрицание было добавлено для того, чтобы понять, как его наличие, не связанное с отрицательным конкордом (как в случае с предложениями, содержащими группы с союзом *ни ... ни*), влияет на оценки приемлемости: в частности, известно, что наличие отрицания усложняет восприятие предложений [Каур et al., 2007]. После четырех конкретных грамматичных филлеров следовали вопросы на понимание, предназначенные для выявления невнимательных респондентов. Неграмматичные предложения строились по аналогичной модели, но содержали ошибку в предикативном согласовании (*И Митя, и Боря загораем на городском пляже. Яша и Максим не гоняю на гоночном велосипеде.*).

Экспериментальные листы были составлены по правилу латинского квадрата, а экспериментальные и филлерные предложения чередовались по следующей схеме: экспериментальное предложение – филлер – экспериментальное предложение – филлер и т.д.

3.3. Участники

Эксперимент был реализован на платформе PCIBex [Zehr, Schwarz, 2018]. Набор респондентов осуществлялся через социальные сети и через платформу TaskPay (taskpay.ru). Второй источник привлек большое количество недобросовестных респондентов. Исключение таких респондентов основывалось на следующих метриках:

- а) значительные отклонения от ожидаемых оценок для грамматичных и неграмматичных филлеров по сравнению с другими участниками;
- б) наличие нескольких быстрых ответов (менее 300 миллисекунд);
- в) пропуск более трех предложений;
- г) неправильные ответы на половину или более контрольных вопросов [Герасимова, 2021].

В эксперименте приняло участие 157 человек, но данные 93 респондентов пришлось исключить из анализа на основании указанных критериев.

Средний возраст участников, чьи ответы были проанализированы, составляет 30 лет ($sd = 7,8$). Гендерное распределение участников следующее: 37 женщин, 27 мужчин.

3.4. Результаты

Для проведения статистического анализа все оценки были нормализованы, чтобы устранить проблему вариативности в использовании шкалы участниками. В качестве основного метода статистического анализа использовались смешанные линейные модели и попарные сравнения с помощью теста Тьюки. Максимальная модель была построена с использованием пакета *buildmer* [Voeten, 2023], а фиксированные эффекты были сокращены вручную.

На рисунке 1 представлен график взаимодействия, показывающий результаты эксперимента.

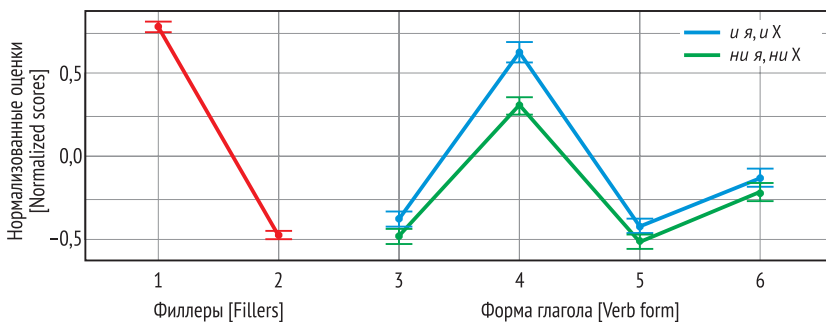


Рис. 1. Результаты эксперимента (график взаимодействия):

1 – грамматические; 2 – неграмматические; 3 – 1 л. ед. ч.; 4 – 1 л. мн. ч.;
5 – 3 л. ед. ч.; 6 – 3 л. мн. ч.

Fig. 1. Experimental results (interaction graph):

1 – grammatical; 2 – ungrammatical; 3 – 1st person singular;
4 – 1st person plural; 5 – 3rd person singular; 6 – 1st person plural

Формула итоговой линейной модели для результатов эксперимента приведена в (23). Хотя фактор типа союза оказался значимым ($\beta = 0,15$, $SE = 0,03688$, $df = 233$, $t\text{-ratio} = 4,068$, $p\text{-value} < 0,0001$), приемлемые стратегии согласования для обоих типов союзов совпадают. Они включают в себя не только согласование по правилам разрешения, но и согласование по 3 лицу множественного числа, которое получило значимо более высокие оценки по сравнению с неграмматичными филлерами ($t\text{-тест}$: для конструкций с *и ... и* и неграмматичных филлеров: $t\text{-ratio} = 4,531$, $df = 373,55$, $p\text{-value} < 0,0001$; для конструкций с *ни ... ни* и неграмматичных филлеров: $t\text{-ratio} = 5,7387$, $df = 357,2$, $p\text{-value} < 0,0001$). Оценки для согласования по 1 лицу единственного числа и 3 лицу единственного числа не показали значимой разницы

при сравнении с оценками для неграмматичных филлеров (*t*-тест для согласования по 1 лицу единственного числа: для конструкций с *и ... и* и неграмматичных филлеров: $t\text{-ratio} = -0,21379$, $df = 413,59$, $p\text{-value} = 0,8308$; для конструкций с *ни ... ни* и неграмматичных филлеров: $t\text{-ratio} = 1,7933$, $df = 395,56$, $p\text{-value} = 0,07369$; *t*-тест для согласования по 3 лицу единственного числа: для конструкций с *и ... и* и неграмматичных филлеров: $t\text{-ratio} = -0,77296$, $df = 419,79$, $p\text{-value} = 0,44$; для конструкций с *ни ... ни* и неграмматичных филлеров: $t\text{-ratio} = 0,99102$, $df = 398,86$, $p\text{-value} = 0,3223$).

(23) оценки $\sim 1 + \text{согласование} + \text{союз} + (1 \mid \text{номер предложения}) + (1 \mid \text{номер респондента})$

Конструкции с союзом *ни ... ни* получили значимо более высокие оценки по сравнению с конструкциями с союзом *и ... и* (данные попарных сравнений Тьюки для всех четырех вариантов согласования: $\beta = -0,15$, $SE = 0,0369$, $df = 233$, $t\text{-ratio} = -4,068$, $p\text{-value} = 0,0017$). Это может быть связано с наличием в предложениях с *и ... и* отрицания, которое, в отличие от конструкций с союзом *ни ... ни*, не мотивировано какими-либо грамматическими требованиями (в частности, отрицательным конкордом).

Как показывает сравнение грамматичных филлеров с подлежащими, сочиненными союзами *и* и *и ... и*, с отрицанием и без, грамматичные филлеры с отрицанием оцениваются значимо ниже грамматичных филлеров без отрицания (*t*-тест: $t\text{-ratio} = -3,6026$, $df = 1014,6$, $p\text{-value} = 0,0003304$). Это может свидетельствовать о том, что добавление приглагольного отрицания могло внести дополнительное усложнение в интерпретацию предложений, что сказалось и на оценках приемлемости. Возможно, более удачным решением для дизайна эксперимента было бы добавление элемента, требующего отрицательное согласование, как в предложениях с *и ... и*, так и в предложениях с *ни ... ни* (например, наречия *никогда*).

4. Обсуждение

Большинство гипотез, выдвинутых нами в разделе 3.2, подтвердились. Согласование по правилам разрешения оказывается наиболее приемлемой стратегией согласования. Как предполагалось на основе корпусных данных, эта стратегия получает более высокие оценки для конструкций с *ни ... ни*. Однако мы полагаем, что это может быть связано с присутствием отрицания в предложениях с союзом *и ... и*, которое, как в случае с предложениями с *ни ... ни*, не связано с требованием отрицательного согласования, что могло вызвать снижение оценок, т.к. наличие отрицания усложняет восприятие предложений. Отметим, что

собранные корпусные предложения с союзом *и ... и* с порядком слов SV, использовавшимся в экспериментальных предложениях (см. разд. 2), не содержат отрицания.

Среди приемлемых стратегий также оказалось согласование по 3 лицу множественного числа. Поскольку эта стратегия фиксируется даже в случаях, когда ни один из конъюнктов не несет признак 3 лица [Белова, 2024] (для сочиненных подлежащих вида *я и ты / ты и я*), можно предположить, что значение 3 лица в этом случае является дефолтным значением для признака лица.

Не подтвердилась гипотеза о возможности согласования с ближайшим конъюнктом при сочиненных подлежащих с союзом *ни ... ни*. Хотя этот вариант согласования встречается в корпусном материале, в экспериментальном исследовании он был оценен на уровне неграмматичных филлеров.

Однако следует отметить, что большинство случаев согласования с ближайшим конъюнктом при конструкциях с *ни ... ни* вызваны тем, что ближайший конъюнкт является неререферентной именной группой (см. разд. 2), такие случаи оказались вне проведенного экспериментального исследования. Предпочтительность согласования с ближайшим конъюнктом в таких случаях можно объяснить, опираясь на идею, предложенную в работе [Willer-Gold et al., 2018], согласно которой, если некоторый признак не имеет интерпретации, то его зондирование может быть отложено и произойти после линеаризации (в цитируемой работе предполагается, что признак рода не имеет интерпретации у неодушевленных существительных). Можно предположить, что неререферентная именная группа обладает некоторыми неинтерпретируемыми признаками, что приводит к тому, что согласование с такими сочиненными подлежащими происходит на этапе постсинтаксической обработки, что, в свою очередь, позволяет согласованию произойти с линейно ближайшим конъюнктом. Тем не менее, возможность согласования с ближайшим конъюнктом для конструкций с *ни ... ни* при порядке слов SV на данном этапе исследования не может быть полностью исключена, т.к. необходимо провести аналогичное исследование для конструкций, у которых ближайшим конъюнктом окажется не именная группа 3 лица, как в нашем исследовании, а личное местоимение, которое, возможно, будет более сильным аттрактором согласования.

5. Заключение

В данной работе представлено корпусное и экспериментальное исследование согласования с подлежащим, сочиненным при помощи союза *ни ... ни*, первым конъюнктом которого является личное местоимение *я*.

Корпусное исследование выявило возможность согласования с ближайшим конъюнктом при порядке слов SV для исследуемых конструкций с *ни ... ни*.

Однако в эксперименте не было выявлено значимых различий в стратегиях согласования между конструкциями с *ни ... ни* и конструкциями с *и ... и*. Эксперимент подтвердил приемлемость только двух стратегий согласования: согласования по правилам разрешения и согласования по 3 лицу множественного числа, которые также были зафиксированы в корпусе. Единственное различие между этими типами конструкций заключается в более низких оценках всех вариантов согласования для конструкций с *и ... и*, что, вероятно, связано с присутствием отрицания в этих экспериментальных предложениях.

В корпусном материале стратегия согласования с ближайшим конъюнктом для конструкций с *ни ... ни* характерна для случаев, когда ближайший конъюнкт является нереферентной именной группой, однако такие случаи остались за рамками экспериментального исследования. Случаи согласования с ближайшим конъюнктом, зафиксированные в корпусе, могут быть объяснены наличием у нереферентных конъюнктов неинтерпретируемых признаков, что приводит к откладыванию согласования на постсинтаксический компонент *и*, как следствие, к согласованию с ближайшим конъюнктом.

Дальнейшие исследования согласования с конструкциями с *ни ... ни* могут быть направлены на изучение влияния различного порядка слов, а также на анализ числового согласования (корпусные данные по числовому согласованию и порядку слов VS уже рассмотрены в статье, но экспериментального исследования по этим вопросам не проводилось).

Библиографический список / References

Белова, Давидюк, 2023 – Белова Д.Д., Давидюк Т.И. Согласование с сочиненным подлежащим, содержащим личное местоимение: экспериментальное исследование на материале русского языка // Рема. Rhema. 2023. № 2. С. 53–88. DOI: 10.31862/2500-2953-2023-2-53-88 [Belova D.D., Davidyuk T.I. Agreement with coordinated subjects containing a personal pronoun: Experimental data from Russian. *Rhema*. 2023. No. 2. Pp. 53–88. (In Rus.). DOI: 10.31862/2500-2953-2023-2-53-88]

Белова, 2024 – Белова Д.Д. Предикативное согласование с местоименными конъюнктами: опыт экспериментального исследования // Труды Института русского языка им. В.В. Виноградова. 2024. № 2. С. 62–72. [Belova D.D. Predicate agreement with pronominal conjuncts: An experimental research. *Proceedings of the V.V. Vinogradov Russian Language Institute*. 2024. No. 2. Pp. 62–72. (In Rus.)]

Герасимова, 2021 – Герасимова А.А. Учебные материалы практикума по экспериментальному синтаксису. Отбор респондентов. На правах рукописи. 2021. URL: https://agerasimova.com/wp-content/uploads/Gerasimova_Practice_Outliers.pdf

[Gerasimova A.A. Uchebnye materialy praktikuma po eksperimentalnomu sintaksisu. Otkor respondentov [Teaching materials on experimental syntax. Respondent selection]. Manuscript copyright. 2021. URL: https://agerasimova.com/wp-content/uploads/Gerasimova_Practice_Outliers.pdf]

Граудина и др., 2008 – Граудина Л.К., Ицкович В.А., Катлинская Л.П. Словарь грамматических вариантов русского языка. 3-е изд. М., 2008. [Graudina L.K., Itskovich V.A., Katlinskaya L.P. Slovar grammaticheskikh variantov russkogo yazyka [Dictionary of grammatical variants of the Russian language]. Moscow, 2008.]

Гриценко, 2024 – Гриценко И.О. Предикативное согласование с сочинительными конструкциями, содержащими синтаксически и семантически разнородные конъюнкты. Доклад, представленный на учебной конференции по экспериментальной лингвистике «Экспериментальные исследования языка», НИВЦ, МГУ имени М.В. Ломоносова, 13–14 июня 2024 г. [Gritsenko I.O. Predikativnoe soglasovanie s sochinitelnyimi konstruksiyami, soderzhashchimi sintaksicheskimi i isemanticheskimi raznorodnymi konyunkty [Predicative agreement with coordinative constructions containing syntactically and semantically heterogeneous conjuncts]. Talk presented at the student conference on experimental linguistics “Experimental Language Research”, Research Computing Center, Lomonosov Moscow State University, June 13–14, 2024.]

Давидюк, 2024 – Давидюк Т.И. Лично-числовое согласование с подлежащими, сочиненными одиночными и повторяющимися союзами, в русском языке // Вестник Московского университета. Серия 9. Филология. 2024. № 4. С. 58–70. [Davidyuk T.I. Person-number agreement with subjects containing coordinating and correlative conjunctions in Russian. *Lomonosov Philology Journal. Series 9. Philology*. 2024. No. 4. Pp. 58–70. (In Rus.)]

Пекелис, 2013а – Пекелис О.Е. «Частичное согласование» в конструкции с повторяющимся союзом: корпусное исследование основных закономерностей // Вопросы языкознания. 2013. № 4. С. 55–86. [Pekelis O.E. “Partial agreement” in the design with a repeated union: A corpus study of the main patterns. *Voprosy yazykoznaniya*. 2013. No. 4. Pp. 55–86. (In Rus.)]

Пекелис, 2013b – Пекелис О.Е. Сочинение. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru/>). На правах рукописи. М., 2013. [Pekelis O.Ye. Sochinenie. Materialy dlya proekta korpusnogo opisaniya russkoy grammatiki (<http://rusgram.ru/>) [Coordination. Materials for the project of corpus description of Russian grammar (<http://rusgram.ru/>)]. As a manuscript. Moscow, 2013.]

РГ, 1980 – Русская грамматика / под ред. Н.Ю. Шведовой. М., 1980. Т. 2. [Russkaja grammatika [Russian grammar]. N.Yu. Shvedova (ed.). Vol. 2. Moscow, 1980.]

Розенталь и др., 1994 – Розенталь Д.Э., Джанджакова Е.В., Кабанова Н.П. Справочник по правописанию, произношению, литературному редактированию. М., 1994. [Rozenal D.E., Dzhandzhakova E.V., Kabanova N.P. Spravochnik po pravopisaniyu, proiznosheniyu, literaturnomu redaktirovaniyu [A guide to spelling, pronunciation, literary editing]. Moscow, 1994.]

Савчук и др., 2024 – Савчук С.О., Архангельский Т.А., Бонч-Осмоловская А.А. Национальный корпус русского языка 2.0: новые возможности и перспективы развития // Вопросы языкознания. 2024. № 2. С. 7–34. [Savchuk S.O., Arkhangel'sky T.A., Bonch-Osmolovskaya. A.A. Russian National Corpus 2.0:

New possibilities and development prospects. *Voprosy yazykoznaniiya*. 2024. No. 2. Pp. 7–34. (In Rus.)]

Санников, 2008 – Санников В.З. Русский синтаксис в семантико-прагматическом пространстве. М., 2008. [Sannikov V.Z. Russkij sintaksis v semantiko-pragmaticheskom prostranstve [Russian syntax in the view of semantics and pragmatics]. Moscow, 2008.]

Adamson, Anagnostopoulou, in press – Adamson L., Anagnostopoulou E. Gender features and coordination resolution in Greek and other three-gendered languages implications for the cross-linguistic representation of gender. *Linguistic Inquiry*. In press.

Baker, 2011 – Baker M. When agreement is for number and gender but not person. *Natural Language & Linguistic Theory*. 2011. Vol. 29. Pp. 875–915.

Corbett, 1983 – Corbett G. Resolution rules: Agreement in person, number, and gender. *Order, Concord and Constituency*. G. Gazdar, E. Klein, G.K. Pullam (eds.). Dordrecht, 1983. Pp. 175–205.

Corbett, 2006 – Corbett G. Agreement. Cambridge University Press, 2006.

Felser, Jessen, 2021 – Felser C., Jessen A. Correlative coordination and variable subject–verb agreement in German. *Languages*. 2021. Vol. 6 (67). Pp. 1–20.

Haspelmath, 2004 – Haspelmath M. Explaining the ditransitive person-role constraint: A usage-based approach. *Constructions*. 2004. Vol. 2. Pp. 1–49.

Himmelreich, Hartmann, 2023 – Himmelreich A., Hartmann K. Agreement with disjoined subjects in German. *Glossa: A Journal of General Linguistics*. 2023. Vol. 8. Pp. 1–44.

Kaup et al., 2007 – Kaup B., Lüdtke J., Zwaan R.A. The experiential view of language comprehension: How is negation represented? *Higher Level Language Processes in the Brain. Inference and Comprehension Processes*. F. Schmalhofer, C.A. Perfetti (eds.). Lawrence Erlbaum Associates, 2007. Pp. 255–288.

Nevins, Weisser, 2019 – Nevins A., Weisser P. Closest conjunct agreement. *Annual Review of Linguistics*. 2019. No. 5. Pp. 219–241.

Peterson, 1986 – Peterson P.G. Establishing verb agreement with disjunctively conjoined subjects: Strategies vs principles. *Australian Journal of Linguistics*. 1986. Vol. 6. Pp. 231–250.

Timmermans et al., 2004 – Timmermans M., Schriefers H., Dijkstra T., Haverkort M. Disagreement on agreement: Person agreement between coordinated subjects and verbs in Dutch and German. *Linguistics*. 2004. No. 42. Pp. 905–929.

Voeten, 2023 – Voeten C.C. buildmer: Stepwise elimination and term reordering for mixed-effects regression. R package, version 2.11. 2023.

Willer-Gold et al., 2018 – Willer-Gold J., Arsenijević B., Batinić M. et al. When linearity prevails over hierarchy in syntax. *Proceedings of the National Academy of Sciences*. 2018. Vol. 115. No. 3. Pp. 495–500.

Zehr, Schwarz, 2018 – Zehr J., Schwarz F. PennController for Internet based experiments (IBEX). 2018. DOI: 10.17605/OSF.IO/MD832

Zwicky, 1977 – Zwicky A. Hierarchies of person. *Papers from the Thirteenth Regional Meeting of the Chicago Linguistic Society*. W.A. Beach, S.E. Fox, S. Philosph (eds.). Chicago, 1977. Pp. 714–733.

Статья поступила в редакцию 19.08.2024
The article was received on 19.08.2024

Сведения об авторе / About the author

Давидюк Татьяна Игоревна – аспирант кафедры теоретической и прикладной лингвистики филологического факультета, Московский государственный университет имени М.В. Ломоносова; программист Научно-исследовательского вычислительного центра, Московский государственный университет имени М.В. Ломоносова; младший научный сотрудник Научного центра по сохранению, возрождению и документации языков России, Институт языкознания РАН, г. Москва

Tatiana I. Davidyuk – PhD student at the Department of Theoretical and Applied Linguistics of the Philological Faculty, Lomonosov Moscow State University; specialist at the Research Computing Center, Lomonosov Moscow State University; junior researcher at the Research Center for the Preservation, Revitalization and Documentation of the Languages of Russia, Institute of Linguistics RAS, Moscow, Russian Federation

ORCID: <https://orcid.org/0000-0003-2276-2439>

E-mail: rachekit@yandex.ru

DOI: 10.31862/2500-2953-2024-3-54-73

УДК 81-114.4

А.В. Прокопьева

Московский государственный университет имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация

Согласование предиката по числу с сочинительными конструкциями с союзами *как ... так и* и *не только ... но и*: экспериментальное исследование

В статье описываются результаты экспериментального исследования, посвященного оценке приемлемости стратегий полного и частичного согласования предиката по числу с сочиненным подлежащим, содержащим союзы *как ... так и* и *не только ... но и*. Согласно сведениям из грамматического описания русского языка и данным экспериментального исследования предложений с союзом *sowohl ... als auch* 'как ... так и' на материале немецкого языка, ожидается, что более предпочтительно согласование по множественному числу. Однако исследование с использованием данных Национального корпуса русского языка демонстрирует, что возможны случаи, когда частичное согласование употребляется чаще. В экспериментальном исследовании проверяется предположение о том, что на вариативность оказывает влияние фактор порядка слов и семантические свойства самих союзов. Результаты показали, что для двух типов предложений (для стимулов с союзом *не только ... но и* и порядком, при котором предикат следует за подлежащим, а также тех, что содержат союз *как ... так и* и порядок, при котором предикат предшествует подлежащему), нет значимых различий в оценке приемлемости в зависимости от числа предиката. Напротив, среди стимулов с порядком, при котором предикат следует за подлежащим, и союзом *как ... так и* значимо более высокие оценки получают предложения с предикатом во множественном числе. Значимая разница также обнаружена в оценках стимулов с порядком, при котором предикат предшествует подлежащему, и союзом *не только ... но и*: предложения с единственным числом предиката оцениваются выше. Данные

© Прокопьева А.В., 2024

Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License



эксперимента можно использовать для дальнейшего анализа порядка слов как фактора, способного повлиять на выбор стратегии предикативного согласования, а также для расширения описания семантики и структуры союзов *как ... так и* и *не только ... но и*.

Ключевые слова: предикативное согласование, сочиненное подлежащее, двухместный союз, порядок слов, экспериментальный синтаксис, корпусное исследование, русский язык

Для ЦИТИРОВАНИЯ: Прокопьева А.В. Согласование предиката по числу с сочинительными конструкциями с союзами *как ... так и* и *не только ... но и*: экспериментальное исследование // Рема. Rhema. 2024. № 3. С. 54–73. DOI: 10.31862/2500-2953-2024-3-54-73

DOI: 10.31862/2500-2953-2024-3-54-73

A. Prokopeva

Lomonosov Moscow State University,
Moscow, 119991, Russian Federation

Predicative number agreement with coordinated subjects containing *kak ... tak i* ‘both ... and’ and *ne tol’ko ... no i* ‘not only ... but also’: An experimental study

This paper presents the results of an experimental study on predicative number agreement with coordinated subjects featuring the conjunctions *kak ... tak i* ‘both ... and’ and *ne tol’ko ... no i* ‘not only ... but also’. According to grammars and data from an experimental study of the German conjunction *sowohl ... als auch* (‘both ... and’), we expect that plural agreement is more preferable. However, analysis of data from the Russian National Corpus reveals instances where partial agreement appears to be more frequent. We propose that the variability is influenced by factors such as word order and the semantic properties of the conjunctions themselves. To test this hypothesis we conducted an experiment using the acceptability judgement task (Likert scale 1–7). The results indicate that, for sentences with SV word order and the conjunction *ne tol’ko ... no i*, there is no significant

difference in the acceptability ratings between singular and plural agreement. This finding is similarly observed for stimuli with VS word order and the conjunction *kak ... tak i*. In contrast, among sentences featuring SV word order and subjects coordinated by *kak ... tak i* stimuli with plural predicates were rated significantly higher. In a different context, when examining stimuli with *ne tol'ko ... no i* and VS word order, sentences with singular agreement had higher acceptability ratings. The experimental data contribute to a deeper understanding of word order as a factor influencing the choice of agreement strategy. Furthermore, the findings can enrich the linguistic description of the semantics and structure of the conjunctions *kak ... tak i* and *ne tol'ko ... no i*.

Key words: predicative agreement, coordinated subject, correlative conjunctions, word order, experimental syntax, corpus study, the Russian language

FOR CITATION: Prokopeva A. Predicative number agreement with coordinated subjects containing *kak ... tak i* 'both ... and' and *ne tol'ko ... no i* 'not only ... but also': An experimental study. *Rhema*. 2024. No. 3. Pp. 54–73. (In Rus.) DOI: 10.31862/2500-2953-2024-3-54-73

1. Введение

В ситуации, когда в качестве подлежащего в предложении выступает сочинительная конструкция, в русском языке наблюдаются две возможные стратегии согласования предиката по числу. Первая стратегия, которая получила название стандартной или же полной, предполагает согласование предиката по множественному числу (1). Вторая стратегия заключается в том, что в качестве контролера согласования выступает только один из конъюнктов в составе подлежащего, часто наиболее линейно близкий к предикату [Пешковский, 1928, с. 451]; такое согласование называется частичным (2).

- (1) Текст и эпилог зависят от параметров встречи. (НКРЯ. Карпов В.Э., Мещерякова Т.В. Об автоматизации нетворческих литературных процессов // Информационные технологии. 2004)
- (2) Среди методов гидробиологического анализа экологического состояния водных объектов сапробиологический анализ и биоиндикация занимает одно из важнейших мест. (НКРЯ. Экологическое состояние р. Москвы на территории Раменского района Московской области // Геоинформатика. 2004)

В данной статье рассматривается вопрос о предикативном согласовании по числу с сочинительными конструкциями с союзами *как ... так и* и *не только ... но и*, относящимися к группе двухместных союзов, которые ранее редко рассматривались на предмет вариативности согласования такого рода. Они обладают особыми семантическими и структурными свойствами: в частности, интерес для исследователей предикативного согласования может представлять семантическое неравноправие конъюнктов, т.к. нельзя исключить вероятность, что разная выделенность конъюнктов [Апресян, Пекелис, 2012] оказывает влияние на выбор согласовательной стратегии. Также на выбор стратегии согласования может оказывать влияние порядок слов: так, предполагается, что препозиция глагола способствует отдельной интерпретации конъюнктов [Пекелис, 2013б], что влечет за собой большую приемлемость частичного согласования.

Для получения новых эмпирических данных используются экспериментальные методы, которые ранее применялись при рассмотрении вариативности в согласовании предиката с сочиненным подлежащим, содержащим одиночный союз *и* [Белова, Давидюк, 2023; Паско, 2023] и повторяющийся союз *и ... и* [Врубель, 2023]. Результаты экспериментального исследования сопоставляются с предварительными выводами, сделанными на основе Национального корпуса русского языка (НКРЯ).

2. Коррелятивы, или двухместные союзы

В разделе 2.1 описываются отличительные признаки *как ... так и* и *не только ... но и*, а также сходных с ними союзов, на основании которых они выделяются в особую группу двухместных союзов. Особое внимание уделяется семантическим свойствам двухместных союзов, на основе которых выдвигается основной вопрос исследования: какая стратегия согласования предиката является более предпочтительной в зависимости от семантики конкретного союза и порядка слов в предложении?

В разделе 2.2 приводятся существующие описания предикативного согласования с сочиненными подлежащими, включающими двухместные союзы.

2.1. Особенности союзов *как ... так и* и *не только ... но и*; проблема выбора термина

При рассмотрении союзов *как ... так и* и *не только ... но и* необходимо упомянуть, что они относятся к отдельному типу союзов, имеющему как структурные, так и семантические особенности. В работе [Haspelmath, 2007] упоминается, что помимо обычного сочинения (*A and B*)

выделяется эмфатическое сочинение (*both A and B*), при котором каждый конъюнкт рассматривается отдельно. Как правило, языки Европы выражают такой тип сочинения бисиндетически (англ. *bisyndetic coordination*), т.е. используют союз, состоящий из двух единиц. Две единицы, образующие подобную конструкцию (нем. *sowohl ... als auch* 'как ... так и', польск. *jak ... tak (i)* 'как ... так и'), согласно [Haspelmath, 2007], называют коррелятами (англ. *correlative coordinators*), поскольку как минимум одна из этих единиц не может употребляться в отсутствие другой.

В русской грамматической традиции, однако, термин «корреляты» охватывает скорее показатели, которые сочетаются с подчинительными союзами [Пекелис, 2018], поэтому возникает необходимость использовать иное обозначение. В [РГ, 1980] группу союзов, образованных коррелятами, к которой относятся как ... *так и* и *не только ... но и*, называют двухместными союзными соединениями; они отличаются от остальных неодноместных союзов (*и ... и ... и*; *то ... то ... то*) тем, что представляют собой соединения двух формально не совпадающих и позиционно разобщенных элементов. Другой используемый термин – «двойные союзы» – появляется, к примеру, в работе В.Ю. Апресян и О.Е. Пекелис (2012). Там же указывается, что двойным союзам свойственна общая семантика – указание на то, что «вторая часть утверждения, вводимая второй частью союза, с точки зрения Говорящего, неожиданна для Адресата или несколько менее очевидна» [Апресян, Пекелис, 2012, п. 3.2]. При этом первая и вторая части могут одинаково соответствовать действительности, или же вторая часть может соответствовать действительности в большей степени – т.н. градационная семантика, которая сближает союзы этой группы с противительными союзами. Градационная семантика с указанием на существенное отклонение второго конъюнкта от нормы свойственна союзу *не только ... но и*, а в значении союза *как ... так и* допустимо не выделять «градационность» [Санников, 2008]. То есть помимо общей семантики в значении союзов есть и важное различие, которое в дальнейшем может повлиять на выбор стратегии согласования.

На основе приведенных выше описаний двухместных союзов можно высказать предположение, что конъюнкты, соединяемые союзами из этой группы, имеют разную степень выделенности; это, вероятно, оказывает влияние на выбор согласовательной стратегии, в частности на возможность согласования с ближайшим конъюнктом. В свою очередь, ближайший конъюнкт определяется взаимным расположением субъекта и предиката, т.е. фактор порядка слов также необходимо

принимать во внимание. Целью исследования становится поиск данных, которые позволили бы ответить на вопрос, как согласование по числу определяется взаимодействием семантических свойств союза и расположением предиката относительно субъекта.

2.2. Согласование предиката по числу: данные грамматик и предыдущих исследований

Согласно [РГ, 1980], двухместные союзы оформляют закрытые сочинительные ряды. Утверждается, что если закрытый ряд словоформ, объединяющий формы единственного и множественного или только единственного числа, занимает место подлежащего, то возможна вариативность (3), однако предпочитается форма множественного числа сказуемого; она же считается единственно верной при порядке SV¹ (4).

(3) К берегу бежали/бежал Ваня и Коля.

(4) Станция и мост сгорели.

Экспериментальные исследования согласования предиката с подлежащим, выраженным сочинительной конструкцией, включающей корреляты / двухместные союзы, немногочисленны. В работе [Felser, Jessen, 2021] рассматриваются союзы *sowohl ... als auch* 'как ... так и', *weder ... noch* 'ни ... ни' и *entweder ... oder* 'или ... или', из которых нас больше остальных интересует *sowohl ... als auch*.

Первый эксперимент предполагал оценку приемлемости предложений и содержал две независимые переменные, не считая тип союза: число предиката и порядок конъюнктов в сочинительной конструкции, где один конъюнкт был в единственном числе, а другой во множественном; второй фактор был назван «близостью», т.к. указывал на совпадение или несовпадение предиката в числе с ближайшим конъюнктом. Также были добавлены два стимула, содержавшие оба конъюнкта в единственном числе. Предикат всегда располагался в постпозиции. В результате было обнаружено, что стимулы с предикатом во множественном числе для всех рассмотренных условий оценивались выше, чем соответствующие предложения с глаголом в единственном числе. Также появление конъюнкта в единственном числе на втором месте повышало приемлемость согласования по единственному числу.

¹ Условные обозначения: SV – порядок слов, при котором предикат следует за подлежащим; VS – порядок слов, при котором предикат предшествует подлежащему; PL – множественное число; SG – единственное число.

В другом эксперименте респондентам предлагалось самостоятельно выбрать нужную форму глагола для завершения предложения, и снова во всех комбинациях конъюнктов (SG-PL – PL-SG – SG-SG) большинство ответов содержали предикат во множественном числе.

Третий эксперимент показал, что при использовании метода категориальной оценки стимула с предикатом во множественном числе и союзом *sowohl ... als auch* более 90% суждений свидетельствуют о приемлемости предложений такого типа, в то время как у стимулов с предикатом в единственном числе максимальная доля положительных суждений составляет 17%, и она достигается для предложений со вторым – ближайшим – конъюнктом также в единственном числе.

Данное исследование демонстрирует, что в немецком языке при подлежащем с союзом *sowohl ... als auch* и порядке SV более предпочтительно полное согласование по множественному числу, хотя близкое расположение конъюнкта в единственном числе может несколько повысить приемлемость частичного согласования.

Сопоставляя результаты исследования [Felser, Jessen, 2021] с суждениями из [РГ, 1980], можно проследить сходство в предпочтении формы множественного числа предиката при порядке SV, которое мы можем ожидать в предложениях с союзами *как ... так и* и *не только ... но и*. Однако также необходимо иметь в виду и не упомянутую в грамматике возможность согласования по единственному числу, которая актуальна для предложений, где ближайшим к предикату является конъюнкт в форме единственного числа.

Поскольку имеющиеся данные описывают выбор стратегии согласования в зависимости от типа союза и порядка слов в предложении, представляется эффективным проверить предварительные выводы о влиянии этих факторов, сделанные на основе [РГ, 1980] и эксперимента [Felser, Jessen, 2021], с использованием НКРЯ. На его материале подсчитывается количество предложений для каждого интересующего нас условия (*не только ... но и*, SV; *не только ... но и*, VS; *как ... так и*, SV; *как ... так и*, VS), где реализуются различные согласовательные стратегии. Полученные таким образом результаты позволяют определить, насколько достоверно существующие исследования описывают данные современного русского языка. Далее выводы на основе корпусных данных можно использовать в качестве отправной точки для формулирования гипотез экспериментального исследования, чтобы проверить их на более единообразных и систематизированных примерах и получить отрицательный языковой материал.

3. Корпусное исследование согласования предиката с сочиненным подлежащим

Исследование осуществляется на материале основного подкорпуса НКРЯ.

Поисковый запрос включает в себя сочинительную конструкцию с исследуемыми союзами и глагол, порядок слов при этом варьируется.

Необходимо отметить, что поиск примеров, включающих сочиненное подлежащее с союзами *как ... так и* и *не только ... но и*, затруднен из-за неопределенности расстояния между частями двойного союза, поэтому в данном исследовании учтены исключительно те предложения, где конъюнктами являются существительные в форме единственного числа без зависимых. Мы не учитываем тип предиката, также на данном этапе игнорируется семантика конъюнктов.

В ходе обработки полученных данных удаляются примеры, где предикатом является глагольная форма *есть*, т.к. форма множественного числа соответствующего глагола *суть* признается устаревшей.

Результаты по состоянию на 12 марта 2024 г. представлены в таблице 1.

Таблица 1

**Предикативное согласование по числу
с сочиненным подлежащим вида «не только X, но и Y»
и «как X, так и Y» (X, Y – существительные в ед. ч.):
данные Национального корпуса русского языка
[Predicate number agreement with a coordinated subject
of the form “not only X but also Y” and “both X and Y”
(where X and Y stand for a singular noun):
Russian National Corpus data]**

Союз [Conjunction]	<i>не только ... но и</i> [not only ... but also]		<i>как ... так и</i> [both ... and]	
	SV	VS	SV	VS
Порядок слов [Word order]	SV	VS	SV	VS
ед. ч. [SG]	54	248	2	31
мн. ч. [PL]	18	15	130	20

Согласно сведениям из грамматик(и) и данным предыдущих исследований, мы ожидаем, что носители будут предпочитать согласование предиката по множественному числу во всех рассмотренных случаях, однако в корпусе для каждого условия находятся примеры,

демонстрирующие частичное согласование по единственному числу с ближайшим конъюнктом, см. (5)–(8).

(5) *не только ... но и, SV*

При этом теперь не только территория, но и культура рассматривается как ресурс развития национального государства. (НКРЯ. А.Г. Санина. Информационное общество и государственная идентичность // Информационное общество. 2013)

(6) *как ... так и, SV*

И не дивно, что у древних римлян как должность, так и благодеяние означалось сим словом – officium. (НКРЯ. Григорий Сковорода. Разговор, называемый Алфавит или Букварь мира (1760–1775))

(7) *не только ... но и, VS*

Интересно, что под горячую руку попал не только мат, но и жаргон. (НКРЯ. Ксения Туркова. Словарный запас. Итоги 2016 года (31.12.2016) // Сноб. 2016)

(8) *как ... так и, VS*

Во внимание принимается как содержание, так и доступность различных услуг. (НКРЯ. Акаткин Ю.М., Ясиновская Е.Д. Современное развитие электронного правительства России в контексте мировых тенденций // Информационное общество. 2017)

Стоит дополнительно упомянуть, что оба примера согласования предиката по единственному числу для порядка *SV* и союза *как ... так и* датированы серединой XVIII в. Если ограничить область поиска подкорпусом, содержащим исключительно тексты, созданные с 1900 г., будут получены данные, которые вкратце представлены в таблице 2 (по состоянию на 8 августа 2024 г.)

Уменьшается общее число предложений, но на примере таблице 2 можно заметить те же закономерности, что и на примере таблице 1. Так, стоит обратить внимание на то, что при порядке слов *VS* для обоих союзов выше число примеров с предикатом в единственном числе, т.е. наблюдается эффект частичного согласования по числу с ближайшим конъюнктом. Данное явление объясняется тем, что препозиция глагола повышает приемлемость частичного согласования с первым конъюнктом [Пешковский, 1928].

При порядке слов SV соотношение количества найденных предложений с предикатом в единственном и во множественном числе зависит от союза в составе сочиненного подлежащего.

Для союза *как ... так и* примеры с предикатом в единственном числе крайне немногочисленны и встречаются только в текстах середины XVIII в. Напротив, среди предложений с сочиненным подлежащим, включающим союз *не только ... но и*, примеры, которые содержат предикат в единственном числе, количественно преобладают. Подобное различие между союзами может объясняться их семантическими характеристиками.

Как было отмечено ранее, предположительно, союз *не только ... но и*, в отличие от *как ... так и*, обладает градационной семантикой. Это может отразиться в большей выделенности второго конъюнкта и, следовательно, большей приемлемости частичного согласования в предложениях с союзом *не только ... но и*.

Таблица 2

**Предикативное согласование по числу
с сочиненным подлежащим вида «не только X, но и Y»
и «как X, так и Y» (X, Y – существительные в ед. ч.):
данные Национального корпуса русского языка по текстам,
созданным не ранее 1900 г.**

**[Predicate number agreement with a coordinated subject
of the form “not only X but also Y” and “both X and Y”
(where X and Y stand for a singular noun):**

Russian National Corpus data, texts written not earlier than 1900]

Союз [Conjunction]	<i>не только ... но и</i> [not only ... but also]		<i>как ... так и</i> [both ... and]	
Порядок слов [Word order]	SV	VS	SV	VS
ед. ч. [SG]	40	203	0	28
мн. ч. [PL]	14	11	80	15

Корпусные данные позволяют выявить некоторые различия в частотности примеров с согласованием предиката по единственному или множественному числу в зависимости от выбора союза и порядка слов. Однако для того, чтобы сделать выводы на основании более единообразного материала, где минимизировано влияние посторонних факторов, мы воспользовались экспериментальными методами.

4. Экспериментальное исследование согласования предиката с сочиненным подлежащим

Чтобы выявить, какая стратегия согласования предиката по числу является более предпочтительной в предложениях с подлежащими, выраженными сочинительными конструкциями с союзами *как ... так и* и *не только ... но и*, мы провели экспериментальное исследование.

4.1. Дизайн и реализация эксперимента

Для проведения экспериментального исследования была выбрана методика оценки приемлемости предложений по шкале Ликерта от 1 до 7, использовался дизайн $2 \times 2 \times 2$. Независимыми переменными выступали:

- 1) союз (*как ... так и / не только ... но и*);
- 2) порядок слов (VS/SV);
- 3) число предиката (ед. ч. / мн. ч.)

Всего эксперимент включал восемь условий, которые требовались, чтобы проверить три гипотезы, выдвинутые на основании результатов корпусного исследования.

Гипотеза № 1: если предикат располагается в препозиции (порядок VS), для обоих союзов наиболее предпочтительной является стратегия частичного согласования по единственному числу с ближайшим конъюнктом. Пример (9) лучше, чем (10).

(9) На первую пару пришел не только Дима, но и Егор.

(10) На первую пару пришли не только Дима, но и Егор.

Гипотеза № 2: если предикат располагается в постпозиции (порядок SV), то при подлежащем, выраженным сочинительной конструкцией с союзом *не только ... но и*, более приемлемой оказывается стратегия частичного согласования по единственному числу с ближайшим – вторым – конъюнктом. Пример (11) лучше, чем (12).

(11) Не только Дима, но и Егор пришел на первую пару.

(12) Не только Дима, но и Егор пришли на первую пару.

Гипотеза № 3: если предикат располагается в постпозиции (порядок SV), то при подлежащем, выраженным сочинительной конструкцией с союзом *как ... так и*, более предпочтительна стратегия согласования по множественному числу. Пример (13) лучше, чем (14).

(13) Как Дима, так и Егор пришли на первую пару.

(14) Как Дима, так и Егор пришел на первую пару.

Подлежащим в стимульных предложениях являлась сочинительная конструкция, где конъюнктами были одушевленные имена собственные мужского рода в форме единственного числа. Предикатом служил глагол совершенного вида в прошедшем времени, переходность/непереходность глагола была сбалансирована. В зависимости от переходности глагола в позиции предиката каждое предложение включало либо прямое дополнение, либо предложную группу с пространственным значением. Также в начале стимулов присутствовало обстоятельство времени. Пример экспериментального блока представлен в таблице 3.

Таблица 3

Пример экспериментального блока
[Stimuli example]

Союз [Conjunction]	Порядок слов [Word order]	Число предиката [Predicate number]	Стимул [Stimulus]
<i>не только ... но и</i> [not only ... but also]	SV	мн. ч. [PL]	В понедельник не только Артём, но и Максим пришли на первую пару.
		ед. ч. [SG]	В понедельник не только Артём, но и Максим пришел на первую пару.
	VS	мн. ч. [PL]	В понедельник на первую пару пришли не только Артём, но и Максим.
		ед. ч. [SG]	В понедельник на первую пару пришел не только Артём, но и Максим.
<i>как ... так и</i> [both ... and]	SV	мн. ч. [PL]	В понедельник как Артём, так и Максим пришли на первую пару.
		ед. ч. [SG]	В понедельник как Артём, так и Максим пришел на первую пару.
	VS	мн. ч. [PL]	В понедельник на первую пару пришли как Артём, так и Максим.
		ед. ч. [SG]	В понедельник на первую пару пришел как Артём, так и Максим.

В материал эксперимента также вошли филлеры, которые включали сочиненную именную группу не в позиции подлежащего (15)–(16). Помимо грамматичных мы создали неграмматичные филлеры, содержащие ошибку в числе предиката (17) или в падеже существительного в предложной группе (18). Часть филлеров выступала в роли тренировочных предложений.

- (15) После занятий преподаватели похвалили не только Толю, но и Валеру.
- (16) Не только в Саше, но и в Гене фанаты разочаровались после проигрыша.
- (17) Из-за скандала ответ возмутили не только Витю, но и Володю.
- (18) Не только Антонова, но и Семёнова следователь выслушал в полицию.

Экспериментальный лист включал 32 стимула (по 4 предложения на каждое условие), 32 филлера и 4 тренировочных предложения, стимулы распределялись по восьми экспериментальным листам согласно правилу латинского квадрата.

К четырем грамматичным филлерам предлагались вопросы с тремя вариантами ответа, которые использовались для того, чтобы проверить уровень внимательности респондента при прохождении эксперимента.

Эксперимент был реализован с использованием платформы PCibex [Zehr, Schwarz, 2018], набор респондентов осуществлялся посредством социальных сетей.

Нормализацию полученных оценок приемлемости и дальнейшую обработку данных мы проводили в программе RStudio [R Core Team, 2024].

4.2. Результаты экспериментального исследования

Всего в ходе исследования были получены 107 ответов от респондентов, их средний возраст был равен 26,9 лет, $SD = 12,41$.

Однако далее некоторые данные были исключены из дальнейшего анализа на основании следующих метрик:

- а) сумма квадратов разностей между оценками, которые респондент выставлял филлерам и тренировочным предложениям, и установленным эталоном, равным 2 для неграмматичных предложений и 6 для грамматичных, более чем на два стандартных отклонения отличалась от среднего значения данной метрики по всем респондентам;
- б) респондент ошибся в ответе на 3 или 4 контрольных вопроса;
- в) большое число стимулов пропущено [Герасимова, 2021].

В дальнейшем анализе использовались ответы 86 респондентов.

Полученные данные были обработаны с использованием смешанных линейных моделей в пакете lmerTest [Kuznetsova, 2017] и попарных сравнений Тьюки. Оптимальной мы признали следующую модель: $zscores \sim 1 + \text{Порядок_слов} + \text{Союз} + \text{Порядок_слов}:\text{Союз} + \text{Число_предиката} + \text{Союз}:\text{Число_предиката} + \text{Порядок_слов}:\text{Число_предиката} + (1 + \text{Число_предиката} + \text{Порядок_слов} + \text{Порядок_слов}:\text{Число_предиката} \mid \text{респондент}) + (1 \mid \text{предложение})$.

Данная модель включает факторы порядка слов, используемого союза и числа предиката, а также все их попарные взаимодействия в качестве фиксированных эффектов, случайный отрезок и угол наклона для факторов порядка слов, числа предиката и их взаимодействия в зависимости от респондента и, помимо этого, случайный отрезок в зависимости от стимульного предложения.

Значимыми факторами, влияющими на оценку приемлемости, оказываются число предиката ($\beta = -0,77$, $SE = 0,09$, $t = -8,87$, $p\text{-value} < 0,0001$), а также взаимодействие числа предиката и типа союза ($\beta = 0,74$, $SE = 0,06$, $t = 11,60$, $p\text{-value} < 0,0001$), числа предиката и порядка слов ($\beta = 0,67$, $SE = 0,08$, $t = 8,11$, $p\text{-value} < 0,0001$). Кроме того, значимым признается и взаимодействие порядка слов и типа союза ($\beta = -0,20$, $SE = 0,06$, $t = -3,18$, $p\text{-value} = 0,002$), и отдельное влияние фактора типа союза ($\beta = -0,11$, $SE = 0,06$, $t = -2,01$, $p\text{-value} = 0,045$).

Выводы о результатах исследования мы делаем с опорой на диаграмму взаимодействия факторов на рисунке 1.

Первая гипотеза, выдвинутая перед началом экспериментального исследования, касалась влияния фактора порядка слов: предполагалось, что вне зависимости от используемого союза при порядке слов VS более высоко будут оцениваться стимулы с согласованием предиката по единственному числу.

Данное предположение подтверждается только для стимулов с союзом *не только ... но и*: среднее значение оценок приемлемости таких предложений значимо выше, чем у стимулов, содержащих предикат во множественном числе (критерий Тьюки, $\beta = -0,64$, $SE = 0,08$, $t = -8,34$, $p\text{-value} < 0,0001$). Более того, это единственная группа стимулов, оценки которых находятся на уровне грамматических филлеров.

Ситуация иного рода наблюдается при анализе предложений с союзом *как ... так и*, где значимых различий в оценках между стимулами с предикатом в форме единственного и множественного числа обнаружено не было (критерий Тьюки, $\beta = 0,10$, $SE = 0,08$, $t = 1,29$, $p\text{-value} = 0,90$).

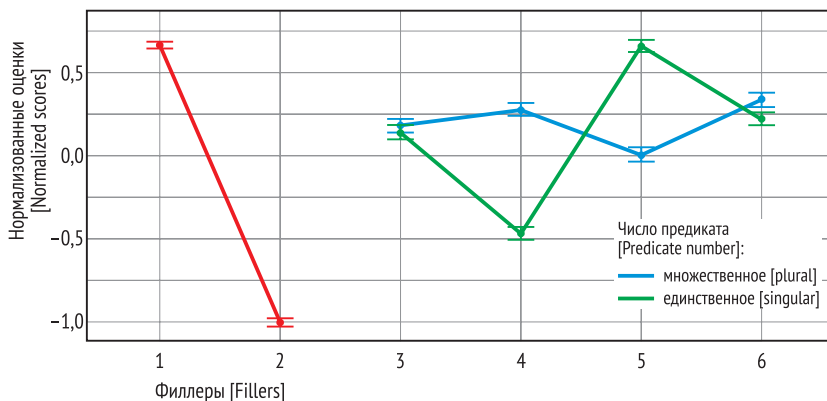


Рис. 1. Диаграмма взаимодействия факторов:

1 – грамматичные; 2 – неграмматичные; 3 – *не только ... но и, SV*;
4 – *как ... так и, SV*; 5 – *не только ... но и, VS*; 6 – *как ... так и, VS*

Fig. 1. Interaction plot

1 – grammatical; 2 – ungrammatical; 3 – *не только ... но и, SV*;
4 – *как ... так и, SV*; 5 – *не только ... но и, VS*; 6 – *как ... так и, VS*

Вторая гипотеза заключалась в предположении, что среди стимулов с союзом *не только ... но и* и порядком *SV* более высокие оценки получают те, что содержат предикат в единственном числе. Однако мы не получили значимых различий (критерий Тьюки, $\beta = 0,03$, $SE = 0,09$, $t = 0,33$, $p\text{-value} = 1,00$). Таким образом, вторая гипотеза не подтверждается.

Третья гипотеза затрагивала предложения с союзом *как ... так и* и порядком слов *SV*, она позволяла ожидать более высокие оценки у стимулов с предикатом во множественном числе. Действительно, приемлемость предложений с глаголом во множественном числе оценивается значимо выше, чем у стимулов с глаголом в единственном числе (критерий Тьюки, $\beta = 0,77$, $SE = 0,09$, $t = 8,87$, $p\text{-value} < 0,0001$). Третья гипотеза подтверждается данными эксперимента.

Экспериментальное исследование также позволяет высказать предположение, что приемлемость стимулов с предикатом во множественном числе оценивается на одном уровне за исключением единственной пары условий (табл. 4). Это исключение состоит в значимых различиях в оценках приемлемости предложений с порядком *VS*: те стимулы, что содержали союз *не только ... но и*, были оценены ниже, чем стимулы с тем же согласованием по множественному числу, но с союзом *как ... так и* (критерий Тьюки, $\beta = 0,31$, $SE = 0,06$, $t = 5,70$, $p\text{-value} < 0,0001$).

Для примеров с порядком VS и союзом *не только ... но и* при этом более приемлемой является стратегия частичного согласования, как было отмечено ранее.

Таблица 4

Результаты применения критерия Тьюки для сравнения стимулов с предикатом во множественном числе
[The results of the application of the Tukey criterion for comparing stimuli with a predicate in the plural]

Условия [Conditions]	Результат [Result]
как ... так и, SV, PL – не только ... но и, SV, PL	$\beta = 0,11$, SE = 0,06, $t = 2,01$, p -value = 0,4754
как ... так и, SV, PL – как ... так и, VS, PL	$\beta = -0,03$, SE = 0,06, $t = -0,54$, p -value = 0,9994
как ... так и, SV, PL – не только ... но и, VS, PL	$\beta = 0,28$, SE = 0,07, $t = 3,96$, p -value = 0,0026
не только ... но и, SV, PL – как ... так и, VS, PL	$\beta = -0,15$, SE = 0,07, $t = -2,05$, p -value = 0,4485
не только ... но и, SV, PL – не только ... но и, VS, PL	$\beta = 0,17$, SE = 0,06, $t = 2,67$, p -value = 0,1391
как ... так и, VS, PL – не только ... но и, VS, PL	$\beta = 0,31$, SE = 0,06, $t = 5,70$, p -value < 0,0001

Приемлемость стимулов с предикатом в единственном числе варьирует в зависимости от союза и порядка слов, причем изменение порядка с SV на VS значительно повышает оценки для предложений с частичным согласованием, как для союза *не только ... но и* (критерий Тьюки, $\beta = -0,50$, SE = 0,07, $t = -7,61$, p -value < 0,0001), так и для союза *как ... так и* (критерий Тьюки, $\beta = -0,70$, SE = 0,07, $t = -10,69$, p -value < 0,0001). Также стимулы с союзом *не только ... но и* и единственным числом предиката оцениваются выше, чем примеры с *как ... так и*, тем же порядком слов и той же стратегией согласования: это наблюдается и для условий с порядком SV (критерий Тьюки, $\beta = -0,63$, SE = 0,06, $t = -11,38$, p -value < 0,0001), и для условий с порядком VS (критерий Тьюки, $\beta = -0,42$, SE = 0,06, $t = -7,71$, p -value < 0,0001). Это наблюдение частично подтверждает выдвинутое перед началом исследования предположение, что порядок VS и союз *не только ... но и* могут повышать приемлемость согласования по единственному числу, но демонстрирует более сложное взаимодействие факторов, чем ожидалось.

5. Обсуждение

Только одна гипотеза из трех, выдвинутых перед началом экспериментального исследования, получила непосредственное подтверждение при анализе результатов. Гипотезы были основаны на корпусных данных: это служит свидетельством того, что результаты корпусных исследований являются хорошим основанием для дальнейшей работы, но после их проверки в ходе эксперимента, где контролируются способные повлиять на оценку грамматические факторы и анализируется отрицательный языковой материал, выводы могут быть переосмыслены.

К примеру, результаты, показывающие большую приемлемость согласования по множественному числу для примеров с союзом *как ... так и* и порядком *SV*, а также согласования по единственному числу в предложениях с союзом *не только ... но и* и порядком *VS*, дополняют корпусные данные, где примеры, соответствующие описанным условиям, встречаются в большем количестве, чем предложения со сходной структурой, но другой формой числа предиката. Напротив, исследованию на материале НКРЯ несколько противоречат выводы эксперимента для стимулов с союзом *как ... так и* и порядком *VS*, а также *не только ... но и* и порядком *SV*, где не возникает ожидаемое предпочтение в приемлемости стратегии частичного согласования.

Кроме того, экспериментальное исследование не подтвердило предположение о том, что для двухместных союзов и порядка *SV* единственно приемлема форма множественного числа предиката, чего можно было бы ожидать, используя как источник сведения из грамматики [РГ, 1980] и исследование схожих союзов в немецком языке [Felser, Jessen, 2021]. Таким образом, комбинируя данные исследований с применением различных методик, мы наблюдаем более комплексную картину и получаем эмпирические данные, потенциально полезные для дальнейших теоретических рассуждений.

Согласно результатам эксперимента, стандартное, или полное, согласование можно назвать приемлемой стратегией для обоих рассмотренных союзов, вне зависимости от порядка слов, поскольку даже для условия *не только ... но и*, *VS*, где подобные стимулы получили самые низкие оценки, приемлемость не обнаруживает значимых отличий от двух из трех других возможных условий, включающих предикат во множественном числе, и значимо отличается от уровня неграмматичных филлеров. Однако частичное согласование также является возможной стратегией, приемлемость которой значимо варьируется в зависимости от порядка слов и типа союза, основанного на семантике. Данные позволяют предположить, что оба этих фактора являются слабыми; эта

идея ранее высказывалась в работе [Пекелис, 2013а]. В экспериментальном исследовании они влияют на приемлемость частичного согласования, повышая его, если наблюдаемый порядок – VS или союз – *не только ... но и*, т.е. когда в стимуле присутствует хотя бы один из двух факторов, способствующих отдельному рассмотрению конъюнктов. При этом возможность стандартного согласования, как было отмечено выше, не блокируется полностью.

Экспериментальное исследование подтверждает, что между союзами *как ... так и* и *не только ... но и* имеются различия, связанные в первую очередь с приемлемостью при них частичного согласования. Это сложно объяснить наличием одной лишь градационной семантики у союза *не только ... но и*, поскольку она способствовала бы большей выделенности второго конъюнкта в силу значения «нарастания» [РГ, 1980], но в полученных результатах значимо более приемлемыми формы единственного числа оказываются только при порядке VS, т.е. когда ближайший конъюнкт, напротив, является менее выделенным. Вероятно, стоит говорить не конкретно про компонент градации в значении союза, а про возможность отдельной интерпретации существительных в составе сочинительной конструкции, которая может быть обусловлена не только семантическими, но и структурными свойствами союза. Суть таких различий между двухместными союзами может стать темой дальнейших исследований.

Библиографический список / References

Апресян, Пекелис, 2012 – Апресян В.Ю., Пекелис О.Е. Сочинительные союзы. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru/>). На правах рукописи. М., 2013. [Apresyan V.Yu., Pekelis O.E. Sochinitelnye soyuzu. Materialy dlya proekta korpusnogo opisaniya russkoi grammatiki (<http://rusgram.ru/>)] [Coordinating conjunctions. Materials for the project of corpus description of Russian grammar (<http://rusgram.ru/>)]. As a manuscript. Moscow, 2012.]

Белова, Давидюк, 2023 – Белова Д.Д., Давидюк Т.И. Согласование с сочиненным подлежащим, содержащим личное местоимение: экспериментальное исследование на материале русского языка // Рема. Rhema. 2023. № 2. С. 53–88. DOI: 10.31862/2500-2953-2023-2-53-88 [Belova D.D., Davidyuk T.I. Agreement with coordinated subjects containing a personal pronoun: Experimental data from Russian. *Rhema*. 2023. No. 2. Pp. 53–88. (In Rus.). DOI: 10.31862/2500-2953-2023-2-53-88]

Врубель, 2023 – Врубель Д.Д. Эффект синкретизма при предикативном согласовании с сочинительными конструкциями с повторяющимся союзом *и* // Рема. Rhema. 2023. № 2. С. 104–118. DOI: 10.31862/2500-2953-2023-2-104-118 [Vrubel D.D. The role of syncretism in the predicate agreement with coordinate constructions with the Russian correlative conjunction *i...i*. *Rhema*. 2023. No. 2. Pp. 104–118. (In Rus.). DOI: 10.31862/2500-2953-2023-2-104-118]

Герасимова, 2021 – Герасимова А.А. Учебные материалы практикума по экспериментальному синтаксису. Отбор респондентов. 2021. URL: https://agerasimova.com/wp-content/uploads/Gerasimova_Practice_Outliers.pdf (дата обращения: 14.05.2024) [Gerasimova A.A. Uchebnye materialy praktikuma po ehkspierimentalnomu sintaksisu. Othor respondentov [Experimental syntax. Practical course. Selecting respondents]. 2021. URL: https://agerasimova.com/wp-content/uploads/Gerasimova_Practice_Outliers.pdf]

Паско, 2023 – Паско Л.И. Против АТВ-анализа частичного согласования в русском языке: экспериментальное исследование // Рема. Rhema. 2023. № 2. С. 89–103. DOI: 10.31862/2500-2953-2023-2-89-103 [Pasko L.I. Against АТВ-analysis of partial agreement in Russian: An experimental study. *Rhema*. 2023. No. 2. Pp. 89–103. (In Rus.). DOI: 10.31862/2500-2953-2023-2-89-103]

Пекелис, 2013а – Пекелис О.Е. Сочинение. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru/>). На правах рукописи. М., 2013. [Pekelis O.Ye. Sochinenie. Materialy dlya proekta korpusnogo opisaniya russkoj grammatiki (<http://rusgram.ru/>) [Coordination. Materials for the project of corpus description of Russian grammar (<http://rusgram.ru/>)]. As a manuscript. Moscow, 2013.]

Пекелис, 2013б – Пекелис О.Е. «Частичное согласование» в конструкции с повторяющимся союзом: корпусное исследование основных закономерностей // Вопросы языкознания. 2013. № 4. С. 55–86. [Pekelis O.E. “Partial agreement” in the design with a repeated union: A corpus study of the main patterns. *Voprosy yazykoznanija*. 2013. No. 4. Pp. 55–86. (In Rus.)]

Пекелис, 2018 – Пекелис О.Е. Корреляты. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru/>). На правах рукописи. М., 2018. [Pekelis O.E. Korrelyaty. Materialy dlya proekta korpusnogo opisaniya russkoj grammatiki (<http://rusgram.ru/>)]. [Correlative pronouns. Materials for the project of corpus description of Russian grammar (<http://rusgram.ru/>)]. As a manuscript. Moscow, 2018.]

Пешковский, 1928/2001 – Пешковский А.М. Русский синтаксис в научном освещении. 8-е изд. (1-е изд. 1928). М., 2001. [Peshkovskii A.M. Russkii sintaksis v nauchnom osveshchenii [Russian syntax in a scientific light]. Moscow, 2001.]

РГ, 1980 – Русская грамматика / под ред. Н.Ю. Шведовой. М., 1980. Т. 2. [Russkaja grammatika [Russian grammar]. N.Yu. Shvedova (ed.). Vol. 2. Moscow, 1980.]

Санников, 2008 – Санников В.З. Русский синтаксис в семантико-прагматическом пространстве. М., 2008. [Sannikov V.Z. Russkij sintaksis v semantiko-pragmaticheskom prostranstve [Russian syntax in the view of semantics and pragmatics]. Moscow, 2008.]

Felser, Jessen, 2021 – Felser C., Jessen A. Correlative Coordination and Variable Subject–Verb Agreement in German. *Languages*. 2021. Vol. 6. P. 67.

Haspelmath, 2007 – Haspelmath M. Coordination. *Language typology and syntactic description*. Vol. II: Complex constructions. 2nd ed. T. Shopen (ed.). Cambridge University Press, 2007. Pp. 1–51.

Kuznetsova et al., 2017 – Kuznetsova A. lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*. A. Kuznetsova, P.B. Brockhoff, R.H.B. Christensen (eds.). 2017. Vol. 82. No. 13. Pp. 1–26.

R Core Team, 2024 – R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria, 2024.

Zehr, Schwarz, 2018 – Zehr J., Schwarz F. PennController for Internet based experiments (IBEX). 2018.

Статья поступила в редакцию 15.08.2024

The article was received on 15.08.2024

Сведения об авторе / About the author

Прокопьева Анна Вадимовна – студент бакалавриата отделения теоретической и прикладной лингвистики филологического факультета, Московский государственный университет имени М.В. Ломоносова

Anna V. Prokopyeva – BA student at the Department of Theoretical and Applied Linguistics of the Philological Faculty, Lomonosov Moscow State University, Moscow, Russian Federation

E-mail: anyaprokopyeva2004@gmail.com

DOI: 10.31862/2500-2953-2024-3-74-86

УДК 81-114.4

А.Г. Анисимова

Московский государственный университет имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация

Предикативное согласование по роду с сочинительными конструкциями с двойным заместительным союзом («X, а не Y»)

Считается, что в случае предикативного согласования по роду с сочинительными конструкциями с двойным заместительным союзом *а не* согласование определено однозначно: предикат согласуется с тем из конъюнктов, который обозначает «фактически действующий субъект». Однако в более поздних работах мы также находим мнение, что в таком случае происходит «столкновение» двух сильных правил согласования: грамматического и семантического. Семантическое диктует согласование с конъюнктом, который не находится под отрицанием, а грамматическое – с ближайшим из них. В данной работе мы исследовали вопрос о том, может ли изменение факторов порядка слов и наличия расстояния между подлежащим и сказуемым влиять на повышение приемлемости предложений, где сказуемое согласовано с конъюнктом, который не обозначает реально действующий субъект, но расположен ближе всего к предикату. Экспериментально мы проверили гипотезу о повышении оценок приемлемости для предложений с согласованием по грамматическому правилу при порядке слов, когда предикат следует за подлежащим, и наличии расстояния между подлежащим и сказуемым соответственно. В эксперименте респонденты оценивали приемлемость предложений по шкале Ликерта от 1 до 7. В ходе исследования заявленная гипотеза не подтвердилась. Стимулы с согласованием по семантическому правилу стабильно оцениваются выше, чем предложения с согласованием по грамматическому. Кроме того, не было найдено значимых отличий в оценках стимулов с согласованием с ближайшим



к предикату конъюнктом при разных порядках слов и наличии/отсутствии расстояния между подлежащим и сказуемым.

Ключевые слова: предикативное согласование, сочиненное подлежащее, заместительные союзы, экспериментальный синтаксис, русский язык

Для ЦИТИРОВАНИЯ: Анисимова А.Г. Предикативное согласование по роду с сочинительными конструкциями с двойным заместительным союзом («X, а не Y») // Рема. Rhema. 2024. № 3. С. 74–86. DOI: 10.31862/2500-2953-2024-3-74-86

DOI: 10.31862/2500-2953-2024-3-74-86

A. Anisimova

Lomonosov Moscow State University,
Moscow, 119991, Russian Federation

Predicate gender agreement with the coordinate constructions with a double substitution conjunction ("X, а не Y")

It is believed that in the case of predicate gender agreement with compositional constructions with double substitution conjunction *a ne* 'and not' the predicate agrees with the conjuncts that denote the "actually acting subject". However, in recent papers it is stated that there is rather a "collision" of two rules of agreement: grammatical and semantic. Semantic rule dictates the agreement with the non-negated conjunct and grammatical – with the nearest of them. In this paper, we examined whether the factors of word order and the presence of a distance between the subject and the predicate increase the acceptability of sentences with agreement with the conjunct, which does not denote a real acting subject, but is the closest one to the predicate. The main hypothesis was that the SV(O) word order and the distance between the subject and the predicate would raise the acceptability scores for the grammatical agreement. This experiment was conducted using 1–7 Likert scale. During the study, the stated hypothesis was not confirmed. Stimuli with semantic agreement are consistently rated

higher than the ones with grammatical agreement. In addition, no significant differences were found in the scores of stimuli where the predicate agrees with the conjunct closest to it for different word orders and the presence/absence of a distance between the subject and predicate.

Key words: predicate agreement, coordinated subject, substitutive conjunctions, experimental syntax, the Russian language

FOR CITATION: Anisimova A. Predicate gender agreement with the coordinate constructions with a double substitution conjunction (“X, a ne Y”). *Rhema*. 2024. No. 3. Pp. 74–86. (In Rus.). DOI: 10.31862/2500-2953-2024-3-74-86

1. Введение

Известно, что согласование сказуемого с подлежащим, представляющим собой сочиненную группу, подчиняется довольно сложным правилам, которые многие исследователи освещали в своих работах (согласование с сочиненной группой с повторяющимся союзом *и* [Врубель, 2023], с одиночным союзом *и* [Белова, Давидок, 2023]). В [Санников, 2008] они описываются как тенденции, в согласии с которыми носители согласовывают подлежащее со сказуемым. Первая из них – это, конечно, согласование по смыслу или семантическое правило. Например, при согласовании по числу, если сказуемому приписывается множественное число, то в ситуации принимают участие несколько объектов и т.п. Иногда этот инструмент передачи смысла оказывается единственным, как в (1), где в зависимости от числа глагола, меняется интерпретация всего предложения:

- (1) а. Здесь *работает* знаменитый писатель и врач [одно лицо].
б. Здесь *работают* знаменитый писатель и врач [два лица].
[Иомдин, 1979, с. 17]

Однако существует и альтернативная тенденция, которая предполагает согласование сказуемого с ближайшим к нему конъюнктом из сочиненного подлежащего (2). Эта закономерность проявлялась еще в древнерусском языке и была вполне обыкновенна, например, при порядке слов VS [Глинкина, 1978]. Далее в работе указанные тенденции будут упоминаться как семантическое и грамматическое правило соответственно.

- (2) Не только отец, но и мать *осталась*.
[Санников, 2008, с. 150]

Данное исследование посвящено проблеме предикативного согласования по роду с двойным заместительным союзом («X, а не Y»). По семантическому правилу согласование возможно только с неотрицаемым конъюнктом (3а), при порядке слов VS¹ вариант согласования только один – ближайший конъюнкт [РГ, 1980, с. 168] (3b), однако при порядке слов SV, когда ближе к предикату становится конъюнкт под отрицанием, может возникать вариативность (см. (3а), (3с)).

- (3) а. Роман, а не повесть будет опубликован в журнале.
[Розенталь, 1974, § 190]
- б. Опубликован роман, а не повесть.
[Розенталь, 1974, § 190]
- с. Самая реальная вещь в рассказе служанки это то что джун из фашистского диктаторского государства под перекрестным огнём и страхом смерти лучшая подруга, а не муж приехал забирать. (Telegram)

Целью работы было проверить экспериментально, каким образом носители русского языка разрешают подобные случаи и рассмотреть факторы, которые могут способствовать выбору «неканонического» пути при согласовании. Корпусное исследование в данном случае не смогло бы показать объективной и релевантной языковой картины, т.к. в письменных текстах подобная вариативность в согласовании чаще всего рассматривается как ошибка и исправляется при редактировании [Розенталь, 1974], однако в устной речи и условиях эксперимента, когда у участников нет возможности долго размышлять над ответом, есть шанс получить данные, показывающие реальное состояние языка и возможные тенденции к изменению нормы.

Стоит отметить, что, ко всему прочему, не все исследователи выделяют заместительные союзы, о которых пойдет речь в статье (*а не (не ... а); не ... но; если бы ... а то*), в отдельный класс, относя их к противительным [Розенталь, 1974, с. 239] или составным [РГ, 1980, § 2073].

Далее в статье будут рассмотрены аргументы за выделение заместительных союзов в отдельный класс, а затем особенности предикативного согласования с подлежащим с такими союзами, выдвинуты и экспериментально проверены предположения по поводу того, что может влиять на приемлемость той или иной стратегии согласования в таком случае.

¹ Условные обозначения: SV – порядок слов, при котором предикат следует за подлежащим; VS – порядок слов, при котором предикат предшествует подлежащему.

2. Заместительные союзы

В грамматике русского языка 1980 г. классификация сочинительных союзов не выделяет заместительные в отдельный класс. Только замечается, что подобные союзы являются составными, имеют несинтагматическую структуру, т.к. не включают в свой состав падежной словоформы и выражают «отрицательно-противительное значение взаимного исключения» [РГ, 1980, & 1677, 2077]. Похожего мнения придерживается Д.Э. Розенталь, определяя предложения типа *Роман, а не повесть будет опубликован в журнале* как предложения с противительным союзом [Розенталь, 1974, § 188].

Иная позиция по этому поводу у Санникова, который утверждает, что заместительные союзы стоит выделить из подкласса противительных, опираясь при этом на факты, некоторые из которых приведены ниже. Так, при определении значения того или иного союза и дальнейшей классификации в первую очередь используются семантические признаки отношения к реальности с точки зрения говорящего (реальность/нереальность/возможность) и противопоставленности/непротивопоставленности компонентов [Санников, 2008].

Сравним примеры (4а) и (4b). В обоих случаях действия ‘слушать’ и ‘есть’ действительно совершаются Машей в реальности. Значит, по этому признаку можно определить оба типа союзов в один класс – соединительные. Однако в (4b) можно выделить и значение противопоставленности компонентов, следовательно, этот критерий позволяет произвести деление уже внутри выделенной ранее группы, образуя подкласс противительных союзов.

- (4) а. Маша слушает и ест.
б. Маша слушает, но ест.
в. Маша слушает, а не ест.

В свою очередь, в примере (4с) в значении подобного союза есть отличия от группы соединительных и, соответственно, противительных. Событие ‘слушает’ – реально, т.е. имеет место, а ‘ест’ – нет. Таким образом, возникает необходимость выделения союзов с подобным набором признаков в отдельный класс.

Еще одно предположение, которое полемизирует с утверждением из [РГ, 1980] заключается в том, что в данном случае мы имеем дело с цельной единицей, а не союзом и отрицательной частицей, т.к. значение ее не выводится по принципу аддитивности, не равно простому объединению значений двух частей [Крейдли, 1979]. Этот факт хорошо виден при сравнении заместительного союза и соединительного, который как раз предполагает соединение значений всех частей для

получения семантики всего предложения. В (5a) мы можем выделить два параллельно происходящих события: ‘Маша не слушает’ и ‘Маша ест’, говорящий утверждает реальность каждого из них. Тогда как в (5b) утверждение говорящего лишь одно и его можно перефразировать как «Маша совершает действие; неверно, что это действие ‘есть’». Верно, что ‘слушать’».

- (5) а. Маша не слушает и ест.
 б. Маша слушает, а не ест.

Кроме того, в пользу теории о том, что *не* в заместительных союзах – это не отрицательная частица, говорит 17 различий между нейтральным и противопоставительным отрицанием, которые можно найти почти на всех грамматических уровнях, которые описал И.М. Богуславский [Богуславский, 1985]. В конце концов, принять мысль об объединении заместительных союзов с противительными в одну группу не дает еще и невозможность взаимной замены этих союзов в примерах из работы [Hurford, 1974]:

- (7) а. Я родился во Франции, но не в Париже.
 б. *Я родился во Франции, а не в Париже.
 [Ibid, p. 1, 4]

В свете приведенных доводов в пользу выделения союзов типа *а не* в отдельный класс заметим, что название «заместительные» используется для подчеркивания их смысла: из двух конъюнктом один заменяется другим, одна информация сменяется другой.

3. Предикативное согласование с двойным заместительным союзом

Несмотря на то, что многие исследователи сходятся на том, что при согласовании с подлежащим, выраженным конструкцией с союзом с замещающей семантикой, вариативность невозможна и предикат согласуется с тем из конъюнктов, который обозначает «фактически действующий субъект» [Розенталь, 1974], но при участии фактора порядка слов, сила которого варьирует от слабой до средней [Врубель, Паско, Студеникина, в печати], может появляться вариативность, которая, в свою очередь, обуславливает повышение приемлемости для предложений, где выбрана «неканоническая» стратегия согласования. В связи с этим логично предположить, что при порядке слов SV согласование с конъюнктом под отрицанием будет вероятнее, чем при порядке VS, т.к. он находится ближе к предикату.

(8) а. Пришел/*Пришла Петя, а не Маша.

б. Петя, а не Маша пришел²пришла.

Рассмотрим альтернативные факторы, способствующие согласованию с ближайшим конъюнктом. Так, в работе [Граудина и др., 1976] утверждается, что на согласование сказуемого с подлежащим, выраженным существительным *ряд*, может влиять их удаленность друг от друга. Это способствует согласованию сказуемого по множественному числу, что можно объяснить как раз правилом согласования с ближайшей к предикату частью подлежащего.

(9) Ряд делегатов от западных стран, в том числе Англии, США, *пытались* воспрепятствовать принятию антирасистских резолюций. (Правда, 15 октября 1974 г.) [цит. по: Граудина, Ицкович, Катлинская, 1976, с. 25]

Руководствуясь этим, мы предположили, что этот же фактор может давать дополнительный стимул для согласования по грамматическому правилу, которое диктует выбор ближайшего к предикату конъюнктура в качестве контролера и в случаях употребления двойного заместительного союза *а не*.

(10) Честно, сердце, а не голова объективно у меня тогда за рулем была. (Telegram)

(11) Командор, а не Анна, был предметом любви Дон Жуана. (НКРЯ. И.Н. Вирабов. Андрей Вознесенский (2015))

Принимая во внимание описанную выше вариативность, далее мы опишем экспериментальное исследование, проверяющее предположение о том, что упомянутые выше факторы будут увеличивать приемлемость предложений с непривычной (грамматической) стратегией согласования по роду с подлежащим, соединенным союзом *а не*, для носителей русского языка.

4. Экспериментальное исследование

Экспериментальное исследование, проведенное в рамках данной работы, проверяло, влияют ли на повышение приемлемости предложений с согласованием с ближайшим конъюнктом факторы порядка слов и расстояния между подлежащим и сказуемым.

Дизайн эксперимента включал в себя три независимые переменные: порядок слов (SV/VS), наличие расстояния между подлежащим и сказуемым (есть/нет), род предиката (мужской, женский). Размер дистанции между сочинительной конструкцией и предикатом во всех стимульных

предложениях составил 5 слов, 2–3 из которых знаменательные части речи, а остальные – служебные. Тип составляющей, которая создает эту дистанцию, выбран в качестве сбалансированной переменной (предложная группа (12a) / деепричастный оборот (12b)).

- (12) а. Навес, а не веранда со времен Второй мировой войны стоял/стояла на даче.
 б. Волк, а не лиса, приюхиваясь к деревьям на пути, пробегал/пробегала по лесу.

Кроме того, под отрицанием в стимулах всегда была именная группа женского рода, а без отрицания – мужского. Время предиката выбрано прошедшим, чтобы согласование по роду всегда могло быть выражено на предикате. Схема стимульного предложения представлена в таблице 1.

Таблица 1

Схема стимульного предложения [Stimuli scheme]

Подлежащее («X, а не Y») [Subject («X, а не 'and not' Y»)]	Глагол [Verb]	Глагольное дополнение [Verb adjunct]	Предложная группа / деепричастный оборот [Prepositional phrase / adverbial phrase]
<i>Навес, а не веранда</i>	<i>стоял/стояла</i>	<i>на даче</i>	<i>со времен Второй мировой войны</i>

Стимулы распределялись по листам по правилу латинского квадрата. Количество лексикализаций составило 32 предложения. Пример экспериментального блока можно увидеть в таблице 2. Отношение стимулов к филлерам 1 : 1. Филлеры делились на две половины: грамматичные и неграмматичные.

Грамматичные филлеры создавались путем переноса заместительного союза из подлежащего в объект/предикат (13a) или постановки обоих сочиняемых в подлежащем существительных во множественное число (13a), чтобы предикативное согласование по роду было невозможно. Неграмматичные филлеры содержали ошибки в предикативном согласовании по числу и лицу (14a) и в падежах существительных (14b). Филлеры чередовались со стимулами в случайном порядке.

- (13) а. Библиотекарь выдал читателю *не книгу, а журнал*, отметив это в его карточке.
 б. *Матери, а не отцы* на собраниях решают, какие шторы будут висеть в кабинете.

- (14) а. *Штатный стекольщик в компании *занимаются* починкой не сантехники, а стекла.
- б. *Торговки на рынке, заворачивая товар, *не улыбаются* о покупателе, а показывали язык.

Таблица 2

Пример экспериментального блока [Stimuli example]

Порядок слов [Word order]	Расстояние (депричастный оборот) [Distance (adverbial clause)]	Род предиката [Predicate gender]	Стимул [Stimulus]
SV	Есть	Мужской	<i>Навес, а не веранда со времен Второй мировой войны стояла на даче.</i>
SV	Есть	Женский	<i>Навес, а не веранда со времен Второй мировой войны стояла на даче.</i>
SV	Нет	Мужской	<i>Навес, а не веранда стояла на даче со времен Второй мировой войны.</i>
SV	Нет	Женский	<i>Навес, а не веранда стояла на даче со времен Второй мировой войны.</i>
VS	Есть	Мужской	<i>На даче стояла со времен Второй мировой войны навес, а не веранда.</i>
VS	Есть	Женский	<i>На даче стояла со времен Второй мировой войны навес, а не веранда.</i>
VS	Нет	Мужской	<i>На даче стоял навес, а не веранда со времен Второй мировой войны.</i>
VS	Нет	Женский	<i>На даче стояла навес, а не веранда со времен Второй мировой войны.</i>

Основными гипотезами исследования были предположения о том, что приемлемость предложений с согласованием по грамматическому правилу значимо возрастет при порядке слов SV и наличии расстояния между подлежащим и сказуемым.

Эксперимент проводился в программе PC Ixex Farm (<https://farm.pcixex.net/>), распространялся в социальных сетях, а также среди семьи,

друзей и знакомых автора. Участие в эксперименте не оплачивалось. Время оценки каждого языкового выражения было ограничено 10 секундами. В ходе эксперимента участники выставляли стимульным и филлерным предложениям оценки от 1 до 7 (по шкале Ликерта). В исследовании приняли участие 123 респондента: 68 женщин и 36 мужчин. Средний возраст составил 36 лет, минимальный возраст – 17 лет, максимальный – 73 года. Из 123 участников ответы 19 были отсеяны по причине большого количества пропусков (от 15 предложений), ошибок при ответе на контрольные вопросы, а также слишком сильного смещения оценок в положительную или отрицательную сторону (в том числе выставление высоких оценок неграмматичным филлерам и т.п.).

Результаты эксперимента были проанализированы с использованием линейной смешанной модели и множественных попарных сравнений Тьюки. Оптимальной оказалась модель регрессии, включающая номер участника в качестве случайного эффекта и род предиката в качестве фиксированного (1 + род предиката + (1 + род предиката | номер участника) + (1 | номер предложения)). Из всех факторов значимым оказался лишь род предиката ($\beta = 1,016$, $SE = 0,06$, $t\text{-value} = 14,9$, $p\text{-value} \ll 0,001$). Этот результат хорошо виден также на графиках прямого и обратного взаимодействия (рис. 1, 2).

Высокие (близкие к грамматичным филлерам) оценки получили предложения с согласованием по мужскому роду, т.е. с первым и неотрицаемым конъюнктом, а низкие (близкие к неграмматичным филлерам) оценки при любых условиях получали предложения с согласованием по женскому роду – с конъюнктом под отрицанием. Так один из представленных в дизайне эксперимента факторов значимо не повлиял на оценки экспериментальных предложений.

Однако анализ с помощью попарных сравнений методом Тьюки показал значимое различие между предложениями с согласованием по женскому роду (с ближайшим конъюнктом) и неграмматичными филлерами ($\beta = -1,02$, $SE = 0,0682$, $p\text{-value} = <0,0001$), что дает основания не полностью приравнивать указанную вариативность в согласовании к ошибкам.

5. Заключение

В ходе представленного экспериментального исследования выдвинутая в начале гипотеза не подтвердилась. Согласование по семантическому правилу оценивается носителями стабильно выше согласования по грамматическому правилу при любой комбинации факторов порядка слов и наличия расстояния между подлежащим и сказуемым. Эти же факторы не оказывают значимого воздействия на оценки согласования по грамматическому правилу и без сравнения с семантическим. Что могло стать причиной таких результатов?

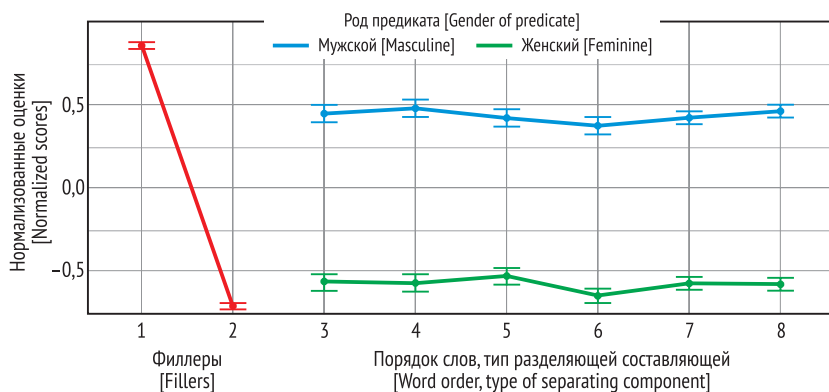


Рис. 1. График взаимодействия:

- 1 – грамматические; 2 – неграмматические; 3 – SV, предложная группа;
 4 – SV, деепричастный оборот; 5 – VS, предложная группа;
 6 – VS, деепричастный оборот; 7 – SV, нет расстояния;
 8 – VS, нет расстояния

Fig. 1. Interaction chart:

- 1 – grammatical; 2 – ungrammatical; 3 – SV, prepositional group;
 4 – SV, adverbial phrase; 5 – VS, prepositional group;
 6 – VS, adverbial phrase; 7 – SV, no distance;
 8 – VS, no distance

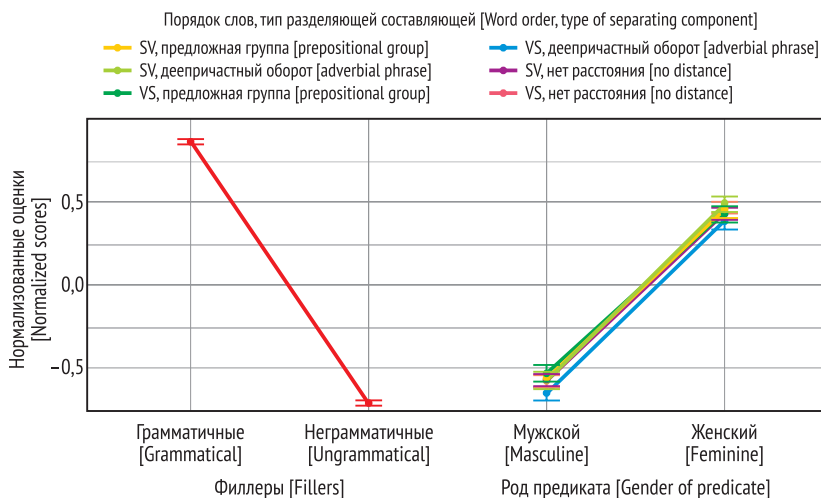


Рис. 2. График обратного взаимодействия

Fig. 2. Reverse interaction chart

Возможно, использованная в экспериментальном исследовании методика оказалась нечувствительной к малейшим изменениям в восприятии носителями стимульных предложений с разными паттернами согласования. Чтобы уловить изменение количества когнитивных усилий, которые затрачиваются на обработку предложений, стоило использовать методику чтения с саморегулирующей скорости и сравнить время, которое уходит у респондента на понимание и выставление оценки разным предложениям. Подобные предположения также обусловлены тем, что достаточно много респондентов пропустили больше 15 предложений эксперимента. Десяти секунд, которые давались участникам на оценку одного предложения, по-видимому, для многих оказалось просто недостаточно, чтобы прочесть и понять стимулы, которые включали деепричастные обороты и предложные группы, усложняющие понимание. Так что предложенная методика могла бы помочь снизить количество респондентов, пропускающих большое число предложений из-за подобных сложностей с восприятием стимулов.

Кроме того, свою роль мог сыграть тот факт, что мужской род в русском языке в прошедшем времени не маркирован на глаголе в отличие от женского. А учитывая то, что в контексте эксперимента согласование по грамматическому правилу и без того противоречит современным правилам литературного языка, то дополнительная когнитивная нагрузка, связанная с выбором окончания, лишь повышает вероятность снижения оценок для таких предложений.

Все озвученные выше предположения мы надеемся проверить в будущих исследованиях, но на данный момент, опираясь на полученные данные, нужно признать, что предикативное согласование по грамматическому правилу с подлежащим, соединенным заместительным союзом *a ne*, остается неприемлемым для современных носителей русского языка и не может претендовать на закрепление в качестве регулярной тенденции.

Библиографический список / References

Богуславский, 1985 – Богуславский И.М. Исследования по синтаксической семантике: Сферы действия логических слов. М., 1985. [Boguslavskij I.M. Issledovaniya po sintaksicheskoj semantike: Sfery dejstviya logicheskikh slov [Research on syntactic semantics: The scope of logical words]. Moscow, 1985.]

Глинкина, 1978 – Глинкина Л.А. Подлежащее // Историческая грамматика русского языка: Синтаксис: Простое предложение. М., 1978. С. 7–40. [Glinkina L.A. Subject. *Istoricheskaya grammatika russkogo yazyka: Sintaksis: Prostoje predlozhenie*. Moscow, 1978. Pp. 7–40. (In Rus.)]

Граудина и др., 1976 – Граудина Л.К., Ицкович В.А., Катлинская Л.П. Грамматическая правильность русской речи: Опыт частотно-стилистического словаря вариантов. М., 1976. [Graudina L.K., Itskovich V.A., Katlinskaya L.P. Grammaticheskaya pravilnost russkoi rechi: Opyt chastotno-stilisticheskogo slovarya variantov [Grammatical correctness of Russian speech: An attempt to compiling frequency-based stylistic dictionary of variants]. Moscow, 1976.]

Иомдин, 1979 – Иомдин Л.Л. Еще раз о синтаксическом согласовании в русском языке // Предварительные публикации Института русского языка АН СССР. Вып. 122. М., 1979. [Iomdin L.L. Once again on syntactic agreement in the Russian language. *Predvaritelnye publikacii Instituta russkogo yazyka AN SSSR*. Vol. 122. Moscow, 1979. (In Rus.)]

Крейдлин, 1979 – Крейдлин Г.Е. Служебные слова в русском языке (семантические и синтаксические аспекты их изучения): дис. ... канд. филол. наук. М., 1979. [Krejdlin G.E. Sluzhebnye slova v russkom yazyke (semanticheskie i sintaksicheskie aspekty ikh izucheniya) [Function words in the Russian language (semantic and syntactic aspects of their study)]. PhD dis. Moscow, 1979.]

РГ, 1980 – Русская грамматика / под ред. Н.Ю. Шведовой. М., 1980. Т. 2. [Russkaya grammatika [Russian grammar]. N.Yu. Shvedova (ed.). Moscow, 1980. Vol. 2.]

Розенталь, 1974 – Розенталь Д.Э. Практическая стилистика русского языка. 3-е изд. М., 1974. [Rozenal D.E. Prakticheskaya stilistika russkogo yazyka [Practical stylistics of the Russian language]. Moscow, 1974.]

Санников, 2008 – Санников З.В. Русский синтаксис в семантико-прагматическом пространстве. М., 2008. [Sannikov V.Z. Russkij sintaksis v semantiko-pragmaticheskom prostranstve [Russian syntax in the view of semantics and pragmatics]. Moscow, 2008.]

Статья поступила в редакцию 02.08.2024
The article was received on 02.08.2024

Сведения об авторе / About the author

Анисимова Александра Геннадьевна – студент бакалавриата отделения теоретической и прикладной лингвистики филологического факультета, Московский государственный университет имени М.В. Ломоносова

Alexandra G. Anisimova – BA student at the Department of Theoretical and Applied Linguistics of the Philological Faculty, Lomonosov Moscow State University, Moscow, Russian Federation

E-mail: anisimova.a.g96@gmail.com

А.Ю. Хоменко^{1, 2}, Д.П. Бальба¹, Д.А. Исаков¹

¹ Национальный исследовательский университет
«Высшая школа экономики»,
603155 г. Нижний Новгород, Российская Федерация

² ООО «Центр экспертизы и оценки «ЕСИН»»,
603093 г. Нижний Новгород, Российская Федерация

Кластеризация лексики для решения задачи диагностики афазии

В данной работе реализуется алгоритм автоматической кластеризации лексики для решения задачи создания компьютерной модели диагностики афазии на основе результатов теста на семантическую вербальную беглость. Данный тест представляет собой называние максимального количества слов, соответствующих некоторой заданной категории, за определенный промежуток времени. При анализе результатов теста часто исследуют количественные и качественные значения: количество правильных ответов, количество переключений – переходов от лексики одной субкатегории, хранящейся в сознании в виде группы лексем, сгруппированных на основе некоторого общего признака, к лексике другой субкатегории. При этом объединения лексем на основе общего признака принято называть кластерами, размеры которых также являются одним из качественных показателей прохождения теста. Была выдвинута гипотеза, согласно которой респонденты, подверженные афазии, дают меньшее количество ответов с меньшим количеством переключений, но похожим размером кластеров, а также имеют менее выраженную семантическую близость в кластерах и между ними. Авторами использовался алгоритм автоматической кластеризации лексики на основе правил, предложенный в работе Н. Лундин и коллег (Psychiatry Research. 2022. Т. 309), а также материалы теста на вербальную беглость, собранные и описанные в работе О.В. Буйволовой (Российский журнал когнитивной науки. 2020. Т. 7). В результате была создана функциональная модель для определения афазии у носителей

© Хоменко А.Ю., Бальба Д.П., Исаков Д.А., 2024



Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License

русского языка. Было также показано, что респонденты с афазией склонны давать менее стандартизированные ответы, формирующие меньшее количество подкатегорий, однако с большей ассоциативной связью внутри кластеров. Также с помощью сравнительного анализа с данными Национального корпуса русского языка было выявлено, что для респондентов обеих исследуемых групп – контрольной группы и группы лиц с афазией – характерно давать стандартные для носителей русского языка ассоциации.

Ключевые слова: диагностика афатических расстройств, кластеризация лексики, вербальная беглость

Благодарности. Исследование осуществлено в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики».

ДЛЯ ЦИТИРОВАНИЯ: Хоменко А.Ю., Бальба Д.П., Исаков Д.А. Кластеризация лексики для решения задачи диагностики афазии // Рема. Rhema. 2024. № 3. С. 87–111. DOI: 10.31862/2500-2953-2024-3-87-111

DOI: 10.31862/2500-2953-2024-3-87-111

A. Khomenko^{1, 2}, D. Balba¹, D. Isakov¹

¹ HSE University,
Nizhny Novgorod, 603155, Russian Federation

² LLC “Center for Expertise and Assessment ‘ESIN’”,
Nizhny Novgorod, 603093, Russian Federation

Vocabulary clustering for solving problems of aphasia diagnostics

In this article an algorithm of automatic vocabulary clustering is implemented to solve the problem of creating a computer model for aphasia diagnostics based on the results of the verbal fluency test. This test involves generating the maximum number of words that fit a specified category within a given time limit. When analyzing test results, researchers often examine both quantitative and qualitative metrics: the total number of correct responses

and the number of “switches” – transitions from one subcategory of lexicon to another. The subcategories are presented as groups of lexemes that are organized based on some common attributes. They are usually referred to as clusters and their sizes serve as additional qualitative indicators of test performance. The study hypothesizes that the respondents who were affected by aphasia produced a lower number of responses with fewer switches, but showed similar cluster size, as well as less distinct semantic similarity in and between clusters. The authors employed a rule-based algorithm of automatic vocabulary clustering, proposed in the work of N. Lundin et al. (Psychiatry Research. 2022. T. 309), along with the materials collected and described in the work of O. Buivolova et al. (Russian Journal of Cognitive Science. 2020. Vol. 7. No. 3). As a result, a functional model for determining aphasia among Russian speakers was created. The findings also indicate that respondents with aphasia tend to give less standardized answers, forming fewer subcategories, but with a greater associative score within clusters. Additionally, comparative analysis with data from The Russian National Corpus revealed that respondents of both groups tend to provide associations, that are typical and standardized for native Russian speakers.

Key words: aphasia diagnostic, vocabulary clustering, verbal fluency

Acknowledgments. This article is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

FOR CITATION: Khomenko A., Balba D., Isakov D. Vocabulary clustering for solving problems of aphasia diagnostics. *Rhema*. 2024. No. 3. Pp. 87–111. (In Rus.). DOI: 10.31862/2500-2953-2024-3-87-111

1. Введение

Ранняя диагностика когнитивных нарушений является крайне важной задачей в психиатрии, неврологии и педагогике, т.к. она позволяет запустить восстановительный процесс вовремя и предупредить осложнения. Клиническая лингвистика позволяет выявить эти нарушения при помощи анализа устной речи как на микролингвистическом (фонетическом, семантическом, лексическом), так и на макролингвистическом уровнях. Данная задача традиционно решается вручную, однако она нуждается в цифровых инструментах, которые помогут быстро и объективно оценить, требуется ли человеку помощь специалиста.

Для быстрой оценки расстройств мышления в различных клинических популяциях используется тест на вербальную беглость, входящий в состав батарей стандартизированных тестов и использующийся при проведении фундаментальных исследований психики. Данный тест характеризуется простотой процедуры и чувствительностью к когнитивным нарушениям и состоит в воспроизведении максимального количества слов, принадлежащих к одной семантической категории (например, «животные» или «фрукты»). Как правило, при оценке результатов теста учитывается только количество правильных ответов, однако современные исследования стремятся учитывать не только количественные, но и качественные показатели порожденного списка слов. Это обусловлено тем фактом, что семантические связи в последовательностях слов пациентов с нарушениями отличаются от семантических связей в группе нормы. Обработка результатов теста традиционно проводится исследователями вручную, однако в данной статье предлагается автоматический способ оценки семантических паттернов выполнения задания при помощи методов векторной семантики и языковых моделей. При этом недавние исследования показывают, что инструменты компьютерной лингвистики дают возможность извлекать детализированную семантическую информацию с большей эффективностью и надежностью, нежели при анализе, проведенном вручную [Corcoran, 2020, p. 163]. До настоящего момента используемый в работе алгоритм не применялся к русскоязычному материалу, полученному после проведения тестов на вербальную беглость для лиц, страдающих афазией, что обуславливает новизну работы.

В существующих аналогичных англоязычных исследованиях на материале тестирования лиц с психическими расстройствами проверяются гипотезы о том, что члены группы носителей с заболеваниями продуцируют меньшее количество ответов, а после кластеризации в этой группе обнаруживается меньшее количество переключений между кластерами, при этом средние размеры кластеров респондентов разных групп не различаются [Elvevåg et al., 2002; Lundin et al., 2020, 2022]. В данном исследовании мы проверяем эту гипотезу на русскоязычном речевом материале респондентов с афатическими расстройствами.

2. Теория категоризации как основа представления кластеров

При поиске семантически связанных групп слов в ответах респондентов опираются на теорию категоризации, согласно которой в сознании носителя для каждой единицы формируется категория и для каждой

категории – определенные единицы [Болдырев, 2014]. Существует три основных подхода к осмыслению данной теории.

Одним из них является прототипический подход, при котором в основе организации категории находится иерархия ее членов, где наиболее значимыми представителями этой категории называют прототипы [Rosch, 1978].

Прототипы отражают признаки категории в наибольшей степени и часто являются первой реакцией в качестве ответа на стимул-категорию. Внутри одной лексической категории вокруг прототипов в сознании человека группируются другие элементы, которые входят в данную категорию. Эти элементы формируют свои субкатегории, объединяясь между собой по какому-то общему характерному признаку, позволяющему опознавать субкатеорию [Elvevåg et al., 2002; Болдырев, 2014; Lundin et al., 2020]. Данные факты дают нам возможность говорить о том, что при продукции речи участники тестов создают связанные группы слов – кластеры [Lundin et al., 2022]. Во время переключения (смены одного кластера другим) говорящий переходит от одного слова к другому, менее связанному семантически.

Различные исследования указывают на то, что данные контрольных групп и групп с различными психоневрологическими нарушениями имеют расхождения, что позволяет опираться на результаты кластеризации при диагностике [Ibid].

3. Метод автоматической кластеризации лексики

Для изучения особенностей прохождения теста лицами из разных клинических групп можно использовать инструменты компьютерной лингвистики, в частности методы векторной семантики и языковые модели, которые позволяют извлекать детализированную семантическую информацию с большей скоростью и надежностью, нежели при анализе, проведенном вручную. Данная работа основана на алгоритме кластеризации, предложенном в исследовании, которое посвящено построению диагностической модели для определения наличия у англоговорящих респондентов психотических расстройств (psychosis) на ранней стадии при помощи теста на семантическую вербальную беглость [Ibid]. Данный алгоритм до настоящего момента не использовался в исследованиях афатических расстройств на русскоязычном материале.

Для реализации данного алгоритма в настоящем исследовании используется датасет, содержащий ответы респондентов на тест на вербальную беглость, который проводился Центром языка и мозга Национального исследовательского университета «Высшая школа экономики» [Буйволова и др., 2020].

При тестировании были собраны результаты по трем семантическим категориям: «животные», «профессии» и «города». После первичной предобработки с удалением пустых строк, возникших в результате ошибки сбора или расшифровки, были удалены элементы, заполняющие хезитационные паузы ([а-, э-, м-]-образные звуки, десемантизированная лексика). При помощи морфологического анализатора PyMorphy2¹ была проведена токенизация и лемматизация материала [Korobov, 2015]. Следующим этапом стал экспертный анализ и ручное тегирование некоторых лексем.

В результате данной предобработки был получен список токенов из ответов респондентов, содержащий существительные, соответствующие заданной категории. К нему применялся описанный в работе [Lundin et al., 2022] алгоритм кластеризации, на основе которого считались статистические метрики, описанные в разделе 4.

3.1. Описание материалов исследования

Используемый датасет представляет собой русскоязычный речевой материал трех групп: ответы контрольной группы неврологически здоровых лиц, ответы лиц с афазией различного генезиса и ответы группы дисфазии – 120, 60 и 20 человек соответственно [Буйволова и др., 2020]. Все участники теста являются носителями русского языка.

Под афазией понимается специфическое нарушение речи, возникающее при органических поражениях головного мозга различной этиологии [Арутюнян, 2013]. Отдельно стоит упомянуть понимание термина дисфазия (dysphasia). В этой работе термин дисфазия выступает как приобретенное нарушение устной и письменной речи в результате травмы. При этом трактовка данного понятия в различных источниках варьируется: в отечественной литературе традиционно под дисфазией понимается нарушение развития речи у детей, а в иностранной литературе понятия дисфазии и афазии часто воспринимаются как синонимичные или взаимозаменяемые. В данной работе термин дисфазия являет собой менее тяжелую форму афазии [Perry, 2013].

Во время первого этапа предобработки был удален речевой материал тех респондентов, ответы которых отсутствовали в результате ошибки сбора или расшифровки. Таким образом, после преобразования сохранились следующие значения: 94 ответа контрольной группы и 69 ответов в группе лиц с афатическими расстройствами и дисфазией. Так как данных группы дисфазии оказалось недостаточно для противопоставления ее двум другим группам, было решено объединить их с группой лиц

¹ URL: <https://pymorphy2.readthedocs.io/en/stable/> (дата обращения: 12.01.2024).

с афазией различного генезиса (с сохранением метки «dys» в качестве типа афатического расстройства).

После первичной предобработки было выделено две группы: контрольная группа и группа лиц с афатическим расстройством. Данные классы использовались в качестве материалов исследования.

Возраст респондентов контрольной группы варьируется от 17 до 86 лет, группы лиц с афазией – от 26 до 79 лет.

Среди лиц с афазией встречаются носители следующих типов заболевания: акустико-мнестическая, моторная афферентная, амнестическая, проводниковая, тотальная сенсомоторная, динамическая, моторная эфферентная, сенсорная афазия и дисфазия. Преобладающим типом оказалась моторная афазия (табл. 1). Также следует учитывать, что многие респонденты подвержены нескольким типам афатического расстройства (например, частое сочетание – моторная афферентная и эфферентная одновременно).

Таблица 1

**Количество респондентов
с различными типами афатического расстройства
[Number of respondents with different types of aphasic disorder]**

Тип афатического расстройства [Type of aphasic disorder]	Количество респондентов [Number of respondents]
Акустико-мнестическая афазия [Acoustic-amnestic aphasia]	8
Моторная афферентная афазия [Motor afferent aphasia]	13
Амнестическая афазия [Amnestic aphasia]	1
Проводниковая афазия [Conduction aphasia]	16
Тотальная сенсомоторная афазия [Total sensorimotor aphasia]	1
Динамическая афазия [Dynamic aphasia]	6
Моторная эфферентная афазия [Motor efferent aphasia]	24
Сенсорная афазия [Sensory aphasia]	15
Дисфазия [Dysphasia]	16

В данной работе мы преследовали цель описания общих особенностей категоризации лексики, свойственных людям, страдающим афатическими расстройствами, и отличающих их от носителей языка группы нормы, что объясняет объединение лиц всех типов афатического расстройства в одну группу для противопоставления контрольной группе.

3.2. Предобработка материалов

Перед кластеризацией ответы респондентов были предобработаны для приведения их к общему машиночитаемому виду: в частности, были удалены вербально заполненные паузы хезитации разного рода: фонетико-фонологические (такие как э-э, *хм*, удлинения звуков и т.д.), лексико-семантические (*так, ну, в общем* и т.д.), синтаксические (непреднамеренные повторы слов и словосочетаний, незаконченные высказывания). Также убирались лексемы, не соответствующие задаче (например, упоминание *колобка* в категории «животные»; такие лексемы затрудняли исследование ассоциаций между словами конкретной категории, при этом, встретившись в малом количестве в данных, не позволяли сделать каких-то дополнительных выводов, мы воспринимали такие случаи как шум в данных), и повторы (в случае с повторами сохранялось только первое упоминание). Следующим этапом была произведена лемматизация ответов при помощи морфологического анализатора PyMorphy2 [Korobov, 2015].

Затем после экспертного анализа материала было произведено ручное тегирование некоторых групп лексики. В частности, для лексем с пояснениями (например *Свердловск, то есть Екатеринбург* или *козел горный, то есть архар*) был введен тег ЛСП, для лексем с уточнениями (*удав, нет, это пресмыкающееся* или *попугай, это ж, хоть и птица, но животное*) – ЛСУ, для семантически восстанавливаемых окказионализмов (*топибара* – семантическая группа «животные», предположительно, восстанавливаемо до *капибара*) – СВО, для семантически невосстанавливаемых окказионализмов (*мовы, стантея*) – СНО. Чтобы избежать утраты полноты информации, были созданы два набора данных – с тегами (т.н. набор «чистых» данных), где лексемы были заменены на соответствующие теги, и набор данных со всеми лексемами, где мы сохранили лексемы СВО и СНО в исходном виде, а для ЛСП и ЛСУ сохранили обе лексемы (главное слово и пояснение/уточнение) в лемматизированном виде.

3.3. Алгоритм кластеризации

Для преобразования ответов в векторное представление использовалась предобученная статическая эмбединговая модель `geowac_lemmas_`

none_fasttextskipgram_300_5_2020 сервиса RusVectōrēs², с помощью которой были получены эмбединги типа fastText [Bojanowski, 2017]. Данная модель представляет собой полносвязную нейронную сеть, обученную на русскоязычном подкорпусе корпуса GeoWAC [Dunn, Adams, 2020] на задачу предсказания по данному (центральному) слову других слов из его контекста (контекст определяется как окно длиной $2k + 1$, где k – количество слов до/после центрального). При этом для создания эмбедингов типа FastText каждое слово перед обучением модели делится на токены – символьные n -граммы, где n – количество символов в одном токене (для данной модели были взяты $n = 3, 4$ и 5). Вектор слова, таким образом, определяется как сумма векторов всех n -грамм, входящих в это слово, и вектора самого слова. Данный тип векторных представлений слов был выбран в связи с тем, что описанный выше алгоритм обучения эмбедингов позволяет кодировать окказионализмы или слова с фонетической эрозией (т.е. слова, которые не могли встретиться в обучающем корпусе текстов, но легко получаются из обученных n -грамм), продуцируемые респондентами с афазией.

При кластеризации использовалось правило, предложенное в работе [Lundin et al., 2022]. Согласно этому правилу, для последовательности слов A, B, C, D переключение находится после слова B в случае, когда косинусная близость $S(X, Y)$ между векторами слов A и B больше косинусной близости между векторами слов B и C , а косинусная близость между векторами слов B и C меньше, чем косинусная близость между векторами слов C и D : $S(A, B) > S(B, C) < S(C, D)$.

Для проведения кластеризации лексики был написан код на языке Python с использованием библиотек Gensim³ и Pandas⁴.

4. Результаты

Были получены наборы кластеров для каждого респондента по каждой из трех семантических категорий: «животные», «профессии» и «города». Также с помощью программного кода на Python для каждого набора кластеров были посчитаны следующие метрики: количество переключений между кластерами, средний размер кластера, среднее расстояние между кластерами, средние значения метрик t -score и silhouette-score. Далее все эти значения были агрегированы по категориям

² RusVectōrēs: семантические модели для русского языка. URL: <https://rusvectors.org/> (дата обращения: 12.01.2024).

³ URL: <https://radimrehurek.com/gensim/> (дата обращения: 12.01.2024).

⁴ URL: <https://pandas.pydata.org/> (дата обращения: 12.01.2024).

и респондентам и были получены средние, медианы и стандартные отклонения метрик для каждой группы респондентов и каждого типа предобработки (с тегированием и без).

4.1. Результаты кластеризации

Результаты применения алгоритма кластеризации лексики к ответам отдельных респондентов для семантической категории «животные» показан на графиках (рис. 1–3).



Рис. 1. Визуализация кластеризации ответа респондента (ID C6) контрольной группы для семантической категории «животные»

Для первого слова в ответе значение косинусной близости между векторами этого слова и предыдущего равно 0

Начало ответа отмечено красным, переключение (первое слово в новом кластере после переключения) – желтым цветом

Fig. 1. Visualization of the clustering of the respondent's answer (ID C6) of the control group for the semantic category "animals"

For the first word in the answer, the value of the cosine similarity between the vectors of this word and the previous one is 0

The beginning of the answer is marked in red, the switch (the first word in the new cluster after the switch) is in yellow

На приведенных примерах (см. рис. 1–3) можно заметить, что у респондентов контрольной группы графики располагаются ближе к единице по оси ординат, что говорит о большей семантической близости слов в пределах кластеров и ответа в целом (т.к. семантическая

близость между последним словом одного кластера и словом-переключением все равно сохраняет высокие значения, что говорит о том, что несмотря на переключение, слова являются семантически близкими).

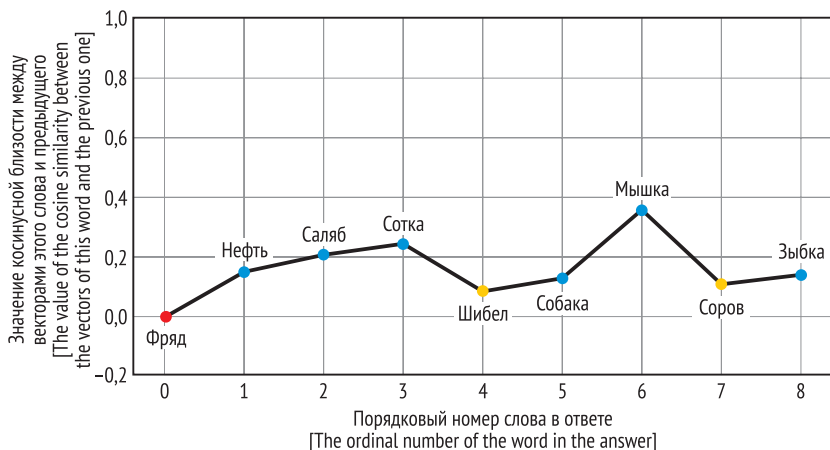


Рис. 2. Визуализация кластеризации ответа респондента (ID P012) группы с афатическим расстройством для семантической категории «животные»

Для первого слова в ответе значение косинусной близости между векторами этого слова и предыдущего равно 0

Начало ответа отмечено красным, переключение (первое слово в новом кластере после переключения) – желтым цветом

Fig. 2. Visualization of clustering of the respondent's answer (ID P012) from the group with aphasic disorder for the semantic category "animals"

For the first word in the answer, the value of the cosine similarity between the vectors of this word and the previous one is 0

The beginning of the answer is marked in red, the switch (the first word in the new cluster after the switch) is in yellow

Для респондента группы с афатическим расстройством можно наблюдать низкие значения у-координаты на всем графике, что говорит о низкой семантической близости слов (рис. 2). В данном случае это также определяется окказионализмами, приведенными респондентом в качестве ответов.

Для сравнения приведен график ответов респондента группы афазии, не включающий в себя окказионализмы (рис. 3). Можно наблюдать присутствие стандартизированных (прототипических) ответов, для

которых значение семантической близости близко к единице, однако график представляет собой более «ломаную» траекторию ответов, чем в группе нормы, что свидетельствует о меньшей семантической близости ответов.



Рис. 3. Визуализация кластеризации ответа респондента (ID P001) группы с афатическим расстройством для семантической категории «животные»

Для первого слова в ответе значение косинусной близости между векторами этого слова и предыдущего равно 0

Начало ответа отмечено красным, переключение (первое слово в новом кластере после переключения) – желтым цветом

Fig. 3. Visualization of clustering of the respondent's answer (ID – P001) from the group with aphasic disorder for the semantic category “animals”

For the first word in the answer, the value of the cosine similarity between the vectors of this word and the previous one is 0

The beginning of the answer is marked in red, the switch (the first word in the new cluster after the switch) is in yellow

4.2. Статистический анализ результатов кластеризации

После работы алгоритма были подобраны такие метрики результатов кластеризации, которые оценивают особенности категоризации лексики в сознании носителей языка из двух исследуемых групп и показывают различия между ними. Так, были выделены следующие метрики:

– количество переключений для каждого человека по каждой категории;

- среднее расстояние между кластерами для каждого человека по каждой категории (считается как расстояние между центроидами кластеров; за центроид берется среднее значение векторов, находящихся в одном кластере);
- средний размер кластера для каждого человека.

Данные метрики уже использовались ранее в аналогичных исследованиях кластеризации лексики, продуцируемой респондентами теста на вербальную беглость, и показывали значимые различия между исследуемыми группами [Lundin et al., 2022]. Также нами были протестированы следующие метрики, способные показывать количественную информацию о свойствах ассоциативных связей ответов респондентов.

1. Среднее значение *t*-score в кластере для каждого человека по каждой категории. *T*-score – это адаптированная корпусная метрика, отображающая то, насколько неслучайной является сила ассоциации между коллокатами, т.е. насколько неслучайно два слова встретились рядом друг с другом в корпусе, насколько сильна их семантическая связь [Захаров, Хохлова, 2010]. В качестве коллокатов в нашем случае берутся все последовательности из двух слов внутри одного кластера. Метрика *t*-score для одной пары слов рассчитывалась следующим образом:

$$t\text{-score} = \frac{f(n, c) - \frac{f(n) \cdot f(c)}{N}}{\sqrt{f(n, c)}},$$

где $f(n)$, $f(c)$ – частоты слов n и c в данной семантической категории соответственно; $f(n, c)$ – частота совместной встречаемости слов n и c (под совместной встречаемостью мы понимаем нахождение двух слов в одном кластере при условии отсутствия других слов между ними); N – количество всех слов в данной семантической категории.

2. Метрика *silhouette-score* отображает, насколько близок объект к своему кластеру по сравнению с другими кластерами: чем больше значение данной метрики, тем больше объект «похож» на объекты своего собственного кластера [Rousseeuw, 1987]. *Silhouette-score* для каждого слова рассчитывается следующим образом:

$$\text{silhouette-score} = \frac{a - b}{\min(a, b)},$$

где a – средняя косинусная близость между данным словом и другими словами в том же кластере, в котором находится данное слово; b – средняя косинусная близость между данным словом и словами из ближайшего кластера к тому, в котором находится данное слово (в нашем

случае ближайший кластер – это следующий или предыдущий кластер в последовательности). Отметим, что *silhouette-score* принимает значения от -1 до 1 . Также экспериментальным путем мы наложили ограничения на значения a и b так, что $|a| > 0,01$ и $|b| > 0,01$.

3. Коэффициент лексического разнообразия – *type-token ratio* (TTR). Данная метрика позволяет оценить лексическое разнообразие выборки, диапазон и вариативность словарного запаса, который говорящий реализует. Нами был применен самый известный способ измерения коэффициента лексического разнообразия, вычисляемый следующим образом:

$$TTR = \frac{V}{N},$$

где V – количество уникальных лексических единиц; N – общее количество словоформ. Данная метрика принимает значения от 0 (не включительно) до 1 [Захарова, 2020]. При этом, т.к. в исследовании мы использовали леммы, N сводится к общему количеству слов, которые названы некоторой выборкой, а V – к уникальным словам, которые эта выборка называет.

4. Средний размер словарного запаса (количество уникальных слов, названных респондентами одной группы). Анализ статистических метрик проводился с помощью языка программирования Python и библиотек Numpy⁵ и Scipy⁶. Был использован статистический t -критерий Стьюдента для сравнения средних, исходя из предположения о независимости выборок и нормальном распределении данных. В качестве уровня значимости было взято стандартное значение $\alpha = 0,05$. После анализа статистических метрик было обнаружено, что значимых различий между значениями их средних для подвыборки с тегированием и подвыборки с исходными лексемами не обнаружено (табл. 2), поэтому далее анализ проводился только на втором наборе данных. Значения метрик можно увидеть в таблице 3.

Значение метрики t -score оказалось выше для контрольной группы. Это манифестирует, что в ней ассоциации между реакциями более стандартизированы, т.е. в ответах контрольной группы чаще встречаются одинаковые устойчивые пары слов-ассоциаций.

Еще одним подтверждением более стандартизированных ассоциаций в группе нормы является большее лексическое разнообразие TTR в группе афазии, которое, кроме менее стандартизированных связей,

⁵ URL: <https://numpy.org/> (дата обращения: 12.01.2024).

⁶ URL: <https://scipy.org/> (дата обращения: 12.01.2024).

мотивировано присутствием большого числа окказионализмов в ответах данной группы. Это объясняется тем, что значение TTR прямо пропорционально количеству уникальных слов в данных, которое оказывается выше для группы лиц с афазией из-за присутствия окказионализмов. Кроме того, меньшее значение лексического разнообразия свидетельствует о прототипических ответах в группе нормы. Несмотря на меньшее лексическое разнообразие, словарный запас (количество уникальных слов) в контрольной группе превышает словарный запас лиц с афазией в 1,7 раз. Можно заметить, что значение silhouette-score выше для лиц с афазией, что свидетельствует о большей связанности слов внутри одного кластера.

Таблица 2

Усредненные метрики кластеризации для подвыборок группы лиц с афазией с тегированием и без по всем категориям
[Average clustering metrics for subsamples of aphasia group with and without tagging across all categories]

Статистические метрики [Statistical metrics]	Контрольная группа с тегированием [Control group with tagging]	Контрольная группа с сохранением исходных лексем [Control group with preservation of original lexemes]
Количество переключений [Number of switches]	2,38	2,39
Расстояние между кластерами [Distance between clusters]	0,65	0,66
<i>t</i> -score	22,01	21,68
Silhouette score	0,25	0,25
Размер кластера [Cluster size]	3,08	3,09

4.3. Анализ частотных коллокаций

Были также проанализированы наиболее частотные коллокации внутри кластеров, которые были отсортированы по относительной частоте внутри своей группы. Для подсчета наиболее частотных коллокаций использовалась относительная частота по корпусу всех коллокаций, которые образовывались в одном кластере. Наиболее частотные коллокации для категории «животные» приведены в таблице 4.

Таблица 3

Усредненные метрики кластеризации по всем категориям
[Average clustering metrics across all categories]

Статистические метрики [Statistical metrics]	Контрольная группа [Control group]	Лица с афазией [People with aphasia]	Статистическая значимость (по <i>t</i> -критерию Стьюдента, $\alpha = 0,05$) [Statistical significance (Student's <i>t</i> -test, $\alpha = 0.05$)]
<i>Количество переключений*</i> [Distance between clusters*]	5,23; 1,88**	2,39; 1,56	$p < 0,001$
Расстояние между кластерами [Distance between clusters]	0,64; 0,04	0,66; 0,06	$p > 0,05$
<i>t-score</i>	51,58; 14,3	21,68; 10,76	$p < 0,001$
<i>Silhouette score</i>	0,16; 0,06	0,25; 0,14	$p < 0,001$
Размер кластера [Cluster size]	3,15; 0,31	3,09; 0,43	$p > 0,05$
TTR (мера лексического разнообразия) [Type-token ratio]	0,23	0,34	
Словарный запас (количество уникальных слов) [Vocabulary (number of unique words)]	412	241	

* Курсивом выделены статистически значимые метрики.

** Указаны среднее значение параметра (по каждому респонденту и категории) и его стандартное отклонение.

[* Statistically significant metrics are highlighted in italics.

** The average value of the parameter (for each respondent and category) and its standard deviation are indicated].

Эти коллокации имеют высокий сочетаемостный уровень в корпусе ответов респондентов. Заметим, что каждая из этих коллокаций входит в 10 наиболее частотных в Национальном корпусе русского языка (НКРЯ) (по состоянию на 13 февраля 2024 г.) (табл. 5). При поиске коллокатов в НКРЯ в качестве запроса использовалось первое слово

из коллокации, а коллокаты сортировались по значению *t*-score. Поиск производился по основному корпусу, для запроса первое слово задавалось как лемма, а для второго слова задавались параметры (существительное, нарицательное). Частотность данных пар говорит о репрезентации стандартных для русского языка коллокаций внутри кластеров. При этом относительная частота этих коллокаций в обеих группах равна, что может быть связано с тем, что разнообразие коллокаций в группе нормы мотивировано большим словарным запасом, а в группе афазии – большим количеством окказионализмов.

Таблица 4

Наиболее частотные коллокации со значением *t*-score для контрольной группы и группы лиц с афазией по семантической категории «животные»
[The most frequent collocations with *t*-score value for the control group and the group of people with aphasia for the semantic category “animals”]

Контрольная группа [Control group]		Лица с афазией [People with aphasia]	
Коллокация [Collocation]	<i>t</i> -score	Коллокация [Collocation]	<i>t</i> -score
<i>кошка–собака</i>	23,93	<i>кошка–собака</i>	17,28
<i>лев–тигр</i>	13,53	<i>волк–медведь</i>	8,58
<i>жираф–слон</i>	9,53	<i>заяц–лиса</i>	7,91
<i>корова–овца</i>	7,53	<i>волк–лиса</i>	7,24
<i>заяц–лиса</i>	7,14	<i>лев–тигр</i>	5,90

Так как значения *t*-score выше для контрольной группы, можно сделать вывод о том, что ассоциации между словами внутри кластеров этой группы в среднем более стабильные, чем в ответах лиц с афазией. Похожую картину можно наблюдать для коллокаций категорий «профессии» (табл. 6, 7) и «города» (табл. 8, 9).

При этом можно заметить, что средние значения частоты коллокации и *t*-score максимальны для категории «животные» в обеих группах, т.е. там присутствуют наиболее стандартизированные ответы. Значение частоты и метрики *t*-score минимальны в категории «профессии», что говорит о наименее стандартизированных ответах в данной категории, а следовательно, и меньшей ее устойчивости. Это можно объяснить тем, что разнообразие ответов в категориях «профессии» и «города» зависит

Таблица 5

Значения *t*-score коллокаций категории «животные» в Национальном корпусе русского языка (поиск по первому слову пары) для контрольной группы и лиц с афазией
[*T*-score values of collocations of the category “animals” in the Russian National Corpus (search by the first word of the pair) for the control group and individuals with aphasia]

	Коллокация [Collocation]	Относительная частота в корпусе ответов респондентов [Relative frequency in the corpus of respondents' answers]	<i>t</i> -score в корпусе ответов респондентов [<i>t</i> -score in the corpus of respondents' answers]	<i>t</i> -score в НКРЯ [<i>t</i> -score in the Russian National Corpus]	Место в списке коллокаций НКРЯ, отсортированном по частотности [Place in the list of collocations of the Russian National Corpus, sorted by frequency]
Контрольная группа [Control group]	<i>кошка–собака</i>	0,026	23,930	36,480	1
	<i>лев–тигр</i>	0,015	13,533	17,150	4
	<i>жираф–слон</i>	0,010	9,535	6,480	8
	<i>корова–овца</i>	0,008	7,535	26,460	1
	<i>заяц–лиса</i>	0,008	7,135	10,090	3
Лица с афазией [People with aphasia]	<i>кошка–собака</i>	0,028	17,280	36,480	1
	<i>волк–медведь</i>	0,014	8,580	22,370	3
	<i>заяц–лиса</i>	0,013	7,911	10,090	3
	<i>волк–лиса</i>	0,012	7,242	13,030	6
	<i>лев–тигр</i>	0,010	5,903	17,150	4

не только от культурного опыта респондента и наличия у него речевого нарушения, но во многом и от личностного опыта (например, респонденты склонны называть города своей родины, профессии врачей, которых они посещают, и т.п.).

Таблица 6

Наиболее частотные коллокации со значением t -score для контрольной группы и лиц с афазией групп по семантической категории «профессии»
[The most frequent collocations with t -score value for the control group and individuals with aphasia groups by the semantic category “professions”]

Контрольная группа [Control group]		Лица с афазией [People with aphasia]	
Коллокация [Collocation]	t -score	Коллокация [Collocation]	t -score
<i>врач–учитель</i>	18,91	<i>врач–учитель</i>	13,04
<i>преподаватель–учитель</i>	10,53	<i>сантехник–электрик</i>	9,19
<i>инженер–строитель</i>	8,14	<i>сантехник–слесарь</i>	7,27
<i>слесарь–токарь</i>	8,14	<i>токарь–фрезеровщик</i>	7,27
<i>врач–инженер</i>	6,94	<i>профессор–учитель</i>	7,27

Однако в целом как респонденты без нарушений, так и респонденты с афазией склонны давать ответы, отражающие стандартные для носителей русского языка ассоциации (например, *кошка–собака*, *заяц–лиса*, *врач–учитель*, *Москва–Санкт-Петербург*), при этом ответы контрольной группы оказываются статистически более стабильными.

5. Выводы

В ходе данного исследования был создан драфт диагностического алгоритма для определения афатических нарушений речи у носителей русского языка. При этом он моделирует некоторые когнитивные особенности носителей языка с афазией и без нее. По результатам работы алгоритма кластеризации проявляются различия между двумя группами по нескольким метрикам: по количеству переключений между кластерами, по связанности на основе *silhouette-score*, по силе ассоциаций по t -score, а также по значениям лексического разнообразия TTR и объему словарного запаса. Также можно сказать, что обе группы склонны давать предсказуемые для носителей русского языка ассоциации,

Таблица 7

Значения *t*-score коллокаций категории «профессии» в Национальном корпусе русского языка (поиск по первому слову пары) для контрольной группы и лиц с афазией
[*T*-score values of collocations of the category “professions” in the Russian National Corpus (search by the first word of the pair) for the control group and individuals with aphasia]

	Коллокации [Collocation]	Относительная частота в корпусе ответов респондентов [Relative frequency in the corpus of respondents' answers]	<i>t</i> -score в корпусе ответов респондентов [<i>t</i> -score in the corpus of respondents' answers]	<i>t</i> -score в НКРЯ [<i>t</i> -score in the Russian National Corpus]	Место в списке коллокаций НКРЯ, отсортированном по частотности [Place in the list of collocations of the Russian National Corpus, sorted by frequency]
Контрольная группа [Control group]	<i>врач–учитель</i>	0,008	18,914	25,020	2
	<i>преподаватель–учитель</i>	0,005	10,533	10,490	9
	<i>инженер–строитель</i>	0,004	8,138	17,800	5
	<i>слесарь–токарь</i>	0,004	8,138	10,680	1
	<i>врач–инженер</i>	0,003	6,941	20,170	4
Лица с афазией [People with aphasia]	<i>врач–учитель</i>	0,009	13,043	25,020	2
	<i>сантехник–электрик</i>	0,007	9,197	6,710	1
	<i>сантехник–слесарь</i>	0,005	7,273	7,680	3
	<i>токарь–фрезеровщик</i>	0,005	7,273	6,080	1
	<i>профессор–учитель</i>	0,005	7,273	10,490	9

однако контрольная группа в большей мере отображает общезыковую норму согласно значениям *t*-score и результатам сравнения с данными НКРЯ. Наиболее стабильной оказалась категория «животные».

Таблица 8

**Наиболее частотные коллокации со значением *t*-score
для контрольной группы и лиц с афазией
по семантической категории «города»**

**[The most frequent collocations with *t*-score value for the control group
and individuals with aphasia for the semantic category “cities”]**

Контрольная группа [Control group]		Лица с афазией [People with aphasia]	
Коллокация [Collocation]	<i>t</i> -score	Коллокация [Collocation]	<i>t</i> -score
<i>Москва–Санкт-Петербург</i>	27,77	<i>Москва–Санкт-Петербург</i>	13,99
<i>Лондон–Париж</i>	11,81	<i>Москва–Питер</i>	12,56
<i>Ленинград–Москва</i>	10,22	<i>Ленинград–Москва</i>	11,13
<i>Севастополь–Симферополь</i>	8,62	<i>Омск–Томск</i>	8,27
<i>Вашингтон–Нью-Йорк</i>	9,42	<i>Киев–Москва</i>	8,27

Можно также заметить, что гипотеза, выдвинутая на основе англоязычных материалов, подтверждается на материалах данного исследования не полностью. Действительно, респонденты, подверженные когнитивным дисфункциям (в нашем случае – афатическому расстройству), в среднем дают меньшее количество ответов, имеют меньшее количество переключений, а также среди ответов лиц с афазией наблюдается менее выраженная семантическая близость. При этом средний размер кластеров и среднее расстояние между ними не имеют значимых различий (табл. 10), тогда как зарубежные ученые на основе исследования группы с шизофреническим спектром подчеркивали менее выраженную близость между кластерами [Elvevåg et al., 2002; Lundin et al., 2020, 2022].

Отметим, что данное исследование имеет ряд ограничений. Во-первых, небольшой объем выборки: для эффективного использования модели с методами машинного обучения и более точной диагностики расстройств желательно иметь в распоряжении несколько сотен объектов для обучения. Во-вторых, неравномерная репрезентация типов афазии: большинство респондентов подвержены моторной афазии, что не позволяет нам на данном этапе исследования изучать особенности отдельных типов речевых нарушений.

Таблица 9

Значения *t*-score коллокаций категории «города» в Национальном корпусе русского языка (поиск по первому слову пары) для контрольной группы и лиц с афазией
[*T*-score values of collocations of the category “cities” in the Russian National Corpus (search by the first word of the pair) for the control group and individuals with aphasia]

	Коллокации [Collocation]	Относительная частота в корпусе ответов респондентов [Relative frequency in the corpus of respondents' answers]	<i>t</i> -score в корпусе ответов респондентов [<i>t</i> -score in the corpus of respondents' answers]	<i>t</i> -score в НКРЯ [<i>t</i> -score in the Russian National Corpus]	Место в списке коллокаций НКРЯ, отсортированном по частотности [Place in the list of collocations of the Russian National Corpus, sorted by frequency]
Контрольная группа [Control groups]	<i>Москва–Санкт-Петербург</i>	0,013	27,766	28,050	14
	<i>Лондон–Париж</i>	0,006	11,814	32,820	2
	<i>Ленинград–Москва</i>	0,005	10,218	38,440	2
	<i>Севастополь–Симферополь</i>	0,004	8,623	7,280	4
	<i>Вашингтон–Нью-Йорк</i>	0,004	9,421	10,950	2
Лица с афазией [People with aphasia]	<i>Москва–Санкт-Петербург</i>	0,010	13,990	28,050	14
	<i>Москва–Питер</i>	0,009	12,560	23,950	25
	<i>Ленинград–Москва</i>	0,008	11,131	38,440	2
	<i>Омск–Томск</i>	0,006	8,271	6,630	2
	<i>Киев–Москва</i>	0,006	8,271	27,780	12

**Полученные после кластеризации данных значения метрик,
предложенных для сравнения в гипотезе
[The values of the metrics proposed for comparison
in the hypothesis obtained after data clustering]**

	Контрольная группа [Control group]	Группа лиц с афазией [Group of people with aphasia]
Среднее количество ответов [Average number of responses]	19,46	10,52
Среднее количество переключений [Average number of switches]	5,23	2,39
Средний размер кластеров [Average cluster size]	3,14	3,08
Среднее расстояние между кластерами [Average distance between clusters]	0,64	0,65

Полученные результаты доказывают эффективность применения методов компьютерной лингвистики для задачи диагностики афазии на русскоязычном материале и открывают потенциал для использования предложенного алгоритма кластеризации лексики для разных типов афазии. Перспективной также выглядит диагностика ментальных расстройств (депрессии, шизофрении и пр.) с применением компьютерных алгоритмов кластеризации.

Библиографический список / References

Арутюнян и др., 2013 – Арутюнян В.Г., Клачек П.М., Кошелева И.Л. Мозг и афазия: нейролингвистические подходы как основа методики для компьютеризации диагностики и реабилитации // Труды XI международной научной конференции «Инновации в науке, образовании и бизнесе – 2013». Калининград, 2013. С. 338. [Arutyunyan V.G., Klachek P.M., Kosheleva I.L. Brain and aphasia: Neurolinguistic approaches as a basis for the methodology for computerization of diagnostics and rehabilitation. *Trudy XI mezhdunarodnoy nauchnoy konferentsii «Innovatsii v nauke, obrazovanii i biznese – 2013»*. Kaliningrad, 2013. P. 338. (In Rus.)]

Болдырев, 2014 – Болдырев Н.Н. Когнитивная семантика. Введение в когнитивную лингвистику: учебное пособие. 4-е изд., испр. и доп. Тамбов, 2014. [Boldyrev N.N. Kognitivnaya semantika. Vvedenie v kognitivnuyu lingvistiku [Cognitive Semantics. Introduction to cognitive linguistics]. Tambov, 2014.]

Буйволова и др., 2020 – Adaptation of the Aphasia Bedside Check for Russian / О.В. Буйволова и др. // Российский журнал когнитивной науки. 2020. Т. 7.

№ 3. С. 45–67. [Buivolova O.V. et al. Adaptation of the Aphasia Bedside Check for Russian. *Russian Journal of Cognitive Science*. 2020. Vol. 7. No. 3. Pp. 45–67.]

Захаров, Хохлова, 2010 – Захаров В.П., Хохлова М.В. Анализ эффективности статистических методов выявления коллокаций в текстах на русском языке // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (Бекасово, 26–30 мая 2010 г.). Вып. 9 (16). М., 2010. С. 137–143. [Zakharov V.P., Khokhlova M.V. Analysis of the effectiveness of statistical methods for identifying collocations in Russian-language texts. *Computational Linguistics and Intellectual Technologies Papers from the Annual International Conference “Dialogue”* (2010). Issue 9 (16). Moscow, 2010. Pp. 137–143. (In Rus.)]

Захарова, Савина, 2020 – Захарова Е.Ю., Савина О.Ю. Лексическое разнообразие текста и способы его измерения // Вестник Тюменского государственного университета. Серия: Гуманитарные исследования. *Humanitates*. 2020. Т. 6. № 1 (21). С. 20–34. DOI: 10.21684/2411-197X-2020-6-1-20-34 [Zakharova E.Yu., Savina O.Yu. Lexical diversity measures' review and classification. *Tyumen State University Herald. Humanities Research. Humanitates*. 2020. Vol. 6. No. 1 (21). Pp. 20–34. DOI: 10.21684/2411-197X-2020-6-1-20-34 (In Rus.)]

Савчук и др., 2024 – Савчук С.О., Архангельский Т.А., Бонч-Осмоловская А.А. Национальный корпус русского языка 2.0: новые возможности и перспективы развития // Вопросы языкознания. 2024. № 2. С. 7–34. [Savchuk S.O., Arkhangelsky T.A., Bonch-Osmolovskaya. A.A. Voprosy yazykoznaniiya. 2024. No. 2. Pp. 7–34. (In Rus.)]

Bojanowski, 2017 – Bojanowski P. et al. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*. 2017. Vol. 5. Pp. 135–146.

Corcoran, 2020 – Corcoran C.M. et al. Language as a biomarker for psychosis: A natural language processing approach. *Schizophrenia Research*. 2020. Vol. 226. Pp. 158–166.

Dunn, Adams, 2020 – Dunn J., Adams B. Geographically-balanced gigaword corpora for 50 language varieties. *Proceedings of the Language Resources and Evaluation Conference (LREC 2020)*. 2020. Pp. 2528–2536.

Elvevåg et al., 2002 – Elvevåg B., Fisher J.E., Gurd J.M., Goldberg T.E. Semantic clustering in verbal fluency: Schizophrenic patients versus control participants. *Psychol. Med*. 2002. No. 32. Pp. 909–917.

Korobov, 2015 – Korobov M. Morphological analyzer and generator for Russian and Ukrainian languages. *Analysis of Images, Social Networks and Texts*. 2015. Pp. 320–332.

Lundin et al., 2020 – Lundin N.B. et al. Semantic search in Psychosis: Modeling local exploitation and global exploration. *Schizophr Bull Open*. 2020. No. 1. Pp. 1–11.

Lundin et al., 2022 – Lundin N.B. et al. Semantic and phonetic similarity of verbal fluency responses in early-stage psychosis. *Psychiatry Research*. 2022. Vol. 309. P. 114404.

Rosch, 1978 – Rosch E. Principles of categorization. *Cognition and Categorization*. Routledge, 1978. Pp. 27–48.

Rousseuw, 1987 – Rousseuw P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*. 1987. No. 20. Pp. 53–65.

Статья поступила в редакцию 19.08.2024

The article was received on 19.08.2024

Сведения об авторах / About the authors

Хоменко Анна Юрьевна – кандидат филологических наук; старший научный сотрудник Центра языка и мозга, научный сотрудник Лаборатории теории и практики систем поддержки принятия решений, приглашенный преподаватель департамента фундаментальной и прикладной лингвистики, Национальный исследовательский университет «Высшая школа экономики», г. Нижний Новгород; эксперт-лингвист, эксперт-авторовед, эксперт-фоноскопист, ООО «Центр экспертизы и оценки “ЕСИН”», г. Нижний Новгород

Anna Yu. Khomenko – PhD in Linguistics; senior researcher at the Center for Language and Brain, Researcher at the Laboratory of Theory and Practice of Decision Support Systems, Visiting Lecturer at the Department of Fundamental and Applied Linguistics, HSE University, Nizhny Novgorod; Expert Linguist, Author Expert, Expert Phonoscopist, LLC “Center for Expertise and Assessment ESIN”, Nizhny Novgorod, Russian Federation

ORCID: <https://orcid.org/0000-0003-3564-6293>

E-mail: akhomenko@hse.ru

Бальба Дарья Петровна – студент бакалавриата образовательной программы «Фундаментальная и прикладная лингвистика» факультета гуманитарных наук, Национальный исследовательский университет «Высшая школа экономики», г. Нижний Новгород; стажер-исследователь Центра языка и мозга, Национальный исследовательский университет «Высшая школа экономики», г. Нижний Новгород

Darya P. Balba – BA student at the educational program “Fundamental and Applied Linguistics” of the Faculty of Humanities, HSE University, Nizhny Novgorod; research intern at the Center for Language and Brain, HSE University, Nizhny Novgorod, Russian Federation

ORCID: <https://orcid.org/0009-0004-1434-6966>

E-mail: dpbalba@edu.hse.ru

Исаков Данила Андреевич – студент бакалавриата образовательной программы «Фундаментальная и прикладная лингвистика» факультета гуманитарных наук, Национальный исследовательский университет «Высшая школа экономики», г. Нижний Новгород; стажер-исследователь Центра языка и мозга, Национальный исследовательский университет «Высшая школа экономики», г. Нижний Новгород

Danila A. Isakov – BA student at the educational program “Fundamental and Applied Linguistics” of the Faculty of Humanities, HSE University, Nizhny Novgorod; research intern at the Center for Language and Brain, HSE University, Nizhny Novgorod, Russian Federation

ORCID: <https://orcid.org/0009-0000-3354-5048>

E-mail: issakov.danila@mail.ru

Все авторы прочитали и одобрили окончательный вариант рукописи

All authors have read and approved the final manuscript

DOI: 10.31862/2500-2953-2024-3-112-129

УДК 81'23:81-114.4

П.А. Пилипец¹, В.Н. Староверова²

¹ Московский государственный университет имени М.В. Ломоносова,
119991 г. Москва, Российская Федерация

² Национальный исследовательский университет «Высшая школа экономики»,
105066 г. Москва, Российская Федерация

Понимание сложных синтаксических конструкций школьниками младших классов

Статья посвящена анализу понимания сложных синтаксических конструкций школьниками 1–4 классов. В исследовании были использованы типы конструкций и стимульные предложения из теста синтаксических компетенций взрослых носителей русского языка, представленного в работе [Chernova et al., 2023]. Мы провели эксперимент, в рамках которого школьники слушали предложения и отвечали на вопросы по их содержанию. Мы обнаружили изменения в интерпретации конструкций с ранним и поздним закрытием: ученики 1–3 классов лучше справлялись с поздним закрытием, а четвероклассники предпочитали раннее закрытие, как и взрослые. Практически во всех классах самой простой для понимания конструкцией оказались сложные предложения с временными клаузами. Было выявлено, что четвероклассники понимают сравнительные конструкции значительно лучше, чем первоклассники.

Ключевые слова: понимание предложений, синтаксическая обработка, младшие школьники, психолингвистика

Благодарности. Исследование осуществлено в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики».

Авторы выражают благодарность участникам исследования и администрации школы № 131 г. Нижнего Новгорода за помощь в организации исследования;

© Пилипец П.А., Староверова В.Н., 2024

Контент доступен по лицензии Creative Commons Attribution 4.0 International License
The content is licensed under a Creative Commons Attribution 4.0 International License



а также рецензентам и коллегам из Центра языка и мозга Национального исследовательского университета «Высшая школа экономики», предложившим ценные замечания.

ДЛЯ ЦИТИРОВАНИЯ: Пилипец П.А., Староверова В.Н. Понимание сложных синтаксических конструкций школьниками младших классов // Рема. Rhema. 2024. № 3. С. 112–129. DOI: 10.31862/2500-2953-2024-3-112-129

DOI: 10.31862/2500-2953-2024-3-112-129

P. Pilipets¹, V. Staroverova²

¹ Lomonosov Moscow State University,
Moscow, 119991, Russian Federation

² HSE University,
Moscow, 105066, Russian Federation

Understanding complex syntactic constructions by primary school students

The study investigates the comprehension of complex syntactic constructions in primary school children. We used the constructions and the stimuli from the test of syntactic competences in healthy Russian-speaking adults presented in the article by [Chernova et al., 2023]. We conducted an experiment: the children listened to the stimuli sentences and answered the questions about their content. We found that the comprehension of the relative clause attachment changes with age. The students in the 1–3 grades better understood the low attachment sentences while the fourth-grade students preferred the high attachment as adults do. Almost in all grades the complex sentences with temporal clauses were the easiest construction to analyze. We also found that the fourth-grade students understand the multiple comparative constructions significantly better than the first-grade students.

Key words: sentence comprehension, syntactic processing, primary school children, psycholinguistics

Acknowledgments. This work is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

The authors express their gratitude to the study participants and the administration of School No. 131 of Nizhny Novgorod for their assistance in organizing the study; as well as to the reviewers and colleagues from the Center for Language and Brain, HSE University, who offered valuable comments.

FOR CITATION: Pilipets P., Staroverova V. Understanding complex syntactic constructions by primary school students. *Rhema*. 2024. No. 3. Pp. 112–129. (In Rus.). DOI: 10.31862/2500-2953-2024-3-112-129

1. Введение

Понимание предложений (syntactic processing) включает в себя извлечение значения из каждой составляющей, структурный и семантический анализ полученной информации и оценку правдоподобности интерпретаций, составленных с помощью локального знания и знаний о мире [Roberts, 2016, p. 227].

Известно, что понимание предложений у детей отличается от понимания предложений у взрослых. Так, дети дошкольного и младшего школьного возраста при интерпретации предложений больше опираются на семантику, чем на структурную информацию, и хуже, чем взрослые, обрабатывают служебные слова [Friederici, 1983; Wu et al., 2016]. Подростки же в большей степени, чем взрослые, опираются на синтаксический контекст и хуже замечают семантическую неправдоподобность [Traxler, 2002]. Тем не менее, даже у взрослых носителей могут наблюдаться трудности с пониманием отдельных предложений.

Так, в исследовании [Chernova et al., 2023] был представлен тест на понимание предложений, который состоит из нескольких типов конструкций, сочетающих в себе разные факторы сложности. Тест оценивает синтаксическую обработку у здоровых взрослых носителей русского языка и является достаточно трудным для среднестатистического носителя. В тест включены следующие типы конструкций, являющиеся самыми сложными для обработки:

- 1) пространственные конструкции;
- 2) сложные предложения с временными клаузами;
- 3) сравнительные конструкции с несколькими объектами сравнения;
- 4) сложные предложения с объектными относительными клаузами;
- 5) предложения с ранним закрытием в причастном обороте;
- 6) предложения с поздним закрытием в причастном обороте.

Эти конструкции являются сложными и для детей, как показывают немногочисленные работы отдельно по каждому типу конструкций. Так, ключевым фактором, усложняющим интерпретацию пространственных конструкций, является нестандартный порядок слов. Порядок «обстоятельство места – дополнение» (см. (1a)) понимается младшими школьниками гораздо хуже, чем порядок «дополнение – обстоятельство места» (см. (1b)) [Statnikov, Fetiskina, 2012].

(1) а. *Мальчик положил в коробку шляпу.*

б. *Мальчик положил шляпу в коробку.*

Порядок смысловых частей оказывает большое влияние и на интерпретацию временных клауз (см. (2)). В исследованиях на материале английского языка было показано, что порядок клауз, соответствующий порядку событий, облегчает синтаксическую обработку у детей [Clark, 1971]. Выбор союза (*перед тем как / после того как*) также влияет на сложность синтаксической обработки [Ibid], однако этот тезис оспаривается в других работах [Townsend, Ravelo, 1980]. В работе [Фёдорова, 2005], где сравнивается понимание временных конструкций с союзами *перед тем как* и *после того как* и с бессоюзной связью у детей, подростков, молодых взрослых и людей среднего возраста, было показано, что дети совершают больше ошибок при интерпретации конструкций, чем другие группы участников. Также для детей, как и для более взрослых носителей, самым сложным оказался порядок *А перед тем, как В*, что может быть связано с низкой частотностью такой конструкции и ее неестественностью для русского языка.

(2) *Перед тем как папа вернулся домой, мама ушла в парк.*

Понимание сравнительных конструкций с несколькими объектами сравнения (см. (3)) у детей усложняет порядок объектов, не соответствующий реальному проявлению признака [Dragou et al., 2015].

(3) *Михаил старше Николая, но младше Ефрема.*

Объектные относительные клаузы (см. (4)) сочетают в себе несколько факторов, усложняющих интерпретацию предложения. Согласно работе [Rakhlin et al., 2016], даже в возрасте 10–11 лет школьники обрабатывают такие конструкции значительно хуже, чем взрослые. Для русскоязычных школьников наибольшую трудность представляют неканонический порядок слов, придаточное предложение, разрывающее главное, и несовпадение тематической роли определяемого слова и относительного местоимения [Rakhlin et al., 2016].

(4) *Старик, которого ровно в полночь ожидал на кладбище колдун, неподвижно стоял у ограды.*

Понимание предложений с ранним и поздним закрытием (см. (5a) и (5b) соответственно) значительно изменяется по мере взросления. Детям легче, чем взрослым, дается интерпретация предложений с поздним закрытием, т.к. она требует меньше ресурсов рабочей памяти. Русскоязычные взрослые, наоборот, предпочитают раннее закрытие и допускают значимо больше ошибок в предложениях с поздним закрытием [Чернова и др., 2016]. Известно, что переход к предпочтению раннего закрытия завершается уже в подростковом возрасте. Так, молодые взрослые чаще, чем подростки, пользуются синтаксической эвристикой при интерпретации причастного оборота (предпочтение раннего закрытия во всех случаях) [Lopukhina et al., 2022].

(5) а. *В штаб доложили о командире солдата, исполнявшем приказ.*

б. *Все сочувствовали переводчику дипломата, попавшего в неловкое положение.*

Итак, отдельные исследования показывают, как дети обрабатывают каждый тип конструкций и какие факторы влияют на обработку. Тем не менее, не существует работ, сравнивающих понимание детьми этих типов конструкций между собой. Цель нашего исследования – изучение синтаксической обработки наиболее сложных синтаксических конструкций у детей младшего школьного возраста. Мы хотим определить наиболее сложные конструкции для каждого из классов и сравнить понимание предложений у школьников разного возраста.

2. Метод

2.1. Участники

Участниками исследования стали 145 русскоязычных детей, учащихся 1–4 классов: 42 первоклассника (средний возраст 7,07, диапазон 6–8 лет), 35 второклассников (средний возраст 8,03, диапазон 7–9 лет), 35 третьеклассников (средний возраст 9,06, диапазон 8–10 лет) и 33 четвероклассника (средний возраст 10,09, диапазон 9–11 лет). Участники исследования были сгруппированы в соответствии с классом обучения, т.к. у учеников одного класса во многом совпадает читательский и когнитивный опыт, что в значительной степени влияет на синтаксическую обработку.

Согласно родительским анкетам, все участники эксперимента имели нормальные или скорректированные до нормальных зрение и слух и не имели речевых или неврологических нарушений.

Родители участников дали письменное согласие на прохождение эксперимента. Эксперимент был одобрен Комиссией по внутриуниверситетским опросам и этической оценке эмпирических исследовательских проектов Национального исследовательского университета «Высшая школа экономики».

2.2. Материалы

Стимулами в эксперименте стали 18 предложений из теста синтаксических компетенций для взрослых носителей русского языка [Chernova et al., 2023]. Среди них были все шесть типов из представленных в тесте конструкций: пространственные конструкции (см. (1a)), сложные предложения с временными клаузами (см. (2)), сравнительные конструкции с несколькими объектами сравнения (см. (3)), сложные предложения с объектными относительными клаузами (см. (4)), предложения с ранним закрытием (см. (5a)) и предложения с поздним закрытием (см. (5b)).

Пространственные конструкции были представлены простыми предложениями с предлогами *под* и *в*. Временные конструкции исследовались на материале сложных предложений с союзами *перед тем как* и *после того как*. Сравнительные конструкции были представлены только простыми предложениями, в каждой конструкции сравнивалось три объекта, которые назывались в порядке, не соответствующем проявлению признака в действительности. Конструкции с ранним и поздним закрытием изучались только на материале простых предложений, осложненных причастным оборотом.

Каждая конструкция была представлена тремя предложениями. Стимульные предложения имели длину от 6 до 13 слов и от 38 до 85 символов. В эксперимент не были включены тренировочные и филлерные предложения.

2.3. Процедура

Экспериментатор читал участникам стимульные предложения вслух и задавал вопрос с двумя вариантами ответа, например:

- (8) *В штаб доложили о командире солдата, исполнявшем приказ.
Кто исполнял приказ: командир или солдат?*

Школьники отвечали устно, их ответы письменно фиксировались.

Порядок предъявления стимулов был одинаковым для всех участников. Участники проходили эксперимент в индивидуальном порядке в тихом кабинете. Эксперимент продолжался примерно 15 минут, после прохождения эксперимента каждый школьник получил небольшой подарок.

2.4. Анализ

Сначала мы проанализировали среднюю точность ответов по каждой конструкции в каждом классе и ее стандартное отклонение. Средняя точность ответов была посчитана по всем ответам данного типа в данном классе. Далее мы посчитали для каждой конструкции среднюю точность по всем классам (среднее средних).

Чтобы сравнить правильность ответов в зависимости от типа конструкции, мы построили в среде R [R Core Team, 2024] смешанные модели логистической регрессии с помощью пакета glmer [Bates et al., 2015]. Зависимой дихотомической переменной являлась точность ответов (0/1), а независимыми переменными – тип конструкции, класс участника и их взаимодействие (фиксированные эффекты), а также участник и стимул (случайные эффекты). Далее мы провели попарные сравнения конструкций внутри классов и между ними с помощью критерия Тьюки [Tukey, 1949].

3. Результаты

Самая высокая точность ответов среди всех классов была отмечена для временных (средняя точность ответов – 71,3%; SD = 45,31%) и пространственных конструкций (60,23%; SD = 49%) (табл. 1). Самая низкая точность ответов была зафиксирована для сравнительных конструкций (47,59%; SD = 50%) и предложений с ранним закрытием (53,79%; SD = 49,91%).

Здесь и далее описываются результаты попарных сравнений с помощью критерия Тьюки [Tukey, 1949].

Самая низкая точность ответов у первоклассников была зафиксирована для сравнительных конструкций, она оказалась значимо меньше, чем для предложений с временными клаузами (Est. = 1,42, SE = 0,34, $z = 4,22$, $p < 0,001$) и поздним закрытием (Est. = 1,46, SE = 0,34, $z = 4,32$, $p < 0,001$) (рис. 1). Между другими типами конструкций не было обнаружено значимой разницы.

У второклассников самая высокая точность ответов была отмечена для сложных предложений с временными клаузами, а самая низкая – для предложений с ранним закрытием. Однако разница между всеми остальными конструкциями, кроме этих двух (Est. = 1,09, SE = 0,35, $z = 3,09$, $p = 0,02$), оказалась незначимой (рис. 2).

У третьеклассников не было значимой разницы между какими-либо конструкциями (рис. 3).

Таблица 1

Средняя точность понимания сложных синтаксических конструкций по классам, %
[The mean accuracy of understanding complex syntactic constructions in each grade, %]

	Классы [Grades]				Средняя точность ответов по конструкции [The mean accuracy of the answers for each construction]
	1	2	3	4	
Раннее закрытие [High attachment]	55,56 (SD = 49,89)	43,8 (SD = 49,85)	49,52 (SD = 50,24)	66,67 (SD = 47,38)	53,79 (SD = 49,91)
Позднее закрытие [Low attachment]	69,84 (SD = 46,08)	60 (SD = 49,22)	62,86 (SD = 48,55)	40,4 (SD = 49,32)	59,08 (SD = 49,23)
Сравнительные конструкции с несколькими объектами сравнения [Complex comparative constructions]	35,71 (SD = 48,11)	51,43 (SD = 50,22)	47,62 (SD = 50,19)	58,59 (SD = 49,51)	47,59 (SD = 50)
Пространственные конструкции [Spatial constructions]	53,17 (SD = 50,1)	57,14 (SD = 49,72)	64,76 (SD = 48)	67,68 (SD = 47)	60,23 (SD = 49)
Сложные предложения с временными клаузами [Complex sentences with temporal clauses]	69,05 (SD = 46,41)	69,52 (SD = 46,25)	68,57 (SD = 46,65)	78,79 (SD = 41)	71,26 (SD = 45,31)
Сложные предложения с объектными относительными клаузами [Complex sentences with object relative clauses]	48,41 (SD = 50,17)	59,05 (SD = 49,41)	61,9 (SD = 48,8)	62,63 (SD = 48,63)	57,47 (SD = 49,5)

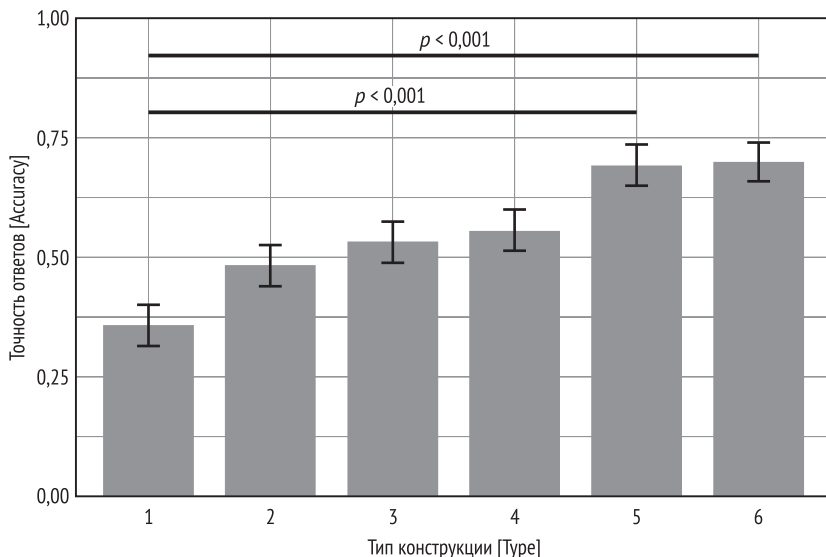


Рис. 1. Понимание сложных синтаксических конструкций учениками 1 класса (точность ответов в эксперименте):

- 1 – сравнительные конструкции с несколькими объектами сравнения;
- 2 – сложные предложения с объектными относительными клаузами;
- 3 – пространственные конструкции; 4 – предложения с ранним закрытием;
- 5 – сложные предложения с временными клаузами; 6 – предложения с поздним закрытием

На концах линий над столбцами расположены типы конструкций, между которыми обнаружена значимая разница по точности ответов

Fig. 1. Understanding complex syntactic constructions by 1st grade students (the accuracy of the answers in the experiment):

- 1 – multiple comparative constructions; 2 – complex sentences with object relative clauses; 3 – spatial constructions; 4 – high attachment; 5 – complex sentences with temporal clauses; 6 – low attachment

At the line ends above the columns there are the types of constructions with a significant difference in the answer accuracy found between them

У четвероклассников самая низкая точность ответов была обнаружена для предложений с поздним закрытием: она оказалась значимо меньше, чем в предложениях с ранним закрытием (Est. = 1,1, SE = 0,36, $z = 3,06$, $p = 0,03$), пространственных (Est. = 1,15, SE = 0,36, $z = 3,19$, $p = 0,02$) и временных конструкциях (Est. = 1,73, SE = 0,38, $z = 4,57$, $p < 0,001$) (рис. 4).

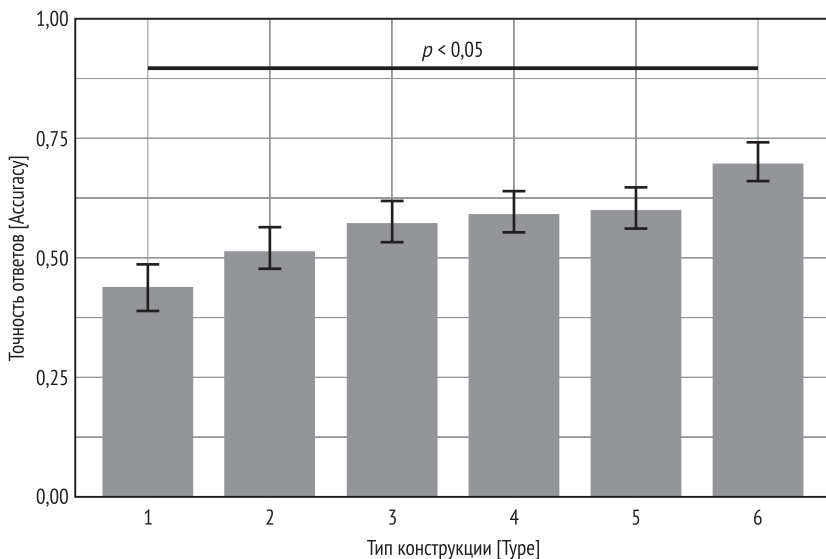


Рис. 2. Понимание сложных синтаксических конструкций учениками 2 класса (точность ответов в эксперименте):

- 1 – сравнительные конструкции с несколькими объектами сравнения;
- 2 – сложные предложения с объектными относительными клаузами;
- 3 – пространственные конструкции;
- 4 – предложения с ранним закрытием;
- 5 – сложные предложения с временными клаузами;
- 6 – предложения с поздним закрытием

На концах линий над столбцами расположены типы конструкций, между которыми обнаружена значимая разница по точности ответов

Fig. 2. Understanding complex syntactic constructions by 2nd grade students (the accuracy of the answers in the experiment):

- 1 – multiple comparative constructions;
- 2 – complex sentences with object relative clauses;
- 3 – spatial constructions; 4 – high attachment;
- 5 – complex sentences with temporal clauses; 6 – low attachment

At the line ends above the columns there are the types of constructions with a significant difference in the answer accuracy found between them

При попарных сравнениях каждого типа конструкций между классами мы не обнаружили никакой разницы между какими-либо классами для временных и пространственных конструкций; сложных предложений с объектными относительными клаузами.

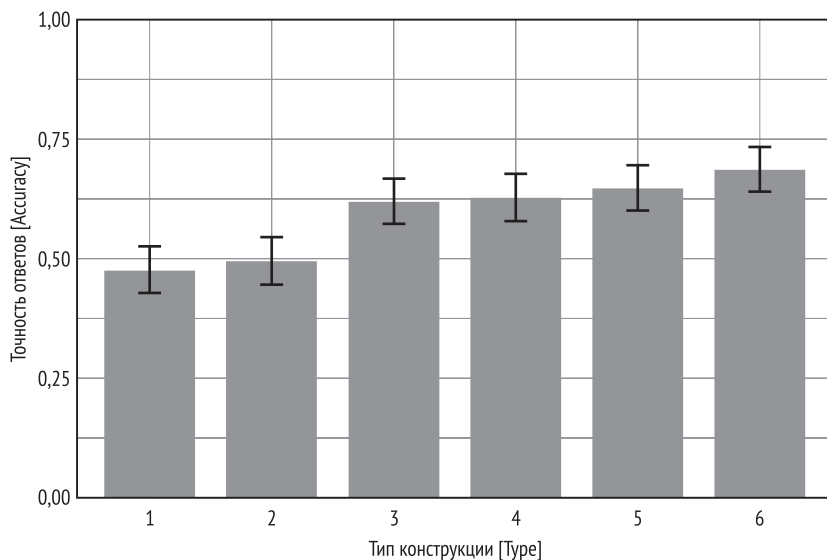


Рис. 3. Понимание сложных синтаксических конструкций учениками 3 класса (точность ответов в эксперименте):

- 1 – сравнительные конструкции с несколькими объектами сравнения;
- 2 – сложные предложения с объектными относительными клаузами;
- 3 – пространственные конструкции; 4 – предложения с ранним закрытием;
- 5 – сложные предложения с временными клаузами; 6 – предложения с поздним закрытием

Fig. 3. Understanding complex syntactic constructions by 3rd grade students (the accuracy of the answers in the experiment):

- 1 – multiple comparative constructions; 2 – complex sentences with object relative clauses; 3 – spatial constructions; 4 – high attachment; 5 – complex sentences with temporal clauses; 6 – low attachment

Мы нашли значимую разницу в обработке сравнительных конструкций между первым и четвертым классами: четвероклассники справлялись с ними лучше, чем первоклассники ($Est. = 0,95$, $SE = 0,28$, $z = 3,4$, $p = 0,004$).

Между этими же классами была обнаружена разница в обработке конструкций с поздним закрытием: первоклассники, наоборот, справлялись с ней лучше, чем четвероклассники ($Est. = 1,25$, $SE = 0,29$, $z = 4,38$, $p < 0,001$). Второклассникам ($Est. = 0,81$, $SE = 0,29$, $z = 2,8$, $p = 0,03$) и третьеклассникам ($Est. = 0,93$, $SE = 0,29$, $z = 3,2$, $p = 0,008$) также было проще обработать конструкции с поздним закрытием по сравнению с четвероклассниками.

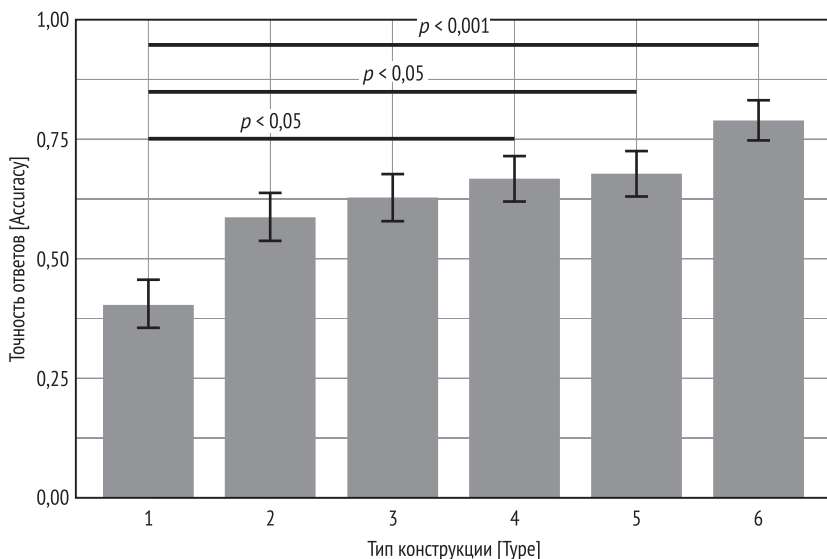


Рис. 4. Понимание сложных синтаксических конструкций учениками 4 класса (точность ответов в эксперименте):

- 1 – сравнительные конструкции с несколькими объектами сравнения;
- 2 – сложные предложения с объектными относительными клаузами;
- 3 – пространственные конструкции; 4 – предложения с ранним закрытием;
- 5 – сложные предложения с временными клаузами; 6 – предложения с поздним закрытием

На концах линий над столбцами расположены типы конструкций, между которыми обнаружена значимая разница по точности ответов

Fig. 4. Understanding complex syntactic constructions by 4th grade students (the accuracy of the answers in the experiment):

- 1 – multiple comparative constructions;
- 2 – complex sentences with object relative clauses; 3 – spatial constructions;
- 4 – high attachment; 5 – complex sentences with temporal clauses;
- 6 – low attachment

At the line ends above the columns there are the types of constructions with a significant difference in the answer accuracy found between them

Напротив, четвероклассники лучше справлялись с конструкциями с ранним закрытием по сравнению со второклассниками ($Est. = 0,95$, $SE = 0,29$, $z = 3,26$, $p = 0,006$). Однако разница в обработке данного типа конструкций у четвероклассников с первоклассниками и третьеклассниками оказалась незначима.

4. Обсуждение

В данном исследовании изучалось понимание сложных синтаксических конструкций младшими школьниками. Задачами нашей работы были выявление наиболее сложных синтаксических конструкций для детей младшего школьного возраста и сравнение понимания предложений у школьников разных классов.

Для первого класса самым простым типом оказались сложные предложения с поздним закрытием, а самым сложным – сравнительные конструкции. Статистически значимая разница была зафиксирована между этими типами конструкций, а также между временными и сравнительными конструкциями. Для второго класса, согласно описательной статистике, самыми простыми конструкциями оказались временные клаузы, а самыми сложными – предложения с ранним закрытием. Статистически значимая разница была отмечена только между этими типами конструкций. Для третьеклассников, по описательной статистике, самым простым типом оказались также предложения с временными клаузами, а самым сложным – сравнительные конструкции, однако статистически значимой разницы между какими-либо конструкциями не было найдено. Наконец, для четвероклассников самыми простыми оказались также сложные предложения с придаточными времени, а самыми сложными – конструкции с поздним закрытием. Предложения с поздним закрытием были для них значимо труднее, чем предложения с придаточными времени, пространственные конструкции и предложения с ранним закрытием.

Одной из самых простых конструкций для всех школьников оказались сложные предложения с временными клаузами. Описательно для них была зафиксирована самая высокая точность ответов во 2–4 классах, статистически значимой разницы в точности ответов между классами не было обнаружено. Результаты школьников совпали с результатами взрослых: в исследовании [Chernova et al., 2023] для временных конструкций была отмечена самая высокая точность ответов. Это может быть связано с особенностями стимульного материала: все предложения эксперимента начинались с временного союза *перед тем как* или *после того как*. В исследовании на материале финского языка [Ruukkönen, Järvikivi, 2012] было обнаружено, что дети 8–12 лет лучше понимают предложения, когда временной союз расположен в начале предложения, а не в его середине. По мнению авторов, в этом случае дети затрачивают меньше когнитивных ресурсов на построение и корректировку ментальной репрезентации описанной в предложении ситуации. Также на результаты могло повлиять, что в одной из трех временных конструкций в эксперименте описанные события шли в хронологическом порядке, что значимо облегчило понимание [Clark, 1971].

Для пространственных конструкций при сравнении между классами не было обнаружено значимой разницы между школьниками разного возраста. В других исследованиях, однако, утверждается, что анализ пространственных и других логико-грамматических конструкций существенно изменяется по мере взросления. Так, в работе [Ахутина и др., 2017] описывается переход от более детской стратегии интерпретации таких конструкций (опора только на порядок слов) к более взрослой (опора и на порядок слов, и на грамматические маркеры) у учеников первого и второго классов, причем для первого класса более характерно использование детской стратегии.

Для объектных относительных клауз также не было найдено значимой разницы в понимании при сравнении между классами. Согласно исследованию [Rakhlina et al., 2016], уже в шесть лет дети обладают грамматическими навыками, требуемыми для интерпретации объектных относительных клауз, и используют те же стратегии анализа, что и взрослые, однако не всегда верно обрабатывают эти конструкции из-за нехватки прагматических навыков.

Мы зафиксировали изменения в интерпретации сравнительных конструкций по мере взросления. При сравнении между классами было выявлено, что четвероклассники интерпретируют такие конструкции значительно лучше, чем первоклассники. Одной из ключевых причин сложности этих конструкций для первоклассников может быть неканонический порядок слов: в работе [Boyle et al., 2013] отмечается, что этот фактор представляет особую трудность для детей 4–6,5 лет. Другой причиной трудности сравнительных конструкций для первоклассников может быть небольшой объем рабочей памяти. Согласно исследованию [Dragoy et al., 2015], сравнительные конструкции такого типа схожи с конструкциями с пространственными предлогами по механизмам обработки. Однако сравнительные конструкции представляют большую сложность для интерпретации, т.к. испытуемому требуется расставить в правильном порядке три объекта, а не два, как в пространственных конструкциях. Увеличение объема рабочей памяти по мере взросления [Chi, 1976] позволяет четвероклассникам лучше справиться с этим заданием. Еще одним фактором, влияющим на понимание сравнительных конструкций первоклассниками, может быть неполная освоенность конструкций такого типа в 6–7 лет, в то время как навыки понимания и порождения сравнительных конструкций у четвероклассников сравнимы с навыками взрослых [Syrett, 2016].

При сравнении между классами было выявлено, что ученики 1–3 класса справляются с поздним закрытием значительно лучше, чем четвероклассники; при этом четвероклассники понимают раннее закрытие значимо

лучше, чем второклассники (значимой разницы с первоклассниками и третьеклассниками не было зафиксировано). Наши результаты схожи с результатами экспериментов, описанных в [Фёдорова, 2020], согласно которым русскоязычные дети 5–6 лет предпочитают позднее закрытие, а дети 10–12 лет – раннее. Предложения с поздним закрытием также были самыми трудными для взрослых в тесте синтаксических компетенций в исследовании [Chernova et al., 2023]. Авторы связывают это с приоритетом раннего закрытия в русском языке [Чернова, 2015], а также с низкой частотностью такой конструкции, наличием причастного оборота и необходимостью анализировать трудный для обработки падежный показатель, чтобы получить верную интерпретацию. Кроме того, на понимание таких предложений может влиять высокая беглость чтения у более старших школьников и взрослых, вследствие которой носители обращают меньше внимания на окончания и ошибочно причисляют конструкции с поздним закрытием к более частотным конструкциям с ранним закрытием. Таким образом, наши результаты могут свидетельствовать о переходе к более взрослой стратегии понимания предложений у четвероклассников. Схожие результаты описаны в работе [Felser et al., 2003]: выбор раннего или позднего закрытия у англоязычных детей 6–7 лет зависит от объема рабочей памяти, в то время как интерпретация таких конструкций детьми 10–11 лет совпадает с интерпретацией у взрослых носителей. Согласно исследованию [Friederici, 1983], именно в 10–11 лет достигается обработка синтаксической информации отдельно от семантической, что также может влиять на анализ конструкций с ранним и поздним закрытием.

Итак, мы обнаружили, что по мере взросления происходит переход ко взрослой стратегии обработки конструкций с ранним и поздним закрытием и к более эффективной обработке сравнительных конструкций. Было выявлено, что одной из самых простых конструкций для всех классов являются временные клаузы.

В связи с тем, что настоящее исследование было частью более крупного исследования чтения, в котором описанный тест применялся для оценки синтаксических компетенций, оно имеет ряд ограничений, которые мы собираемся учесть в следующем исследовании.

Так, планируется улучшить статистическую мощность эксперимента – увеличить количество стимулов и участников. В будущем мы планируем использовать другой дизайн, чтобы снизить вероятность угадывания ответа школьниками: участники эксперимента будут выбирать из нескольких предложенных картинок одну, соответствующую смыслу предложения.

Наконец, перспективой работы также является проведение лонгитудного исследования, в котором можно было бы проследить развитие синтаксических навыков каждого школьника по мере взросления, и разработка стандартизированного теста для оценки синтаксических компетенций младших школьников.

Библиографический список / References

Ахутина и др., 2017 – Ахутина Т.В., Корнеев А.А., Матвеева Е.Ю. Возрастная динамика понимания логико-грамматических конструкций у младших школьников и ее мозговые механизмы // *Специальное образование*. 2017. № 3. С. 15–31. [Akhutina T.V., Korneev A.A., Matveeva E.Ju. The developmental evolution of comprehension of the logical-grammatical constructions in primary school children and its brain principles. *Specialnoe obrazovanie*. 2017. No. 3. Pp. 15–31. (In Rus.)]

Фёдорова, 2005 – Фёдорова О.В. ПЕРЕД или ПОСЛЕ: что проще? (понимание сложноподчиненных предложений с придаточными времени) // *Вопросы языкознания*. 2005. № 6. С. 44–58. [Fedorova O.V. BEFORE or AFTER: What is simpler? (The understanding of complex sentences with temporal clauses). *Voprosy yazykoznaniya*. 2005. No. 6. Pp. 44–58. (In Rus.)]

Фёдорова, 2020 – Фёдорова О.В. Психоллингвистика. М., 2020. [Fedorova O.V. *Psiholingvistika* [Psycholinguistics]. Moscow, 2020.]

Чернова, 2015 – Чернова Д.А. Синтаксический анализ предложения в процессе восприятия речи: экспериментальное исследование обработки синтаксически неоднозначных конструкций в русском языке // *Вестник Пермского университета. Российская и зарубежная филология*. 2015. № 1 (29). С. 36–44. [Chernova D.A. Syntactic analysis of sentence during the speech perception: Experimental study of the processing of syntactically ambiguous constructions in Russian. *Vestnik Permskogo universiteta. Rossiiskaya i zarubezhnaya filologiya*. 2015. No. 1 (29). Pp. 36–44. (In Rus.)]

Чернова и др., 2016 – Экспериментальные исследования грамматики: синтаксический анализ неоднозначных предложений / Д.А. Чернова, Н.А. Слюсарь, В.К. Прокопья и др. // *Вопросы языкознания*. 2016. № 6. С. 36–50. [Chernova D.A., Slyusar N.A., Prokopenya V.K. et al. Experimental studies of the grammar: Syntactic analyses of ambiguous sentences. *Voprosy yazykoznaniya*. 2016. No. 6. Pp. 36–50. (In Rus.)]

Bates et al., 2015 – Bates D., Mächler M., Bolker B., Walker S. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 2015. Vol. 67. No. 1. Pp. 1–48.

Boyle et al., 2013 – Boyle W., Lindell A.K., Kidd E. Investigating the role of verbal working memory in young children's sentence comprehension. *Language Learning*. 2013. Vol. 63. No. 2. Pp. 211–242.

Chernova et al., 2023 – Chernova D., Novozhilov A., Slioussar N. Sentence comprehension test for Russian: A tool to assess syntactic competence. *Frontiers in Psychology*. 2023. Vol. 14.

Chi, 1976 – Chi M.T.H. Short-term memory limitations in children: Capacity or processing deficits? *Memory & Cognition*. 1976. Vol. 4. No. 5. Pp. 559–572.

Clark, 1971 – Clark H.H. The primitive nature of children's relational concepts. *Cognition and the development of language*. J.R. Hayes (ed.). New York, 1971. Pp. 269–278.

Dragoy et al., 2015 – Dragoy O., Bergelson M., Iskra E. et al. Comprehension of reversible constructions in semantic aphasia. *Aphasiology*. 2015. No. 30 (1). Pp. 1–22.

Felser et al., 2003 – Felser C., Marinis T., Clahsen H. Children's processing of ambiguous sentences: a study of relative clause attachment. *Language Acquisition*. 2003. Vol. 11. No. 3. Pp. 127–163.

Friederici, 1983 – Friederici A.D. Children's sensitivity to function words during sentence comprehension. *Linguistics*. 1983. Vol. 21 (5). Pp. 717–739.

Lopukhina A. et al., 2022 – Lopukhina A., Laurinavichyute A., Malyutina S. et al. Reliance on semantic and structural heuristics in sentence comprehension across the lifespan. *Quarterly Journal of Experimental Psychology*. 2022. Vol. 75. No. 7. Pp. 1367–1381.

Poulsen et al., 2016 – Poulsen M., Gravgaard A.K.D. Who did what to whom? The relationship between syntactic aspects of sentence comprehension and text comprehension. *Scientific Studies of Reading*. 2016. Vol. 20. No. 4. Pp. 325–338.

Pyykkönen, Järvikivi, 2012 – Pyykkönen P., Järvikivi J. Children and situation models of multiple events. *Developmental Psychology*. 2012. Vol. 48. No. 2. Pp. 521–529.

R Core Team, 2024 – R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2024.

Rakhlin et al., 2016 – Rakhlin N., Kornilov S., Kornilova T., Grigorenko E. Syntactic complexity effects of Russian relative clause sentences in children with and without developmental language disorder. *Language Acquisition*. 2016. No. 23 (4). Pp. 333–360.

Roberts, 2016 – Roberts L. Syntactic processing. *The Cambridge Handbook of Linguistic Multi-Competence*. Cambridge Handbooks in Language and Linguistics. V. Cook, L. Wei (eds.). Cambridge University Press, 2016. Pp. 227–247.

Statnikov, Fetiskina, 2012 – Statnikov A., Fetiskina V. Understanding logical-grammatical constructions: Hierarchy of complexity. *Moscow International Congress dedicated to the 110th anniversary of A.R. Luria's birth*. Book of abstracts. 2012.

Statnikov, Akhutina, 2013 – Statnikov A., Akhutina T. Logical-grammatical constructions comprehension and serial organization of speech: Finding the link using computer-based tests. *Procedia-Social and Behavioral Sciences*. 2013. No. 86. Pp. 518–523.

Syrett, 2016 – Syrett K. Acquisition of comparative and degree constructions. *The Oxford Handbook of Developmental Linguistics*. J. Lidz, W. Snyder, J. Pater. (eds.). Oxford University Press, 2016. Pp. 463–497.

Townsend, Ravelo, 1980 – Townsend D.J., Ravelo N. The development of complex sentence processing strategies. *Journal of Experimental Child Psychology*. 1980. Vol. 29. No. 1. Pp. 60–73.

Traxler, 2002 – Traxler M.J. Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. *The Quarterly Journal of Experimental Psychology: Section A*. 2002. No. 55 (1). Pp. 75–96.

Tukey, 1949 – Tukey J.W. Comparing individual means in the analysis of variance. *Biometrics*. 1949. Vol. 5. No. 2. Pp. 99–114.

Wu et al., 2016 – Wu C.Y., Vissienon K., Friederici A.D., Brauer J. Preschoolers' brains rely on semantic cues prior to the mastery of syntax during sentence comprehension. *NeuroImage*. 2016. Vol. 126. Pp. 256–266.

Статья поступила в редакцию 19.08.2024

The article was received on 19.08.2024

Сведения об авторах / About the authors

Пилипец Полина Андреевна – студент отделения теоретической и прикладной лингвистики филологического факультета, Московский государственный университет имени М.В. Ломоносова

Polina A. Pilipets – student at the Department of Theoretical and Applied Linguistics of the Philological Faculty, Lomonosov Moscow State University, Moscow, Russian Federation

ORCID: <https://orcid.org/0009-0008-9397-8273>

E-mail: polina.pilipets@list.ru

Староверова Владислава Николаевна – младший научный сотрудник Центра языка и мозга, Национальный исследовательский университет «Высшая школа экономики», г. Москва

Vladislava N. Staroverova – junior research fellow at the Center for Language and Brain, HSE University, Moscow, Russian Federation

ORCID: <https://orcid.org/0000-0002-4856-9429>

E-mail: vstaroverova@hse.ru

Все авторы прочитали и одобрили окончательный вариант рукописи

All authors have read and approved the final manuscript

Издание
подготовили
к печати:
редактор
А. А. Козаренко,
корректор
А. А. Алексеева,
обложка, макет,
компьютерная
верстка
Н. А. Попова

Rhema. Рема

2024.3

Сайт журнала: rhema-journal.com
E-mail: rhema.pema@gmail.com

Авторы статей несут полную ответственность за точность приводимой информации, цитат, ссылок и списка литературы.

Перепечатка материалов, опубликованных в журнале, невозможна без письменного разрешения редакции.

Подписано в печать 29.09.2024.
Формат 60×90 1/16. Гарнитура «Liberation Serif».
Объем 8,13 п. л. Тираж 1000 экз.