

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Институт русского языка имени В. В. Виноградова

RUSSIAN ACADEMY OF SCIENCES
V. V. Vinogradov Russian Language Institute

РОССИЙСКАЯ АКАДЕМИЯ НАУК
Институт русского языка имени В. В. Виноградова

Труды
Института русского языка
им. В. В. Виноградова

№ 4

МОСКВА
2024

ISSN 2311-150X

Журнал основан в 2013 г.

Главный редактор

А. М. Молдован, доктор филол. наук, академик РАН (Москва, Россия)

Ответственный редактор номера:

В. А. Плунгян, академик РАН, профессор (Москва, Россия)

Редакционная коллегия номера:

Е. В. Рахилина, доктор филол. наук (Москва, Россия);

С. О. Савчук, канд. филол. наук (Москва, Россия)

Редакционный совет:

А. Е. Аникин, доктор филол. наук, академик РАН (Новосибирск);

Ю. Д. Апресян, доктор филол. наук, академик РАН, профессор (Москва, Россия);

Синтия Вакарелийская, PhD, профессор (Орегон, США);

Ж. Ж. Варбот, доктор филол. наук, профессор (Москва, Россия);

Бьёрн Вимер, доктор филологии, профессор (Майнц, Германия);

Марчелло Гардзанини, PhD, профессор (Флоренция, Италия);

А. А. Гиппиус, доктор филол. наук, академик РАН, профессор (Москва, Россия);

Ольга Йокояма, PhD, профессор (Лос-Анджелес, США);

М. Л. Каленчук, доктор филол. наук, профессор (Москва, Россия);

А. П. Майоров, доктор филол. наук, профессор (Улан-Удэ, Россия);

Ольга Младенова, PhD, профессор (Калгари, Канада);

Туре Нессет, доктор филологии, профессор (Тромсё, Норвегия);

В. А. Плунгян, доктор филол. наук, академик РАН, профессор (Москва, Россия);

Ахим Рабус, Dr. habil., профессор (Фрайбург, Германия);

Ф. Б. Успенский, доктор филол. наук, член-корреспондент РАН, профессор (Москва, Россия);

Майкл Флайер, PhD, профессор (Кембридж, США);

Вацлав Чермак, доктор филологии (Прага, Чехия);

А. Д. Шмелев, доктор филол. наук, член-корреспондент РАН, профессор (Москва, Россия)

Ответственный секретарь

канд. филол. наук А. Е. Журавлёва

Выходит 4 раза в год

Издание зарегистрировано Федеральной службой по надзору
в сфере связи, информационных технологий и массовых коммуникаций.

Свидетельство о регистрации ПИ № ФС 77-76037

Адрес редакции:

119019, Москва, ул. Волхонка, д. 18/2

E-mail: ruslang@ruslang.ru

© Институт русского языка им. В. В. Виноградова РАН, 2024

© Авторы, 2024

RUSSIAN ACADEMY OF SCIENCES
V. V. Vinogradov Russian Language Institute

**Proceedings
of the V. V. Vinogradov
Russian Language Institute**

№ 4

MOSCOW
2024

ISSN 2311-150X

The Journal was founded in 2013

Editor-in-Chief

Alexandr M. Moldovan, D. Sc., Full Member of the RAS (Moscow, Russia)

Chief editor of the issue

Vladimir A. Plungian, D.Sc., Professor, academician of the RAS (Moscow, Russia)

Editorial Board of the issue

Ekaterina V. Rakhilina, D. Sc. (Moscow, Russia);
Svetlana O. Savchuk, Ph. D. (Moscow, Russia)

Editorial Board

Aleksandr E. Anikin, D. Sc., Full Member of the RAS (Novosibirsk, Russia);
Yury D. Apresyan, D. Sc., Full Member of the RAS, Professor (Moscow, Russia);
Václav Čermák, Ph. D., Professor (Prague, Czech Republic);
Michael S. Flier, Ph. D., Professor (Cambridge, USA);
Marcello Garzaniti, D. Sc., Professor (Florence, Italy);
Alexey A. Gippius, D. Sc., Full Member of the RAS, Professor (Moscow, Russia);
Maria L. Kalenchuk, D. Sc., Professor (Moscow, Russia);
Alexandr P. Mayorov, D. Sc., Professor (Ulan-Ude, Russia);
Olga Mladenova, Ph. D., Professor em. (Calgary, Canada);
Tore Nasset, D. Sc., Professor (Tromsø, Norway);
Vladimir A. Plungian, D. Sc., Full Member of the RAS, Professor (Moscow, Russia);
Achim Rabus, D. Sc., Professor (Freiburg, Germany);
Alexey D. Shmelev, D. Sc., Corresponding Member of the RAS, Professor (Moscow, Russia);
Fjodor B. Uspensky, D. Sc., Corresponding Member of the RAS, Professor (Moscow, Russia);
Cynthia M. Vakareliyska, Ph. D., Professor (Oregon, USA);
Zhanna Zh. Varbot, D. Sc., Professor (Moscow, Russia);
Björn Wiemer, D. Sc., Professor (Mainz, Germany);
Olga T. Yokoyama, Ph. D., Distinguished Professor (Los Angeles, USA)

Executive secretary

Alexandra E. Zhuravleva

The journal is registered by the Federal Service for Supervision
of Communications, Information Technology, and Mass-Media.

Registration certificate ПИ № ФС 77-76037

Address:

18/2, Volkhonka street, Moscow, 119019

E-mail: ruslang@ruslang.ru

СОДЕРЖАНИЕ

<i>Плунгян В. А.</i>	
Предисловие.....	11

I. Национальный корпус русского языка: исследования и разработки

<i>Масевич А. Ц.</i>	
Памяти В. П. Захарова.....	13
<i>Масевич А. Ц., [Захаров В. П.]</i> (Санкт-Петербург)	
Разметка слов-цветообозначений в Национальном корпусе русского языка.....	15
<i>Маркасова Е. В.</i> (Пекин, Китай), <i>Митрофанова О. А.</i> (Санкт-Петербург)	
Геминация (тройной повтор) по корпусным данным.....	41
<i>Добровольский Д. О., Анна А. Зализняк</i> (Москва)	
Использование параллельного корпуса в семантическом анализе русской лексики: case study <i>приспичило</i> и <i>норовить</i>	60
<i>Хохлова М. В.</i> (Санкт-Петербург)	
Высокочастотные глагольные словарные коллокации с существительными в Национальном корпусе русского языка	78
<i>Рахилина Е. В., Плунгян В. А.</i> (Москва)	
ПРЕЛЕСТЬ: История слова.....	89
<i>Рыжаченков И. И.</i> (Москва)	
Толк и толки: диахронический портрет имени с семантикой ментальной сферы.....	102
<i>Кустова Г. И.</i> (Москва)	
О некоторых свойствах неопределенно-количественных выражений с местоименным компонентом	119
<i>Богуславский И. М., Дяченко П. В., Иншакова Е. С., [Иомдин Л. Л.],</i> <i>Лазурский А. В., Митюшин Л. Г., Мовсисян А. А., Рыгаев И. П.,</i> <i>Сизов В. Г., Тимошенко С. П., Фролова Т. И., Чага А. В.</i> (Москва)	
Современное состояние корпуса СинТагРус.....	141
<i>Савчук С. О.</i> (Москва)	
Обучающий корпус в составе НКРЯ как инструмент учителя-словесника	170
<i>Бонч-Осмоловская А. А.</i> (Москва)	
Дистрибутивные модели в проекции теоретических подходов лексической семантики (по данным НКРЯ)	200

II. Слово и жест: вопросы мультимодальной коммуникации

<i>Янко Т. Е.</i> (Москва)	
Просодия вопроса при опущении показателя иллюкутивной силы.....	243
<i>Федорова О. В.</i> (Москва)	
«Репортажи о грушах»: итоги первых исследований и дальнейшие перспективы	258

<i>Киосе М. И., Леонтьева А. В., Агафонова О. В.</i> (Москва)	
Полиmodalное конструирование референтов в экспозиторном русскоязычном дискурсе: СЛОВО vs. ОБЪЕКТ.....	273
<i>Переверзева С. И.</i> (Москва)	
Невербальная манипуляция и проблемы её выявления: жест «положить (свою) руку на (чью-либо) руку»	293
<i>Шестакова Л. Л., Кулева А. С.</i> (Москва)	
Слово жест в поэтическом дискурсе	311

III. Рецензии

<i>Хохлова М. В.</i>	
[Рец. на / Rewiew of:] Doing Linguistics with a Corpus. Methodological Considerations for the Everyday User. Jesse Egbert, Tove Larsson, Douglas Biber. Cambridge University Press, 2020. 88 p.....	330

CONTENTS

<i>Plungian V. A.</i>	
Preface	11

I. The Russian National Corpus: Research and Development

<i>Andrei Ts. Masevich.</i>	
In memory of V. P. Zakharov	13
<i>Andrei Ts. Masevich, [Victor P. Zakharov]</i> (Saint Petersburg)	
Tagging of colour words in the Russian National Corpus	15
<i>Markasova E. V.</i> (Beijing, China), <i>Mitrofanova O. A.</i> (Saint Petersburg)	
Gemination (triple repeat) according to the corpus data	41
<i>Dobrovol'skij D. O., Anna A. Zalizniak</i> (Moscow)	
Parallel corpus in the semantic analysis of Russian lexical units: case study <i>prispichilo</i> 'have an urge' and <i>norovit</i> 'to be eager'	60
<i>Maria V. Khokhlova</i> (Saint Petersburg)	
High-frequency verb-noun collocations from dictionaries in the Russian National Corpus	78
<i>Rakhilina E. V., Plungian V. A.</i> (Moscow)	
PRELEST': The history of word	89
<i>Ryzhachenkov I. I.</i> (Moscow)	
<i>Tolk</i> and <i>tolki</i> : diachronic portrait of a noun with the semantics of the mental sphere	102
<i>Kustova G. I.</i> (Moscow)	
On some properties of indefinite-quantitative expressions with a pronoun component	119
<i>Igor M. Boguslavsky, Alexandra V. Chaga, Pavel V. Djachenko,</i> <i>Tatyana I. Frolova, Evgenia S. Inshakova, [Leonid L. Iomdin],</i> <i>Alexander V. Lazurski, Leonid G. Mityushin, Andrey A. Movsesyan,</i> <i>Ivan P. Rygaev, Victor G. Sizov, Svetlana P. Timoshenko</i> (Moscow)	
The current state of the SynTagRus corpus	141
<i>Savchuk S. O.</i> (Moscow)	
The Educational Corpus of the RNC as valuable resource for educators	170
<i>Bonch-Osmolovskaya A. A.</i>	
Distributive models in the projection of theoretical approaches of lexical semantics (according to the RNC)	200

II. Word and Gesture: Multimodal communication issues

<i>Yanko T. E.</i> (Moscow)	
Prosody of a question when omitting the illocutinary proper component	243
<i>Fedorova Olga</i> (Moscow)	
"Pears reportages": the results of the first researches and further perspectives	258
<i>Kiose M. I., Leonteva A. V., Agafonova O. V.</i> (Moscow)	
Multimodal construal of referents in expository discourse: WORD vs. OBJECT	273

Pereverzeva S. I. (Moscow)

Nonverbal manipulation and the problems of its detection: the case of the gesture
“to lay one’s hand on somebody’s hand”293

Shestakova L. L., Kuleva A. S. (Moscow)

Gesture in poetic discourse311

III. Reviews

Maria V. Khokhlova

[Review of:] *Doing Linguistics with a Corpus. Methodological Considerations
for the Everyday User.* Jesse Egbert, Tove Larsson, Douglas Biber.
Cambridge University Press, 2020. 88 p.330

ПРЕДИСЛОВИЕ

В настоящем выпуске «Трудов Института русского языка им. В. В. Виноградова» читатель найдет две подборки статей, связанных с Национальным корпусом русского языка. Публикации на тему Корпуса, отражающие как этапы его поступательного развития, так и полученные с его помощью результаты исследований, — наша традиционная тематика (последняя по времени публикация появилась в «Трудах» № 2 (32) за 2022 г.), но те статьи, что печатаются сегодня, имеют важную особенность: они отражают очень существенные изменения, произошедшие в составе и функционале Корпуса как раз в течение 2022–2023 годов. За это время состав Корпуса был существенно пополнен (в настоящее время он превышает 2 млрд словоупотреблений и включает целый ряд новых модулей), а архитектура — модернизирована, в том числе с применением нейросетевых технологий, позволивших внедрить в Корпус современные средства дистрибутивно-статистического анализа лексики. Существенный шаг вперед был также сделан в отношении инструментов визуализации данных, ставших сегодня привычным элементом крупных корпусов во всем мире. Подробнее об этих нововведениях можно прочесть в обстоятельной статье [Савчук и др. 2024], а в публикуемой ниже подборке представлены как описания тех изменений, которые произошли с Корпусом в недавнее время, так и — даже в большей степени — результаты исследований, полученных с помощью Корпуса (в том числе — и в его нынешнем обновленном виде, с учетом возросшего объема и более разнообразного и гибкого функционала).

Статьи настоящего номера разделены на две части. В первой части собраны работы, непосредственно отражающие работу над отдельными корпусными модулями и представляющие научные результаты корпусных исследований русской грамматики и лексики (впрочем, в некоторых статьях оба формата сочетаются). Описание отдельных проектов по развитию Корпуса читатель найдет в статье большого коллектива авторов (И. М. Богуславский и др.), касающейся современного состояния экспериментального корпуса разметки синтаксических зависимостей СинТагРус, созданного в ИППИ РАН и претерпевшего за последние годы ряд значимых изменений. Другой публикацией этого жанра является статья С. О. Савчук, в которой дается характеристика относительно нового (и важного) для Корпуса ресурса — Обучающего корпуса, предназначенного для работы с текстами школьной программы по русскому языку и литературе; впрочем, как показано в статье, этот подкорпус имеет не только педагогическое значение, но может с успехом применяться и в исследованиях, например, русского языка XIX века. Особенно интересной для тех, кто следит за обновлениями корпусных инструментов, может оказаться статья А. А. Бонч-Осмоловской, подробно обсуждающая так называемый «портрет слова» — новый лексикологический ресурс, основанный на применении векторных моделей.

Разные подходы к семантической разметке корпуса (на материале слов-цветообозначений) обсуждаются в статье А. Ц. Масевича и В. П. Захарова¹.

«Исследовательская» подборка статей в этой части достаточно разнообразна. Отметим традиционные для корпусного подхода синхронные и диахронические исследования лексики: полисемичной лексемы *прелесть* и ее более позднего французского аналога *шарм* (Е. В. Рахилина, В. А. Плунгян), интересный и неожиданный диахронический анализ лексем *толк* и *толки* (И. И. Рыжаченков), анализ русских лингвоспецифических лексем *приспичить* и *норовить* с применением уже давно и успешно зарекомендовавшего себя для решения таких задач параллельного корпуса (Д. О. Добровольский, Анна А. Зализняк).

Синтаксически ориентированные исследования включают чрезвычайно своеобразный материал конструкций с тройным повтором (Е. В. Маркасова, О. А. Митрофанова), анализ частотных коллокаций глагола и существительного с использованием теперь широко доступных в Корпусе инструментов (М. В. Хохлова), а также нестандартной конструкции с «неопределенно-количественным» значением вида *мало кто / много какой* и т. п. (Г. И. Кустова). Именно такой — сложный и необычный — материал является прекрасной иллюстрацией тех разнообразных возможностей, которые Корпус сегодня предоставляет исследователю.

Что касается второй части данного номера, то в ней собраны статьи, отражающие проблематику ставших традиционными «Гришинских чтений» — небольшой ежегодной конференции, на которой обсуждается проблематика мультимодальных исследований русского языка, включающих анализ жестикуляции на материале мультимодального подкорпуса МУРКО. Эта проблематика была, как известно, введена в пионерской работе [Гришина 2017]; идеи Е. А. Гришиной, создавшей этот подкорпус, и сегодня являются востребованными. В статьях этой части исследуется как центральная тема чтений, т. е. собственно жесты, сопровождающие спонтанную речь (большинство публикуемых статей), так и некоторые смежные вопросы, как, например, русская просодия (Т. Е. Янко) или семантический анализ лексем, обозначающих жесты, в том числе и самого слова *жест* (Л. Л. Шестакова, А. С. Кулева).

Литература

Гришина Е. А. Русская жестикуляция с лингвистической точки зрения. М.: ЯСК 2017.

Савчук С. О., Архангельский Т. А., Бонч-Осмоловская А. А., Донина О. В., Кузнецова Ю. Н., Ляшевская О. Н., Орехов Б. В., Подрядчикова М. В. Национальный корпус русского языка 2.0: новые возможности и перспективы развития // *Вопросы языкознания*. 2024. № 2. С. 7–34.

В. А. Плунгян

¹ Следует отметить, что это одна из последних публикаций Виктора Павловича Захарова, безвременно ушедшего из жизни в начале 2024 года. Виктор Павлович был ярчайшим и бескорыстнейшим энтузиастом корпусной лингвистики в России и большим другом коллектива Национального корпуса русского языка; эта публикация — благодарная дань его памяти.

I. НАЦИОНАЛЬНЫЙ КОРПУС РУССКОГО ЯЗЫКА: ИССЛЕДОВАНИЯ И РАЗРАБОТКИ

Памяти В. П. Захарова

Седьмого января 2024 года скончался Виктор Павлович Захаров, один из авторов опубликованной ниже статьи. В Интернете можно легко найти сведения об учебных степенях, званиях и должностях Виктора Павловича — кандидат филологических наук, доцент кафедры математической лингвистики Санкт-Петербургского университета. Статья в Википедии сообщает о его научных интересах: корпусная лингвистика, автоматизированная обработка текста, компьютерная лексикография и семантика, информационный поиск. Могу к этому добавить — еще и машинный перевод, библиотечные информационные системы, диахрония частотности разных лексических групп, узус и семантика русских предлогов и многое другое.

Меня связывают с Виктором Павловичем многолетние (более тридцати лет) дружеские отношения. Мы вместе работали в библиотеке Академии наук России — Виктор Павлович возглавлял отдел информатики и автоматизации, а я в этом отделе заведовал сектором. В соавторстве у нас опубликовано 35 статей, одна монография (учебное пособие для студентов СПбГУ). Эти публикации относятся и к библиотечной сфере, и к корпусной лингвистике, а также к методике преподавания разных дисциплин.

Около года назад мы начали работу над статьей, завершающей цикл исследований частотного поведения цветообозначений в русском языке и по корпусно-ориентированной разработке их классификации. Планировали закончить в первом полугодии 2024 года. Придется доводить этот текст до конца без Виктора Павловича. Без него будет очень трудно.

Филологическая наука и образование с уходом Виктора Павловича потеряли многогранного ученого и преподавателя с обширнейшим опытом: он преподавал в средней школе и в нескольких университетах, читал лекции в Испании и в Китае, в Чехии и в Германии.

Ушел широко образованный специалист, интересовавшийся помимо своей основной области, еще и чешской культурой, владевший английским и чешским языками, член общества братьев Чапек, писавший о спорте и о кино. В личной

библиотеке Виктора Павловича есть уникальная коллекция книг с авторскими автографами.

А я потерял друга и коллегу, у которого в любой момент можно было найти поддержку и помощь, поделиться с ним мыслями, планами и сомнениями. Это очень горькая утрата.

Поэт Василий Андреевич Жуковский в коротком стихе «Воспоминание о милых спутниках» дал нам формулу утешения: «Не говори с тоской: их нет;/ Но с благодарностию: были». Эти слова, да и никакие слова и формулы не снимут тяжести потери. Но чувство благодарности за то, что жизнь дала годы близкого общения с таким человеком, сделает память о нем светлой и долгой. Спасибо тебе за все Виктор Павлович!

А. Ц. Масевич

А. Ц. Масевич

*Санкт-Петербургский институт культуры
(Россия, Санкт-Петербург)
andmasev@mail.ru*

В. П. Захаров

*Санкт-Петербургский государственный университет
(Россия, Санкт-Петербург)
v.zakharov@spbu.ru*

РАЗМЕТКА СЛОВ-ЦВЕТООБОЗНАЧЕНИЙ В НАЦИОНАЛЬНОМ КОРПУСЕ РУССКОГО ЯЗЫКА

Описываются ошибки разметки слов-цветообозначений в основном и поэтическом корпусах НКРЯ. Основное внимание уделено семантической разметке. Ошибки можно разделить на две группы — отсутствие тега ‘цвет’ в описании лексических единиц, имеющих признак ‘цвет’ и нерелевантное приписывание этого тега. Помимо этого выявлена ошибка в грамматической разметке слов (наречий и кратких прилагательных), имеющих тег. Показана непоследовательность в разметке цветообозначений квазицветов и аналоговых цветообозначений. Предложены некоторые способы устранения этой непоследовательности. Выявлено несколько частотных прилагательных, которым тег приписан нерелевантно, что создает при использовании их значительный информационный шум. Произведен анализ цветообозначений глаголов и существительных. Установлено, что абсолютная частота цветообозначений, имеющих тег ‘цвет’, в НКРЯ значительно выше, чем цветообозначений, не имеющих тега. Поэтому, несмотря на выявленные ошибки, сохраняется возможность эффективного использования НКРЯ для исследований. Тем не менее, ошибки могут серьезно повлиять на результаты исследований и их интерпретацию. В этой связи устранение ошибок представляется целесообразным.

Ключевые слова: семантическая разметка, НКРЯ, цветообозначения, ошибки разметки, части речи, тег

Введение

В течение нескольких лет мы использовали семантическую разметку НКРЯ в исследовании цветообозначений в русском языке [Масевич, Захаров 2019–2022]. В ходе этой работы мы столкнулись с некоторыми проблемами, ошибками и недостатками, касающимися семантической и частеречной (морфологической) разметки.

Сделанные наблюдения и наша попытка их осмыслить являются содержанием настоящей статьи.

Под разметкой текстов корпуса понимают приписывание лексическим единицам текста кодов грамматических, семантических, культурно-контекстуальных (метатекстовых) и других признаков. Теги цветообозначений на основе англоязычной нотации являются единицами метаязыка лексико-семантических признаков.

Области описания слова относятся к разным лексическим уровням: грамматическая разметка описывает как лексему (например, прилагательное), так и словоформу в контексте, т. е. словоупотребление (например, женский род, родительный падеж, единственное число).

Семантическая разметка описывает исключительно лексему. Семантическая разметка делится на основную и дополнительную (рис. 1). Покажем вариант разметки на следующем примере:

*Есть у известного современного скульптора изваянная фигура Сергея Радонежского, установленная на зеленом холме среди **зеленой** поляны возле Сергиева Посада.* [Виктор Астафьев. Затеси (1999) // «Новый Мир», 2000].

зеленой	
ОСНОВНОЙ РАЗБОР	омонимия снята
Лемма	зеленый (см. в словаре)
Грамматика	прил, ж, род, ед
Семантика основная	t:physq:color, r:qual
Семантика дополнительная	r:qual, der:s, dt:time:age, dt:physq:color, t:humq, ev:neg, der:shift, t:time:age, r:rel, t:physq:color
Синтаксические признаки	amod, partp
Доп. признаки	gendered, numred, casered

Рис. 1. Описание словоформы *зеленой* в основном корпусе НКРЯ

Лексико-семантическая информация имеет иерархическую фасетную структуру. Фасеты верхнего уровня — это знаменательные части речи: существительные, прилагательные, числительные, местоимения, глаголы и наречия. При этом существительным соответствуют три фасета (разряда) — имена предметные, не-предметные и собственные. Внутри фасетов информация, приписываемая словам, состоит из трех групп признаков: разряд, лексико-семантические характеристики, деривационные (словообразовательные) характеристики.

Лексемам, которые являются цветообозначениями (колоративами) в НКРЯ присваивается тег *цвет*: *t:color* (существительные), *t:physq:color* (прилагательные), *dt:physq:color* (отадъективные наречия) (*der:a*), *dt:physq:color* (глаголы).

Многозначность плана содержания, многовариантность плана выражения, а также зависимость значения лексической единицы (в нашем случае — цветообозначения) от контекста употребления — это те факторы, что могут обуславливать ошибки семантической разметки. Ошибки, как правило, состоят либо в отсутствии тега при наличии у слова данного семантического признака (в нашем случае

‘цвет’), либо в присвоении тега лексической единице, не несущей такого признака (т. е. нерелевантного тега).

Разметка представляет собой информационно поисковый язык, язык индексирования. С ее помощью пользователь должен иметь возможность извлечь из корпуса текстов некоторые массивы примеров (словоупотреблений) с заданными семантическими, грамматическими или иными признаками.

С одной стороны, практически любое цветообозначение всегда многозначно. Так, прилагательное *красный* имеет 12 значений [Масевич, Захаров 2022]. Слова-цветообозначения могут образовывать символ или метафору. В переносном значении эти слова часто входят в состав устойчивых сочетаний.

С другой стороны, присутствие у лексемы значения ‘цвет’ не всегда очевидно. Значение ‘цвет’ может быть преобладающим или периферийным (см. ниже, аналоговые цветообозначения, квазицвета).

В реальности абсолютная полнота и точность разметки недостижимы, будем считать эти показатели условными идеальными ориентирами при оценке качества разметки НКРЯ.

Существует ряд публикаций, посвященных семантической разметке НКРЯ [Козера 2018, Кустова и др. 2005]. Эти авторы описывают принципы и теоретическую основу семантической разметки.

Однако нам пока не встретилось публикаций, описывающих реальное использование этого инструмента для решения конкретных исследовательских задач, и проблемы, при этом возникающие.

В ряду слов-цветообозначений превалируют прилагательные, поэтому значительная часть статьи посвящена им (разд. 3). Далее рассматривается морфологическая и семантическая разметка других частей речи (разд. 4–6). В разделе 7 обсуждается проблема основного и периферийного значения прилагательных цвета. Для ее решения предлагается более точно и широко использовать параметры «семантическая разметка основная» и «семантическая разметка дополнительная». Особенности разметки и статистическая оценка ее корректности даются в разрезе классов слов-цветообозначений (разд. 1 и 2).

1. Классификация цветообозначений

Для обозначений цвета, чаще всего прилагательных, существуют различные классификации, см., например, [Герасимов 1969; Герасимов 1978; Василевич и др. 2005]. Как мы уже отмечали [Масевич, Захаров 2021: 17], отличие предлагаемой нами классификации от других состоит в том, что, с одной стороны, она составлена *a posteriori*, то есть на основе исследуемого материала — слов-цветообозначений, содержащихся в НКРЯ. С другой стороны, мы ставили задачу создания универсальной классификации, под которую должны подпадать любые цветообозначения русского языка.

Классификация, которую мы предлагаем, имеет разные основания выделения классов. Это связано с тем, что классифицируемые объекты (цветообозначения) имеют признаки, относящиеся к различным категориям, иначе говоря, к разным

уровням и аспектам языка. Один и тот же объект может иметь разные аспекты рассмотрения. Например, прилагательное *желтенький* рассматривается с точки зрения семантики как базовое цветообозначение собственно цвета, а также с точки зрения морфологии как прилагательное-дминутив.

Основания предлагаемой классификации следующие:

- 1) **семантика** — собственно цвета (базовые — небазовые), квазицвета, аналоговые (с вторичной номинацией), объектные (с ограниченной сочетаемостью);
- 2) **морфология** (аттенуативы, дминутивы, сложные слова);
- 3) **части речи** (прилагательные, наречия, глаголы, существительные);
- 4) **сочетаемость** (коллокации).

Поскольку относящиеся к разделам классификации разные признаки могут присутствовать в одном и том же объекте классификации, мы вводим понятие фасетов, позволяющее комбинировать разделы в индексе одного объекта. Каждому цветообозначению может быть сопоставлено значение из каждого фасета.

Такая подробная («объемная») классификация в настоящее время готовится к публикации. В то же время нам представляется, что имеет право на существование и усеченная «плоская» классификация на базе одного или двух фасетов. Фрагмент такой классификации, а именно классификация прилагательных цвета по их семантике и морфологии, с примерами, был опубликован в [Масевич, Захаров 2019: 33–45; Масевич, Захаров 2021: 17–18].

В кратком представлении она выглядит следующим образом.

1. Прилагательные собственно цвета (*алый, бежевый, белый, голубой* и др.).
2. Прилагательные квазицвета, характеризующие разные визуальные характеристики предметов, как интенсивность цвета или света, оттенок, просто наличие цвета (*бледный, пёстрый, тёмный, цветной* и др.).
3. Объектные прилагательные — слова с ограниченной сочетаемостью [Дюпина 2013], сочетающиеся с ограниченным множеством существительных (*гнедой, пегий, русый, саврасый* и др.).
4. Аналоговые прилагательные цвета — слова вторичной номинации [Дюпина 2013], отражающие цветовую картину какого-либо реального предмета (*вишнёвый, золотой, малиновый, янтарный* и др.).
5. Дминутивы (*аленький, беленький, бледненький, малиновенький, русенький, саврасенький, черненький, краснехонький* и др.).
6. Атенуативы — прилагательные подобия или неполноты качества (*бледноватый, зеленоватый, темноватый* и др.)¹.
7. Сложные (составные) прилагательные (*белогривый, нежно-голубой, тёмно-синий, черноволосый* и др.)².

¹ Отметим, что позиции 5–6 образуются, как правило, от классов 1–3.

² Позиции 5–7 могут рассматриваться как подклассы класса «дериваты» с основанием деления «морфология».

2. Количественная характеристика цветообозначений в НКРЯ

Несложно подсчитать абсолютное число словоупотреблений (контекстов) слов, имеющих тег 'цвет'. Для этого достаточно выполнить несколько запросов по семантическому признаку 'цвет' (табл. 1).

Таблица 1. Распределение слов-цветообозначений, имеющих тег 'цвет', по частям речи

Часть речи	Запрос по тегам	Основной корпус (ОК), число словоупотреблений	Поэтический корпус (ПК), число словоупотреблений
Прилагательные	A, t:physq:color	1 110 125	148 952
Существительные	S, r:abstr & t:color	57 331	17 362
Наречия	ADV, der:a & dt:physq:color	24 102	6 570
Глаголы	V, dt:physq:color	81 644	13 349
Всего		1 273 202	186 233

В основном корпусе НКРЯ (ОК) число употреблений слов, размеченных как цветообозначения, составляет 0,34% от всего объема корпуса. В поэтическом корпусе (ПК) эта доля равна 1,43%.

В новой версии ОК НКРЯ имеется функция «Частотность», показывающая для найденных в корпусе вхождений (поисковых элементов) статистические данные каждого вхождения в общем массиве найденных словоупотреблений: число вхождений, его долю, долю с доверительным интервалом, IPM, а также примеры употребления. Выдается 100 самых частотных вхождений (см. табл. 2, удалены столбцы, избыточные для нашего анализа).

Таблица 2. Фрагмент частотных характеристик прилагательных по результатам поиска по тегу 'цвет' (A, t:physq:color) в ОК НКРЯ

Ранг	Лемма	Число вхождений	IPM
1.	белый	138 183	369,03
2.	черный	132 420	353,64
3.	красный	106 775	285,15
4.	темный	66 891	178,64
...			
98.	цветастый	753	2,01
99.	багряный	744	1,99
100.	рыженький	739	1,97

Ранжированный список из 100 лексем с тегом 'цвет' мы разделили на три множества: 1) «реальные» цветообозначения, разбитые на классы в соответствии

с предложенной нами классификацией [Масевич, Захаров 2019: 34; Масевич, Захаров 2021: 17]; 2) лексические единицы, в которых значение ‘цвет’ периферийно; 3) лексические единицы, в которых значение ‘цвет’ отсутствует.

Ниже приводится список цветообозначений в этих трех множествах.

- 1) «Реальные» цветообозначения (с разбиением по вышеуказанным классам (разд. 1), по убыванию количества лексем в каждом классе) (всего 84):
 - цветообозначения собственно цвета (26): *алый, багровый, багряный, белесый, белый, бурый, голубой, желтый, зеленый, коричневый, красный, лазурный, лиловый, оранжевый, половой, пунцовый, пурпурный, розовый, рыжий, серый, сивый, сизый, синий, фиолетовый, червонный, черный*;
 - квазицвета (18): *бесцветный, бледный, клетчатый, колоритный, красочный, матовый, мутный, пестрый, пятнистый, разноцветный, рябой, светлый, сумрачный, темный, тусклый, цветастый, цветной, яркий*;
 - аналоговые цветообозначения (14): *бирюзовый, дымчатый, золотой, золотистый, золотой, изумрудный, кофейный, лимонный, молочный, огненный, румяный, сахарный, серебристый, снежный*;
 - аттенуативы (8): *голубоватый, желтоватый, зеленоватый, красноватый, розоватый, рыжеватый, сероватый, синеватый*;
 - сложные цветообозначения (7): *белоснежный, темно-зеленый, темно-красный, темно-синий, черно-белый, черноволосый, ярко-красный*;
 - объектные цветообозначения (6): *вороной, гнедой, карий, пегий, русский, седой*;
 - диминутивы (5): *беленький, красненький, рыженький, серенький, черненький*.
- 2) Слова, значение цвета у которых является периферийным (6): *грязный, мертвенный, пшеничный, свинцовый, стальной, хрустальный*.
- 3) Слова, которые цветообозначениями не являются, но имеют тег ‘цвет’ (10): *защитный, песчаный, спокойный, суровый, теплый, химический, холодный, чайный, ядовитый, ярый*.

Приведенный анализ дает приблизительное представление о доле каждого класса и каждого типа (1–3) в общем массиве тегированных цветообозначений³.

Ниже проанализирована семантическая разметка слов — членов некоторых классов, выделяемых по основаниям «семантика» и «морфология» и отличающихся заметной непоследовательностью разметки (разд. 3), а также цветообозначений, относящихся к разным *частям речи* — это наречия, глаголы и существительные (разд. 4–6).

³ Представление приблизительное, так как не учитывается реальное ранговое положение прилагательных в частотном списке; может оказаться, что меньшее число прилагательных какого-либо класса с высокими рангами составит более значительную долю в общем массиве, чем большее число прилагательных с низкими рангами. Кроме того, функция «Частотность» выдает лишь 100 самых частотных вхождений прилагательных и все остальные вхождения в наши подсчеты не попадают.

3. Семантическая разметка прилагательных

3.1. Цветообозначения собственно цвета, не имеющие тега цвет

Практически все колоративы класса ‘собственно цвет’ имеют несколько значений. Присвоение тегов осуществляется независимо от того, в прямом или переносном значении употреблено слово (т. е. независимо от контекста), и поэтому в коллокациях *красная девица*, *красная рубаха*, *красный командир*, *красный нос*, *красный уголок*, *красное лето* прилагательное *красный* во всех случаях имеет тег ‘цвет’.

Для выявления лексем-цветообозначений, не имеющих тега ‘цвет’, были использованы указатель цветообозначений из монографии Н. Б. Бахилиной «История цветообозначений в русском языке» [Бахилина 1975], а также разные списки цветообозначений в Интернете. В ОК нам удалось выявить 78 лексем-цветообозначений, не имеющих тега, в ПК — 39. В основном это малоупотребительная, часто архаичная лексика. Встречаются слова с композитами, заимствованными из иностранных языков, чаще из французского (ср. *вердепомовый*, *гридеперлевый*), или аналоговые цветообозначения, первое значение которых часто требует консультации со специальными словарями (ср. *гагатовый*, *селадоновый*, *серизовый*). Приведем фрагмент списка таких слов (рис. 2). Мы условно отнесли эту лексику к цветообозначениям собственно цвета.

	A	B	C	D	E	F
1	слово	абс ОК	абс ПК	IPM ОК	IPM ПК	
2	авроровый	3	0	0,0080	0,000	
3	адамантовый	27	19	0,0721	1,461	
4	алебастровый	285	9	0,7615	0,692	
5	ализариновый	17	1	0,0454	0,077	
6	амиантовый	2	1	0,0053	0,077	
7	аспидный	224	26	0,5985	2,000	
8	бакановый	2	0	0,0053	0,000	
9	берилловый	7	2	0,0187	0,154	
10	бистровый	1	0	0,0027	0,000	
11	бланжевый	50	0	0,1336	0,000	
12	бледый	1	0	0,0027	0,000	
13	блукитный	1	0	0,0027	0,000	
14	блондовый	50	2	0,1336	0,154	
15	вердепомовый	3	0	0,0080	0,000	
16	гагатовый	3	0	0,0080	0,000	
17	гелиотроповый	15	1	0,0401	0,077	
18	гридеперлевый	2	0	0,0053	0,000	
19	двуличневый	11	1	0,0294	0,077	
20	изабеловый	1	0	0,0027	0,000	
21	индиговый	30	5	0,0802	0,385	
22	инкарнатный	3	0	0,0080	0,000	
23	карамазый	1	0	0,0027	0,000	

Рис. 2. Фрагмент списка цветообозначений
(условно собственно цвета), не имеющих тега ‘цвет’

Особенности классификационной принадлежности, семантики и использования этих слов требуют специального исследования. Существенным наблюдением для настоящей работы является отсутствие у них тега ‘цвет’. Суммарное количество словоупотреблений таких лексем-цветообозначений без тега в ОК составляет 5042 (средний ИРМ 0,17), а в ПК 1042 (средний ИРМ 1,03).

Уделим специальное внимание одной из лексем этого списка. *Рдяный* (*красный, алый*) относится к классу собственно цвета. Таким цветообозначениям, как правило, тег в НКРЯ присвоен, даже если они не относятся к базовым цветовым категориям (например, *алый, багровый, багряный, бежевый* и др.). Прилагательное *рдяный* в поэтическом корпусе имеет заметную частоту — 246 словоупотреблений (ИРМ 18,35) (в основном корпусе меньше — 55 словоупотреблений). Это прилагательное и его производные являются довольно частотными словами у авторов XVIII — начала XIX века. Так, в авторском подкорпусе Г. Р. Державина *рдяный* встречается 12 раз, ИРМ 128,44. Однако в НКРЯ это слово тега ‘цвет’ не имеет.

3.2. Квазицвета

Под квазицветами мы понимаем слова, которые не указывают на цвет, но характеризуют его, а также описывают визуальные свойства объекта, кроме его объема и формы. Нет единого мнения о том, правомерно ли относить эту группу слов к цветообозначениям. Р. М. Фрумкина пишет, что нет таких имен цвета (ИЦ), которые бы содержали информацию о яркости и насыщенности без указания об основном тоне (цвете). «Слова *светлый, яркий, тусклый, блестящий, пёстрый* мы не считаем ИЦ» [Фрумкина 1984].

В статье О. Н. Ляшевской с соавторами говорится: «Группа слов цвета также включает прилагательные, относящиеся к яркости и насыщенности (saturation) цвета: тёмный ‘dark’, пёстрый ‘motley’, тусклый ‘dull’» [Lyashevskaya et al. 2021] (статья опубликована на английском языке, цитата в нашем переводе).

Кроме того, если посмотреть область «Похожие слова» в портрете слова *светлый*, то ясно видна семантическая связь слов-квазицветов с цветообозначениями (рис. 3).

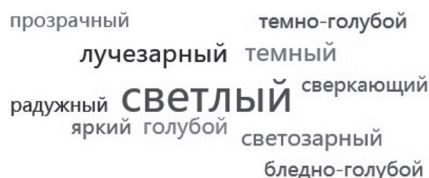


Рис 3. Область «Похожие слова» в портрете слова *светлый*

В группе квазицветов также выявлена непоследовательность разметки. Одна часть этой лексики имеет тег ‘цвет’, другая нет. Так, прилагательные *светлый, темный, бледный, пестрый, мутный, сумрачный, цветастый* имеют тег ‘цвет’, а прилагательные *блеклый, матовый, узорчатый, радужный, блестящий* тега ‘цвет’ не имеют. Обратите внимание, что *блеклый* тега ‘цвет’ не имеет, а у прилагательного *поблеклый* этот тег присутствует.

3.3. Аналоговые цветообозначения

Аналоговыми цветообозначениями мы называем такие, которые обозначают цвет по аналогии с цветом хорошо известного предмета. Цветообозначения *аметистовый, дымчатый, золотой, изумрудный, малахитовый, молочный, опаловый, рубиновый, румяный* имеют тег ‘цвет’ (в ПК 3960 словоупотреблений, в ОК 14 914). В то же время цветообозначения *аспидный, васильковый, вишневый, горчичный, канареечный, коралловый, кровавый, кумачовый, малиновый, медовый, морковный, огненный, пепельный, порфиновый, салатный, свекольный, серебряный, сиреневый, фиалковый, чернильный, янтарный* тега ‘цвет’ не имеют. При этом число этих словоупотреблений совокупно составляет 8002 в ПК и 61 841 в ОК. Приведенный здесь список аналоговых цветообозначений без тега ‘цвет’ не является исчерпывающим.

Аналоговые цветообозначения всегда имеют не менее двух значений, одним из которых является обозначение цвета. Встает вопрос, насколько велика частота употребления этих лексем в функции цветообозначений.

Рассмотрим прилагательное *малиновый*. В НКРЯ оно не имеет тега ‘цвет’.

Абсолютное число словоупотреблений в ПК составляет 306.

При просмотре 112 случайных контекстов в ПК выделены следующие значения:

Растение (малиновые заросли, малиновый куст): 5 контекстов (4,5 %);

Еда (малиновое варенье): 5 контекстов (4,5 %);

Звук (малиновый звон, малиновый грей): 9 контекстов (8%);

Другие переносные значения (малиновая любовь): 13 контекстов (11,6%);

Цвет (малиновое солнце, малиновые губы, малиновая юбка): 80 контекстов (71,4 %).

Абсолютное число словоупотреблений в ОК — 3988.

Просмотрено 150 контекстов, выделены значения:

Растение (малиновый куст, малиновые заросли): 5 контекстов (3,3%);

Еда (малиновое варенье, малиновый сироп, малиновая наливка): 24 (16%);

Звук (малиновый звон): 5 контекстов (3,3%);

Аромат, вкус (малиновый вкус): 3 контекста (2%);

Собственное имя, топоним (село Малиново): 5 контекстов (3,3%);

Другие переносные значения (малиновый пар, малиновое безвременье): 3 контекста (2%);

Цвет (малиновая рубашка, малиновый торс, малиновые ногти): 105 контекстов (70%).

Как видим, доля значения ‘цвет’ в обоих корпусах существенно больше доли других значений.

Это противоречит данным разметки (рис. 4), где категория ‘цвет’ не указана вовсе.

малиновый	
Лемма	малиновый (см. словарь)
Грамматика	прилагательное
Семантика основная	der:s dt:food dt:fruit r:rel

Рис. 4. Область грамматических и семантических признаков слова *малиновый*

Частотность значения ‘цвет’ слова *малиновый* также подтверждается списком его коллокатов в портрете слова *малиновый* (Табл. 3).

Таблица 3. Ранжированный список наиболее частотных коллокатов прилагательного *малиновый* из портрета слова

	Определяемые слова	Сила связи
1.	варенье	10,24
2.	бархат	9,66
3.	звон	8,81
4.	пиджак	8,29
5.	берет	8,28
6.	кант	7,84
7.	кафтан	7,39
8.	околыш	7,36
9.	сукно	7,35
10.	джем	7,18

Как видно, в семи коллокациях прилагательное *малиновый* явно относится к категории ‘цвет’, определяя цвет предметов, в двух случаях — к категории ‘еда’ (1,10) и в одном — к категории ‘звук’ (3), образуя фразеологизм.

Анализ контекстов прилагательного *красавый* в ПК показал 23% употреблений в значении ‘цвет’ (*красавый закат*, *красавый подбой*, *красавый флаг*), в основном — 18% (*красавый закат*, *красавый подбой*, *красавая мантия*).

Однако получить эти частотные характеристики без сплошного просмотра затруднительно. Возможно, всем аналоговым цветообозначениям необходимо тег ‘цвет’ приписать в форме альтернативной разметки, в области «семантика дополнительная». При этом при поиске следует дать возможность искать отдельно по признакам «семантика основная» и «семантика дополнительная». Существующая функция поиска «первое и/или другие значения» в отношении цветообозначений работает неустойчиво.

3.4. Сложные цветообозначения

Существует два варианта написания сложных цветообозначений — в одном случае композиты пишутся через дефис (*темно-синий*), в другом — слитно (*белоснежный*, *рыжеволосый*, *голубоглазый*).

Встречаются оба варианта написания одного слова. Например:

*В гранит оделася Нева;
Мосты повисли над водами;
Темнозелеными садами
Ее покрылись острова.*
[А. С. Пушкин. Медный всадник (1833)]

Кровать была крепкая, но аляповато выкрашенная темно-зеленою масляною краской с розами на столбиках. [Н. С. Лесков. Дама и фефёла (1894)].

В НКРЯ выявлено четыре варианта разметки сложных прилагательных-колоративов при написании через дефис.

- (1) Сложный колоратив из двух композитов рассматривается как одна лексема (например, *желто-зеленый*) и имеет тег 'цвет' (рис. 5).

4. Г. Р. Державин. Евгению. Жизнь Званская : «Блажен, кто менее зависит от людей...» (05.1807-07.1807) [омо]

Иль смотрим, как бежит под черной тучей тень
По кошнам, по снопам, коврам желто-зеленым
И сходит солнышко на нижнюю степь
К холмам и рощам сине-темным.

[Г. Р. Державин. Евгению. Жизнь Званская : «Бл

5. С. С. Бобров. Херсонида, или Картина лучи
снята]

Тебе в дождях явили призму
И в поясе желто-зеленом
Те показали нити света,

желто-зеленым	
Лемма	желто-зеленый (см. в словаре)
Грамматика	прип, дат, мн, полн прип, твор, м, полн, ед прип, твор, с, полн, ед
Семантика основная	dt:physq:color, r:rel, der:comp, t:physq:color
Доп. признаки	genderred, numred, casered, norm, rhyme

Рис. 5. Г. Р. Державин. Евгению. Жизнь Званская (05.1807-07.1807)

- (2) Сложный колоратив рассматривается как две лексемы, каждый композит имеет тег 'цвет' (рис. 6).

Зелено-красный желузь в светофоре —

Зелено	
Лемма	зеленый (см. в словаре)
Грамматика	прип, кратк, с, ед
Семантика основная	t:physq:color, r:qual
Семантика дополнительная	t:physq:color, dt:time:age, der:shift, dt:phys: t:time:age, t:humq, ev:neg, r:qual
Доп. признаки	acomma, amark, bhyphen, bmark, capital, f, numred, norm

Зелено-красный желузь в светофоре —

красный	
Лемма	красный (см. в словаре)
Грамматика	прип, вин, неод, м, полн, ед прип, м, им, полн, ед
Семантика основная	t:physq:color, r:qual
Семантика дополнительная	r:rel, t:species, der:shift, dt:physq:color, m, r:qual
Лемма	красный (см. в словаре)
Грамматика	сущ, одуш, м, им, ед
Семантика дополнительная	r:propn, r:concr, der:subst, t:fam, t:hum, i
Доп. признаки	ahyphen, amark, posred, genderred, numr, casered, animad, norm

Рис. 6. Л. Н. Мартынов. «И снова осень... / Велосипедист...» (1952)

- (3) Сложный колоратив рассматривается как две лексемы, при этом один композит имеет тег ‘цвет’, а другой нет (рис. 7).

Бархатно-черная, с теплым отливом сливы созревшей, вот распахнула сладостно свет вдоль круговой		Бархатно-черная, с теплым отливом сливы созревшей, вот распахнула сладостно свет вдоль круговой	
Бархатно		черная	
Лемма	бархатный (см. в словаре)	Лемма	черный (см. в словаре)
Грамматика	прил, кратк, с, ед	Грамматика	прил, ж, им, полн,
Доп. признаки	bhyphen, bmark, first, capital, posred, numred, norm	Семантика основная	t:physq:color, r:qual
		Семантика дополнительная	r:rel, t:species, t:ph dt:physq:color, t:hu
Сообщить об ошибке...		Лемма	черная (см. в словаре)
		Грамматика	сущ, одуш, ж, фем
		Семантика основная	r:propn, t:hum, t:fan
		Доп. признаки	ahyphen, amark, bc norm

Рис. 7. В. В. Набоков. Бабочка (1917–1922)

- (4) Сложный колоратив рассматривается как две лексемы, при этом ни один композит не имеет тега ‘цвет’ (рис. 8).

Встал орех с листвою мокро-ржавой; Вялых гроздьев и вина на стол Садовод поставит величавый.		Встал орех с листвою мокро-ржавой; Вялых гроздьев и вина на стол Садовод поставит величавый.	
мокро		ржавой	
Лемма	мокро (см. в словаре)	Лемма	ржавый (см. в словаре)
Грамматика	н	Грамматика	прил, ж, предл, полн, ед прил, дат, ж, полн, ед прил, ж, род, полн, ед прил, ж, твор, полн, ед
Лемма	мокрый (см. в словаре)	Семантика основная	dt:changest, t:physq, r:q
Грамматика	прил, кратк, с, ед	Доп. признаки	amark, ahyphen, bmark, numred, casered, norm,
Семантика основная	t:physst, r:qual		
Доп. признаки	bmark, bhyphen, posred, n		

Рис. 8. С. В. Шервинский. Вечером на Норке (1921–1984)

Если сложное слово рассматривается как одна лексическая единица (пример 1, рис. 5), то при обсчете результатов оно учитывается как одна единица.

Если в сложном слове из двух и более композитов все композиты имеют тег, то слово учитывается как несколько единиц, т. е. каждый композит учитывается как самостоятельная единица (пример 2, рис. 6), что, на наш взгляд, неправильно.

Встречаются случаи (пример 3, рис. 7) когда композиты размечены отдельно, при этом один из них цветообозначением не является и тега ‘цвет’ не имеет. Второй композит является цветообозначением и имеет тег. Благодаря наличию тега у второго композита, такое слово при обсчете будет учтено как цветообозначение. Однако мы считаем, что методически более правильным было бы разметать такие слова как одну лексему.

Наконец, есть случаи, достаточно редкие, когда каждый композит в отдельности цветообозначением не является (пример 4, рис. 8), но их сочетание в одном сложном слове выполняет эту функцию. Разметка подобных лексем может быть осуществлена только вручную.

Мы предлагаем все сложные прилагательные с написанием через дефис всегда считать одним словом, при этом тег ставится при наличии признака ‘цвет’ хотя бы у одно из компонентов.

3.5. Сложные слова со слитным написанием

Некоторые из таких слов имеют тег 'цвет' (рис. 9). Однако большинство их тега 'цвет' не имеет (*белокаменный, белогривый, белопенный*).

.. А. Ахматова. Творчество : «Бывает так: какая-то истома...» (05.11.1936) [омонимия не снята]

Но вот уже послышались слова
И легких рифм сигнальные звоночки, —
Тогда я начинаю понимать,
И просто продиктованные строчки
Ложатся в белоснежную тетрадь. 5 ноября 1936 Ленинград, Фонтанный Дом [А. А. Ахматова. Творч

I. А. Тарловский. Пре	белоснежную	
	Лемма	белоснежный (см. в словаре)
	Грамматика	прил, вин, ж, полн, ед
	Семантика основная	der:comp, r:rel, t:physq:color
Особенно, — он	Доп. признаки	gendered, numred, casered, norm
Когда скиталец, б		
На бесконцовке б		
И сам при этом о		

ый страны

прочим...» [Bei

Рис. 9. Тег 'цвет' в описании сложного прилагательного *белоснежный*

Обращает на себя внимание непоследовательность присвоения тегов в разметке слов этого типа, включая и однокоренные слова. Так, *черноволосый, рыжеволосый* имеют тег 'цвет', но *желтоволосый, златоволосый* тега не имеют.

Слова с корнем *-глаз* (*черноглазый, кареглазый*) тега не имеют. Добавим сюда и восторженное обращение Г. Р. Державина к *агатовоокой* деве.

Слова с корнем *-бород* тега не имеют за исключением слова *рыжебородый*. Словам с корнем *-волос, -кудр* тег присвоен непоследовательно (*рыжеволосый* имеет тег, *рыжекудрый* нет).

Еще пример непоследовательности — слова *чернокожий, желтокожий, краснокожий* имеют тег 'цвет', слова *белокожий, темнокожий* и *светлокожий* не имеют. К этому списку можно добавить слова с корнем *-щек*: *краснощекий, розовощекий, сизощекий*, встретилось даже *золотистощекий*.

Таких слов немало. Некоторые из них более употребительны, другие менее. Безусловно, их выявление и разметка представляют собой определенную трудность. Тем не менее, видно, что число слов, построенных по такой модели, без тега больше, чем с тегом. На наш взгляд, в данном случае следует придерживаться принципа, предложенного для описания слов через дефис: если один из корней имеет признак *цвет*, следует присваивать тег в области «семантика основная».

4. Наречие vs краткое прилагательное

Наречия, образованные от качественных прилагательных с помощью суффикса *-о* или *-е*, по форме очень напоминают краткие формы среднего рода прилагательных. Для их различения требуется анализ контекста. Наречие, как правило, это обстоятельство, и в предложении зависит от сказуемого (глагола). А краткое прилагательное связано с существительным или местоимением и является определением или именной частью сказуемого.

Основная проблема разметки таких слов следующая: в значительном числе случаев наречиям-цветообозначениям присваивается тег ‘краткое прилагательное’. При сплошном просмотре 203 контекстов словоформы *красно* в ПК выявлено, что в 69 случаях (33,99%) слово, являющееся наречием, имеет тег ‘краткое прилагательное’ (рис. 10, 11). Еще в 75 случаях (36,94%) тег ‘краткое прилагательное’ присвоен слову *красно* релевантно (*красно солнышко*, *красно лето*, *красно вино*). Обратим внимание, употребления краткой формы прилагательного почти исключительно имеют переносное значение.

*Износившуюся тему
Красно выкрасив опять,
Под написанной поэмой
Ставят круглую печать.*
[М. А. Светлов. «Поздно,
почти на самой заре...»
[Ночные встречи, 2]
(1924–1925)]

М. А. Светлов. «Поздно, почти на самой заре...» [Ночные встречи, 2] (1924–1925)

Износившуюся тему Красно выкрасив опять, Под написанной поэмой Ставят круглую печать. [М. А. Светлов. «Поздно, почти на самой заре...» [Ночные встречи, 2] (1924–1925)]	
Красно	
Лемма	красный (см. в словаре)
Грамматика	прил, кратк, с, ед
Семантика основная	t:physq:color, g:qual
Семантика дополнительная	ev:posit, g:qual, g:rel
Доп. признаки	capital, numred, norm

Рис. 10. Пример наречия *красно* с ошибочным тегом ‘краткое прилагательное’

*Меж кочек, в медном лунном свете,
Красно сияет медь воды.*
[Максим Горький.
В Финляндии : «Под медным
оком злой луны...» (10.02.1919)]

1. Максим Горький. В Финляндии : «Под медным оком злой луны...» (10.02.1919)

Меж кочек, в медном лунном свете, Красно сияет медь воды. [Максим Горький. В Финляндии : «Под медным оком злой луны...» (10.02.1919)]	
Красно	
Лемма	красный (см. в словаре)
Грамматика	прил, кратк, с, ед
Семантика основная	t:physq:color, g:qual
Семантика дополнительная	ev:posit, g:qual, g:rel
Доп. признаки	amark, acomma, capital, numred, norm

Рис. 11. Пример наречия *красно* с ошибочным тегом ‘краткое прилагательное’

В 49 случаях (24,14%) форма *красно* является композитом сложных прилагательных (*красно-рыжий*, *красно-пурпурный*). К оставшимся 10 случаям (4,94%) мы отнесли повторы, а также лексику, в частеречной идентификации которой мы не уверены, например, *Песнь моя к тому есть безместна: // Красно все пред тобой стыдится... В. К. Тредиаковский. (Императрице Елизавете Петровне в день ее коронавания).*

Таким образом, поисковый аппарат НКРЯ рассматривает наречие *красно* не как отдельную лексему, а как словоформу лексемы *красный* и поиск по наречию становится нерелевантным.

Сходный результат получен при исследовании форм *бело*, *желто*, *зелено*, *сине*, *розово*, *коричнево*. В поэтическом корпусе эти формы вне зависимости от их реальной

частеречной принадлежности помечены как краткое прилагательное, причем многие из них являются композитами сложных слов.

Формы *ало*, *серо*, *черно* имеют тег 'наречие'. Отметим, что форма *ало* помечена исключительно как наречие, а формы *серо*, *черно* имеют альтернативную разметку — и тег 'наречие', и тег 'краткое прилагательное' (рис. 12).

поляна пустая
летит нарастая —
во мне вырастая —
пустая блистая
мухоморы мухоморы
выспучило вспучило
серо сыро хмуро хмара
мхом обозначило
встретило нас чучело
и расчеловечило
 [Г. В. Сапгир. Песня тоннеля
 от станции «Белорусской»
 до «Новослободской»
 «свист и скорость хорошо...»
 (1999)]

выспучило вспучило серо сыро хмуро хмара	
серо	
Лемма	серый (см. в словаре)
Грамматика	прил, кратк, с, ед
Семантика основная	t:physq:color, r:qual
Семантика дополнительная	dt:humq, dt:physq:color, ev:neg, r:qual, t:humq, der:shift, t:appearance
серо	
Лемма	серо (см. в словаре)
Грамматика	н
Семантика основная	ev:neg, dt:physq:color, der:a
Семантика дополнительная	ev:neg, dt:physq:color, der:a
Доп. признаки	posred, genderred, numred, norm

Сообщить об ошибке

Рис. 12. Пример наречия *серо* с тегом 'краткое прилагательное'

Разметка становится еще более непоследовательной, когда цветообозначение является композитом сложного слова. Сложное слово при этом не рассматривается как единая лексема. В одних случаях левый композит помечается в НКРЯ как краткое прилагательное, в других — как наречие (рис. 13).

Как неяркие бутоны
превращают лепестки
В ярко-радостные розы,
ало-красные цветы, —
Так твой ропот затаенный,
стоны девичей тоски,
Стали — сладостные грезы,
жадно-страстные мечты!
 [В. Я. Брюсов. Как неяркие
 бутоны : «Как неяркие бутоны
 превращают лепестки...»
 (16.11.1914)]

В ярко-радостные розы, ало-красные цветы, — Так твой ропот затаенный, Стали — сладостные грезы	
ало	
Лемма	ало (см. в словаре)
Грамматика	н
Семантика основная	dt:physq:color, der:a
Доп. признаки	acomma, amark, bmark, bhyphen

Сооб

Рис. 13. Пример формы *ало*, являющейся композитом сложного прилагательного *ало-красный* с тегом 'наречие'

Ошибки в присвоении тега ‘наречие’ vs ‘краткое прилагательное’ обуславливают неполноту и неточность результатов при корпусно-статистических исследованиях. Допустим, для решения некоторой задачи требуется выявить число наречий-колативов в поэтических текстах. С этой целью естественно выполнить поиск с использованием семантической разметки, сформулировав запрос следующим образом: часть речи — наречие, семантический признак — цвет. При таком запросе в результат не войдут наречия, ошибочно размеченные как краткие прилагательные, а их немало.

5. Глаголы-цветообозначения

Присвоение тега ‘цвет’ глаголам также отличается непоследовательностью. Была изучена разметка глаголов в результатах поиска по серии запросов (всего 38 запросов в ОК и ПК), построенных из глаголов с корнями, имеющими семантический признак ‘цвет’.

Примеры запросов:

- 1) обелеть | очернить | окраснить | ожелтеть | озеленить | осинить | орозоветь | отемнить | осветлеть | обледнеть | отускнить |
- 2) обелеться | очерниться | окрасниться | и т. д.
- 3) обелить | очернить | окраснить | ожелтить | озеленить | осинить | орозовать | отемнить | осветлить | обледнить | отускнить |
- 4) обелиться | очерниться | окрасниться | и т. д.

Мы попытались понять, зависит ли наличие тега ‘цвет’ от грамматических характеристик глаголов, модели деривации (с помощью суффиксов и/или префиксов) и класса исходного прилагательного (по нашей классификации). 38 запросов были выбраны так, чтобы были представлены различные варианты этих характеристик. В таблице 4 приведены примеры результатов поиска.

Даже из этого небольшого числа примеров видна непоследовательность семантической разметки глаголов. В каждом из 38 примеров вручную просматривались контексты и вычислялся процент словоупотреблений с тегами. Анализ показал, что корреляции между характеристиками глаголов и наличием/отсутствием тега ‘цвет’ не наблюдается.

У большинства глаголов тег ‘цвет’ отсутствует. В то же время в абсолютном значении число глагольных словоупотреблений, имеющих этот тег, суммарно значительно больше, чем тех, у которых тега нет. Так, в исследованном массиве в ОК число словоупотреблений, имеющих тег, составляет 79 159 словоупотреблений, не имеющих — 11 593 словоупотреблений, доля цветообозначений без тега — 12,7%. В ПК число глаголов без тега 2582 словоупотребления, доля глаголов без тега — 19,21%.

Видно, что чаще не имеют тега малоупотребительные слова, а небольшое количество глаголов, частотность которых высока, обеспечивает высокий процент общей «тегированности» (наличия тега ‘цвет’). Можно предположить, что семантическая разметка глаголов опирается на словарь разметки.

Таблица 4. Примеры описаний и контекстов глаголов-цветообозначений в ОК и ПК

Лемма и грамматика	Префикс	Пример контекста и семантика
Белеть непереходный, несовершенный	без префикса	Снизу смутно <i>белело</i> его запрокинутое лицо. Семантика основная <i>dt:physq:color, ca:noncaus, der:a, t:changest</i>
Краснить переходный, несовершенный	без префикса	Но так как раствор чистого нашатыря всегда <i>краснит</i> лакмус, то ясно ..., Семантика основная <i>dt:physq:color, t:changest, der:a, ca:noncaus</i>
Синеться непереходный, несовершенный	без префикса	Множество низеньких домиков, разбросанных по берегу Терека <...>, мелькали из-за деревьев, а дальше <i>синелись</i> зубчатой стеной горы, из-за них выглядывал Казбек... Семантика основная <i>ca:noncaus</i>
Высинить переходный, совершенный	вы	Когда же Лори <...> сорвал с лица девушки синий пластырь, оказалось, что краска с вуали порядочно-таки <i>высинила</i> лицо. Семантика основная <i>der:a, d:pref, der:v</i>
Зажелтеть непереходный, совершенный	за	Палуба была вся мокрая, серая и вдруг <i>зажелтела</i> от солнца. Семантика основная <i>d:pref</i>

6. Существительные-цветообозначения

Существительные-кolorативы помечаются тегом *t:color* и в поисковом интерфейсе для семантики имеется соответствующий признак.

В ходе работы над классификацией цветообозначений (публикация в подготовке) мы выделили 6 морфологических моделей существительных-цветообозначений:

noun1 — *белизна, желтизна, голубизна* и др.

noun2 — *чернота, краснота, зеленота* и др.

noun3 — *розовость, фиолетовость, синесть, красность, яркость, бледность* и др.

noun4 — *чернь, зелень, синь, желть, жолтень (исключение), краснь* (вариант *крась*), *серь, бель* и др. Эта форма встречается также с префиксом про-: *пробель, прочернь, прозелень* и т. д.

noun5 — *желтинка, красинка, сининка, чернинка, серебринка, сединка, темнинка* и др.

Кроме того, имеется большое число других малопродуктивных типов, которые мы объединили в одну общую группу.

noun 6 — *багрец, лазурь, киноварь, охра, сурик, золото, серебро, медь, савраска, буланка, русак, беяк, чернобурка* и др. Эта модель охватывает лексемы различных словообразовательных типов и семантических классов, в том числе аналоговые и объектные.

Эту классификацию мы использовали для отбора существительных для эксперимента, а именно, было отобрано 48 существительных так, чтобы в наборе присутствовали слова, относящиеся к разным моделям и семантическим классам. Эксперимент заключался в поиске контекстов с отобранными существительными и ручной проверке присутствия/отсутствия тега 'цвет' в семантической разметке.

Из 48 существительных 16 слов имеют тег 'цвет'. Это следующие существительные (по убыванию частотности): *синь, лазурь, синева, зелень, белизна, румянец, бледность, голубизна, желтизна, багрец* (IPM в ПК от 129,7 до 11,6, в ОК от 13,7 до 2,2), а также *бель, прозелень, сединка, краснота, прочернь, желтинка* (IPM в ПК от 7,5 до 0,01, в ОК от 0,46 до 0,03). Эту последовательность в ОК несколько нарушает *зелень* (IPM 30,1, больше чем у *лазури* и *синева*, что можно объяснить его частым употреблением в значении 'листва').

32 существительных не имеют тега 'цвет'. Это слова: *золото, серебро, медь, чернь, пурпур, чернота, яркость, охра, русак, кармин, желть, савраска, беляк* (IPM в ПК от 204,8 до 1,5, в ОК от 13,7 до 0,1 (*желть*)), а также *зеленость, золотинка, розовость, янтарность, фиолетовость, пробель, серь, изумрудность, краснота, краснь, зеленота, синесть, буланка, краснинка, оранжевость, серебринка, белота, чернинка, малиновость* (IPM в ПК от 0,77 до нуля, в ОК от 0,21 (*розовость*) до нуля (*янтарность*)).

Слова *золото, серебро, медь* иногда используются в значении 'цвет' (как аналоговые цветообозначения), например, *золото волос, золото ржи, золото мимозы, серебро реки, серебро чешуи, медь зари, медь загара* и т. д. Поэтому для таких слов кажется целесообразным размещение тега 'цвет' в области «семантика дополнительная» при обязательном наличии опции включения/отключения этой области при поиске.

Наличие/отсутствие тега хорошо коррелирует с частотой лексем. Соотношение среднего числа словоупотреблений с тегом/без тега на одну лемму составляет 487 против 60 в ПК в первой группе и 2084 против 397 в ОК (без учета высокочастотных *золото, серебро, медь*) во второй группе. Причем корреляция намного больше выражена в поэтическом корпусе.

Обратим также внимание, что слова, относящиеся к модели *noun1*, как правило, имеют тег, в то время как существительные моделей *noun2* и *noun3* за единичными исключениями тега не имеют.

7. Семантика основная и семантика дополнительная

Существует ряд прилагательных, имеющих тег 'цвет', но значение 'цвет' у них является периферийным. Например, в списке из 100 прилагательных, генерируемых функцией «Частотность» на запрос по тегу 'цвет' в ОК, мы выявили 9 таких аналоговых прилагательных (*грязный, свинцовый, стальной, хрустальный* и др.).

Рассмотрим в качестве примера из этого списка прилагательное *грязный* (рис. 14).

Лемма	грязный (см. портрет слова, см. словарь)
Грамматика	прилагательное, средний, именительный, единственное
Семантика основная	t:physst, dt:stuff, der:s, r:qual, ev:neg
Семантика дополнительная	dt:stuff, t:physq:color, der:s, r:qual, ev:neg, r:rel
Похожие слова	неопрятный (0.72532), вонючий (0.684122), грязноватый (0.68096), прегрязный (0.666575), немывтый (0.665285)
Синтаксические признаки	атрибутивный модификатор, главная клауза, глагольная клауза
Доп. признаки	genderred, numred, casered

Рис. 14. Фрагмент описания словоформы *грязное* в ОК НКРЯ

При просмотре 274 контекстов со словом *грязный* обнаружено всего 3 употребления в значении ‘цвет’, что составляет 1,1%.

Примеры:

*Теперь у нас полный праздник
на огородах напрасных
под небом **грязным** и ветреным.*

[Н. Байтов. «Какие кислые доски...» (2009–2010)]

*... черты лица у нее были какие-то закругленные, кожа заметно грубовата, белокурые волосы, как бывает это у горничных, какого-то **грязного** цвета.*
[А. Ф. Писемский. Люди сороковых годов (1869)].

В этом случае можно поставить под сомнение целесообразность присвоения слову тега ‘цвет’. Об этом же свидетельствуют списки частых коллокатов (*белье, тряпка, посуда, пол, снег, лужа*) и семантически связанных (похожих) слов (рис. 15).

грязнейший зловонный мокрый
неопрятный вонючий
пыльный **грязный** немывтый
рваный грязноватый
прегрязный

Рис. 15. Область «Похожие слова» в портрете слова *грязный*

Из 10 наиболее частотных коллокаций нет ни одной, в которой присутствовал бы признак ‘цвет’. Среди парадигматических ассоциаций (рис. 15) также отсутствуют слова с этим признаком.

Другой случай — прилагательное *свинцовый*. Из 260 просмотренных контекстов в ПК выявлено 55 употреблений в значении ‘цвет’ (21,2%). Правда, в ОК контекстов с признаком ‘цвет’ у этого прилагательного гораздо меньше.

Таким образом, в семантике прилагательных этого типа имеет место широкая шкала присутствия/отсутствия признака ‘цвет’. Повторим уже высказанную идею:

таким прилагательным целесообразно всегда приписывать тег ‘цвет’ в области «семантика дополнительная» с возможностью включать и отключать поиск по этому полю.

Обратим внимание, что эти аналоговые прилагательные также часто бывают композитами сложных цветообозначений, например: *синеваато-свинцовый, аспидно-свинцовый, грязно-рыжий, зелено-грязный*.

Приведем весьма выразительный пример непоследовательности семантической разметки:

*И стала грязно-белой грязь,
И стала грязно-синей грязь,
И стала грязно-желтой грязь
Под кистью маляра.
[А. А. Галич. Кадиш: «Сэн-Луи блюз —
ты во мне как боль, как ожог...» (1974)]*

В первом случае словоформа *грязно-белой* размечена как одна лексема и имеет тег ‘цвет’. Во втором случае словоформа *грязно-синей* также размечена как одна лексема, но тега не имеет. В третьем случае словоформа *грязно-желтой* размечена как две лексемы, при этом первый композит *грязно*, помеченный как краткое прилагательное, тега ‘цвет’ не имеет.

Рассмотрим далее прилагательные *теплый, холодный, спокойный, суровый, химический* и *чайный*. Все эти слова имеют тег ‘цвет’. В отличие от прилагательных, рассмотренных выше, которые приписывают признак ‘цвет’ широкому кругу коллокатов (*грязное небо, свинцовые тучи, пшеничные волосы*), эти прилагательные при употреблении в значении ‘цвет’ всегда относятся к существительным особого семантического подкласса, несущим признак цвета имманентно, таким как *цвет, оттенок, колорит, отлив* и т. п. Анализ контекстов в ИПС Государственного Эрмитажа (см. пример ниже) и в НКРЯ (рис. 16) показал, что эти прилагательные никогда не используются в качестве цветообозначений самостоятельно, но всегда в сочетании со словами вышеуказанного класса.

*Динамика композиции, подчеркнутая объемность форм, **теплый золотисто-коричневый колорит** и широкая живописная манера характеризуют Йорданса как последователя Рубенса и одного из ведущих мастеров фламандского барокко.* (ИПС Государственного Эрмитажа)

Чаще всего эти прилагательные в качестве цветовых характеристик используются в специальных текстах по искусствоведению. Доля их употреблений в значении характеристики цвета в обоих корпусах мизерна, сотые доли процента. Нам представляется, что слова *теплый, холодный* и т. п. в значении цветовой характеристики можно условно считать искусствоведческими терминами. Учитывая высокую частотность этих слов в различных значениях, тег ‘цвет’ для этих слов в семантической разметке является избыточным.

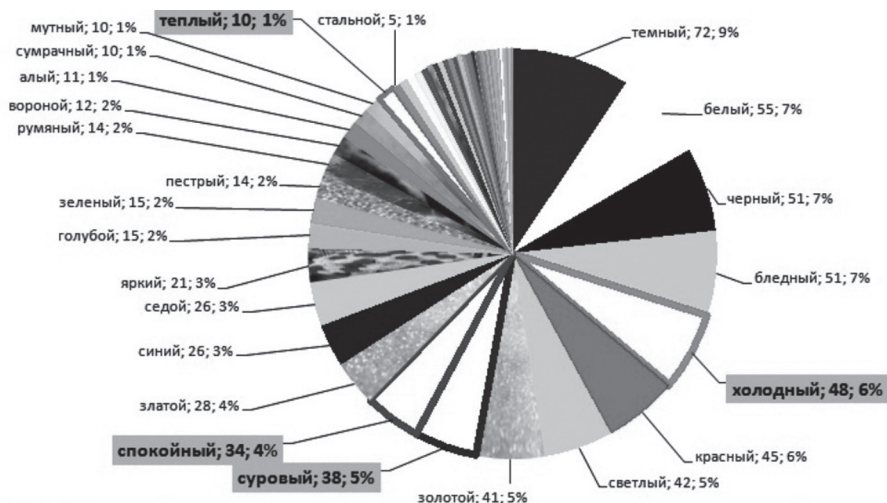


Рис. 17. Доля прилагательных-цветообозначений с тегом 'цвет' в подкорпусе текстов А. С. Пушкина в ПК НКРЯ

Если сопоставить число словоупотреблений цветообозначений, имеющих тег 'цвет', с числом словоупотреблений, не имеющих тега, то получим, что цветообозначений с тегом примерно в пять с половиной раз больше, чем без тега.

Сформулируем основные задачи, которые, на наш взгляд, следует решить для корректировки разметки цветообозначений.

- 1) Выявление с помощью словарей и разметка редко встречающейся и архаичной лексики.
- 2) Корректировка грамматической разметки в части сходных по написанию наречий и кратких прилагательных.
- 3) Корректировка разметки квазицветов и аналоговых цветообозначений. Здесь можно выделить несколько основных задач.
- 4) Определение критерия присвоения тега квазицветам, т. е. словам, характеризующим цвет, но не содержащим указания на тон. По-видимому, она может быть решена только на конвенциональной основе, то есть на основе договоренности между специалистами, занимающимися этой областью.
- 5) Определение критерия присвоения тега 'цвет' аналоговым цветообозначениям. Здесь приписывание тега должна определять частота употребления данного слова в значении 'цвет'. Остается вопрос, как ее определить и как велика должна быть пороговая частота (например, ИРМ или доля значений)?
- 6) Другой, более реалистичный вариант — присвоение этому классу цветообозначений альтернативного тега в области «семантика дополнительная». При этом необходимо в запросах иметь опцию «искать по основной или по дополнительной семантике». В настоящее время в поисковом интерфейсе

существует возможность задать первое и/или другие значения, однако в отношении цветообозначений функция работает неустойчиво.

- 7) Для сложных цветообозначений, на наш взгляд, должен быть применен простой принцип: любое сложное слово должно рассматриваться как одна лексема. При этом, если хотя бы один из композитов (слово или корневая морфема) несет значение 'цвет', этот тег должен быть приписан в области «семантика основная».
- 8) Сложной является задача корректировки разметки глаголов и существительных. Вопрос требует дополнительного изучения.
- 9) Выявлены случаи нерелевантного приписывания тега, что создает избыточный информационный шум (прилагательные *теплый, холодный, спокойный, суровый, химический* и *чайный*). Доля употребления этих слов в значении 'цвет' очень невелика. Кроме того, они никогда не определяют предмет прямо, но всегда опосредованно, через существительные типа *цвет, тон, оттенок* и т. п. На наш взгляд, тег 'цвет' должен быть удален из описаний этих слов.

Подводя итоги, отметим, что разметка, в частности семантическая, является ценным инструментом исследований, основанных на статистике. Некоторые выявленные нами ошибки в ряде случаев могут повлиять на достоверность результатов исследования, поэтому всегда нужно учитывать особенности разметки и решаемой задачи. Нам кажется, что устранение выявленных нами конкретных ошибок повысит достоверность исследований на базе НКРЯ.

Литература

- Бахилина Н. Б. История цветообозначений в русском языке. М.: Наука, 1975. 288 с.
- Василевич А. П., Кузнецова С. Н., Мищенко С. С. Цвет и название цвета в русском языке / под общ. ред. А. П. Василевича. М.: КомКнига, 2005. 216 с.
- Герасимов Г. И. Прилагательные, обозначающие цвет // Русский язык за рубежом. 1969. № 3. С. 26–35.
- Герасимов Г. И. Снежно-белый, изумрудно-зеленый: О сложных прилагательных, обозначающих цвет // Русская речь. 1978. № 2. С. 79–84.
- Дюпина Ю. В. Классификации цветообозначений в лингвистической литературе [Электронный ресурс] / Ю. В. Дюпина, Т. В. Шакирова, Н. А. Чуманова // Молодой ученый. 2013. № 1. URL: <https://moluch.ru/archive/48/6090/> (дата обращения: 20.10.2023).
- Захаров В. П., Масевич А. Ц. Категоризация прилагательных цвета в русских поэтических текстах (корпусное исследование) // Труды Института русского языка им. В. В. Виноградова. 2019. № 3 (21). С. 164–180. URL: <http://ruslang.ru/doc/trudy/vol21/9-zacharova.pdf> (дата обращения: 31.08.2023).
- Козера И. (Kozera Izabela). Семантическая характеристика русского глагола по данным НКРЯ // Acta Universitatis Lodziensis. Folia Linguistica Rossica. 2018. № 16. С. 75–92.

Кустова Г. И., Ляшевская О. Н., Падучева Е. В., Рахилина Е. В. Семантическая разметка лексики в Национальном корпусе русского языка: принципы, проблемы, перспективы // Национальный корпус русского языка: 2003–2005. Результаты и перспективы. М.: Индрик, 2005. С. 155–174.

Масевич А. Ц., Захаров В. П. Диахронический аспект семантики прилагательного красный в русских поэтических текстах: корпусное исследование // Труды Института русского языка им. В. В. Виноградова. 2022. № 2 (32). С. 107–123.

Масевич А. Ц., Захаров В. П. Частотное поведение прилагательных цвета в русских поэтических текстах // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2019. № 1 (17). С. 21–48.

Масевич А. Ц., Захаров В. П. Диахронические изменения частотности цвето-обозначений в русских поэтических текстах // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2021. № 4 (19). С. 5–35. URL: <https://lingngu.elpub.ru/jour/article/view/289> (дата обращения: 31.08.2023).

Рахилина Е. В. Когнитивный анализ предметных имен: семантика и сочетаемость. М.: Русские словари, 2008. 416 с.

Фрумкина Р. М. Цвет, смысл, сходство: Аспекты психолингвистического анализа. М.: Наука, 1984. 175 с.

Lyashevskaya O., Litvintseva K., Vlasova E., Sechina E. A Data Analysis Tool for the Corpus of Russian Poetry // HSE Working papers WP BRP 77/LNG/2018, Moscow: National Research University Higher School of Economics Press, 2018. [Электронный ресурс]. <https://wp.hse.ru/data/2018/12/29/1143017363/77LNG2018.pdf> (дата обращения: 31.08.2023).

Andrei Ts. Masevich

*St. Petersburg Institute of Culture
(Russia, St. Petersburg)
andmasev@mail.ru,*

Victor P. Zakharov

*St. Petersburg State University
(Russia, St. Petersburg)
v.zakharov@spbu.ru*

TAGGING OF COLOUR WORDS IN THE RUSSIAN NATIONAL CORPUS

The paper describes errors in tagging of color terms in the Main and Poetry corpora of the Russian National Corpus. We focus our attention on the semantic annotation. The errors could be divided in two basic groups that are the absence of the corresponding semantic tag in the descriptions of color terms or irrelevant ascribing of the tag to words which do not have the meaning of color. In addition, a mistake in grammatical tagging of adverbs and adjectives in short form which have the tag of color was

found. We discovered a few adjectives of high frequency to which the tag was ascribed irrelevantly, which causes considerable informational noise. We represent the analysis of verbs and nouns used as color terms. It was established that frequency of tagged color terms is higher than those that are not tagged. Therefore, despite the founded tagging errors the Russian National Corpus could be used as an efficient instrument of scientific researches. However, these errors can noticeably influence the results and interpretations of researches. That is why the correction of the tagging errors is advisable.

Key words: semantic annotation, Russian National Corpus, color terms, annotation errors, parts of speech, tag

References

Bakhilina N. B. *Istoriya tsvetooboznachenii v russkom yazyke* [History of colour denotations in the Russian language]. Moscow, Nauka Publ, 1975. 288 p.

Dyupina Yu. V. [Classifications of color terms in linguistic literature] (Electronic resource) / Yu. V. Dyupina, T. V. Shakirova, N. A. Chumanova. *Molodoi uchenyi* [Young scientist], 2013, no. 1. (In Russ.) Available at: <https://moluch.ru/archive/48/6090/> (accessed 20.10.2023).

Frumkina R. M. *Tsvet, smysl, skhodstvo: Aspekty psikholingvisticheskogo analiza* [Colour, Meaning, Similarity: Aspects of Psycholinguistic Analysis]. Moscow, Nauka Publ, 1984. 175 p.

Gerasimov G. I. [Adjectives denoting color]. *Russkij yazyk za rubezhom*. 1969, no. 3, pp. 26–35. (In Russ.)

Gerasimov G. I. [Snow-white, emerald-green: on complex adjectives denoting color]. *Russkaya rech'*. 1978, no. 2, pp. 79–84. (In Russ.)

Kozera I. [Semantic characterisation of the Russian verb according to RNC data]. *Acta Universitatis Lodzensis. Folia Linguistica Rossica*. 2018, no. 16, pp. 75–92. (In Russ.)

Kustova G. I., Lyashevskaya O. N., Paducheva E. V., Rakhilina E. V. [Semantic annotation of lexicon in the Russian National Corpus: principles, problems, prospects]. *Nacional'nyj korpus russkogo yazyka: 2003–2005. Rezul'taty i perspektivy* [Russian National Corpus: 2003–2005. Results and prospects]. Moscow, Indric Publ., 2005, pp. 155–174. (In Russ.)

Lyashevskaya O., Litvintseva K., Vlasova E., Sechina E. A Data Analysis Tool for the Corpus of Russian Poetry. *HSE Working papers WP BRP 77/LNG/2018*, Moscow, National Research University Higher School of Economics Press, 2018. Available at: <https://wp.hse.ru/data/2018/12/29/1143017363/77LNG2018.pdf> (accessed 31.08.2023).

Masevich A. Ts., Zakharov V. P. [Diachronic aspect of the semantics of the adjective 'red' in Russian poetic texts: a corpus study]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova* [Proceedings of the V. V. Vinogradov Russian Language Institute]. 2022, no. 2 (32), pp. 107–123. (In Russ.)

Masevich A. Ts., Zakharov V. P. [Frequency behaviour of colour adjectives in Russian poetic texts]. *Vestnik NGU. Seriya: Lingvistika i mezhkul'turnaya kommunikatsiya* [NSU Vestnik. Series: Linguistics and Intercultural Communication]. 2019, no. 1 (17), pp. 21–48. (In Russ.)

Masevich A. Ts., Zakharov V. P. [Diachronic changes in the frequency of colour denotations in Russian poetic texts]. *Vestnik NGU. Seriya: Lingvistika i mezhkul'turnaya kommunikatsiya* [NSU Vestnik. Series: Linguistics and Intercultural Communication]. 2021, no. 4 (19), pp. 5–35. Available at: <https://lingngu.elpub.ru/jour/article/view/289> (accessed 31.08.2023). (In Russ.)

Rakhilina E. V. *Kognitivnyi analiz predmetnykh imen: semantika i sochetaemost'* [Cognitive analysis of content names: semantics and combinability]. Moscow, Russkie Slovari Publ., 2008. 416 p.

Vasilevich A. P., Kuznecova S. N., Mishchenko S. S. *Tsvet i nazvanie tsveta v russkom yazyke* [Color and color names in the Russian language] / ed. by A. P. Vasilevich. Moscow: KomKniga Publ., 2005. 216 p. (In Russ.)

Zakharov V. P., Masevich A. Ts. [Categorisation of adjectives of colour in Russian poetic texts (corpus study)]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova* [Proceedings of the V. V. Vinogradov Russian Language Institute]. 2019, no. 3 (21), pp. 164–180. Available at: <http://ruslang.ru/doc/trudy/vol21/9-zacharova.pdf> (accessed 31.08.2023). (In Russ.)

Маркасова Елена Валерьевна

Пекинский университет

(Китай, Пекин)

markasovaelena@yandex.ru

Митрофанова Ольга Александровна

Санкт-Петербургский государственный университет

(Россия, Санкт-Петербург)

o.mitrofanova@spbu.ru

ГЕМИНАЦИЯ (ТРОЙНОЙ ПОВТОР) ПО КОРПУСНЫМ ДАННЫМ

С позиций грамматики конструкций геминация представляет собой особую лексико-грамматическую конструкцию, на лексическом уровне определяемую единственной фиксированной лексической единицей, заполняющей первый и два последующих слота, на синтаксическом уровне организованную как союзный/бессоюзный сочиненный ряд. В статистико-комбинаторном аспекте геминации представляют собой триграммы, к которым применимы корпусные методы оценки степени устойчивости и методы определения степени тяготения лексических единиц к слотам данной конструкции. В риторике геминация — вид контактного повтора, отличающийся от редупликации количеством элементов и функциями. Эта фигура может служить для описания большого количества однородных объектов, длительности или монотонности процессов, а также отражать погруженность лирического героя в определенное состояние, сфокусированность внимания на чем-либо. Современный уровень развития корпусной лингвистики позволяет рассмотреть поэтику геминации как фигуры на фоне троекратных повторов в языке в целом. В данном исследовании представлен анализ употребления геминаций в разножанровых корпусах текстов на русском языке, изучены лексико-грамматические классы слов, тяготеющих к конструкциям с геминациями. Сделанные наблюдения подкреплены количественной информацией о распределении геминаций в исследовательских корпусах.

Ключевые слова: повторная номинация, контактный повтор, тройной повтор, геминация, редупликация, интертекстуальность, корпус текстов, лексико-грамматические конструкции

1. Введение. Терминология, методология, материал

1.1. Геминация как термин риторики

Термин *геминация* используется в современной лингвистике (конкретнее — в фонетике и словообразовании) для обозначения удвоения звука, в риторике — для обозначения лексического повтора. Как показывают данные корпуса Роэлли, *геминация* в латиноязычных риториках и грамматиках обозначала любые виды повтора. В отечественной лингвистике к классической традиции примыкают толкования геминации как лексического контактного повтора [Хазагеров 2009: 242] и лексического, звукового, слогового контактного повтора [Москвин 2006: 81]. Однако возможно и именованное этим термином двукратного дистантного повтора [Шмараков 2015: 244].

Функции геминации перечислены в словарях риторических терминов. Эта риторическая фигура может служить для описания множества однотипных объектов и тем самым показывать масштаб наблюдаемого, передавать ощущение длительности или монотонности процессов, погруженность лирического героя или автора в определенное состояние и др. Считается, что фигура «особенно пригодна для подчеркивания стабильности, непреложности конечного вывода... причем вывод этот часто носит пессимистическую и отрицательную окраску» [Хазагеров, Ширина 1994: 289]. Выделяют «эмотивное значение» (задумчивость, сожаление, ламентация), взволнованность, и кумулятивное значение, «заставляющее воспринимать ФТП (формулу тройного повтора — Е. М., О. М.) как подобие градации с формально совпадающими членами» [Свиридов 2012]. Есть мнение, что геминация используется «для конкретизации, уточнения, актуализации изображаемого» [Коняшкин, Байкова, 2015: 338].

В отличие от анафоры, эпифоры, дистинкции, обозначающих конкретные виды повтора, геминацию можно считать термином второго порядка, не имеющим «специализации». В русистике XXI века появляется трактовка геминации как *троекратного* контактного повтора [Матвеева 2010, Копнина 2014], что не менее удивительно, чем включение в определение геминации количественной характеристики, противоречащей этимологии (лат. *geminio* «удваиваю»). Можно предположить, что переосмысление термина обусловлено распространением экспрессивного повтора именно в литературе XX века [Кожевникова. 1994: 198]. Спецификация в употреблении термина требует изучения процессов, происходящих в первую очередь в языке, и лишь во вторую — в риторической терминологии.

Цель нашей работы — описать геминацию на основе корпусного подхода, принимая во внимание три параметра: частеречный состав, функционально-стилистический аспект, исторические изменения на уровне корпусов текстов.

Далее речь пойдет лишь о случаях типа «было, было, было», «слова, слова, слова», а в качестве синонимов будут использованы термины геминация, тройной повтор, троекратный повтор. Термины эпизевксис, анадиплосис, дистинкция обозначают двойной повтор и не применяются нами как синонимы геминации. Случаи дистантного тройного повтора не учитываются.

При работе с корпусными данными контексты типа «Вьюн над водой, ой, вьюн над водой,/ Вьюн над водой расстилается» мы в выборку не включали, поскольку, во-первых, междометие «ой» делает повтор дистантным, во-вторых, в примерах такого рода повторяющиеся элементы не являются собой единство и группируются как $1 + 2$ или как $2 + 1$. Также не рассматриваются примеры типа: «Тем, чей путь таинственно суров, Чья душа тоскою *осиянна*, Вы — цветы *осенних* вечеров, Поздних зорь далекая *Осанна*» (М. Волошин), хотя в лингвистике они описываются как «тройной повтор» [Кожевникова 2009: 310]. Вопрос о «песенном характере» тройного повтора в статье не рассматривается.

1.2. Геминация как лексико-грамматическая конструкция

Геминации как троекратные повторы лексических единиц представляют большой интерес для изучения сочетаемости. В этом аспекте геминация — зона пересечения парадигматических и синтагматических отношений, поскольку представляет собой комбинацию форм — представителей одной лексической парадигмы в синтагматической цепочке — сочиненном ряде. В формальном отношении сочетаемость отражает комбинаторный потенциал корпуса текстов, а в содержательном — требования к лексическому значению, морфосинтаксическим свойствам и дискурсивно-прагматическим маркерам контекстного окружения целевого слова.

В рамках нашего исследования применен конструкционный подход к анализу геминаций, опирающийся на работы отечественных и зарубежных представителей лингвистики конструкций.

Одним из тезисов конструкционного подхода является взаимообусловленность лексического значения слова и его сочетаемостных свойств. В исследованиях Д. Н. Шмелева утверждается, что употребление слов, представляющих определенную лексико-семантическую группу, в тех или иных конструкциях можно объяснить общностью семантических признаков целевого слова и элементов контекста [Шмелев 1973]. В. Г. Гак в этой связи говорил о синтагматически обусловленных номинациях [Гак 1998: 317], Ю. Д. Апресян — о неаддитивном сложении смыслов и об отношениях семантического согласования в конструкциях с повторением семантических признаков в сочетающихся словах [Апресян 1995]. Эти идеи развиваются в работах по семантическим и синтаксическим валентностям Московской семантической школы Е. В. Падучевой, Е. В. Рахилиной и коллег: слова Y ($y_1, y_2, \dots, y_n$), заполняющие валентность в конструкции с целевым словом X , образуют допустимые сочетания при наличии в их толкованиях общего семантического признака S [Рахилина, Плунгян 1998; Рахилина 2010]. Из этого следует, что наряду с сочетаемостными предпочтениями могут быть описаны и сочетаемостные ограничения слов. Более явно это утверждение сформулировано И. А. Мельчуком и Л. Н. Иорданской при рассмотрении фразем как особой разновидности конструкций: «Фразему можно определить как такое сочетание, означаемое и означающее которого не могут быть построены без ограничений и регулярно» [Мельчук, Иорданская 2007: 226]. Характеристика геминации как проявления нестандартной сочетаемости в синтагматической последовательности сформулирована в исследованиях

Л. Л. Иомдина, посвященных интегральному описанию микросинтаксических конструкций: синтаксических фразем русского языка, к которым относятся и сочинительные конструкции с лексически тождественными частями [Иомдин 2006].

Это определение согласуется с пониманием конструкции, принятым в работах Ч. Филлмора и его последователей А. Голдберг, М. Томаселло и других [Fillmore 1988; Fillmore et al. 1988; Goldberg 1995, 2006; Tomasello 2003]: конструкция — это «языковое выражение, у которого есть аспект плана выражения или плана содержания, не выводимый из значения или формы составных частей» [Goldberg 1995]. Тем самым, лингвистика конструкций, ориентированная на взаимодействие явлений разных уровней языка в процессе их функционирования в речи, позволяет перейти от описания атомарных, изолированных компонентов словаря к их комплексам как целостным языковым знакам. Этот переход осуществлен в корпусно-ориентированных исследованиях конструкций, проводимых А. Стефановичем и С. Грисом, которые ввели для исследования сочетаемости понятие коллострукции (collostruction) и разработали комплексную методику коллострукционного анализа, включающую три этапа: колексемный анализ, различительный колексемный анализ и ковариация колексем [Stefanowitsch, Gries 2003]. Коллострукционный анализ основывается на предположении о неслучайности сочетаемости и выбора лексем в коллострукциях, а также о конкуренции сходных по смыслу конструкций по отношению к той или иной лексеме. Реализация коллострукционного анализа предполагает построение таблиц сопряженности и оценку данных по точному критерию Фишера и другим мерам ассоциации [Букия и др. 2015].

В количественных описаниях конструкций как устойчивых n-грамм большее внимание уделялось двухкомпонентным сочетаниям (биграммам), тогда как трехкомпонентные сочетания (триграммы) до последнего времени оставались менее изучены. Наше исследование восполняет существующие пробелы в отношении геминаций как разновидностей триграмм. В свете содержательных и статистических подходов к изучению конструкций и их трактовки «исключения как правила» [Копотев, Стеклова 2016] можно высказать следующие аргументы в пользу того, что геминации целесообразно рассматривать как неслучайные троекратные повторы в рамках конструкции:

- 1) для геминаций есть ограничения на сферу употребления: они по-разному функционируют в текстах разной жанрово-стилевой принадлежности и в разных типах дискурса,
- 2) не все лексемы склонны к троекратным повторам: в словаре выделяется ряд лексико-семантических групп слов, значения которых предрасположены к троекратному повтору в составе геминаций.

Эти тезисы нашли подтверждение в процессе проведенного нами корпусного эксперимента.

1.3. Общая характеристика используемых корпусов

Для оценки регулярности употребления геминаций и определения жанрово-стилевой специализации контекстов с геминациями было использовано шесть

подкорпусов (поэтический, устный, прозаический, юридический, научный и публицистический), разработанных в ходе реализации проектов по изучению лексико-грамматических конструкций в русском языке [https://vintagenteleman.github.io/qt_prep_gram/corpora/].

Объем корпусов составляет от 1,2 до 1,26 млн с/у. Случайность отбора текстов и их объем, пригодный для оценки параметров корпусов с точки зрения количественных показателей ipm (item per million), обеспечивает репрезентативность материала и объективность полученных результатов. В подкорпусах представлены основные функциональные стили русскоязычных текстов, при этом художественный стиль представлен двумя разновидностями: поэзия и проза. Корпусы собраны из источников второй половины XX в. — начала XXI в. и фиксируют актуальные тенденции в употреблении языковых средств. Репрезентативность корпусов обусловлена их равновесностью, разнообразием тематики и авторства текстов, равнообъемностью текстов внутри подкорпусов. Подкорпус поэтических текстов собран на основе материалов портала Стихи.ру [URL: <https://stihi.ru/>] и включает в себя около 8 тыс. стихотворений современных авторов. Его хронологические рамки — конец 1990-х годов — первое десятилетие XXI в. По мнению А. А. Бонч-Осмоловской и Б. В. Орехова, специфика данного подкорпуса определяется тем, что в него включены тексты «наивной» поэзии, которые передают черты поэтического канона, в то же время демонстрируя его модификации и живо откликаясь на тематические доминанты времени [Бонч-Осмоловская, Орехов 2013; Орехов 2022]. Подкорпус прозаических текстов содержит оригинальные тексты русскоязычных писателей объемом примерно от 20 до 40 тыс. с/у. Подкорпус устных текстов содержит расшифровки передач «Эхо Москвы», кинесубтитры и фрагменты из «Сбалансированной аннотированной текстотеки» [URL: <https://ord-corpus.spbu.ru/>], представляющие собой спонтанные монологи. Научный подкорпус включает специальные и популярные тексты по точным, естественным и гуманитарным наукам. В юридическом подкорпусе содержатся тексты федеральных законов. Публицистический подкорпус основан на материалах российских газет и журналов с конца 1990-х до 2020-х гг.

Кроме перечисленных корпусов мы использовали Национальный корпус русского языка НКРЯ [URL: <http://ruscorpora.ru>] и информационно-поисковую систему Яндекс. Поиск в НКРЯ позволяет находить не только геминации, но и целые тексты с большим количеством геминаций, а также авторов и группы авторов, склонных к использованию этой конструкции.

1.4. Автоматическое выделение геминаций

Для выделения геминаций была разработана и применена следующая процедура анализа корпусных данных:

- 1) предобработка текстов, фильтрация нетекстовых элементов, ссылок, элементов типографского оформления,
- 2) определение точных объемов подкорпусов в словоупотреблениях,

- 3) токенизация и выделение триграмм с помощью библиотеки `spacy` [URL: <https://spacy.io/>], определение триграмм с контактными повторами (истинных геминаций),
- 4) частеречная разметка геминаций, количественная оценка геминаций разных частей речи, соотношение геминаций со знаменательными и служебными словами,
- 5) расчет коэффициентов геминантности подкорпусов текстов:
 коэффициент геминантности $G_{\text{общ}} = n / N$, где n — общее число геминаций в подкорпусе, N — объем подкорпуса в словоупотреблениях,
 коэффициент геминантности $G_{\text{знам}} = n_{\text{знам}} / N$, где $n_{\text{знам}}$ — общее число геминаций со знаменательными словами в подкорпусе, N — объем подкорпуса в словоупотреблениях,
 коэффициент геминантности $G_{\text{служ}} = n_{\text{служ}} / N$, где $n_{\text{служ}}$ — общее число геминаций со служебными словами в подкорпусе, N — объем подкорпуса в словоупотреблениях.

2. Состав геминаций по данным корпусов

2.1. Общие количественные данные

В Таблице 1 приводятся данные о геминациях с учетом функционально-стилистического аспекта и частеречной принадлежности структурообразующих единиц.

Таблица 1. Количественные данные о геминациях в подкорпусах текстов разных функциональных стилей

	Поэтический подкорпус	Устный подкорпус	Прозаический подкорпус	Научный подкорпус	Юридический подкорпус	Публицистический подкорпус
Объем в с/у	1 202 259	1 260 852	1 245 766	1 246 841	1 260 972	1 256 086
Общее число геминаций	509	468	180	43	38	22
Число знаменательных геминаций	136	233	54	8	0	10
Число служебных геминаций	373	235	126	35	38	12
Коэффициент геминантности $G_{\text{общ}}$	$4,23 \times 10^{-4}$	$3,71 \times 10^{-4}$	$1,44 \times 10^{-4}$	$0,34 \times 10^{-4}$	$0,3 \times 10^{-4}$	$0,17 \times 10^{-4}$
Коэффициент геминантности $G_{\text{знам}}$	$1,13 \times 10^{-4}$	$1,85 \times 10^{-4}$	$0,43 \times 10^{-4}$	$0,06 \times 10^{-4}$	$0,00 \times 10^{-4}$	$0,07 \times 10^{-4}$
Коэффициент геминантности $G_{\text{служ}}$	$3,1 \times 10^{-4}$	$1,86 \times 10^{-4}$	$1,01 \times 10^{-4}$	$0,28 \times 10^{-4}$	$0,3 \times 10^{-4}$	$0,1 \times 10^{-4}$
Распределение по частям речи						
Существительные	54	32	13	7	0	2

Окончание табл. 1

	Поэтический подкорпус	Устный подкорпус	Прозаический подкорпус	Научный подкорпус	Юридический подкорпус	Публици- стический подкорпус
Глаголы	46	97	17	0	0	6
Прилагательные	6	5	3	1	0	1
Наречия	21	34	7	0	0	1
Числительные	1	5	1	0	0	0
Предикативы	8	60	13	0	0	0
Местоимения	5	29	6	0	0	1
Частицы	3	23	2	1	0	0
Союзы	3	2	1	0	0	0
Междометия и звукоподражания	82	181	96	0	0	8
Элементы с неопреде- ленной частеречной принадлежностью	280	0	21	34	38	3

Было установлено, что геминация — это общезыковое явление, поскольку оно зарегистрировано во всех подкорпусах. Полученные данные свидетельствуют о неравномерности употребления геминаций в текстах разных стилей. Три из шести подкорпусов (поэтический, устный и прозаический) содержат геминации в значимых объемах, что было ожидаемо. Для других трех подкорпусов (научного, юридического и публицистического) геминации нетипичны, но если для научного и юридического подкорпуса это было предсказуемо, то для публицистического ожидалось другие выводы. Далее наше внимание сосредоточено на поэтическом, устном и прозаическом корпусах.

2.2. Анализ геминаций в поэтическом, устном и прозаическом корпусах

Самые высокие значения коэффициентов геминантности были зарегистрированы в поэтическом подкорпусе ($G_{\text{общ}} = 4,23 \times 10^{-4}$, $G_{\text{знам}} = 1,13 \times 10^{-4}$, $G_{\text{служ}} = 3,1 \times 10^{-4}$). Обращает на себя внимание то, что доля геминаций со служебными словами в три раза выше, чем доля соответствующих конструкций со знаменательными. Это вызвано большим числом троекратных повторов междометий и звукоподражательных слов (*ай-ай-ай*; *буль-буль-буль*; *ла-ла-ла*; *привет-привет-привет*; *тук-тук-тук*; *тьфу-тьфу-тьфу* и др.). Рассмотрим более детально реализации геминаций со знаменательными словами в поэтическом подкорпусе. Существительные в поэтических геминациях обозначают участников ситуации (*доктор, доктор, доктор*; *мама, мама, мама* и др.), время (*весна, весна, весна*; *зима, зима, зима*; *февраль, февраль, февраль* и др.), место (*карусель, карусель, карусель*; *место, место, место* и др.), типы событий (*праздник, праздник, праздник*; *рождение, рождение, рождение* и др.). Геминации с глаголами встречаются в поэтическом подкорпусе реже, чем конструкции с существительными. Глагольная лексика в геминациях

представлена глаголами движения (*бегу, бегу, бегу; иди, иди, иди; лети, лети, лети* и др.), мыслительной деятельности (*помню, помню, помню* и др.), восприятия (*вижу, вижу, вижу; слышишь, слышишь, слышишь* и др.), речевых действий (*зовет, зовет, зовет; молчите, молчите, молчите* и др.), эмоционального состояния (*любить, любить, любить* и др.). Приведенные примеры свидетельствуют о лексическом разнообразии глаголов и существительных. Доля наречий в составе геминаций выше, чем доля прилагательных, при этом наречия обозначают направление и периодичность движения (*вперед, вперед, вперед; мимо, мимо, мимо; глубоко, глубоко, глубоко* и др.). Предикативы представлены словами *да, да, да; нет, нет, нет* и др.

По представленности геминаций за поэтическим следует устный подкорпус ($G_{\text{общ}} = 3,71 \times 10^{-4}$, $G_{\text{знам}} = 1,85 \times 10^{-4}$, $G_{\text{служ}} = 1,86 \times 10^{-4}$). В этом подкорпусе доли геминаций со знаменательными и служебными словами приблизительно равны. Среди существительных много имен собственных (*Ривьера, Ривьера, Ривьера; Сакраменто, Сакраменто, Сакраменто* и др.), в том числе и личных имен, выступающих в функции обращения (*Марина, Марина, Марина; Оль, Оль, Оль; Рома, Рома, Рома* и др.). Глаголы встречаются в геминациях чаще, чем существительные, при этом частотны геминации с глаголами движения (*встать, встать, встать; едем, едем, едем; идешь, идешь, идешь; уходи, уходи, уходи* и т. д.), физиологического действия (*дышать, дышать, дышать; пей, пей, пей* и др.), а также конструкции со словом *давай*, частеречная принадлежность которого дискурсивно зависима. Специфичны для устной речи геминации *Господи, Господи, Господи; пожалуйста, пожалуйста, пожалуйста (ха-ха-ха; кхе-кхе-кхе* и др.). Как и в поэтическом корпусе, геминации с наречиями более частотны, чем геминации с прилагательными, сохраняется тенденция к употреблению в геминациях наречий образа действия (*бегом, бегом, бегом; быстро, быстро, быстро, вперед, вперед, вперед; тихо, тихо, тихо* и др.) и времени (*сегодня, сегодня, сегодня; сейчас, сейчас, сейчас* и др.). Доли геминаций с существительными и местоимениями примерно равны. Из местоимений в конструкциях с геминациями встречаются преимущественно *всё, всё, всё; то, то, то; я, я, я* и др. Среди предикативов в геминациях употребляются помимо *да, да, да; нет, нет, нет* слова категории состояния с эмоциональным компонентом в значении: *больно, больно, больно; спокойно, спокойно, спокойно* и др. По сравнению с конструкциями в поэтическом подкорпусе геминации в устных текстах более разнообразны с точки зрения их частеречной принадлежности и в то же время более специфичны по своему лексическому составу.

С точки зрения распространенности геминаций после поэтического и устного подкорпусов следует прозаический подкорпус, в котором троекратных повторов меньше ($G_{\text{общ}} = 1,44 \times 10^{-4}$, $G_{\text{знам}} = 0,43 \times 10^{-4}$, $G_{\text{служ}} = 1,01 \times 10^{-4}$). По соотношению знаменательных и служебных частей речи в геминациях прозаический корпус схож с поэтическим: геминаций со знаменательными частями речи существенно меньше, чем конструкций со служебными словами. Междометные геминации и звукоподражания в прозаических текстах более разнообразны, чем в поэтическом и устном подкорпусах (*а-а-а; бух-бух-бух; о-о-о; фиу-фиу-фиу; хи-хи-хи; э-э-э* и др.).

Существительные в прозаических геминациях весьма разнородны по семантике (*ветер, ветер, ветер; дикарь, дикарь, дикарь; мама, мама, мама; окуньки, окуньки, окуньки; пыль, пыль, пыль* и др.). Среди глаголов чаще всего встречаются глаголы движения (*выйти, выйти, выйти; движутся, движутся, движутся; ехали, ехали, ехали; идут, идут, идут; падали, падали, падали* и др.). Прилагательные и наречия в троекратных повторах обозначают скорость и направление движения (*быстрее, быстрее, быстрее; вниз, вниз, вниз; шагом, шагом, шагом* и др.). Предикативы представлены широко распространенными геминациями *да-да-да, нет-нет-нет*, а также эмоционально окрашенными словами категории состояния (*пусто, пусто, пусто; страшно, страшно, страшно; холодно, холодно, холодно* и др.). В целом геминации в поэтических и прозаических текстах однородны, однако прозаические геминации сложнее поддаются систематизации.

Поскольку геминации относятся к эмфатическим конструкциям, в исследуемых подкорпусах они в основном используются как для выделения отдельных аспектов ситуаций, описываемых в текстах, так и для усиления эмоционально-оценочных компонентов значений троекратно повторяющихся слов. Это позволяет предположить, что слова, тяготеющие к использованию в геминациях, имеют явный или скрытый субъективно-модальный компонент в значении, который и формирует семантику геминации как особой лексико-грамматической конструкции. Корпусной подход дает возможность обнаруживать события, описываемые с помощью этой фигуры. Во-первых, это события частной жизни или исторические события, сопровождающиеся утратами и их предчувствием или осознанием невозможности благополучного исхода. Во-вторых, это события, предполагающие продолжительность во времени или протяженность в пространстве. События обоих типов предполагают высокий уровень концентрации внимания их участника или наблюдателя на повторяемом элементе. Имеет смысл разграничивать геминации, сохранившие эмоционально-оценочный компонент, и геминации, его утратившие. К последним, например, относится «да-да-да» (в отличие от «Да! Да! Да!»).

2.3. Сравнение геминаций в поэтических корпусах разных хронологических периодов

Следуя гипотезе, выдвинутой в [Маркасова 2002], мы признаем, что употребление геминаций в текстах изменяется во времени и подвержено влиянию общественно-политических и историко-культурных факторов. Наши данные позволяют сравнить употребления геминаций в поэтическом подкорпусе текстов конца 1990-х гг. — первого десятилетия XXI в. и в подкорпусе с более широкими хронологическими рамками — начало XIX в. — конец XX в. (далее подкорпусы 1 и 2 соответственно). В качестве подкорпуса 2 был использован сегмент поэтического корпуса И. Гусева [URL: <https://github.com/IlyaGusev/PoetryCorpus>], включающий в себя примерно то же число контекстов с геминациями, что и в современном поэтическом подкорпусе. Объем подкорпуса 2 примерно вдвое больше объема подкорпуса 1. Количественные данные по подкорпусам, которые получены согласно схеме, описанной в разделе 2.3, представлены в Таблице 2.

Таблица 2. Количественные данные о геминациях в поэтических подкорпусах

	Поэтический подкорпус 1	Поэтический подкорпус 2
Объем в с/у	1 202 259	2 289 658
Общее число геминаций	509	605
Число знаменательных геминаций	136	335
Число служебных геминаций	373	246
Коэффициент геминантности $G_{\text{общ}}$	$4,23 \times 10^{-4}$	$2,53 \times 10^{-4}$
Коэффициент геминантности $G_{\text{знам}}$	$1,13 \times 10^{-4}$	$1,46 \times 10^{-4}$
Коэффициент геминантности $G_{\text{служ}}$	$3,1 \times 10^{-4}$	$1,07 \times 10^{-4}$
Распределение по частям речи		
Существительные	54	113
Глаголы	46	138
Прилагательные	6	25
Наречия	21	31
Числительные	1	3
Предикативы	8	25
Местоимения	5	20
Частицы	3	3
Союзы	3	0
Междометия и звукоподражания	82	232
Элементы с неопределенной частеречной принадлежностью	280	9

Следует обратить внимание на следующие факты.

1. В подкорпусах 1 и 2 различное соотношение геминаций со знаменательными и служебными частями речи: в подкорпусе 1 больше конструкций со служебными словами, в подкорпусе 2 преобладают конструкции со знаменательными частями речи. Этот факт можно объяснить вторжением разговорной стихии в язык поэзии последних 30–40 лет. Именно незнаменательные слова являются строительным элементом троекратных повторов в спонтанной речи [Шерстинова 2017].

2. В подкорпусах 1 и 2 различное соотношение глагольных и именных конструкций с геминациями: в подкорпусе 1 несколько больше геминаций с существительными, в подкорпусе 2 чаще встречаются геминации с глаголами. Это наблюдение позволяет связать наши данные с концепцией именных и глагольных стилей в поэзии [Иванов Вяч. Вс. 1978; Лотман 1982] Хотя мы не имеем права возводить это частное наблюдение в ранг вывода, все-таки можно предположить, что современная поэзия тяготеет к именному стилю, в отличие от поэзии предшествующего периода.

3. В подкорпусах 1 и 2 различные доли междометных геминаций: их в подкорпусе 2 почти в четыре раза больше, чем в подкорпусе 1.

4. В подкорпусе 2 велико разнообразие контекстов с троекратными повторами прилагательных, в основном цветообозначений (*багряна, багряна, багряна*;

белый, белый, белый; синий, синий, синий, и др.), большее (по сравнению с подкорпусом 1) разнообразие лексико-семантических групп существительных с абстрактным (*горе, горе, горе; мгновение, мгновение, мгновение; тревога, тревога, тревога; чудо, чудо, чудо* и др.) и конкретным (*город, город, город; деревня, деревня, деревня; сады, сады, сады* и др.) значением, специфичность имен собственных (*Пушкин, Пушкин, Пушкин; Россия, Россия, Россия* и др.). В глагольной лексике отмечается преобладание двух лексико-семантических групп: глаголы движения (*движутся, движутся, движутся; идут, идут, идут; мчится, мчится, мчится* и др.) и эмоционального отношения (*любит, любит, любит* и др.). Среди геминаций с предикативами (*боязно, боязно, боязно; жалобно, жалобно, жалобно* и др.) также преобладают единицы, выражающие эмоции.

Количественные данные подкорпуса 1 по сравнению с характеристиками подкорпуса 2 указывают на повышение частотности геминаций в поэзии конца XX в. — начала XXI в., повышение номинативности геминаций и снижение лексического разнообразия конструкций с геминациями.

3. Перспективы изучения геминации на основе полученных данных

3.1. Геминация как конструкция и геминация как фигура

Благодаря собственно грамматическому (формальному) подходу к выделению геминации вычлениются все контактные повторы, состоящие из трех элементов. Все случаи тройного повтора делятся на две группы: сохранившие/не сохранившие изобразительно-выразительный потенциал, экспрессивные/не экспрессивные. Можно говорить о стертых и живых геминациях (подобно стертым и живым метафорам). Риторический потенциал сохраняют в основном знаменательные части речи, но и среди них есть не сохранившие экспрессивность и не имеющие риторической нагрузки. К стертым геминациям можно причислить повторы частиц, междометий, слов-предложений (типа *вот-вот-вот, ай-ай-ай, да-да-да*), описанные на материале ОРД [Шерстинова 2017]. Нельзя принять утверждение автора, что «трехэлементные контактные лексические повторы в русской устной речи по популярности близки к двухэлементным» [Шерстинова 2017: 165]. Сравнение распространенности двух разных конструкций в условиях невозможности четко дифференцировать случаи с разными видами группировки элементов в каждом примере делает этот вывод сомнительным. Зафиксированные в ОРД случаи доведения повторов до 5–8 обусловлены желанием говорящего усилить их позицию за недостаточностью уже автоматизированного тройного повтора. Это попытка стертых геминаций восстановить былую экспрессивность. В исследованиях Е. А. Гришиной, посвященных устной речи, геминации как разновидность повторов последовательно анализируются наравне с разного рода интенсификаторами [Гришина 2007]. Эксперименты, проведенные на материале корпуса МУРКО, показали, что для повторов предикативов *да* и *нет* в устной речи характерно усиление интенсивности реакции коммуникантов, при этом такие повторы в речевой ситуации носят векторный характер, то есть

направлены на определенное лицо внутри или вне ситуации [Гришина 2011: 183]. В этом заключается главное отличие повторов *да* и *нет* от их одиночного употребления, где предикативы маркируют пропозициональное содержание предшествующего контекста.

У корпусного подхода есть и ограничения, связанные с архитектурой корпус-менеджеров. Повторяющиеся единицы, представленные существительными, прилагательными, глаголами, местоимениями, наречиями, числительными, предлогами, союзами, междометиями, в корпусах находить легко, однако более крупные единицы (словосочетания, предложения, фрагменты текста, состоящие из нескольких предложений) обнаружить с помощью поиска может быть затруднительно.

3.2. Уникальные и не уникальные геминации

Все геминации, не относящиеся к стертым, можно условно разделить на а) уникальные и б) не уникальные.

Строительным материалом уникальной геминации может быть словоформа, словосочетание и более сложные образования (предложение или фрагмент текста). Например: *Пыль подо мною! Пыль подо мною. Пыль подо мною...* (А. Шмелев)). Уникальные геминации, представленные большими языковыми единицами, встречаются (особенно в песенных жанрах) чаще, чем уникальные геминации, представленные отдельным словом. Например: *Нам жарко так с тобой, как в Чили-Чили-Чили* (б.а., б.г. б.м.). Случаи, представленные троекратным повтором одной словоформы можно считать уникальными, пока не доказано другое.

Не уникальными можно считать любые случаи, обнаруженные более двух раз в разные исторические периоды. Например, таковы геминации «слова, слова, слова» (более 100 примеров), «я, я, я» (более 200 примеров). Существование не уникальных геминаций сигнализирует о необходимости выявления интертекстуальных связей.

Можно говорить о больших группах геминаций, имеющих общий прототип. Примеры группируются в «гнезда» и «цепочки», восходящие к общему прецедентному тексту. Отношения внутри такой группы примеров представляют собой три типа связи (как в лексической семантике): **радиальную (геминации восходят к одному тексту)**, **цепочечную** и **радиально-цепочечную**, подобно тому, как это описано для полисемии в [Апресян 1974: 182].

3.3. Факторы, влияющие на восприятие геминации

В процессе обработки материала было обнаружено, что геминация может открывать или закрывать предложение, строфу, текст. Если структурообразующими являются знаменательные части речи, от их положения может зависеть восприятие геминации как «позитивной» или «негативной». Для стертых геминаций, выраженных частицами, междометиями, словами «да» или «нет» (особенно в устной речи), это положение непринципиально. Для поэтических текстов добавляется фактор ритма и метра, наличия или отсутствия анжамбемана.

Восприятие геминации может быть обусловлено и другими факторами, а именно:

- а) лексическое значение и грамматическая семантика повторяющихся элементов;
- б) иллокутивная модальность;
- в) пунктуационное оформление (запятая, точка, восклицательный знак, вопросительный знак, многоточие, сочетание разных знаков);
- г) интертекстуальные связи конкретного примера;
- д) включенность повторяющейся единицы в различные дискурсы (политический, социальный, исторический и др.);
- е) положение геминации в рамках более крупной языковой единицы (предложение, текст).

В 2002 году нами был проведен эксперимент. Группе студентов, не слышавших песню Б. Окуджавы «По Смоленской дороге...», предложили продолжить текст, начинающийся таким образом:

- 1) Столбы, столбы, столбы по Смоленской дороге...
- 2) По Смоленской дороге столбы, столбы, столбы...

В первом случае продолжением стали патетические рассуждения о достижениях электрификации и перспективах реформ на селе. Во втором случае — рассуждения об одиночестве, дальней дороге и безысходности.

Чем была обусловлена такая реакция? Мы предполагаем, что в первом случае троекратный повтор не имеет семантики неизвестного и недосказанного, предложение читается как известное *столбы* (на котором фокусируется внимание) плюс известное *по Смоленской дороге*. Чтение разворачивается слева направо, читатель ждет продолжения истории. Во втором случае работает поэтика недосказанного: читатель намерен узнать подробности, но после троекратного повтора (в сильной позиции) и многоточия картинка не обретает законченности.

Начальная геминация создает разбег для развертывания текста. Так, в тексте В. Маяковского «Вставайте, вставайте, вставайте, /Работники и батраки», мы видим, что строка продолжается, и предугадываем следующее после глагола обращение. Предсказуемость порождает у читателя ощущение уверенности, поскольку дает ему почувствовать себя «проницательным». Это поэтика определенности и завершенности. Если троекратный повтор оказывается в конце строфы или завершает текст, внимание смещается вправо и оказывается в пустоте, в неизвестности, в пугающей многозначности. Русские тексты с геминацией основаны именно на этом принципе. «Начальное» или «конечное» положение тройного контактного повтора — важный дифференцирующий признак для условно оптимистических и пессимистических текстов.

Чтобы установить, какие из этих факторов главные, а какие второстепенные, необходимо провести ряд экспериментов на основе собранного нами материала, а также привлечь представительные корпусные источники в составе НКРЯ.

Заключение

Благодаря корпусному подходу мы можем с уверенностью говорить, что геминация — общеязыковое явление, неравномерно распространенное в разных стилях. Частеречное наполнение геминации различно в поэзии и прозе, имеет специфику в каждом функциональном стиле. Высокий уровень геминантности характерен для поэзии и устной речи, проза же занимает третье место в этой иерархии. Существуют лексико-семантические группы слов, склонных к участию в троекратном повторе.

Повторяющиеся единицы могут быть представлены отдельными словоформами, существительным, прилагательным, глаголом, местоимением, наречием, числительным, предлогом, союзом, междометием, а также более крупными языковыми единицами (словосочетаниями, синтаксемами, предложением, фрагментом текста).

Соотношение частотности контекстов целевых слов с геминациями и без таковых указывает на специфичность геминации как языкового явления и на ее отличие от рядов как синтаксического явления. Различия в частотности геминаций для разных слов указывают на существующие ограничения в употреблении данной конструкции.

Исследование позволяет считать, что обновление термина «геминация» в русистике обусловлено ростом употребления тройных повторов в поэзии и прозе конца XX — начала XXI вв. и изменениями состава повторяемых элементов. Живые процессы дали импульс к изменению терминологии, что естественно, поскольку терминологические системы зависимы от появления новых объектов исследования и возможности наблюдения за ними.

Благодарности. Авторы выражают благодарность студенту кафедры математической лингвистики СПбГУ А. В. Белому за помощь в программной реализации алгоритма.

Литература

Апресян Ю. Д. Избранные труды. Том 1. Лексическая семантика (синонимические средства языка). М.: Языки Славянской Культуры, 1995. 481 с.

Бонч-Осмоловская А. А., Орехов Б. В. Некоторые применения корпусных методов к наивной поэзии // Статьи на случай: Сборник в честь 50-летия Р. Г. Лейбова. Объединенное гуманитарное издательство Ruthenia.ru, 2013. С. 22–36.

Букия Г. Т., Протопопова Е. В., Митрофанова О. А. Корпусная оценка степени близости единиц в лексических конструкциях // Структурная и прикладная лингвистика, вып. 11. СПб, 2015. С. 252–270.

Гак В. Г. Повторная номинация и ее стилистическое использование // Вопросы французской филологии: Сб. трудов / отв. ред. В. Г. Гак; Моск. гос. пед. ин-т им. В. И. Ленина. М., 1972.

Гак В. Г. Типология лингвистических номинаций // Языковые преобразования / В. Г. Гак. М., 1998.

Гаспаров М. Л. Метр и смысл. Об одном из механизмов культурной памяти. М.: РГГУ, 2000. 289 с.

Гришина Е. А. Да в русском устном диалоге // *Russian Linguistics*. Т. 35. Springer, 2011. С. 169–207.

Гришина Е. А. О маркерах разговорной речи (Предварительное исследование подкорпуса кино в Национальном корпусе русского языка, www.ruscorpora.ru) // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог–2007». М.: Изд-во РГГУ, 2007. С. 147–157.

Иванов Вяч. Вс. Чет и нечет: Асимметрия мозга и знаковых систем. М.: Сов. радио, 1978. 184 с.

Иомдин Л. Л. Русские микросинтаксические элементы, мотивированные словом «вид»: корпусное исследование семантики // Труды международной конференции «Корпусная лингвистика — 2019». СПб, 2019. С. 189–201.

Иомдин Л. Л. Многозначные синтаксические фраземы: между лексикой и синтаксисом // Компьютерная лингвистика и интеллектуальные технологии. Труды Международной конференции «Диалог–2006». М.: Изд-во РГГУ, 2006. С. 202–206.

Иорданская Л. Н., Мельчук И. М. Смысл и сочетаемость в словаре. М.: Языки славянских культур, 2007. 665 с.

Ковтунова И. И. Поэтический синтаксис [Текст] / И. И. Ковтунова. М., 1986. 205 с.

Кожеевникова Н. А. Избранные работы по языку художественной литературы / Сост. Е. В. Красильникова, Е. Ю. Кукушкина, З. Ю. Петрова / Под общ. ред. З. Ю. Петровой. М.: Знак. 2009. 896 с.

Коняшкин А. М., Байкова Ю. С. Стилистические особенности афоризмов Е. Евтушенко. Мир науки, культуры, образования. Т. 2. Вып. 51. 2015. С. 336–338.

Копнина Г. А. Геминация. Эффективное речевое общение (базовые компетенции): словарь-справочник / под ред. А. П. Сковородникова. 2-е изд., Красноярск: Сиб. федер. ун-т, 2014. 681 с.

Лотман М. Ю. Именной и глагольный стили // Учебный материал по анализу поэтических текстов. Таллин, 1982. С. 146–165.

Маркасова Е. В. Представления о фигурах речи в русских риториках XVII–XVIII вв. Петрозаводск, Изд-во ПетрГУ, 2002. 179 с.

Матвеева Т. В. Полный словарь лингвистических терминов. Ростов-на-Дону, 2010. С. 68–69.

Москвин В. П. Выразительные средства современной русской речи: тропы и фигуры. Общая и частные классификации. Терминологический словарь. М.: URSS, Ленанд, 2006. 373 с.

Орехов Б. В. Рецепция классических поэтических формул у наивных авторов в акцентологическом корпусе НКРЯ // Труды Института русского языка им. В. В. Виноградова. 2022. № 2. С. 124–132.

Очерки истории языка русской поэзии XX века. Поэтический язык и идиостиль: Общие вопросы. Звуковая организация текста // М.: Наука, 1990. 304 с.

Рахилина Е. В. (отв. ред.) Лингвистика конструкций. М.: Азбуковник, 2010. 584 с.

Рахилина Е. В., Плунгян В. А. Парадоксы валентностей // Семиотика и информатика. 1998. Т. 36. С. 108–119.

Свиридов С. В. Формула тройного повтора в трехстопном амфибрахии. Альманах современной науки и образования. 2012. № 12 (67) ч. 1. С. 111–115.

Хазагеров Г. Г. Риторический словарь. М., 2009; ОРСП. Ростов н/Д., 1999. 432 с.

Хазагеров Т. Г., Ширина Л. С. Общая риторика / отв. ред. Е. Н. Ширяев. Ростов н/Д : Изд-во Рост. ун-та, 1994. 190 с.

Шерстинова Т. Ю. Да-да-да-да-да! и другие конструкции повседневной русской речи, состоящие из контактно повторяющихся слов, в контексте исследования просодических моделей языка // Труды Института русского языка им. В. В. Виноградова. Вып. 17. Фонетика. М., 2018. С. 162–175.

Шмелев Д. Н. Проблемы семантического анализа лексики (на материале русского языка). М.: Наука, 1973. 280 с.

Fillmore C. J. The Mechanisms of “Construction Grammar” // Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society. Berkeley, MA, 1988. P. 35–55.

Fillmore C. J., Kay P., O'Connor M. C. Regularity and Idiomaticity in Grammatical Constructions: The Case of ‘Let Alone’ // Language. Vol. 64, 1988. P. 501–538.

Goldberg A. Constructions at Work: The Nature of Generalization in Language. Oxford University Press, 2006. 280 p.

Goldberg A. Constructions: A Construction Grammar Approach to Argument Structure. University of Chicago Press, 1995. 271 p.

Stefanowitsch A., Gries S. Th. Collostructions: Investigating the interaction of words and constructions // International Journal of Corpus Linguistics. Issue 8 (2). 2003. P. 209–243.

Tomasello M. Constructing a Language: A Usage-Based Theory of Language Acquisition. Harvard University Press, 2003. 388 p.

E. V. Markasova

Peking University

(China, Beijing)

markasovaelena@yandex.ru

O. A. Mitrofanova

St. Petersburg State University

(Russia, St. Petersburg)

o.mitrofanova@spbu.ru

GEMINATION (TRIPLE REPEAT) ACCORDING TO THE CORPUS DATA

From the perspective of Construction Grammar, gemination is a specific lexical and grammatical structure defined at the lexical level by a single, fixed lexical unit that fills the first and subsequent slots, and organized at the syntactic level as a series composed of union/non-union elements. In the statistical and combinatorial aspects, geminations are considered trigrams, which can be analyzed using corpus methods to assess the degree

of stability of these structures and determine the degree of attraction of lexical units to the slots of this design. In rhetoric, gemination is a type of repetition that differs from reduplication in terms of the number of elements and functions. This figure can be used to describe a large number of similar objects, the duration or monotonous nature of processes, or to reflect the immersion of a lyrical character in a particular state or the focus on a particular subject. The current level of corpus linguistics development allows us to consider the poetics of gemination as a figure against the background of three-fold repetitions in language as a whole. This study presents the analysis of the use of gemination in multi-genre Russian text corpora, focusing on lexical and grammatical classes of words that are prone to constructions involving geminations. Observations made are supported by quantitative data on the distribution of gemination occurrences in research texts.

Keywords: Repeated nomination, contact repetition, triplication, gemination, reduplication, intertextuality, text corpus, lexical and grammatical constructs.

References

Apresyan Yu. D. *Izbrannye trudy. Tom 1. Leksicheskaya semantika (sinonimicheskie sredstva yazyka)* [Selected works. Volume 1. Lexical semantics (synonymous means of language)]. Moscow, Yazyki Slavyanskoi Kul'tury Publ., 1995. 481 p. (In Russ.)

Bonch-Osmolovskaya A. A., Orekhov B. V. [Some applications of corpus methods to naive poetry]. *Stat'i na sluchai: Sbornik v chest' 50-letiya R. G. Leibova* [Articles for the occasion: Collection in honor of the 50th anniversary of R. G. Leibov]. United Humanitarian Publishing House Ruthenia.ru, 2013. pp. 22–36. (In Russ.)

Bukiya G. T., Protopopova E. V., Mitrofanova O. A. [A corpus-driven estimation of association strength in lexical constructions]. *Strukturnaya i prikladnaya lingvistika, vyp. 11* [Structural and Applied Linguistics, vol. 11]. St. Petersburg, 2015. pp. 252–270. (In Russ.)

Fillmore C. J. The Mechanisms of “Construction Grammar”. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, MA, 1988, pp. 35–55.

Fillmore C. J., Kay P., O'Connor M. C. Regularity and Idiomaticity in Grammatical Constructions: The Case of ‘Let Alone’. *Language*. Vol. 64, 1988, pp. 501–538.

Gak V. G. [Repeated nomination and its stylistic use]. *Voprosy frantsuzskoi filologii: Sb. st.* [Questions of French philology: Collection of articles] / Moscow, Moscow State V. I. Lenin Pedagogical Institute Publ., 1972. (In Russ.)

Gak V. G. [Typology of linguistic nominations]. *Yazykovye preobrazovaniya* [Language Transformations]. Moscow, 1998. (In Russ.)

Gasparov M. L. *Metr i smysl. Ob odnom iz mekhanizmov kul'turnoi pamyati* [Meter and meaning. About one of the mechanisms of cultural memory]. Moscow, Russian State University for the Humanities Publ., 2000. 289 p. (In Russ.)

Goldberg A. *Constructions at Work: The Nature of Generalization in Language*. Oxford University Press, 2006. 280 p.

Goldberg A. *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press, 1995. 271 p.

Grishina E. A. [On markers of spoken speech (Preliminary study of the film subcorpus in the Russian National Corpus, www.ruscorpora.ru)]. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: Trudy mezhdunarodnoi konferentsii "Dialog-2007"* [Computer linguistics and intellectual technologies: Proceedings of the international conference "Dialogue-2007"]. Moscow, RSUH Publ., 2007, pp. 147–157. (In Russ.)

Grishina E. A. [Vocal Gesture YES in Spoken Russian]. *Russian Linguistics*. T. 35. Springer, 2011, pp. 169–207. (In Russ.)

Ivanov Vyach. Vs. *Chet i nechet: Asimmetriya mozga i znakovykh sistem* [Odd and even: Asymmetry of the brain and sign systems]. Moscow, Sovetskoe Radio Publ., 1978. 184 p. (In Russ.)

Iomdin L. L. [Polysemous syntactic phrasemes: between vocabulary and syntax]. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii. Trudy Mezhdunarodnoi konferentsii "Dialog-2006"* [Computer linguistics and intellectual technologies. Proceedings of the International Conference "Dialogue-2006"]. Moscow, RSUH Publ., 2006, pp. 202–206. (In Russ.)

Iomdin L. L. [Russian microsyntactic elements motivated by the word "view": a corpus study of semantics]. *Trudy mezhdunarodnoi konferentsii "Korpusnaya lingvistika-2019"*. [Proceedings of the international conference "Corpus Linguistics-2019"]. St. Petersburg, 2019, pp. 189–201. (In Russ.)

Iordanskaya L. N., Mel'chuk I. M. *Smysl i sochetaemost' v slovare* [Meaning and compatibility in the dictionary]. Moscow, Yazyki Slavyanskikh Kul'tur Publ., 2007. 665 p. (In Russ.)

Khazagerov G. G. *Ritoricheskii slovar'* [Rhetorical Dictionary]. Moscow, 2009; Rostov-on-Don, ORSRP Publ., 1999. 432 p. (In Russ.)

Khazagerov T. G., Shirina L. S. *Obshchaya ritorika* [General rhetoric] / ed. by E. N. Shiryayev; Rostov-on-Don, Rostov State University Publ., 1994. 190 p. (In Russ.)

Konyashkin A. M., Baikova Yu. S. [Stylistic features of Ye. Yevtushenko aphorisms]. *Mir nauki, kul'tury, obrazovaniya* [The world of science, culture, and education]. Vol. 2 (51), 2015, pp. 336–338. (In Russ.)

Kopnina G. A. *Geminatsiya. Effektivnoe rechevoe obshchenie (bazovye kompetentsii): slovar'-spravochnik* [Gemination. Effective verbal communication (basic competencies): dictionary-reference book] / ed. by A. P. Skovorodnikov. 2nd edition. Krasnoyarsk, Siberian Federal University Publ., 2014. 681 pp.

Kovtunova I. I. *Poeticheskii sintaksis* [Poetic syntax]. Moscow, 1986. 205 pp.

Kozhevnikova N. A. *Izbrannye raboty po yazyku khudozhestvennoi literatury* [Selected works on the language of fiction] / comp. by E. V. Krasil'nikova, E. Yu. Kukushkina, Z. Yu. Petrova / ed. by Z. Yu. Petrova. Moscow: Znak Publ.. 2009. 896 pp. (In Russ.)

Lotman M. Yu. [Nominal and verbal styles]. *Uchebnyi material po analizu poeticheskikh tekstov* [Training material on the analysis of poetic texts]. Tallin, 1982, pp. 146–165. (In Russ.)

Markasova E. V. *Predstavleniya o figurakh rechi v russkikh ritorikakh XVII–XVIII vv.* [Ideas about figures of speech in Russian rhetoric of the 17th–18th centuries]. Petrozavodsk, Petrozavodsk State University Publ., 2002. 179 p. (In Russ.)

Matveeva T. V. *Polnyi slovar' lingvisticheskikh terminov* [Complete dictionary of linguistic terms]. Rostov-on-Don, 2010, pp. 68–69. (In Russ.)

Moskvin V. P. *Vyrazitel'nye sredstva sovremennoi russkoi rechi: tropy i figury. Obshchaya i chastnye klassifikatsii. Terminologicheskii slovar'* [Expressive means of modern Russian speech: tropes and figures. General and private classifications. Terminological dictionary]. Moscow, URSS, Lenand, 2006. 373 p. (In Russ.)

Ocherki istorii yazyka russkoi poezii XX veka. Poeticheskii yazyk i idiosstil': Obshchie voprosy. Zvukovaya organizatsiya teksta [Essays on the history of the language of Russian poetry of the twentieth century. Poetic language and idiosstyle: General issues. Sound organization of the text]. Moscow, Nauka Publ., 1990. 304 pp. (In Russ.)

Orekhov B. V. [Reception of classical poetic formulas by naive authors in the accentological corpus of the National Corpus of the Russian Language]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova* [Proceedings of the V. V. Vinogradov Russian Language Institute]. 2022, no. 2, pp. 124–132. (In Russ.)

Rakhilina E. V. *Lingvistika konstruktivnii* [Linguistics of constructions]. Moscow, Azbukovnik Publ., 2010. 584 p. (In Russ.)

Rakhilina E. V., Plungyan V. A. [Paradoxes of valences]. *Semiotika i informatika. T. 36* [Semiotics and computer science. Vol. 36]. 1998, pp. 108–119. (In Russ.)

Sherstinova T. Yu. [“Yes-yes-yes-yes-yes!” and other constructions of everyday Russian speech, consisting of contactally repeated words, in the context of the study of prosodic models of language]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova. Fonetika* [Proceedings of the V. V. Vinogradov Russian Language Institute. Phonetics]. 2018, no. 17, pp. 162–175. (In Russ.)

Shmelev D. N. *Problemy semanticheskogo analiza leksiki (na materiale russkogo yazyka)* [Problems of semantic analysis of vocabulary (based on the material of the Russian language)]. Moscow, Nauka Publ., 1973. 280 p.

Stefanowitsch A., Gries S. Th. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics*. Issue 8 (2), 2003, pp. 209–243.

Sviridov S. V. [Formula for triple repetition in amphibrachic trimeter]. *Al'manakh sovremennoi nauki i obrazovaniya*. 2012, no. 12 (67), p. 1, pp. 111–115. (In Russ.)

Tomasello M. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press, 2003. 388 p.

Д. О. Добровольский

Институт русского языка им. В. В. Виноградова РАН

Институт языкознания РАН

(Россия, Москва)

dobrovolskij@gmail.com

Анна А. Зализняк

Институт языкознания РАН

(Россия, Москва)

anna.a.zalizniak@gmail.com

ИСПОЛЬЗОВАНИЕ ПАРАЛЛЕЛЬНОГО КОРПУСА В СЕМАНТИЧЕСКОМ АНАЛИЗЕ РУССКОЙ ЛЕКСИКИ: CASE STUDY *ПРИСПИЧИЛО* И *НОРОВИТЬ*

В статье излагаются результаты применения «монофокусного» метода контрастивного анализа на материале немецкого параллельного подкорпуса НКРЯ и построенной на его основе надкорпусной базы данных конструкций с немецкими модальными глаголами и их русских соответствий. Предлагается описание двух русских слов сложной семантики — маргинальных глаголов желания/намерения *приспичило* и *норовить*. Показано, что анализ «стимулов» появления этих слов в переводах с немецкого на русский и «моделей» их переводов с русского на немецкий позволил выявить, помимо модальной составляющей, также другие неочевидные компоненты их значения, детерминирующие особенности их семантики и употребления.

Ключевые слова: параллельный корпус, модальные глаголы, перевод, семантика, русский язык, немецкий язык, оценка, говорящий

Анализ переводных соответствий как способ изучения семантики языковых единиц широко применяется в современной лингвистике (ср. в частности [Johansson, Oksefjell 1998; Noël 2003; Сичинава 2014; Bonola 2017]). В данном исследовании применялся «монофокусный» метод контрастивного анализа (в [Зализняк 2015] он был назван «унидирекциональным», в [Шмелев 2020] — «моноцентричным»), суть которого состоит в том, что сопоставление оригинального текста с его переводом используется как инструмент анализа интересующей нас единицы одного из сопоставляемых языков: перевод на другой язык рассматривается как своего

рода толкование для анализируемой единицы языка оригинала, — и наоборот, исследование «стимулов» появления той или иной языковой единицы в переводе может выявить неочевидные компоненты значения этой единицы.

В исследованиях последних лет используется дополнительный эффективный инструмент для работы с материалом параллельных корпусов — надстроенные над параллельным корпусом («надкорпусные») базы данных; см. в частности [Кружков 2015; Зализняк 2016; Инькова 2017]. В наших работах [Добровольский, Зализняк 2022а, б] мы исследуем семантику единиц как немецкого, так и русского языков, опираясь на материал созданной нами надкорпусной базы данных конструкций с немецкими модальными глаголами и их русских соответствий (далее: База данных); о структуре Базы данных см. [Гончаров и др. 2023].

В данном исследовании нас будут интересовать единицы русского языка, которые зафиксированы в Базе данных в качестве вариантов перевода немецких модальных глаголов (*dürfen, können, mögen, müssen, sollen, wollen*) и тем самым включают в свое значение модальный компонент. Имеющийся на данный момент материал Базы данных позволяет сделать вывод, что модальным потенциалом обладают следующие единицы русского языка, чья способность выражать модальные значения не очевидна¹:

- глаголы с личным субъектом: *вздумать, годиться* (кому-л. в кого-л.), *не выносить* (чего-л.), *не обойтись* (без чего-л.), *не терпеть, не удержаться, норовить, позволить себе, постесняться, (не) решиться, собраться, справиться, успеть, ухитриться*;
- безличные глаголы и предикативы: в конструкции <(Dat) + Inf>: *вздумалось, довелось/доводилось, осталось/остаётся, случилось/случается, пришлось/приходится, удалось, приспичило; хорошо бы, лучше (бы), не мешает, стоит; дано, суждено; пора, впору, самое время*; в конструкции <Acc + Inf>: *тянет/потянуло, угораздило*²;
- существительные: *черед, очередь*;
- наречия: *иногда, обычно, невольно, напрасно, по-хорошему, трудно, легко*;
- предлоги: *вместо, для*;
- союз: *чтобы*;
- единицы различных грамматических классов и конструкции: *с трудом, вот-вот, едва не, чуть (ли) не, не иначе как, (никак) не PraesPf* и др.

Для некоторых из этих единиц значение, включающее модальный компонент, отмечается словарями (ср. один из «оттенков» значения б в [МАС: IV, 172] для

¹ Здесь мы не рассматриваем слова и выражения со значением эпистемической модальности.

² В первых двух пунктах этого списка обращает на себя внимание тот факт, что сюда входят почти все глаголы, выражающие специфические не полностью контролируемые установки относительно собственного действия, которые были описаны в [Зализняк, Левонтина 1996] (а именно: *собираться, успеть, удалось, довелось, угораздило*). Тот факт, что эти глаголы встречаются в качестве переводных эквивалентов для немецких модальных глаголов, объясняется, по-видимому, тем, что определенные типы модальных значений имплицируют неполный контроль над обозначенным инфинитивом действием.

собраться «принять намерение отправиться куда-либо»; глагол *собираться* помечен при этом просто как «несов. к *собраться*»)³. Для других такое значение является вовсе неожиданным — как, например, для слов *очередь* и *черед*; оно, однако, очевидно имеет место: чья-либо *очередь* нечто делать означает, что этот человек *может* или *должен* это сделать, ср. примеры появления этих слов в переводах немецких предложений с модальным глаголом⁴:

- (1) “Würdest du <...> abermals auf mich schießen?” <...> “Nein, nächstes Mal **darfst** du mich **abknallen**.”

«<...> ты стал бы стрелять в меня снова?» <...> «Ладно уж, на этот раз твоя **очередь мочить**».

[Günter Grass. Im Krebsgang (2002) | Гюнтер Грасс. Траектория краба (Б. Н. Хлебников, 2002)]

- (2) **Wollen** Sie mir nun auch *beichten*?

Ну что же, теперь ваш **черед исповедоваться**.

[Hermann Hesse. Peter Camenzind (1904) | Герман Гессе. Петер Каменцинд (Р. Эйвадис, 1995)]

Контексты, в которых модальный компонент обнаруживается у единиц *дано*, *суждено*, *трудно*, *вместо*, *пора*, *впору*, *вот-вот*, были описаны в [Добровольский, Зализняк 2020], у единиц *хорошо бы*, *лучше (бы)*, *не мешает*, *стоит*, *напрасно* — в [Добровольский, Зализняк 2022б]. В данной статье мы рассмотрим две единицы из приведенного выше списка: безличный глагол *приспичило* и финитный *норовить*. Оба эти слова лингвоспецифичны, т. е. не имеют прямых аналогов в других языках (в частности, немецком). Анализ таких единиц путем сопоставления их со «стимулами» и «моделями» перевода позволяет выявить те самые дополнительные семантические компоненты, которые утрачиваются — или, наоборот, эксплицируются — в переводе на другой язык и, соответственно, появляются в качестве «лишних» в переводном тексте. Как уже было сказано, глаголы *приспичило* и *норовить* были обнаружены в качестве возможных переводных эквивалентов для немецких модальных глаголов (прежде всего, *wollen* и *müssen*) — что позволило выявить в их семантике модальную составляющую. Дальнейшая процедура состояла в том, что были проанализированы прочие (не модальные) «стимулы» появления в русском переводе этих слов и, наоборот, возможные способы передачи их значения в переводах с русского (такая процедура в [Шмелев 2020] была названа «челночным» методом).

³ О специфическом типе намерения, выражаемом в русском языке глаголом *собираться*, см. [Зализняк, Левонтина 1996].

⁴ Здесь и далее в примерах использована следующая система нотации: анализируемый глагол русского языка и его немецкий эквивалент (в обоих направлениях перевода) выделяется **прямым п/ж** шрифтом, обозначение ситуации-объекта — *светлым курсивом*, прочие элементы контекста, влияющие на выбор эквивалента — *п/ж курсивом*.

приспичило

Глагол *приспичило* включает компоненты сильного желания и «внешней» отрицательной оценки. Эта оценка принадлежит говорящему, которому возникшее у кого-то острое желание немедленно что-то сделать представляется неуместным, глупым, абсурдным⁵. Эта оценка может, в том числе, отражать негативную реакцию говорящего на доставленное ему неудобство; так, можно сказать *Моим соседям приспичило делать ремонт*, чтобы выразить свое неудовольствие происходящим из-за этого шумом.

Слово *приспичило* обозначает настолько сильное желание, что оно уже становится необходимостью — что подтверждается материалом параллельного корпуса: наиболее частотные «стимулы» появления в русском переводе слова *приспичило* — глаголы *wollen* (+ единичное *möchte*) ‘хотеть’ и *müssen* ‘должен’.

На месте *wollen* слово *приспичило* появляется в тех случаях, когда в предложении имеются какие-либо дополнительные указания:

- на то, что это желание очень сильное, переходящее в необходимость, ср. *unbedingt* ‘непреренно’ в примере (3):

(3) Ich dachte, einer von euch hätte eine Idee, wo ihr doch ***unbedingt mitkommen wolltet***.

— Я думал, у кого-нибудь из вас есть план, раз уж вам так **приспичило идти за мной**.

[Cornelia Funke. Tintenblut (2005) | Корнелия Функе. Чернильная кровь (М. М. Сокольская, 2007)]

- на непонятность причины возникновения желания:

(4) Außerdem **wolltest** Du, wie die Wirtin vom Herrenhofwirt erfahren hat, ***aus irgendwelchen Gründen*** damals im Herrenhof *übernachten* und das war allerdings überhaupt nicht anders als durch mich zu erlangen.

Кроме того, от хозяина «Господского подворья» трактирщица узнала, что тебе в тот вечер **неведомо почему приспичило** там *переночевать*, а добиться этого ты, кроме как с моей помощью, не мог.

[Franz Kafka. Das Schloss (1922) | Франц Кафка. Замок (М. Л. Рудницкий, 2006)]

- на неуместность желания:

(5) „Hast schon so was g’hört, Ferencz?“ stupft ihn der Jozsi an, „***jetzt um halber eins, wo die Bude schließt, möchte*** er noch *einen Tarock anfangen*!“

— Как тебе это нравится, Ференц? — подталкивает его Йожи. — Теперь, **в половине первого, когда лавочка закрывается**, ему **приспичило играть в тарок!**

[Stefan Zweig. Ungeduld des Herzens (1939) | Стефан Цвейг. Нетерпение сердца (Н. Н. Бунин, 1961)]

⁵ Отрицательная оценка отсутствует лишь в особом значении, в котором этот глагол не имеет дополнения: ‘срочно нужно в туалет’; его мы рассматривать не будем.

Предложение в оригинале может содержать также явно выраженное основание для отрицательной оценки. Так, в (6) *alles gleich auf einmal* ‘все и сразу’ указывает на нерациональность желания: в норме «все и сразу» не бывает; тот, кто этого хочет, достоин осуждения.

- (6) Warum **willst** du *alles gleich auf einmal*? Ein geachteter Automobilhändler werden und heiraten?

— С чего вдруг тебе **приспичило все и сразу**? И степенным коммерсантом стать, и жениться?

[Erich Maria Remarque. Der Himmel kennt keine Günstlinge (1961) | Эрих Мария Ремарк. Жизнь взаимы, или У неба любимчиков нет (М. Л. Рудницкий, 2018)]

На месте *müssen* ‘должен’ слово *приспичило* обычно также появляется в тех случаях, когда оно сопровождается словом *unbedingt* ‘непрерменно’ (ср. (7), (8)) или когда речь идет о заведомо отрицательно оцениваемой ситуации, ср. (8), (9):

- (7) Und Sie haben dafür gesorgt, dass es noch mehr geworden sind! Aber Sie **muss-ten** sie ja *unbedingt* zu diesem haarsträubenden Abenteuer überreden! Ein nächtlicher Ausflug zur Isola Segreta.

— А вы позаботились о том, чтобы у них было их еще больше! Но вам же **приспичило** подбить их на такое захватывающее, такое жуткое приключение! Ночной набег на остров Тайн!

[Cornelia Funke. Herr der Diebe (2002) | Корнелия Функе. Король воров (М. Л. Рудницкий, 2004)]

- (8) Wie auf meiner Hochzeitsfeier kam mir Blaubarts letzte Frau in den Sinn, die *unbedingt* das verbotene Zimmer inspizieren **mußte**, obgleich sie, wie ich, die Katastrophe witterte.

Как и на свадьбе, мне вспомнилась последняя жена Синей Бороды, которой тоже **приспичило** проникнуть в запретную комнату, хотя она, как и я сейчас, ничего хорошего там увидеть не ждала.

[Ingrid Noll. Die Apothekerin (1994) | Ингрид Нолль. Аптекарьша (Г. А. Шевченко, 2001)]

- (9) „So, so“, sagte der Alte, „ihr **mußtet** also Bomben schmeißen! Ihr **mußtet** euch verschwören <...>.“

— Так, так, — сказал старик, — значит, вам **приспичило** бросать бомбы! **Приспичило** устраивать заговоры <...>.

[Heinrich Böll. Billard um halbzehn (1959) | Генрих Бёлль. Бильярд в половине десятого (Л. Б. Черная, 1961)]

Помимо модальных глаголов *wollen* и *müssen*, «стимулом» появления слова *приспичило* в переводах являются выражения, соответствующие различным отдельным компонентам значения русского *приспичило*.

Прежде всего, это выражения, указывающие на то, что что-то надо сделать немедленно, что это не терпит отлагательств (ср. ‘прямо сейчас’ в (10), ‘спешно’ в (11), (12)):

- (10) So? **Gerade jetzt?** Und das soll ich glauben?

— Вон что? **Приспичило?** Так я и поверил!

[Erich Maria Remarque. *Zeit zu leben und Zeit zu sterben* (1954) | Эрих Мария Ремарк. *Время жить и время умирать* (Н. Каринцева и В. Станевич, 1956)]

- (11) Stefanie Clary schnippt mit dem Finger. Sie weiß auch etwas und **hat es eilig damit**.

Стефани тянет вверх руку и щелкает пальчиками. Она, видите ли, тоже что-то знает про домашних кур, и ей **приспичило именно сейчас рассказать об этом**.

[Erwin Strittmatter. *Tinko* (1954) | Эрвин Штриттматтер. *Тинко* (Вс. Розанов, 1956)]

- (12) „Ihr **habt es aber eilig**“, sagte K., der diesmal sehr zufrieden mit ihnen war, „wir sind noch hier und Ihr müßt schon einrücken?“

— Эк вам **приспичило**, однако, — сказал К., хотя на сей раз был служанкам почти рад, — мы еще не съехали, а вы уже пожаловали.

[Franz Kafka. *Das Schloss* (1922) | Франц Кафка. *Замок* (М. Л. Рудницкий, 2006)]

Это может быть указание на очень сильное желание. Так, в (13) выражение *ihm war daran gelegen* значит ‘ему было важно (сделать что-то)’. Перевод этого выражения с помощью *приспичило* добавляет «лишний», по сравнению с оригиналом, компонент отрицательной оценки говорящим этого желания.

- (13) Ihm, der sich sonst so steif und zurückhaltend gab, und der mich einmal gedemütigt hatte, **war daran gelegen, mit mir zu sprechen**.

Ему, всегда такому чопорному и сдержанному, однажды уже унизившему меня, **приспичило поговорить со мной**.

[Siegfried Lenz. *Schweigeminute* (2008) | Зигфрид Ленц. *Минута молчания* (Г. М. Косарик, 2011)]

В (14) появление в переводе слова *приспичило* обусловлено выражением *auf Teufel kommt raus* «во что бы то ни стало», означающим настоятельную необходимость и включающим оттенок отрицательной оценки.

- (14) <...> nicht weil er eitel war und **auf Teufel komm raus ein adeliges Bettgemahl besitzen musste**, sondern weil er eine Dynastie gründen und seine Nachkommenschaft auf ein Geleise setzen wollte, welches zu höchstem gesellschaftlichem Ansehen und politischem Einfluss führte.

<...> не потому что он был тщеславен и ему **приспичило иметь в постели супругу-аристократку**, но потому, что он желал основать династию

и наставить свое потомство на путь, ведущий к высокому общественному положению и политическому влиянию.

[Patrick Süskind. *Das Parfum: Die Geschichte eines Mörders* (1985) | Патрик Зюскинд. *Парфюмер: История одного убийцы* (Э. Венгерова, 1992)]

В (15) буквально сказано «похоже, что тебе это очень важно». Выбрав слово *приспичило*, переводчик привносит элемент отрицательной оценки со стороны говорящего, который, возможно, не признает важность обсуждаемого обстоятельства: *приспичило* — это как бы дурь какая-то, разумный человек так бы себя не вел.

(15) „Es **scheint** dir *sehr* wichtig zu sein, zu deiner verflorenen Geliebten zurück-zukommen“, schnurrte er.

— Похоже, тебе **сильно приспичило** вернуться к твоей бывшей подруге, — промурлыкал он.

[Cornelia Funke. *Reckless. Steinernes Fleisch* (2003) | Корнелия Функе. «Бесшабашный» (М. Л. Рудницкий, 2010)]

«Стимулом» для появления слова *приспичило* может быть эксплицитная отрицательная оценка говорящим обсуждаемой ситуации. Так, в (16) слово *lästig* ‘обременительный’ указывает на то, что смерть тетушки доставила говорящему неудобства.

(16) Sie hatten drüben Manuela fast vergessen. Sie schickten regelmäßig den Scheck und lebten in Bogotá, und Manuela lebte hier — seit fünf Jahren — und schrieb monatlich einen Brief. <...> Alles war gut — bis sie **so lästig wurde** zu sterben. Da mußte man natürlich kommen, der Reputation wegen.

Там, у себя, они про девчонку считай что забыли. Ну, посылали регулярно чек и жили себе в своей Боготе, а Мануэла была здесь, раз в месяц писала им письмо, и так уже пять лет. <...> Все было замечательно — покуда ей *умирать* не **приспичило**. Тут, конечно, им пришлось срочно сюда отправляться, хотя бы приличия ради.

[Erich Maria Remarque. *Der Himmel kennt keine Günstlinge* (1961) | Эрих Мария Ремарк. *Жизнь взаимы, или У неба любимчиков нет* (М. Л. Рудницкий, 2018)]

В (17) неодобрение еще резче: *eklig hinterher* означает что-то вроде ‘отвратительно (настаивает) на этом’.

(17) Präsident von Wulckow **eklig hinterher**, daß Sie verkaufen, sonst kein Geschäft mit ihm zu machen.

— Президенту фон Вулкову **приспичило...** настаивает на немедленной продаже, иначе он знать вас не знает.

[Heinrich Mann. *Der Untertan* (1914) | Генрих Манн. *Верноподданный* (И. Горкина, 1971)]

В переводах с русского языка на немецкий модальные глаголы вообще не появляются. Очевидно, переводчик стремится найти какой-то другой способ более

точной передачи сложной семантики этого слова. Так, в (18) употреблено безличное *es drängte*, которое является достаточно точным эквивалентом для *приспичило*: глагол *drängen* ‘напирать, теснить’ содержит идею, что что-то на тебя как бы давит и поэтому ты должен прямо сейчас это сделать.

(18) — Типичный кулачок по своей психологии, — заговорил Иван Николаевич, которому, очевидно, **приспичило** обличать Рюхина, — и притом кулачок, тщательно маскирующийся под пролетария.

„Psychologisch gesehen, ist er ein typischer kleiner Kulak“, sagte Iwan, den **es** sichtlich **drängte**, *Rjuchin zu entlarven*, „aber er tarnt sich sorgfältig als Prolet.“

[Михаил Булгаков. Мастер и Маргарита (1929–1940) | Michail Bulgakow. Der Meister und Margarita (Thomas Reschke, 1968)]

В (19) *приспичило* переведено на немецкий наречием *dringend* ‘срочно’, которое по внутренней форме связано с глаголом *dringen* и, соответственно, с *drängte*.

(19) Пошел я, господа, третьего дня утром к здешнему купчине Самсонову занимать у него три тысячи денег под вернейшее обеспечение, — это **вдруг приспичило**, господа, **вдруг приспичило**...

Ich ging also vorgestern Vormittag zu dem hier ansässigen Kaufmann Samsnow, meine Herren, um mir von ihm gegen vollständige Sicherheit dreitausend Rubel zu borgen. Das **war plötzlich dringend geworden**, meine Herren, das **war dringend**... «

[Ф. М. Достоевский. Братья Карамазовы (ч. 3–4) (1878) | Fjodor Dostojewski. Die Brüder Karamasow (3.–4. Teile) (Hermann Röhl, 1923)]

Приспичило может быть также передано чисто оценочным выражением; так, в (20) идиома *welcher Teufel ist in dich gefahren* буквально означает ‘какой черт в тебя вселился’.

(20) Через несколько секунд он робко поскребся в дверь. Сэр Джуффин Халли удивился. — Кому это **приспичило**?

Ein paar Sekunden später klopfte er schüchtern an die Tür. Sir Juffin sah mich staunend an. „**Welcher Teufel ist in dich gefahren**, Max?“

[Макс Фрай. Чужак (2004) | Max Frei. Der Fremdling (Anna Serafin, 2007)]

норовить

В современном языке глагол *норовить* в своем основном типе употребления означает внутреннюю установку (generic state; это тот же аспектуальный тип значения, что у глагола *нравиться*, образованного от того же *нрав*, *норов*) — а именно, чью-то склонность поступать некоторым осуждаемым (социально или лично говорящим) способом, т. е. делать нечто предосудительное или наоборот, не делать то, что принято или должно делать. Ср. типичные контексты (в 3-м лице): *Он норовит улизнуть / пролезть без очереди / лучший кусок себе захватить / нас подслушать, чтобы потом доносы писать* и т. п. Во 2-м лице этот глагол используется

для разоблачения или интерпретации чужого намерения, ср. известные строчки Высоцкого: *А ты придёшь домой, Иван, / Поешь — и сразу на диван, / Иль, вон, кричишь, когда не пьян... / Ты что, Иван? / — Ты, Зин, на грубость нарываешься, / Всё, Зин, обидеть **норовишь!***

Однако отрицательная оценка появилась у этого глагола не раньше середины XX в. До этого *норовить* обозначало просто ‘хотеть, стараться, иметь склонность к чему-л.’

(21) А сейчас какой-нибудь граф, — мелкий, лицо нездоровое, в глазах — сок, слеза, и он сам **норовит** с тобой познакомиться, прибежит к тебе в пансион. [А. Н. Толстой. Черная пятница (1924)]

(22) Только по Москве она, вероятно, скучает и со случайной машиной всё **норовит** съездить туда, сделать маникюр, побывать в театре. [Л. К. Чуковская. Спуск под воду (1949–1957)]

Глагол *норовить* может употребляться также в значении актуального намерения; часто это намерение сопровождается другим действием, указывая на его дополнительную цель. Эта цель обычно (хотя и не обязательно) включает нечто дурное — ср. (23) и (24):

(23) Говорят, героически вела себя красавица Стелла, обороняя своего кавалера, сняла туфлю и лупила нападающих каблуком, **норовя** попасть по очкам. [Юрий Трифонов. Дом на набережной (1976)]

(24) Анче же, наоборот, едва перебирала ногами, **норовя** то коснуться пальцами громко храпевшего мужика в драном бешмете, то присесть у картонной коробки с дырчатыми стенками: внутри что-то явственно ворочалось и мяукало. [Гузель Яхина. Дети мои (2018)]

При этом даже в актуальном значении глагол *норовить* не может обозначать некоторое сознательно планируемое заранее единичное действие. Нельзя, например, сказать *Он норовит прогулять завтрашний семинар / пообедать сегодня за мой счет*. Такого рода цели могут выражаться словом *норовить* только в генерическом смысле, т. е. как установка. С другой стороны, *норовить* может относиться только к какому-то незначительному неблагоприятному поступку, т. е. *норовить* можно только «по мелочи»: где-то сжульничать, недоплатить, подставить подножку, погулять за чужой счет и т. п. Но, например, ограбить банк или убить старушку — это действия, описание которых плохо сочетается с *норовить*, даже если речь идет о чьей-то склонности к такого рода дурным поступкам. Можно сказать *Он так и норовит нажраться на халяву*, но только в шутку: *Он так и норовит банк какой-нибудь ограбить / какую-нибудь старушку убить*.

Анализ переводных соответствий позволяет выявить все основные компоненты значения этого слова.

Прежде всего, обращает на себя внимание тот факт, что в весьма значительной части случаев глагол *норовить* в переводе возникает как бы «ниоткуда»,

т. е. в оригинале стоит просто глагол в личной форме. В этом случае глагол *норовить* вносит дополнительный элемент намерения и повторяемости, ср.:

- (25) Aber früher *hatte er auch ihre Hühner gerissen*.

Раньше он **норовил** *душить* у нее кур.

[Cornelia Funke. Tintentod (2007) | Корнелия Функе. Чернильная смерть (М. М. Сокольская, 2012)]

Рассмотрим «ненулевые стимулы» появления в русском переводе слова *норовить*.

Среди модальных глаголов «стимулом» для появления в переводе *норовить* является глагол *wollen* ‘хотеть’. Дополнительные, по сравнению с просто желанием, семантические компоненты, присутствующие в русском *норовить*, извлекаются переводчиком из контекста, а именно, *норовить* появляется в переводе, когда речь идет о чем-то дурном — социально неодобряемом или просто неприятном говорящему (ср. *облапошить*, *втянуть*, *навязать* в (26), (27), (29)) — и при этом эксплицитно или имплицитно выражена идея, что речь идет о чем-то постоянном, повторяющемся (ср. употребление наречия *immer* в (28), (29)).

- (26) Wahrscheinlich kannte sich der Großvater in den heutigen Preisen nicht aus und lebte in dem Wahn, alle Welt **wolle ihn übers Ohr hauen**.

Видимо, дед и вправду несколько оторвался от жизни и реальных цен на рынке труда, пребывая в безумном убеждении, что весь белый свет только и **норовит** *его облапошить*.

[Ingrid Noll. Die Apothekerin (1994) | Ингрид Нолль. Аптекарьша (Г. А. Шевченко, 2001)]

- (27) Lüge! Die Ratten! Murksen an sich selbst herum und **wollen dann andere hereinreißen**.

— Все ложь! Подлюги проклятые. Сами черт знает что себе делают, а потом **норовят** еще и *других втянуть*.

[Erich Maria Remarque. Arc de Triomphe (1945) | Эрих Мария Ремарк. Триумфальная арка (И. М. Шрайбер, 1978)]

- (28) „Alle **wollten mich immer erziehen**“, sagte sie sehr schläfrig. „Du nicht.“

— Все **норовили** *меня воспитывать*, — пробормотала она уже совсем сквозь сон. — Только ты — нет.

[Erich Maria Remarque. Das gelobte Land (1965–1970) | Эрих Мария Ремарк. Земля обетованная (М. Л. Рудницкий, 2000)]

- (29) Aber vielleicht hätten wir weniger Kriege, wenn jeder den andern nicht **immer so dringend von seiner eigenen Wahrheit überzeugen wollte**.

Хотя, может статься, войн было бы меньше, если бы каждый не **норовил непременно** *навязать другим свою правду*.

[Erich Maria Remarque. Zeit zu leben und Zeit zu sterben (1954) | Эрих Мария Ремарк. Время жить и время умирать (Н. Н. Федорова, 2017)]

Глагол *норовить* появляется в переводе с немецкого также в тех случаях, когда в оригинале есть единица со значением ‘пытаться’ (чаще всего — ср. (30), (31)), ‘нацелиться’ (32), ‘стремиться’ (33), ‘постоянно прилагать усилия’ (34), ‘(делать) с удовольствием’ (35). То, что в немецком языке разлито в контексте, в русском лексикализовано.

- (30) Es schien ihr ähnlich, als wenn Kinder auf der Straße **versuchten**, *vor heranfahrenden Autos noch rasch zur andern Seite hinüberzulaufen*.

Вот так же детишки **норовят** *перебежать дорогу перед самым носом проезжающей автомашины*.

[Erich Maria Remarque. Der Himmel kennt keine Günstlinge (1961) | Эрих Мария Ремарк. Жизнь взаимы, или У неба любимчиков нет (М. Л. Рудницкий, 2018)]

- (31) <...> er leckte ihre Hände und **versuchte**, *ihr Gesicht zu lecken*, er jaulte und heulte und bellte so, daß ein Kreis von erstaunten Menschen sie umstand.

<...> пес лизал ей руки, **норовил** *лизнуть в лицо*, скулил, подвывал и так громко лаял, что вокруг сразу начали собираться зеваки.

[Erich Maria Remarque. Der Himmel kennt keine Günstlinge (1961) | Эрих Мария Ремарк. Жизнь взаимы, или У неба любимчиков нет (М. Л. Рудницкий, 2018)]

- (32) Der rote Kater **hatte es auf das eine Auge des schwarzen abgesehen**.

Красный кот все время **норовит** *попасть когтистой лапой в единственный глаз* черного.

[Otfried Preußler. Krabat (1980) | Отфрид Пройслер. Крабат: легенды старой мельницы (А. Исаева и Э. Иванова, 1985)]

- (33) <...> ergingen sich auch sauber gekleidete Männer und Frauen, mit ihren Feierabendgesichtern **strebten** *sie auf die Rampe, den drohenden Schildern zu*, hinter denen verborgen biedere Baubuden dem Tod zu spotten schienen;

<...> вдоль площадки для игры в гольф прогуливались хорошо одетые мужчины и женщины с праздничными лицами, они **норовили** *подойти вплотную к реке, к страшным щитам*, за которыми прятались прозаические бараки строителей

[Heinrich Böll. Billard um halbzehn (1959) | Генрих Бёлль. Бильярд в половине десятого (Л. Б. Черная, 1961)]

- (34) Meist lief er völlig nackt auf allen vieren in der Zelle herum, bellte hündisch, fraß aus seiner Schüssel wie ein Hund und **legte es immer wieder darauf an**, *Quangel in die Beine zu beißen*.

Обычно он бегал на четвереньках вокруг всей камеры, лаял, лакал из своей миски, как собака, и **постоянно норовил** *укусить Квангеля за ногу*.

[Hans Fallada. Jeder stirbt für sich allein (1947) | Ганс Фаллада. Каждый умирает в одиночку (Н. Касаткина, В. Станевич, И. Татарина, 1971)]

- (35) Sie *hinterlassen* allzu **gern** ihre Fußabdrücke auf meinen Bildern!

Они вечно **норовят** наступить на невысохший рисунок!

[Cornelia Funke. Tintentod (2007) | Корнелия Функе. Чернильная смерть (М. М. Сокольская, 2012)]

В переводах с русского глагол *wollen* вообще не встретился ни разу; видимо отсутствие в нем остальных семантических компонентов русского *норовить* слишком сильно искажает исходный смысл. Используются в основном глаголы и выражения со значением попытки и приложения усилий: то, что человек *норовит* сделать, требует приложения усилий и может не удалиться.

- (36) И опять лезут иноземцы, **норовят** по локоть засунуться в русский карман...

Und nun drängen sich die Fremden uns wieder auf, **versuchen**, die Hand bis an den Ellbogen in die russische Tasche zu versenken...

[А. Н. Толстой. Петр Первый (1929–1943) | Alexei Tolstoi. Peter der Erste (Maximilian Schick, 1944)]

- (37) <...> Пантелей, швейцар, милиционер, официант и поэт Рюхин выносили спеленатого, как куклу, молодого человека, который, заливаясь слезами, плевался, **норовя** попасть именно в Рюхина, давился слезами и кричал: — Сволочь!

<...> Pantelej, der Portier, ein Milizmann, ein Kellner und der Lyriker Rjuchin aus der Tür des Gribojedow einen jungen Mann geschleppt brachten, eingewickelt wie eine Puppe. Dieser war tränenüberströmt, spuckte um sich, wobei er Rjuchin zu treffen **trachtete**, und schrie, daß es über den Boulevard schallte: „Du Miststück!“

[Михаил Булгаков. Мастер и Маргарита (1929–1940) | Michail Bulgakow. Der Meister und Margarita (Thomas Reschke, 1968)]

В значении установки глагол *норовить* в переводе часто просто опускается, подчиненный инфинитив передается формой в узуальном значении (ср. появление конструкции *норовить* + инф. на месте личной глагольной формы в немецком оригинале в примере (25) выше):

- (38) С голоду дохнет, а возьмите его в работники наймите — он вам **норовит** напортить, да еще к мировому судье.

So ein Bauer kriecht beinah vor Hunger; aber nun nehmen Sie den Menschen einmal als Arbeiter an: zuerst **verdirbt** er Ihnen alles mögliche, und dann verklagt er Sie noch beim Friedensrichter.

[Л. Н. Толстой. Анна Каренина (ч. 1–4) (1873–1877) | Lew Tolstoi. Anna Karenina (1.–4. Teile) (Hermann Röhl, 1913)]

* * *

Итак, анализ переводных эквивалентов в обоих направлениях перевода позволил нам выявить следующие семантические компоненты рассматриваемых глаголов: непреодолимость желания и его неуместность для *приспичило*; константность

намерения, приложение усилий к осуществлению определенных действий и их желательность для субъекта — для *норовить*.

В обоих глаголах обнаруживается компонент «внешней» отрицательной оценки: в случае *приспичило* говорящий считает возникновение желания немедленно совершить некоторое действие нерациональным и выражает свое неодобрение человека, который этому желанию поддается; в случае *норовить* говорящий осуждает сам тот тип действий, склонность к которому обнаруживает субъект. Для обоих глаголов затруднено употребление в 1-м л. (для безличного *приспичило* — с дативным субъектом *мне*).

Предложения типа *Мне приспичило купить новую шляпку* возможны лишь в том случае, когда это рассказ о прошлом, т. е. интерпретация своего состояния постфактум. Так, в основном корпусе НКРЯ нашлось всего 12 примеров на *мне приспичило*, и это всегда самоирония, оценка постфактум собственного желания как несколько нерационального, необъяснимого или неуместного, ср. автокомментарий *Вот уж охота пуще неволи* в (39).

(39) В самом первом своем макете мне **приспичило**, чтобы текст обтекал сложную кривую — видел такое в журнале «Vogue». Вот уж охота пуще неволи: пришлось нарезать строчки верстки лапшой и с помощью пинцета выклеивать их по одной — как сейчас помню, их было 82 штуки. [Вячеслав Глазычев. Журнал в моей жизни // «Знание — сила», 2006]

(40) И тут мне **приспичило** Левитанского сборник купить. [Наталья Василькова. Книжная охота (2002) // «Вечерняя Москва», 10.10.2002]

Однако сказать *Мне приспичило купить новую шляпку/сборник Левитанского, давай зайдем сейчас в магазин и купим* можно разве что в шутку.

Ограничение на употребление в 1-м л. глагола *норовить* обусловлено следующим. Субъект этой установки либо скрывает свою склонность к совершению данных действий, либо ее не осознает; именно поэтому формы 2-го и 3-го л. представляют собой своего рода разоблачения (ср. выше). А в форме 1-го л. компонент сокрытия своего намерения субъектом приходит в прямое противоречие с компонентом разоблачения или осознания его говорящим, поскольку субъект и говорящий — это одно и то же лицо.

По корпусным данным, однако, такое употребление не вовсе исключено. Как уже говорилось, являющееся предметом нашего анализа современное значение глагола *норовить* появилось не ранее середины XX в. В основном корпусе НКРЯ форма *норовлю* встречается 42 раза, из них лишь 15 — после 1980 г. В этих примерах это либо самоирония (ср. (41)), либо саморефлексия (ср. (42)), либо реликт устаревшего значения без отрицательной оценки (ср. (43)).

(41) Я вообще-то играю на рояле <...> но кипучий лентяй Виктор Михалыч Полесов во мне не дремлет — все **норовлю** других поучить, как надо ввинчивать лампочки и класть кирпич. [Полина Осетинская. Газонокосильщица // «Русская жизнь», 2012]

- (42) Я стал упорно, почти со спортивной страстью думать: «Зачем бабушка в магазине рассказывает какую-то чушь несусветную людям, с которыми она никогда прежде не виделась и не увидится? Чего это я этой личности всегда улыбаюсь? (Какие у меня цели?) А мимо этой физиономии как-то **норовлю** проскочить?» [М. А. Захаров. Театр без вранья (2007)]
- (43) Я, женщина преклонного возраста, в общественном транспорте **норовлю** встать около такого «лица», оно почти наверняка тут же уступит место. [Несмешная «Комната смеха» // Известия, 2004.02.09]

В книге [Падучева 1985: 141] аномальность таких высказываний, как *Я выпендриваюсь*, *Я повадился*, *Я не почесался* и т. п., объясняется тем, что отрицательная оценка действия входит не в основное содержание лексемы, а в модальную рамку (это утверждение опирается на принцип, сформулированный в [Апресян 1978]). Однако здесь необходимо некоторое уточнение. Так, глагол *угораздило*, содержащий отрицательную оценку в модальной рамке (ср. толкование этого слова в [Апресян 2009: 514]), прекрасно употребляется с субъектом 1-го л., ср. *Меня угораздило простудиться в такую жару*. Действительно, если признание в поступках, обозначение которых содержит отрицательную оценку в своей ассерции, языковой аномалии не порождает, то с оценкой, принадлежащей модальной рамке, говорящий вынужден солидаризироваться — что возможно только в том случае, если он дистанцируется от себя как субъекта описываемого действия (как в случае *меня угораздило*). Для глаголов, оценивающих некоторый поступок постфактум, такое дистанцирование возможно. Но для установок типа *норовить* или *повадиться* (а также для *приспичило* в речевом режиме) это невозможно, поскольку они описывают состояния, имеющие место в момент речи. Аналогичным образом устроенная оценка в глаголах *польститься*, *позариться*, *удосужиться* и ряде других (см. [Левонтина 2012, 2022]), для которых употребление в 1-м лице также возможно лишь в случае переоценки собственных действий постфактум.

Авторы выражают благодарность участникам обсуждения нашего доклада на конференции «Маргиналии-2023» (Арзамас, 22–24 сентября 2023), по материалам которого написана настоящая статья.

Литература

- Апресян Ю. Д. Языковая аномалия и логическое противоречие // Text. Język. Poetyka. Wrocław etc.: Ossolineum, 1978.
- Апресян Ю. Д. Исследования по семантике и лексикографии. Т. 1. Парадигматика. М.: Языки славянских культур, 2009.
- Гончаров А. А., Добровольский Д. О., Зализняк Анна А. База данных конструкций с немецкими модальными глаголами и их русских соответствий // Корпусная лингвистика — 2023. СПб., 2023.
- Добровольский Д. О., Зализняк Анна А. Русские конструкции с потенциально модальным значением по данным параллельных корпусов // Труды Института

русского языка им. В. В. Виноградова. От семантических кварков до вселенной в алфавитном порядке. К 90-летию академика Ю. Д. Апресяна. 2020. № 25. С. 35–48.

Добровольский Д. О., Зализняк Анна А. Немецко-русский параллельный корпус как инструмент исследования модальности долженствования: о семантических различиях между *müssen* и *sollen* // Вестник Санкт-Петербургского университета. Язык и литература. 2022а. № 19 (4). С. 799–820.

Добровольский Д. О., Зализняк Анна А. Оценка как источник деонтической модальности // Критика и семиотика. 2022б. № 1. С. 52–72.

Зализняк Анна А. Лингвоспецифичные единицы русского языка в свете контрастивного корпусного анализа // Компьютерная лингвистика и интеллектуальные технологии. По материалам международной конференции Диалог'2015. М.: Изд. РГГУ, 2015. С. 651–662.

Зализняк Анна А. База данных межъязыковых эквиваленций как инструмент лингвистического анализа. // Компьютерная лингвистика и интеллектуальные технологии. По материалам международной конференции Диалог'2016. М.: Изд. РГГУ, 2016. С. 763–775.

Зализняк Анна А., Левонтина И. Б. Отражение «национального характера» в лексике русского языка (размышления по поводу книги: Anna Wierzbicka. *Semantics, Culture, and Cognition. Universal Human Concepts in Culture-Specific Configurations*. N. Y., Oxford, Oxford Univ. Press, 1992) // Russian Linguistics. 1996. Vol. 20. P. 237–264.

Инькова О. Ю. Принципы определения степени лингвоспецифичности коннекторов // Компьютерная лингвистика и интеллектуальные технологии. По материалам международной конференции Диалог'2017. М.: Изд. РГГУ, 2017. С. 139–149.

Кружков М. Г. Информационные ресурсы контрастивных лингвистических исследований: электронные корпуса текстов // Системы и средства информатики. 2015. Т. 25. № 2. С. 140–159.

Левонтина И. Б. Удосузиться и сподобиться: заметки о двух необычных русских глаголах // Смыслы, тексты и другие захватывающие сюжеты. Сборник статей в честь 80-летия И. А. Мельчука. М.: Языки славянских культур, 2012. С. 368–377.

Левонтина И. Б. Хорошо спрятанная оценка // Lifetime Linguistic Inspirations. To Igor Mel'cuk from Colleagues and Friends for his 90-th Birthday. Wiener Slawistischer Almanach, Linguistische Reihe, Herausgegeben von Tilmann Reuther, Sonderband 101, Berlin, 2022.

МАС — Малый академический словарь = Словарь русского языка в четырех томах. 3-е, стереотип. изд. М.: Русский язык, 1985–1988.

Падучева Е. В. Высказывание и его соотнесенность с действительностью. М.: Наука, 1985. 271 с.

Сичинава Д. В. Использование параллельного корпуса для количественного изучения лингвоспецифичной лексики // Язык, литература, культура: Актуальные проблемы изучения и преподавания. Вып. 10. М.: Макс Пресс, 2014. С. 37–44.

Шмелев А. Д. Перевод как инструмент анализа семантики лингвоспецифичных слов: *страх* и *трепет* в русском языке в свете данных параллельных корпусов //

Русский язык и культура в зеркале перевода: X Международная научная конференция; Материалы конференции. М.: Изд-во Московского университета, 2020. С. 238–250.

Bonola A. The parallel Italian-Russian-Italian corpus as an instrument for the contrastive study of discourse and textual markers: the case of *comunque* and its Russian translations // Съпоставително езикознание. 2017. № 4. София. С. 89–104.

Johansson S., Oksefjell S. (eds). Corpora and Cross-linguistic Research: Theory, Method, and Case Studies. Amsterdam: Rodopi, 1998.

Noël D. Translations as evidence for semantics: an illustration // Linguistics. 2003. № 41 (4). P. 757–785.

D. O. Dobrovol'skij

*V. V. Vinogradov Russian Language Institute
Institute of Linguistics of the Russian Academy of Sciences
(Russia, Moscow)
dobrovolskij@gmail.com*

Anna A. Zalizniak

*Institute of Linguistics of the Russian Academy of Sciences
Institute of Informatics Problems of the Federal Research Center
“Computer Science and Control” of the Russian Academy of Sciences
(Russia, Moscow)
anna.zalizniak@gmail.com*

**PARALLEL CORPUS IN THE SEMANTIC ANALYSIS
OF RUSSIAN LEXICAL UNITS:
CASE STUDY *PRISPICHILO* ‘HAVE AN URGE’
AND *NOROVIT* ‘TO BE EAGER’**

The article presents the results of applying the “monofocal” method of contrastive analysis. The research is based on data from the German parallel subcorpus of the Russian National Corpus (RNC) and on material of a supra-corpus database of constructions with German modal verbs and their Russian equivalents. The article provides a description of two Russian words with complex semantics — *prispichilo* (“have an urge” or “suddenly want to”) and *norovit* (“to be eager” or “to be inclined to”), which are marginal verbs of desire/intention. Analysis of the “stimuli” for the appearance of these words in translations from German to Russian and the “models” of their translations from Russian to German has allowed for the identification of not only the modal component in their semantics, but also other non-obvious components of their meaning that determine the characteristics of their semantics and usage.

Key words: parallel corpus, modal verbs, translation, semantics, Russian, German, evaluation, speaker

References

- Apresjan Ju. D. [Linguistic anomaly and logical contradiction]. *Text. Język. Poetyka*. Wrocław etc.: Ossolineum, 1978. (In Russ.)
- Apresjan Ju. D. *Issledovaniya po semantike i leksikografii* [Research on semantics and lexicography]. Vol. 1. *Paradigmatika* [Paradigmatics]. Moscow: Yazyki slavyanskikh kul'tur Publ., 2009. (In Russ.)
- Bonola A. The parallel Italian-Russian-Italian corpus as an instrument for the contrastive study of discourse and textual markers: the case of *comunque* and its Russian translations. *S''postavitelno ezikoznanie*. 2017, no. 4, pp. 89–104.
- Dobrovol'skij D. O., Zalizniak Anna A. [Russian constructions with potentially modal meaning: an analysis based on parallel corpus data]. *Proceedings of the V. V. Vinogradov Russian Language Institute. From semantic quarks to the universe in alphabetical order. To the 90th anniversary of Ju. D. Apresjan*. 2020, no. 25, pp. 35–48. (In Russ.)
- Dobrovol'skij D. O., Zalizniak Anna A. [Evidentiality and epistemic modality in the semantics of the German verbs *sollen* and *wollen* (based on the data from the German-Russian parallel corpus)]. *Vestnik of Saint Petersburg University. Language and Literature* [Bulletin of St. Petersburg University. Language and literature]. 2022a, no. 19 (4), pp. 799–820. (In Russ.)
- Dobrovol'skij D. O., Zalizniak Anna A. [Evaluation as a source of deontic modality]. *Kritika i semiotika* [Critique and Semiotics]. 2022b, no. 1, pp. 52–72. (In Russ.)
- Goncharov A. A., Dobrovol'skij D. O., Zalizniak Anna A. [Database of constructions with German modal verbs and their Russian correspondences]. *Korpusnaja lingvistika — 2023* [Corpus linguistics — 2023]. Saint Petersburg, 2023. (In Russ.)
- In'kova O. Ju. [Principles for determining the degree of linguistic specificity of connectors]. *Komp'yuternaya lingvistika i intellektual'nye tehnologii. Po materialam mezhdunarodnoi konferentsii Dialog'2017* [Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue” (2017)]. Moscow, RSUH Publ., 2017, pp. 139–149. (In Russ.)
- Johansson S., Oksefjell S. (eds). *Corpora and Cross-linguistic Research: Theory, Method, and Case Studies*. Amsterdam: Rodopi, 1998. 376 p.
- Kruzhkov M. G. [Information resources of contrastive linguistic research: electronic corpora of texts]. *Sistemy i sredstva informatiki* [Systems and means of informatics]. 2015, vol. 25, no. 2, pp. 140–159. (In Russ.)
- Levontina I. B. [*Udosuzhit'sja* and *spodobit'sja*: notes on two unusual Russian verbs]. *Smysly, teksty i drugie zakhvatyvayushchie syuzhety. Sbornik statei v chest' 80-letiya I. A. Mel'chuka* [Meanings, texts and other exciting subjects. Collection of articles in honor of the 80th anniversary of I. A. Mel'čuk]. Moscow: Yazyki slavyanskikh kul'tur Publ., 2012, pp. 368–377. (In Russ.)
- Levontina I. B. [Well-hidden evaluation]. *Lifetime Linguistic Inspirations. To Igor Mel'čuk from Colleagues and Friends for his 90-th Birthday. Wiener Slawistischer Almanach, Linguistische Reihe*. Ed. by T. Reuther, Sonderband 101, Berlin, 2022. (In Russ.)

MAS — *Malyy akademicheskii slovar'* = *Slovar' russkogo yazyka v chetyrekh tomakh* [Small Academic Dictionary = Dictionary of the Russian language in four volumes]. 3rd edn. Moscow, Russkii yazyk Publ., 1985–1988. (In Russ.)

Noël D. Translations as evidence for semantics: an illustration. *Linguistics*, 2003, no. 41(4), pp. 757–785.

Paducheva E. V. *Vyskazyvanie i ego sootnesennost' s dejstvitel'nost'ju* [The utterance and its correlation with reality]. Moscow: Nauka Publ., 1985. 271 p. (In Russ.)

Shmelev A. D. [Translation as a tool for analyzing the semantics of linguistically specific words: *strax* and *trepety* in the Russian language in the light of data from parallel corpora]. *Russkii yazyk i kul'tura v zerkale perevoda* [Russian language and culture in the mirror of translation: Proceedings of the Xth international scientific conference]. Moscow: Moscow University Publishing House, 2020, pp. 238–250. (In Russ.)

Sichinava D. V. [Using a parallel corpus for the quantitative study of language-specific vocabulary]. *Jazyk, literatura, kul'tura: Aktual'nye problemy izucheniya i prepodavaniya* [Language, literature, culture: Current problems of studying and teaching]. Issue 10. Moscow: Max Press Publ., 2014, pp. 37–44. (In Russ.)

Zalizniak Anna A. [Language-specific Russian words in the light of contrastive analysis]. *Komp'yuternaya lingvistika i intellektual'nyeologii. Po materialam mezhdunarodnoi konferentsii Dialog'2015* [Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue” (2015)]. Moscow, RSUH Publ., 2015, pp. 651–662. (In Russ.)

Zalizniak Anna A. [Database of cross-linguistic equivalences as a tool for linguistic analysis]. *Komp'yuternaya lingvistika i intellektual'nyeologii. Po materialam mezhdunarodnoi konferentsii Dialog'2016* [Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue” (2016)]. Moscow, RSUH Publ., 2016, pp. 763–775. (In Russ.)

Zalizniak Anna A., Levontina I. B. [Reflection of “national character” in the vocabulary of the Russian language (reflections on the book: Anna Wierzbicka. *Semantics, Culture, and Cognition. Universal Human Concepts in Culture-Specific Configurations*. N. Y., Oxford, Oxford Univ. Press, 1992)]. *Russian Linguistics*, 1996, vol. 20, pp. 237–264. (In Russ.)

М. В. Хохлова

Санкт-Петербургский государственный университет
(Россия, Санкт-Петербург)
m.khokhlova@spbu.ru

ВЫСОКОЧАСТОТНЫЕ ГЛАГОЛЬНЫЕ СЛОВАРНЫЕ КОЛЛОКАЦИИ С СУЩЕСТВИТЕЛЬНЫМИ В НАЦИОНАЛЬНОМ КОРПУСЕ РУССКОГО ЯЗЫКА¹

Корпусы текстов и ставший доступным в них статистический аппарат дали новый импульс дальнейшему изучению сочетаемости лексических единиц. При этом устойчивые словосочетания достаточно подробным образом представлены в толковых и специализированных словарях. Статья посвящена исследованию того, как, с одной стороны, словарные коллокации (единицы, зафиксированные в лексикографических источниках) отражены в корпусе текстов (на материале Национального корпуса русского языка) и как, с другой стороны, частотные словосочетания, полученные на материале корпуса, соответствуют словарным данным. Материалом послужили коллокации, построенные по модели «глагол + существительное», которые были отобраны из словарей русского языка, со следующими существительными: *жизнь, сила, дело, слово, работа, время, взгляд, вопрос, возможность, право*. Ранжирование словосочетаний в корпусе было произведено при помощи частоты совместной встречаемости, а также статистической метрики logDice. Результаты показывают, что высокочастотные единицы из корпуса довольно полно отражены среди коллокаций, представленных в словарях (около 67%). И наоборот, среди найденных частотных словосочетаний больше половины (55%) составляют словарные коллокации. Мера logDice демонстрирует результаты, схожие с ранжированием по частоте совместной встречаемости.

Ключевые слова: высокочастотные коллокации, словарные коллокации, Национальный корпус русского языка, словари, русский язык

Введение

Корпусные технологии постоянно развиваются и становятся ориентированными на более широкий круг пользователей — теперь к ним относятся не только

¹ Исследование выполнено за счет гранта Российского научного фонда № 22-18-00189, <https://rscf.ru/project/22-18-00189/>

специалисты-лингвисты, но и изучающие язык, в том числе иностранцы, школьники и преподаватели. Можно сказать, что корпусы текстов прошли путь от программ, которые демонстрируют контексты по запросу, а также выдают результаты, соответствующие определенным критериям лингвистической разметки и/или метаразметки, до более сложных систем, которые могут интегрировать, например, лексикографическую информацию.

Устойчивые словосочетания достаточно подробным образом представлены в разнообразных словарях, однако корпусы текстов открывают новые возможности для получения более подробной информации (в том числе для составления новых ресурсов, например, [Кустова 2008]). Так, например, на основе количественных показателей можно маркировать те словосочетания, которые являются устойчивыми и воспроизводимыми в речи, — это определенной степени пересекается с традицией разработки лексических минимумов для изучающих тот или иной язык. Однако хотя статистический аппарат может служить некоторым вспомогательным средством для оценки результатов автоматической выдачи примеров, ценность лексикографического отбора материала не поддается сомнению. По этой причине, несмотря на то, что объемы корпусов увеличиваются, не теряет своей актуальности следующий вопрос: как соотносятся корпусные данные со словарями (более ранние попытки соотнесения материалов двух типов были описаны в работах [Khokhlova 2008; Khokhlova 2016]). Поэтому применительно к исследованию сочетаемости нами ставится задача проанализировать распределение частотных коллокаций, зафиксированных в лексикографических источниках, в корпусе текстов.

Материал и методика исследования

До последнего времени для русского языка существовало не так много систем, в которых присутствовал бы статистический аппарат для оценки устойчивости словосочетаний (краткий обзор некоторых программных продуктов дан в [Хохлова 2023]). Национальный корпус русского языка (НКРЯ) [НКРЯ] за последние годы претерпел существенные изменения не только по своему объему, но также в плане тех технических возможностей, которые стали доступны пользователям. Для исследования сочетаемости лексических единиц особый интерес представляют сервис «Портрет слова» (виджет «Скетчи»), улучшенная демонстрация частотности в табличном виде и отдельно реализованный поиск коллокаций. Два последних инструмента были нами использованы при проведении исследования.

Материалом послужили примеры из Базы данных коллокаций русского языка [Хохлова, Мамаев 2021], в которой представлены устойчивые словосочетания, извлеченные из словарей русского языка, — так называемые словарные коллокации. В настоящее время в ней содержится 40 тыс. единиц, которые были зафиксированы в девяти толковых словарях и словарях устойчивых словосочетаний [там же], репрезентирующих следующие модели: «прилагательное + существительное» (*крупный авторитет*), «причастие + существительное» (*подчёркнутая вежливость*) «наречие + прилагательное» (*безумно популярный*) и «глагол + существительное» (*протить свет*). Можно заметить, что словосочетания обладают разной степенью

устойчивости и, следовательно, в этом случае о коллокациях можно говорить как о «зонтичном» термине. Словосочетания дополнены значениями статистических метрик (MI, MI3, log-likelihood, logDice, t-score), которые позволяют оценить устойчивость и вычислены на основе корпусов текстов (НКРЯ и Araneum Russicum Maximum [Benko 2014]). Дополнительно введен словарный индекс — количество словарей, в которых зафиксирована рассматриваемая единица. Глагольные словосочетания приведены в шести словарях [Регинина и др. 1980; МАС; Дерибас 1983; Мельчук, Жолковский 1984; Борисова 1995; Бирюк, Гусев, Калинина 2008], то есть теоретически его максимальное значение можно быть равно шести, однако ни одно устойчивое словосочетание не было приведено одновременно во всех перечисленных источниках, поэтому на практике оно оказалось равным пяти². Наиболее многочисленным по количеству примеров оказался словарь [Дерибас 1983], в котором зафиксировано 8607 единиц.

В работе нами были рассмотрены глагольные словосочетания, зарегистрированные в базе данных, и проведено их сравнение с результатами из НКРЯ. Подкорпус основного корпуса был ограничен текстами с 1950 по 2023 гг. (общий объем 197 159 944 слов). Согласно [Khokhlova 2023], наиболее частотными глаголами, которые встречаются в составе коллокаций, являются следующие (приведены в порядке уменьшения частот): *быть, давать/ дать, получить, сделать, иметь, получать, делать, принять, брать/взять, провести/проводить, вести, вызвать/вызывать, видеть, проявлять, принимать*. К частотным существительным, образующим наибольшее количество словарных коллокаций, относятся следующие³: *жизнь, сила, дело, слово, работа, время, взгляд, вопрос, возможность, право*. Нами были отобраны глагольные сочетания с приведенными выше существительными общим объемом 875 единиц. Максимальное значение словарного индекса для коллокаций, в которых выступают данные существительные, как было упомянуто выше, так же равно 4.

Цель исследования заключается в том, чтобы проанализировать, как представлены в таком большом и представительном корпусе текстов, как НКРЯ, словарные коллокации, насколько полным и точным является их покрытие — как корпус отражает тот материал, который был подобран вручную, и, наоборот, как в словарях представлены наиболее частотные устойчивые словосочетания, приведенные в корпусе.

Инструмент «Частотность»

Для каждого из приведенных выше существительных нами были рассмотрены 100 коллокаций, полученных при помощи инструмента «Частотность», и проведено сравнение со словарными коллокациями из базы данных. Результаты для полноты

² В то же время для атрибутивных словосочетаний максимальное значение для словарного индекса равно шести [Хохлова 2021].

³ Половина из приведенных ниже существительных принадлежит также к десяти наиболее частотным согласно словарю [Ляшевская, Шаров 2009].

и точности приведены в Табл. 1. В случае если один из глаголов в видовых парах присутствовал в коллокации, то сочетание с глаголом иного вида также нами рассматривалось как совпадающее (это решение было продиктовано существованием разных подходов, реализованных в разных словарях, а также наличием возможных ошибок в морфологической разметке в корпусе). Например, среди результатов в корпусе приводится *рисковать жизнью*, однако в словарных коллокациях присутствуют сочетания с глаголами обоих видов (*рискнуть жизнью* и *рисковать жизнью*), поэтому нами учитывались оба примера.

Таблица 1. Полнота и точность относительно выдачи инструмента «Частотность»

	Словарные коллокации (полнота)	Выдача стат. метрики (точность)
жизнь	0,57	0,56
сила	0,56	0,50
дело	0,60	0,56
слово	0,68	0,62
работа	0,81	0,65
время	0,57	0,41
взгляд	0,69	0,49
вопрос	0,60	0,49
возможность	0,86	0,62
право	0,84	0,58

Наиболее полно корпусные примеры были представлены среди словарных коллокаций для существительных *возможность*, *право* и *работа*. На Рис. 1 показано, как отражены устойчивые словосочетания из словарей в корпусе.

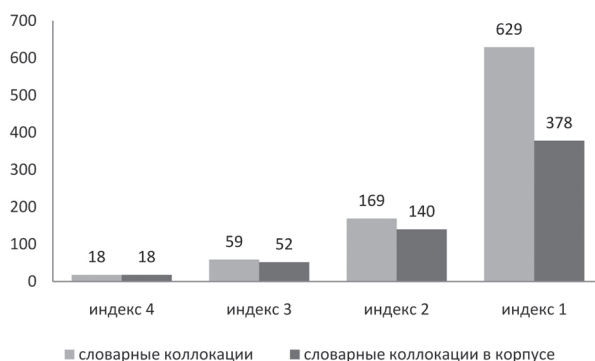


Рис. 1. Распределение коллокаций в НКРЯ согласно значениям словарного индекса (инструмент «Частотность»)

В полученных списках были найдены все словарные коллокации с индексом 4 (всего 18 единиц). Также было зафиксировано большинство словосочетаний с индексом 3 (90%). Среди словосочетаний, отмеченных в трех словарях, но отсутство-

вавших среди результатов, отметим следующие: *занимать время, лишить слова, нарушать слово, отдавать силы, провести / проводить жизнь*. Большая часть коллокаций со словарным индексом 2 также присутствует в частотной выдаче (83%). Интересно отметить, что к отсутствующим в данной выдаче единицам относятся те, которые зафиксированы в словаре [Дерибас 1983]. Данный словарь описывает довольно широкий круг единиц, чем может быть объяснено включение устойчивых, хотя и менее частотных единиц. Например, для *взвешивать / взвесить слово* было найдено 12 примеров в корпусе, *брать / взять время* представлено 43 вхождениями, из которых не все относятся к данному устойчивому словосочетанию, и *выдвинуть / выдвигать вопрос* репрезентировано 9 вхождениями). Наиболее многочисленная группа коллокаций со словарным индексом 1 также характеризуется большим процентом вхождений в корпус (60% от общего количества словосочетаний, приведенных только в одном источнике). Выводы соотносятся с теми, которые были ранее получены для атрибутивных коллокаций [Хохлова 2021]: большее значение словарного индекса свидетельствует о том, что словосочетание обладает большей воспроизводимостью и, как следствие, будет присутствовать среди высокочастотных результатов в корпусе.

Среди выдаваемых частотных результатов можно отметить словосочетания, которые не маркированы в словарях как устойчивые. Например, *определить силу* представлено в корпусе большим количеством примеров из учебников:

Определить силы сопротивления разгону полностью груженного легкового автомобиля ГАЗ-3110 «Волга» на первой, второй и на прямой передачах. [Б. М. Тверсков. Задачи по теории автомобиля (2010)]

Пользуясь шкалой Бофорта, ***определяйте силу*** ветра в баллах и его скорость в течение текущей недели. [Основы безопасности жизнедеятельности. 7 класс (2008)]

В данном случае существительное *сила* употребляется в значении «способность производить какую-л. работу; энергия, мощность» [МАС], в то время как данные терминологические словосочетания отсутствовали среди рассмотренных нами примеров в словарях.

Отметим, что результаты для ряда существительных демонстрируют большой разброс частотностей вхождения сочетаний в корпус: наблюдается несколько высокочастотных коллокаций, в то время как иные встречаются значительно реже (это также объясняется сочетаемостью конкретного существительного с высокочастотным глаголом). Например, *иметь возможность, дать / давать возможность, быть возможность* и *получить возможность* представлены в корпусе большим количеством примеров. *Иметь право* является своего рода статистическим выбросом (48,89 ipm) по сравнению с остальными примерами, относительная частота которых не превышает 3,47 ipm. Для существительного *дело* найденные словосочетания отличаются большой дисперсией: частотности *иметь дело, быть дело* и *обстоять дело* составляют 21,04 ipm, 13,00 ipm и 7,42 ipm

соответственно, в то время как остальные словосочетания встречаются значительно реже (менее 2,60 ipm).

При анализе результатов были отмечены примеры, которые формально удовлетворяют критериям поиска словосочетания, однако являются ошибочными. Например, было найдено сочетание с глаголом *утопать*, однако все 29 примеров не относятся к упомянутому существительному (при этом словоформа *утопающий* также ошибочно размечена как форма глагола):

Как говорится, спасение *утопающих* — *дело* рук самих утопающих... [П. Л. Проскурин. Число зверя. Часть вторая // Роман-газета, 1999]

Памятуя, что «спасение *утопающих* — *дело* рук самих утопающих», владельцы дорогих авто приобретают ещё и полис КАСКО. [Алиса Гаевская. Этот нестрашный страховой брокер // «Наука и жизнь», 2009]

В целом, высокочастотные сочетания в корпусе довольно полно отражены среди коллокаций, представленных в базе данных (около 67% от общего количества). В то время как среди найденных частотных словосочетаний больше половины (55%) составляют словарные коллокации.

Инструмент «Коллокации»

В ходе данного эксперимента были рассмотрены коллокации для существительных, полученные при помощи инструмента «Поиск коллокаций» с использованием меры logDice, которая была предложена в работе [Rychly 2008] с целью отбора коллокаций для решения лексикографических задач. Большая часть словарных коллокаций (в среднем 69%) совпадает с корпусным материалом (см. Табл. 2). В списках, выданных согласно мере logDice, присутствует 57% данных словосочетаний, представленных в словарях. Наибольшее количество словарных коллокаций характерно для существительных *возможность*, *право* и *работа*.

Таблица 2. Полнота и точность относительно выдачи инструмента «Коллокации»

	Словарные коллокации (полнота)	Выдача стат. метрики (точность)
жизнь	0,57	0,54
сила	0,60	0,55
дело	0,59	0,58
слово	0,69	0,62
работа	0,82	0,69
время	0,63	0,44
взгляд	0,63	0,49
вопрос	0,62	0,49
возможность	0,87	0,62
право	0,91	0,64

Мера logDice показывает ранжирование, схожее с ранжированием по частоте совместной встречаемости (значение коэффициента корреляции Спирмена превышает 0,9, $p\text{-value} \ll 0,05$). В целом списки найденных словарных коллокаций пересекаются. Однако встречаются единицы, которые не были представлены среди результатов, выдаваемых согласно частоте, однако содержатся среди результатов, полученных при помощи данного инструмента. Например, *лишиться сил, коверкать слова, путать слова, занять время, отнимать время*.

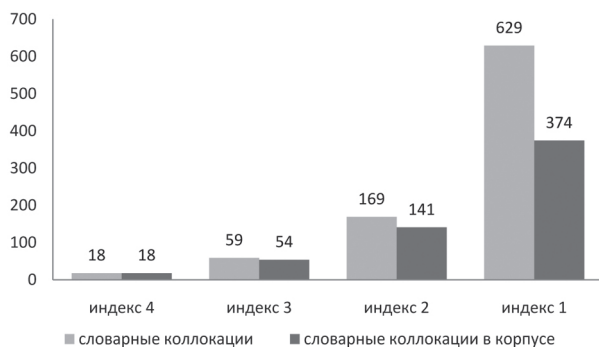


Рис. 2. Распределение коллокаций в НКРЯ согласно значениям словарного индекса (инструмент «Коллокации»)

Сравнение того, как отражены словосочетания разных словарных индексов, полученные при помощи инструментов «Частотность» и «Коллокации», показывает отсутствие статистически значимых отличий ($\chi^2 < 0$, $p\text{-value} \gg 0,05$).

Заключение

Проведенный анализ позволяет прийти к следующим выводам: во-первых, словарные коллокации достаточно полно отражены в корпусе, а во-вторых, частотные словосочетания, полученные на материале корпуса, также во многом пересекаются с данными, зафиксированными в словарях. Таким образом, можно говорить о том, что оба типа источника показывают схожие результаты. Наиболее полно в корпусе отражены глагольные коллокации, представленные в двух, трех и четырех словарях, в то время как словосочетания, имеющие словарный индекс 1, представлены реже и в целом характеризуются низкими частотами.

Результаты, полученные при помощи инструментов «Частотность» и «Коллокации» (мера logDice) демонстрируют большой процент совпадений между собой, а также одинаковое ранжирование полученных списков устойчивых словосочетаний. Это позволяет сделать предположение о взаимозаменяемости метрик (частоты взаимной встречаемости и logDice) применительно к задаче выявления устойчивых словосочетаний, описанных в словарях.

В качестве дальнейших перспектив отметим использование полученных данных применительно к задаче формирования лексических минимумов, а также составления словарей сочетаемости.

Литература

Бирюк О. Л., Гусев В. Ю., Калинина Е. Ю. Словарь глагольной сочетаемости непредметных имен русского языка. М., 2008. [Электронный ресурс]. URL: http://dict.ruslang.ru/abstr_noun.php (дата обращения: 15.01.2024).

БТС — Большой толковый словарь русского языка: А-Я / Сост., гл. ред. канд. филол. наук С. А. Кузнецов. СПб.: Норинт, 1998. 1534 с.

Борисова Е. Г. Слово в тексте. Словарь коллокаций (устойчивых сочетаний) русского языка с англо-русским словарем ключевых слов. М.: Филология, 1995. 148 с.

Дерибас В. М. Устойчивые глагольно-именные словосочетания русского языка. М.: Русский язык, 1983. 256 с.

Кустова Г. И. Электронный словарь степенной сочетаемости на базе Национального корпуса русского языка // Труды международной конференции «Корпусная лингвистика–2008». СПб.: Изд-во СПбГУ, 2008. С. 229–239.

Кустова Г. И. Словарь русской идиоматики. Сочетания слов со значением высокой степени. М., 2008. URL: <http://dict.ruslang.ru/magn.php> (дата обращения: 15.01.2024).

Ляшевская О. Н., Шаров С. А. Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009. [Электронный ресурс]. URL: <http://dict.ruslang.ru/freq.php> (дата обращения: 15.01.2024).

Мельчук И. А., Жолковский А. К. Толково-комбинаторный словарь современного русского языка. Опыты семантико-синтаксического описания русской лексики. Вена: Wiener Slavistischer Almanach, 1984. 992 с.

МАС — Словарь русского языка: В 4-х т. / Под ред. А. П. Евгеньевой. 4-е изд., стер., М.: Рус. яз., 1999.

НКРЯ — Национальный корпус русского языка [Электронный ресурс]. URL: <http://ruscorpora.ru> (дата обращения: 15.01.2024).

Регина К. В., Тюрина Г. П., Широкова Л. И. Устойчивые словосочетания русского языка: Учеб. пос. для студентов-иностранцев / Под ред. Л. И. Широковой М., 1980. 296 с.

Убин И. И. Словарь усилительных словосочетаний русского и английского языков. М.: Рус. яз., 1987. 295 с.

Хохлова М. В. Атрибутивные коллокации в золотом стандарте сочетаемости русского языка и их представление в словарях и корпусах текстов // Вопр. лексикографии. 2021. № 21. С. 33–68. DOI: 10.17223/22274200/21/2

Хохлова М. В. Статистические методы в лексикографических исследованиях: представление частотной лексики // Terra Linguistica. 2023. Т. 14. № 3. С. 80–93.

Хохлова М. В., Мамаев И. Д. Разработка базы данных коллокаций: обзор золотого стандарта на примере атрибутивных словосочетаний // Труды международной конференции «Корпусная лингвистика–2021». СПб.: Скифия-принт, 2021. С. 370–379.

Benko V. Aranea: Yet Another Family of (Comparable) Web Corpora // Text, Speech and Dialogue. 17th International Conference, TSD 2014, Brno, Czech Republic, September 8–12, 2014. Proceedings. LNCS 8655. Springer International Publishing Switzerland, 2014. P. 257–264.

Khokhlova M. Extracting Collocations in Russian: Statistics vs. Dictionary // JADT 2008: actes des 9es Journées Internationales d'Analyse Statistique des Données Textuelles, Lyon, 12–14 mars 2008: Proceedings of 9th International Conference on Textual Data statistical Analysis, Lyon, March 12–14, 2008. Lyon, 2008. P. 613–624.

Khokhlova M. V. Large Corpora and Frequency Nouns // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2016”. Vol. 15 (22). Moscow: RSUH. P. 224–238.

Rychlý P. A lexicographer-friendly association score // Proceedings of Recent Advances in Slavonic Natural Language Processing, RASLAN. Brno, 2008. P. 6–9.

Maria V. Khokhlova
St. Petersburg State University
(Russia, St. Petersburg)
m.khokhlova@spbu.ru

HIGH-FREQUENCY VERB-NOUN COLLOCATIONS FROM DICCTIONARIES IN THE RUSSIAN NATIONAL CORPUS

Corpora and statistical tools have provided new opportunities for further studying the collocability of lexical units. In this context, stable word combinations are described in detail in explanatory and specialized dictionaries. This paper analyzes how, on the one hand, dictionary collocations (units fixed in lexicographical sources) are reflected in the corpus (based on the Russian National Corpus) and, on the other hand, how frequent word combinations retrieved from the corpus correspond to dictionary data. The material includes a list of collocations constructed on the “verb + noun” model, selected from a number of dictionaries of the Russian language, with the following nouns: *zhizn'* ‘life’, *sila* ‘power’, *delo* ‘business’, *slovo* ‘word’, *rabota* ‘work’, *vremja* ‘time’, *vzglyad* ‘glance’, *vorpos* ‘question’, *vozmozhnost'* ‘opportunity’, *pravo* ‘right’. The ranking of word combinations in the corpus was done using the frequency of co-occurrence, as well as logDice. The results show that high-frequency units from the corpus are quite well represented among the collocations presented in dictionaries (about 67%). Conversely, more than half of the identified frequent word combinations (55%) are dictionary collocations. The logDice measure demonstrates results similar to ranking by co-occurrence frequency.

Keywords: high-frequency collocations, dictionary collocations, Russian National Corpus, dictionaries, Russian language

Acknowledgements

The study was supported by the Russian Science Foundation, project No. 22-18-00189, <https://rscf.ru/project/22-18-00189/>

References

Benko V. Aranea: Yet Another Family of (Comparable) Web Corpora. *Text, Speech and Dialogue. 17th International Conference, TSD 2014, Brno, Czech Republic, September 8–12, 2014. Proceedings. LNCS 8655*. Springer International Publishing Switzerland, 2014, pp. 257–264.

Biriuk O. L., Gusev V. Yu., Kalinina E. Yu. *Slovar' Glagol'noj Sochetaemosti Nepredmetnykh Imen Russkogo Yazyka. Slovar' na osnove Natsional'nogo Korpusa Russkogo Yazyka* [Dictionary of Russian Abstract Nouns' Verbal Collocability. A Dictionary based on the Russian National Corpus], Moscow, 2008. Available at: http://dict.ruslang.ru/abstr_noun.php. (accessed 15.01.2024) (In Russ.)

Bol'shoy tolkovyy slovar' russkogo yazyka [Big Explanatory Dictionary of Russian], S. A. Kuznetsov (ed.). St. Petersburg, Norint Publ., 1998. 1534 p.

Borisova E. G. *Slovo v tekste. Slovar' kollokatsiy (ustoychivyykh sochetaniy) russkogo yazyka s anglo-russkim slovarem klyuchevykh slov* [A Word in a Text. A Dictionary of Russian Collocations with English-Russian Dictionary of Keywords]. Moscow, Filologiya Publ., 1995. 148 p.

Deribas V. M. *Ustoichivyye glagol'no-imennyye slovosochetaniya russkogo yazyka* [Verb-Noun Collocations in Russian]. Moscow, Russkiy yazyk Publ., 1983. 256 p.

Slovar' russkogo yazyka v 4 tomakh [Dictionary of the Russian Language in 4 volumes] (MAS) / ed. by A. P. Yevgen'yeva. Vol. 1–4, 4th edition, revised and supplemented. Moscow, Russkiy yazyk Publ., 1999.

Khokhlova M. Extracting Collocations in Russian: Statistics vs. Dictionary. *JADT 2008: actes des 9es Journées Internationales d'Analyse Statistique des Données Textuelles, Lyon, 12–14 mars 2008: Proceedings of 9th International Conference on Textual Data statistical Analysis*. Lyon, 2008, pp. 613–624.

Khokhlova M. V. Large Corpora and Frequency Nouns. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2016"*. Vol. 15 (22). Moscow, RSUH Publ., 2016, pp. 224–238.

Khokhlova M. V. [Attributive Collocations in the Gold Standard of Russian Language Collocability and Their Representation in Dictionaries and Text Corpora]. *Voprosy leksikografii*, 2021, no. 21, pp. 33–68. (In Russ.)

Khokhlova M. V. [Statistical Methods in Lexicographical Research: Representation of Frequency Lexicon]. *Terra Linguistica*, 2023, vol. 14, no. 3, pp. 80–93. (In Russ.)

Khokhlova M. V., Mamaev I. D. [Development of the Russian Collocations Database: Review of the Gold Standard Using Attributive Word Combinations as an Example]. *Trudy mezhdunarodnoy konferentsii "Korpusnaya lingvistika–2021"* [Proc. Int. Conf. "Corpus Linguistics–2021"]. St. Petersburg, Skifiya-print Publ., 2021, pp. 370–379. (In Russ.)

Kustova G. I. [Electronic dictionary of collocability based on the Russian National Corpus]. *Trudy mezhdunarodnoy konferentsii "Korpusnaya lingvistika–2008"* [Proc. Int. Conf. "Corpus Linguistics–2008"]. St. Petersburg, St. Petersburg. Univ. Publ., 2008, pp. 229–239. (In Russ.).

Kustova G. I. *Slovar' russkoy idiomatiki. Sochetaniya slov so znacheniyem vysokoy stepeni* [Dictionary of Russian Idiomatic Expressions]. Moscow, 2008. URL: <http://dict.ruslang.ru/magn.php> (accessed 15.01.2024)

Lyashevskaya O., Sharoff S. *Chastotnyy slovar' sovremennogo russkogo yazyka (na materialakh Natsional'nogo Korpusa Russkogo Yazyka)* [The Frequency Dictionary of Modern Russian based on the Russian National Corpus data]. Moscow, Azbukovnik Publ., 2009. Available at: <http://dict.ruslang.ru/freq.php> (accessed 15.01.2024)

Mel'čuk I., Zholkovsky A. *Tolkovo-kombinatornyy slovar russkogo yazyka* [Explanatory Combinatorial Dictionary of Modern Russian]. Vienna, Wiener Slavistischer Almanach, 1984. 992 p.

Oubine I. I. *Slovar' usilitel'nykh slovosochetaniy russkogo i angliyskogo yazykov* [Dictionary of Intensifying Word Combinations in Russian and English]. Moscow, Russkiy Yazyk Publ., 1987. 295 p.

Reginina K. V., Tjurina G. P., Shirokova L. I. *Ustoychivye slovosochetaniya russkogo yazyka: Uchebnoye posobiye dlya studentov-inostrantsev* [Set Expressions of the Russian Language. A Reference Book for Foreign Students] / ed. by L. I. Shirokova. Moscow, Russkiy yazyk Publ., 1980. 296 p.

Russian National Corpus. Available at: <http://ruscorpora.ru> (accessed 15.01.2024)

Rychlý P. A lexicographer-friendly association score. *Proceedings of Recent Advances in Slavonic Natural Language Processing*, RASLAN, 2008, pp. 6–9.

Е. В. Рахилина

Национальный исследовательский университет

«Высшая школа экономики»

(Россия, Москва)

rakhilina@gmail.com

В. А. Плунгян

Институт русского языка им. В. В. Виноградова РАН

(Россия, Москва)

plungian@gmail.com

ПРЕЛЕСТЬ: История слова¹

Незабываемой Нине Давидовне — к 100-летию

Лингвистам известна загадка слова *прелесть*, совмещавшего в себе два на первый взгляд противоположных значения: ‘соблазн, искушение’ и ‘красота, очарование’. При обсуждении этой необычной полисемии (которая иногда даже квалифицировалась как энантиосемия, что, с нашей точки зрения, не вполне обоснованно) в лингвистических работах главным образом обращалось внимание на сосуществование двух значений в разных сферах — «сакральной» и «светской», однако вопрос об обстоятельствах такого семантического перехода при этом оставался в тени. В статье, описывающей результаты корпусного исследования (с использованием новых возможностей панхронического поиска, доступных в Национальном корпусе русского языка), обращается внимание на некоторые дополнительные аспекты семантической истории этого слова — прежде всего, на влияние французского языка, способствовавшего появлению «светского» значения у *прелесть* / *преlestи* в качестве семантической кальки с франц. *charme(s)*. Интерес представляет также появление и дальнейшее развитие прямого французского заимствования — слова *шарм*, активно употребляющегося и в современном русском языке.

Статья посвящена памяти Нины Давидовны Арутюновой, любившей это слово — и сложность естественного языка.

Ключевые слова: русский язык, французский язык, семантические кальки, полисемия, энантиосемия, корпусное исследование

¹ Работа поддержана Министерством образования Республики Корея и Национальным Научным Фондом Кореи (NRF-2021S1A5A2A03065528).

1. Введение. Загадка прелести

Прелесть было любимым словом Нины Давидовны Арутюновой, всегда выражавшим у нее яркую положительную оценку: *Ну прелесть, просто прелесть!* В ее речи оно звучало необычно и нестандартно — хотя, как показывает НКРЯ, так оно в русском языке и до сих пор достаточно часто используется, в том числе предикативно (как оценка места, вещи, события, поведения человека или каких-то внешних обстоятельств):

- (1) *Никаких очередей в магазинах, тихо на пляже — прелесть!*» [Lenta.ru, 2020.08.27]²

Часто слыша от нее это слово, мы не вполне осознавали тогда, что оно скрывает одну лингвистическую загадку. Она касается двух чрезвычайно интересовавших Нину Давидовну как лингвиста *сюжетов* (еще одно ее любимое слово!): оценки [Арутюнова 1988; Арутюнова 1999: 130–274] и полисемии (уже в [Арутюнова 1976]).

Проблема, о которой пойдет речь, была замечена прежде всего пушкинистами — на примере текстов Пушкина. Первым был В. В. Виноградов, отметивший в ряде публикаций 1940–60-х гг. нетривиальное поведение слова *прелестный* у Пушкина (см. сводные данные о них в [Виноградов 1999: 982–983]); затем на нее обратил внимание Н. В. Перцов (2000) — его статья появилась в одном из сборников серии «Логический анализ языка» под ред. Н. Д. Арутюновой и содержала обсуждение последнего сонета Пушкина, и в том числе известной строки из него, *чистейшей прелести чистейший образец*, интерпретация смысла которой оказалась отнюдь не тривиальной задачей. И только потом обсуждение перекочевало в собственно лингвистические работы (в частности, [Анна Зализняк, Шмелев 2021: 131], [Мельничук 2017] и др.; подробнее мы их рассмотрим ниже).

Речь во всех случаях шла о необычной полисемии этого слова. Действительно, с одной стороны, *прелесть* — это ‘грех, соблазн, обольщение’, то есть нечто *прелестительное*, ср.:

- (2) <...> *оную погибельную прелесть, в которой увязают грешники* <...> [архиепископ Платон (Левшин). Нравоучение седьмое (1757)],

а с другой — ‘очарование, привлекательность’, как в более позднем примере из Лермонтова:

- (3) <...> *за молодыми казачками, — которых прелесть трудно постигнуть, не видав их, он никогда не волочил.* [М. Ю. Лермонтов. Герой нашего времени (1839–1841)]

Загадка состоит в том, почему слово *прелесть*, которое восходит к идее обмана (ср. бесприставочное *лесть*) и исходно значило ‘соблазн’ и ‘грех’, *изменило оценку*

² Здесь и далее примеры взяты из Национального корпуса русского языка (ruscorpora.ru).

с отрицательной на положительную и вдруг стало значить ‘очарование’ — при этом не утратив и возможность употреблений в исходном значении.

В. В. Виноградов усматривал у этих двух значений омонимические отношения и считал, что «*прелестный* в XVIII в. распадается на два слова» [Виноградов 1999: 983]; другими словами, происходит семантический сдвиг от одного значения к другому. При этом «<...> омонимия преодолевается устранением старого церковно-книжного употребления» (там же). В классических текстах — в особенности, в текстах Пушкина — он усматривал намеренный авторский «стилистический синтез таких значений слова, которые уже готовы были разойтись как омонимы», как в примерах типа *Прелестный опыт упреждая, // Мы только счастью вредим* [Евгений Онегин 1, 9, черновая рукопись] (там же).

Поддержке и развитию идей В. В. Виноградова относительно слов *прелесть* и *прелестный* на корпусных данных русского литературного языка XVIII–XXI вв. посвящена специальная статья [Мельничук 2017]. В ней утверждается развитие у этого слова энантиосемии — и распределение значений в зависимости от тематики текста: с отрицательной оценкой оно выступает в религиозных и нравоучительных текстах, постепенно смещаясь на периферию узуса. Как кажется, термин «энантиосемия» в отношении к слову *прелесть* не совсем точен: выделенные значения не антонимичны в строгом смысле, а лишь оценочно противопоставлены. С этим уточнением мы возвращаемся к вопросу о том, какие процессы привели к такому противопоставлению.

Дополнительную трудность составляет направление перехода: как более ранняя засвидетельствована отрицательная оценка — следовательно, она должна признаваться исходной, а положительная — производной. Между тем, вообще говоря, естествен и ожидаем противоположный переход: от положительной оценки к отрицательной. Он обладает чрезвычайно высокой частотностью, потому что опирается на универсальный механизм иронии. Таким образом, практически любая позитивная оценка может в определенном — ироническом — контексте меняться на негативную (см. об этом, например, в [Найман 1998] или, в другом теоретическом контексте, в [Шатуновский 2016]); характерный контекст таких употреблений представлен в (4).

- (4) <...> *учтиво раскланявшись, оставил знаменитого дельца, который, кажется, задумал подражать английским адвокатам и брать деньги даже и за советы. Пятьсот рублей тотчас и столько же по окончании процесса! Нечего сказать, молодец!* [С. П. Жихарев. Записки современника (1806–1809)]

Мена же оценки с отрицательной на положительную — редка и всегда представляет особый интерес. Хорошим примером здесь может служить существительное *гордость*, которое, как показано в [Шмелев 2000], демонстрирует такой переход не за счет семантического сдвига в лексеме, а за счет плюрализма этических

систем. Ср. отрицательную оценку в примерах XVIII века (5–6) и положительную в русском XXI века (7), отражающих разные этические нормы:

- (5) *Паде столп Вавилонский: ибо его созидала гордость*. [митрополит Стефан (Яворский). Проповеди (1700–1722)]
- (6) *Грубость бывающая отъ гордости достойна ненависти, а отъ глупости происходящая презрѣнія*. [С. С. Волчков. Придворной человекъ [перевод книги Грациана с французского] (1742)]
- (7) *Это очень просто, но для ребенка — настоящее чудо и гордость, если он сделал все сам*. [О. Л. Литвиненко. Рекомендации для учащихся по проведению простых биологических экспериментов в домашних условиях (2021)]

2. Секрет прелести

Ключ к этой загадке могут, как и во многих других случаях, дать инструменты корпусного поиска — в данном случае, на помощь приходит недавно внедренный в НКРЯ панхронический поиск (подробнее о нем см. [Сичинава 2024]), который отчетливо показывает, что второе, положительное, значение у слова *прелесть* возникает в XVIII веке первоначально в переводах с французского, где оно калькирует французское *charme* (8–10):

- (8) *Глас дружества приносит некую отраду душе, не чувствующей ни единыя прелести природы; в прежесточайшем своем бедствии Иосиф тяжко-серд не был и не возмог от сообщества людей бежати навеки*. [Д. И. Фонвизин. Иосиф [перевод поэмы П. Ж. Битобе с французского] (1769)].
- (9) — *С того времени, как познал я прелести добродетели, все для меня переменилось*. [Д. И. Фонвизин. Иосиф [перевод поэмы П. Ж. Битобе с французского] (1769)]
- (10) *Победители стали ученики побежденных, и обучились языку, которой Омиры, Пиндары, Фукидиды, Ксенофонты, Демосфены, Платоны, Еврипиды и пр., украсили всеми прелестями своего разума*. [А. Н. Радищев. Размышления о греческой истории или о причинах благоденствия и несчастья Греков [перевод книги Г. де Мабли с французского] (1773)]

Французское *charme* с исходным (оценочно нейтральным) значением ‘колдовство, чары’ развило положительную оценку и приобрело значение ‘очарование, привлекательность’. Оно фиксируется уже с XII в. и этимологически восходит к лат. *carmen* ‘песня; заклинание’. Интересно, что аналоги этим семантическим переходам есть в русском, ср. *чары* > *очаровательный*, а также *обаяние* (< *баять*). Французскому языку хорошо известно и употребление слова *charme* во множественном числе (как в примерах типа *les charmes opulents des modèles de Rubens* ‘пышные формы / прелести моделей Рубенса’), с распространенным метонимическим сдвигом.

Именно в этом последнем значении (упущенном у В. В. Виноградова) оно и было калькировано первоначально: слово *прелести* в таких контекстах выступает

как предметное имя со значением ‘соблазнительные части (женского) тела’. Такие контексты регистрируются в корпусе чуть раньше оценочных (и встречаются, естественно, тоже в переводах и переложениях с французского):

- (11) *Тотъ наилутче сего избудеть, кто у Улисса научится, какъ отъ **прелестей** Цирциныхъ убѣгать*; [С. С. Волчков. Придворной человѣкъ [перевод книги Грациана с французского] (1742)]
- (12) *Такъ ли пріятною луна бываетъ ночью заблудившимся, какою мнѣ сія **прекрасная безпримѣрную** свою бѣлостъ со всѣми **прелѣстями** соединенную въ черное облекши сукно, показалась?* [Ф. А. Эмин. Непостоянная fortuna или похождение Мирамонда (1763)]

Любопытно, что и первые оценочные употребления слова *прелесть* в НКРЯ тоже встречаются во множественном числе, как бы вторя более ранним предметным:

- (13) **прелести** добродетели (1769);
- (14) **прелести** разума (1769);
- (15) <...> украсили всеми **прелестями** своего разума (1773);
- (16) <...> **прелести** любви и нежность любезного пола (1775)
- (17) *Сколь щастливъ человѣкъ, когда онъ видитъ предъ собою часть вселенной, украшенную всѣми весенними **прелестями*** <...> (1792).

Единственное число в оценочных контекстах по НКРЯ регистрируется чуть позже (18):

- (18) *Въ самомъ дѣлѣ, она соединяла въ себѣ умъ великаго царя съ **прелестью** и слабостями любезной женщины*. [А. С. Шишков. Записки (1780–1814)]

Учитывая сказанное, приходится предположить несколько иное развитие семантики слова *прелесть*, чем то, что принято в литературе благодаря В. В. Виноградову. А именно, что в русском происходило не **расхождение** (или «расщепление») смыслов у исходного церковнославянского *прелесть*, а приобретение этим словом нового оценочного значения «извне» в качестве семантической кальки с французского. Кроме того, если говорить строго, в XVIII веке сосуществовали не два, а **три** типа употреблений лексемы *прелесть*, распределенные по типам текстов и числовой форме: славянское ‘соблазн’, калькированное из французского ‘соблазнительные части тела’ (только множественное) и калькированное из французского ‘очарование’; последние два, как мы уже отмечали, связаны метонимически. Славянское *прелесть* с отрицательной оценкой довольно быстро исчезло из общеязыкового узуса и осталось в церковном употреблении. Предметное значение ‘части тела’ к XX веку тоже стало малоупотребительно.

3. Судьба прелести

Между тем «светское» значение слова *прелесть* постепенно меняет свою сочетаемость. Прежде всего, сужается его связь с носителем оценки — обычно женщиной, как в (19). Такого рода примеры становятся крайне редки: так, (20) — как

кажется, единственный контекст в корпусе современных СМИ (речь в нем идет, вопреки ожиданиям, всего лишь про сложные пути похудения):

- (19) *Во мне пробудились и ум, и красноречие, и вкус — все, что составляет **прелесть женщины***. [П. Д. Боборыкин. Жертва вечерняя (1868)]
- (20) ***Прелесть женщины** — это мгновения жизни, вспышки, игра светотени*. [Комсомольская правда, 13.04.2002]

С этим контекстом метонимически связан другой, тоже некогда частотный, а теперь совершенно исчезнувший, и тоже с генитивной конструкцией: сейчас нельзя сказать *прелесть лица*, а до середины XIX именно так говорили в значении ‘красота / миловидность лица’ (21–23):

- (21) *Не смотря на **прелести лица** сей любезной женщины, она обладает всем тем, что может сделать беседу ея приятною*. [Н. Я. Озерецковский. Дневник (1782)]
- (22) ***Прелесть лица**, сладость речей очаровывали всех, кто ни встречал, ни слушал Веру*. [А. О. Корнилович. Андрей Безыменный (Старинная повесть) (1832)]
- (23) *Между сих разговоров София одевается в новое, еще в богатейшее первого платье, и новые **прелести лица** еще более увеличиваются от блеска камней и золота*. [Н. И. Зряхов. Битва русских с кабардинцами... (1840)]

Как видим, связь с носителем оценки нарушается не только семантически — вытеснением наиболее естественных в этой роли имен, но и синтаксически — сокращением генитивных конструкций. Последующее развитие оценочных употреблений связано с появлением в первой половине XIX в. конструкций *прелесть что за / какой X* (24–25), когда носитель оценки перемещен в другую клаузу:

- (24) *Да они ей-богу, **прелесть что** за люди!* [Д. Т. Ленский. Хороша и дурна, и глупа и умна (1833)]
- (25) *Прекраснѣйшее семейство... **Прелесть, что** такое!...* [«Современник», 1850]

В дальнейшем на их базе возникают уже синтаксически изолированные употребления, ср. (26–27):

- (26) *<...> но какой перевод! **прелесть!***. [С. П. Жихарев. Записки современника (1806–1809)]
- (27) *2 строфа — **прелесть!*** [А. С. Пушкин. Письмо П. А. Вяземскому (14.–15.08.1825)]

В принципе, с конца XIX века наблюдается постепенное снижение употреблений *прелесть*, ср. график по основному корпусу на рис. 1:



Рис. 1. График изменения частотности существительного *прелесть*, sg. (НКРЯ)

Новый импульс к семантическому развитию (хотя и не изменивший частотности слова) дал стандартный сдвиг с положительной оценки на отрицательную — как раз такой, о котором мы говорили в первом разделе, на базе иронии; но он произошел уже в XXI веке, ср. (28), пример из интернет-форума:

(28) *А Вася Пупкин из 4А уже курит. Какая прелесть!* [Наши дети: Подростки (2004)]

С тем же оценочным сдвигом *прелесть* употребляется и во множественном числе, но обычно в конструкции с родительным падежом существительного-носителя оценки — точно так же, как это было в начальных, калькированных контекстах с положительной оценкой: *(все) прелести X-а* ‘все отрицательные свойства X’. Корпусные контексты демонстрируют большое разнообразие контекстов: встречаются *прелести жизни / кризиса / отдыха / эвакуации / материнства / соседства / капитализма / войны / курения* и мн. др. Таким образом, в этой зоне узуса и генитивная конструкция, и множественное число снова возродились. Существенно, что эти контексты всегда отрицательны, независимо от встроенной оценки имени, ср.: *прелести кризиса* (*кризис* сам по себе несет отрицательную оценку) vs. *(все) прелести материнства* (исходно *материнство* содержит положительную оценку, но в данной конструкции оценка сочетания в целом только отрицательная: имеются в виду исключительно трудности и проблемы, с которыми сталкивается мать, пока ребенок растет).

4. Другая прелесть: *шарм*

В первом разделе мы говорили о факте калькирования с французского, которое привело к появлению дополнительного значения у слова *прелесть*. Калькирование произошло на пике французского влияния на русский (на рубеже XVIII–XIX вв.), и в то время оно было лишь одним из многих аналогичных эпизодов, свидетельствующих о силе этого влияния и его значимости для русского языка. Однако как известно, ко второй половине XIX века французское влияние идет на спад (из недавних исследований см., например, [Offord et al. 2015]) — и тем не менее, именно к этому периоду относятся первые засвидетельствованные НКРЯ случаи прямого заимствования из французского того самого слова, которое раньше послужило источником семантической кальки — существительного *шарм*. С ним французское *charme(s)* как бы входит в русский язык во второй раз, как новый лексический элемент, начиная несколько теснить свой прежний синоним *прелесть*. Первые примеры всегда особенны. Так, в идиолекте И. С. Тургенева, фактически билингва, русское *шарм* оказывается освоенным настолько, что не просто склоняется, а принимает форму партитива (29). В чуть более позднем примере из М. Е. Салтыкова-Щедрина (30), напротив, *charmes* заимствуется «вместе» с формой множественного числа (повторяя, как мы помним, путь слова *прелесть* на начальном этапе семантического калькирования):

(29) *Обаяния нет, шарму; не то что в нас с тобой.* [И. С. Тургенев. Накануне (1859)]

(30) *И действительно, как ни грустно, а приходится сознаться, что шармы телесные решительно подавляют и, вероятно, долго еще будут подавлять шармы умственные.* [М. Е. Салтыков-Щедрин. Письма о провинции (1868–1870)]

К XX веку *шарм* выглядит уже привычно для авторов и читателей, ср. (31) как один из многих примеров; ср. также график на рис. 2:

(31) *Сколько простоты, грации, уменья и шарма у этой удивительной актрисы.* [В. А. Теляковский. Дневники директора Императорских театров (1908)]

В то же время, до 1960-х годов это заимствование преимущественно встречается только в текстах представителей «старой» культуры (или у писателей-эмигрантов), например, у Андрея Белого, А. Куприна, А. Бенуа, Б. Зайцева, В. Набокова, Б. Поплавского, Г. Иванова, С. Эфрона, М. Цветаевой — то есть по большей части у тех, кто владел французским и жил во Франции. В Советской России в то время слово *шарм*, видимо, не было распространено. Исключением оказывается разве что М. Булгаков (32–33):

(32) *Посылая им улыбку за улыбкой, исполненные коварного шарма, Александр <...> указывал юнкерам на Бородинские полки.* [М. Булгаков. Белая гвардия (1923–1924)]



Рис. 2. График изменения частотности существительного *шарм* (НКРЯ)

- (33) — *Потрясающе!* — затрещал Коровьев, — все очарованы, влюблены, раздавлены, сколько такта, сколько умения, обаяния и *шарма!* [М. Булгаков. Мастер и Маргарита (1929–1940)]

Как видно из примеров НКРЯ, освоение нового слова происходило, как всегда, постепенно и началось уже в послевоенное время: сначала оно объяснялось и даже заключалось в кавычки (34), давались ссылки на французский источник (35) — то есть именно французы представлялись носителями «правильного» представления о шарме. И только в 1990-е, как видно из графика, начался реальный рост его частотности.

- (34) *Надо понять, что я говорю о красоте, не касаясь того, что называется в разных случаях очарованием, обаятельностью, «шармом», того, что может быть (и чаще бывает) сколько угодно у некрасивых.* [И. А. Ефремов. Лезвие бритвы (1959–1963)]
- (35) <...> *но был в нем тот, как говорят французы, шарм, что заставляет женское сердце биться более учащенно.* [Тайник // «Огонек». №№ 5–6, 8–13, 1970]

Самое замечательное, конечно, в том, что калька и заимствование, имеющие общий источник, существенно расходятся в значении: в русском замена *шарм* на *прелесть* практически никогда не дает хорошего результата — по крайней мере для одушевленных носителей оценки. Из примеров видно, что *прелесть* всегда создается, скорее, внешним впечатлением, то есть совершенной красотой — совсем не обязательной для *шарма* (и это скорее соответствует французскому прототипу,

значение которого заимствование тем самым передает точнее), ср. (36), в контексте которого разговор о внешней красоте неуместен, или (37) о кошке с нестандартной внешностью. При этом, в отличие от *шарма*, *прелесть* (будучи исходно метонимически связана с женскими прелестями — см. раздел 2) исходит от особы женского пола, и никогда — от мужчины (38):

- (36) *Стиль. Шарм. Но она — соучастница преступления.* [Виктория Токарева. Мой мастер (1964–1994)]
- (37) *Кошка имела дополнительные пальцы на лапах, ее язык всегда был высу- нут, из-за отсутствия зубов. Но именно этот необычный вид придавал ей шарм.* [Московский комсомолец, 03.12.2019]
- (38) *Попрощавшись с лишними килограммами, Морозов сохранил свои глав- ные достоинства — природное обаяние и шарм / *прелесть.* [Vesti.ru, 14.10.2020]

По-видимому, последнее обстоятельство дает возможность объяснить, почему парикмахерская может называться «Шарм» — но не «Прелесть»:

- (39) *Сегодня главные трудности уже позади и парикмахерская «Шарм» понемногу набирает обороты, завоевывая доверие клиентов.* [«Наша жизнь», 26.11.2014]

5. Заключение. Прелести прелести

Как видим, история слова *прелесть* и соответствующего концепта достаточ- но причудлива: на каждом этапе исследователя ожидают подводные камни разной сложности (можно ли говорить, что это *прелести исследования*?). Вместе с тем, она отлично вкладывается в более общие теоретические положения, которые касаются диахронической семантики и языковых изменений. Обозначим здесь только два.

- (1) Многозначность системна, она основана на моделях, которые воспроизво- дятся — ср. механизм метонимии в случае со значением, связанным с ча- стями тела или механизм иронии. Отступление от регулярной многознач- ности (в нашем случае, необычное направление при изменении оценки) может стать знаком интерференции: вмешательства другой системы.
- (2) Чужая лексическая система может оказаться так привлекательна, что язык может обратиться к ней повторно. Раньше для этого нужны были кальки: они позволяли легче инкорпорировать заимствованный смысл в другой язык как бы маскируя его под уже существовавший. Современный язык осваивает заимствования в разы быстрее, однако и освоение кальки, и освое- ние заимствования предполагают некоторый семантический процесс, в ре- зультате которого и калька, и заимствование удаляются от первоисточника, в той или иной степени теряя с ним связь.

И первое, и второе обстоятельство для русского языка пока недооценены и дают простор для исследователя — и в этом, конечно, есть своя научная прелесть.

Литература

- Арутюнова Н. Д. Предложение и его смысл: Логико-семантические проблемы. М.: Наука, 1976. 383 с.
- Арутюнова Н. Д. Типы языковых значений: Оценка, событие, факт. М.: Наука, 1988. 338 с.
- Арутюнова Н. Д. Язык и мир человека. Изд. 2, испр. М.: ЯРК, 1999. 896 с.
- Виноградов В. В. История слов. М.: ИРЯ РАН, 1999. 1200 с.
- Зализняк Анна А., Шмелёв А. Д. Исследования по русской и компаративной семантике. М.: ЯСК, 2021. 552 с.
- Мельничук В. А. Об одной особенности оценочной динамики церковнославянизмов в русском языке (на примере слова *прелесть*) // Филологические науки. Вопросы теории и практики. 2017. № 11–2 (77). С. 113–117.
- Перцов Н. В. О последнем сонете Пушкина // Логический анализ языка: Языки этики / ред. Н. Д. Арутюнова, Т. Е. Янко, Н. К. Рябцева. М.: ЯРК, 2000. С. 399–405.
- Сичинава Д. В. Панхронический корпус: интеграция исторических и современных корпусных ресурсов // Труды Института русского языка им. В. В. Виноградова. 2024. № 2 (40). С. 336–354.
- Шатуновский И. Б. Речевые действия и действия мысли в русском языке. М.: ЯСК, 2016. 480 с.
- Шмелёв А. Д. Плюрализм этических систем в свете языковых данных // Логический анализ языка: Языки этики / ред. Н. Д. Арутюнова, Т. Е. Янко, Н. К. Рябцева. М.: ЯРК, 2000. С. 380–389.
- Haiman John. Talk is cheap: sarcasm, alienation, and the evolution of language. Oxford: Oxford University Press. 1998. 220 p.
- Offord Derek, Larissa Ryazanova-Clarke, Vladislav Rjéoutski & Gesine Argent (eds.). French and Russian in imperial Russia. Edinburgh: Edinburgh U. Press, 2015. 288 p.

E. V. Rakhilina

National Research University "Higher School of Economics"

(Russia, Moscow)

rakhilina@gmail.com

V. A. Plungian

V. V. Vinogradov Russian Language Institute

Russian Academy of Sciences

(Russia, Moscow)

plungian@gmail.com

PRELEST': THE HISTORY OF THE WORD

Unforgettable Nina Davidovna — for the 100th anniversary

The puzzle of the Russian word *prelest'* is widely recognized in linguistic studies, pointing at the fact that it combines two seemingly opposite meanings: both 'temptation, enticement' and 'beauty, charm'. When discussing this peculiar polysemy (sometimes even treated as a kind of enantiosemy, which, in our opinion, is not entirely justified), linguists have mainly drawn attention to the coexistence of the two meanings in different spheres — "sacred" and "secular"; however, the exact conditions of this semantic change remained in the shade. The present paper provides the results of a corpus study (using advanced functionalities of panchronic search, currently available within Russian National Corpus) and highlights some additional aspects of the semantic history of this word. Of primary importance is here the influence of French, which contributed to the emergence of the "secular" meaning of *prelest'* / *prelesti* as a semantic calque from French *charme(s)*. Of interest is also the emergence and further development of a direct French borrowing — the word *šarm*, which is still actively used in modern Russian.

The article bears a dedication to the memory of Nina Davidovna Arutyunova, who was particularly fond of this word — as well as the complexity of natural language.

Keywords: Russian, French, semantic calques, polysemy, enantiosemy, corpus study

Acknowledgements

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2021S1A5A2A03065528).

References

Arutyunova N. D. *Predlozhenie i ego smysl: Logiko-semanticheskie problemy* [Sentence and its sense]. Moscow, Nauka Publ., 1976. 383 p. (In Russ.)

Arutyunova N. D. *Tipy yazykovykh znachenii: Otsenka, sobytie, fakt* [Types of linguistic meanings: Assessment, event, fact]. Moscow, Nauka Publ., 1988. 338 p. (In Russ.)

Arutyunova N. D. *Yazyk i mir cheloveka* [Language and human world]. 2nd ed. M.: YaRK, 1999. 896 p. (In Russ.)

Haiman, John. *Talk is cheap: sarcasm, alienation, and the evolution of language*. Oxford: Oxford University Press, 1998. 220 p.

Mel' nichuk V. A. [On one aspect of evaluative dynamics of Church Slavonic lexemes in Russian (case study of the word *prelest*)]. *Filologicheskie nauki. Voprosy teorii i praktiki*. 2017, no. 11–2 (77), p. 113–117. (In Russ.)

Offord, Derek, Larissa Ryazanova-Clarke, Vladislav Rjéoutski & Gesine Argent (eds.). *French and Russian in imperial Russia*. Edinburgh: Edinburgh U. Press, 2015. 288 p.

Pertsov N. V. [About Pushkin's last sonnet]. *Logicheskii analiz yazyka: Yazyki etiki* / N. D. Arutyunova, T. E. Yanko, N. K. Ryabtseva (eds.). M.: YaRK Publ., 2000, p. 399–405. (In Russ.)

Sichinava D. V. [Panchronic Corpus: integration of historical and contemporary corpus resources]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova*. 2024, no. 2 (40), p. 336–354. (In Russ.)

Shatunovskii I. B. *Rechevye deistviya i deistviya mysli v russkom yazyke*. [Speech acts and mental acts in Russian]. Moscow, YaSK Publ., 2016. 480 p. (In Russ.)

Shmelev A. D. Plyuralizm eticheskikh sistem v svete yazykovykh dannykh [Plurality of ethic systems in the light of linguistic data]. *Logicheskii analiz yazyka: Yazyki etiki* / N. D. Arutyunova, T. E. Yanko, N. K. Ryabtseva (eds.). Moscow, YaRK Publ., 2000, p. 380–389. (In Russ.)

Vinogradov V. V. *Istoriya slov* [The history of words]. Moscow, IRL RAS Publ., 1999. 1200 p. (In Russ.)

Zaliznyak Anna A.; Shmelev A. D. *Issledovaniya po russkoi i komparativnoi semantike* [Studies in Russian and comparative semantics]. Moscow, YaSK Publ., 2021. 552 p. (In Russ.)

И. И. Рыжаченков

Национальный исследовательский университет

«Высшая школа экономики»

(Россия, Москва)

ryzhachenkov@mail.ru

ТОЛК И ТОЛКИ: ДИАХРОНИЧЕСКИЙ ПОРТРЕТ ИМЕНИ С СЕМАНТИКОЙ МЕНТАЛЬНОЙ СФЕРЫ

В статье исследуется процесс диахронического развития полисемии русского *толк*, привлёкшего наше внимание своим нетривиальным распределением лексем внутри числовой парадигмы. Несмотря на её формальную двучленность, в современном русском языке *толк* и *толки* образуют два семантически несвязанных класса конструкций. Для выявления отношений семантической производности в структуре многозначности *толк* семантические деривации были рассмотрены в диахронической перспективе на материале корпусных данных, а также данных картотеки Института русского языка им. В. В. Виноградова РАН. Было показано, что исконно имя *толк* в русском языке не обладало речевой семантикой, а получило её лишь на рубеже XVIII–XIX вв. в результате участия мотивирующего предиката *толковать* в новом, несвойственном ему классе конструкций. При этом само развитие полисемии *толк* происходило за счёт метонимических переходов и затем лексикализации, позволившей этому имени закрепиться в узусе, несмотря на постепенное выветривание исконной лексической структуры. Были отмечены и некоторые семантические сдвиги в наиболее частотных конструкциях с лексемами *толк* и *толки*, которые повлияли на их синхронную сочетаемость; исследован процесс лексической конкуренции с ближайшими коррелятами — заимствованиями и кальками. Рассмотренный путь семантической эволюции слов с корнем *толк* позволяют говорить о явлении конструкционного «принуждения» (coercion) как об одном из паттернов диахронического развития русской абстрактной лексики.

Ключевые слова: диахронические исследования, корпусная лингвистика, конструкции, метонимические переносы, лексикализация

Ум, разум, разумение, смысл, рассудок, рассуждение, дарование, понятие, воображение, толк. Все сии названия, изображающие качества души, не имеют никакого определенного знаменования. Всякий произносит оные, как сам понимает. Один умом называет дарование; другой чрез дарование разумет понятие; иной смысл мешает с толком; иной толк именует разумением. Словом, надобно из содержания всей речи распознавать, в каком означении употреблено таковое название [Д. И. Фонвизин. Опыт российского ссоловника (1783–1784)]

1. Введение

Числовая семантика русских имён неоднократно становилась предметом исследования, но с акцентом на синхронную дистрибуцию внутри числовой парадигмы [Арбатский 1954; Вежбицкая 1988; Поливанова 1983; Мельчук 1985; Ляшевская 2004; Рахилина, Казкенова 2018]. Важно отметить, что нетривиальное соотношение форм $SG \Leftrightarrow PL$ не всегда укладывается в паттерны регулярной полисемии¹, нередко оно связано с процессами лексикализации или, например, семантического калькирования. И хотя с формальной точки зрения такие имена могут обладать стандартной двучленной оппозицией форм сингуляриса и плюралиса, их собственно семантическая числовая парадигма часто оказывается дефектна.

Примером такой семантически дефектной числовой парадигмы является структура многозначности уходящего из языка, но всё ещё довольно частотного имени *толк*: в единственном числе оно чаще всего трактуется как ‘смысл, разумное содержание чего-н.’, ‘разновидность какого-л. учения,’ ‘прок, польза’, тогда как множественное *толки* называет ‘разговоры, пересуды’ [Ожегов, Шведова 2006], [БТС], [МАС]². Сегодня две числовые формы этого имени настолько далеки семантически, что в некоторых словарях [см. Ожегов, Шведова 2006] даже имеют отдельные словарные статьи.

Заметим, что отношения семантической производности всех этих лексем не могут быть установлены с помощью синхронного семантического анализа, поскольку ни одна из них уже не является лексемой в том смысле этого слова, которое было принято МСШ. Дело в том, что в языке XXI века *толк* встречается только в устойчивых конструкциях, каждая из которых обладает своей уникальной семантикой, специфичным набором грамматических свойств и ограничений на сочетаемость³.

В одних случаях это имя входит в состав таких фразеологических сочетаний, целостный смысл которых не выводится с помощью традиционного лексического

¹ Перечень семантических сдвигов в этой области представлен в [Ляшевская 2004]. Исследование О. Н. Ляшевской показывает, что регулярная полисемия связана с меной таксономического класса имён, при этом семантические сдвиги чаще всего основываются на метонимии.

² В академических словарях выделяются и другие значения, но они содержат помету Устар. См. ‘толкование, объяснение’ в [МАС].

³ По крайней мере часть из них представлена на портале <https://constructicon.ruscorpora.ru/>

анализа лексем-компонентов (*сбить с толку NP-Acc, NP-Inst* ‘запутать, привести в замешательство’; *знать/понимать толк в NP-Loc* ‘разбираться, хорошо знать что-либо’; *добиться толку от NP-Gen* ‘узнать, выведать, получить желаемое’; *взять NP-Acc в толк* ‘уяснить, усвоить’ и т. п.).

В других — *толк* функционирует как семантически несвязанная единица, однако в отличие от абстрактных слов близкой семантики уже не обладает полной грамматической парадигмой форм. Другими словами, выражение того или иного значения *толк* требует выбора определенных предложно-падежных и числовых показателей и синтаксических позиций⁴. При этом определение точного значения *толк* в каждом конкретном случае является довольно трудной задачей для носителя новой нормы, поскольку исконная семантика этого имени «стёрлась» из узуса ввиду высокой степени конструкционализации. Во всех современных примерах со словом *толк* оно встраивается в общую семантику конструкции и потому за её пределами уже не осознаётся как самостоятельная лексическая единица.

В этой небольшой заметке мы покажем, почему две числовые формы одного и того же имени — *толк* и *толки* — в современном языке осознаются носителями как омонимы и не обнаруживают явной семантической связи.

Цель нашей работы — определить те паттерны семантической деривации, которые позволили двум числовым формам имени разойтись по своему лексическому значению и образовать два разных класса конструкций, отличающихся не только семантикой (ср. ‘мышление’ vs ‘речевой акт’), но и степенью лексикализации⁵.

Для этого мы проследим путь диахронического развития слова *толк* и связанных с ним однокоренных слов по следам дошедших до нас конструкций⁶: определим отношения семантической производности в структуре многозначности этого имени, выявим семантические сдвиги в некоторых конструкциях с лексемами *толк* и *толки*. Очевидно, что синхронная структура многозначности этого имени — это лишь «вершина айсберга», а его лексикализованные конструкции — «навигационные огни», указывающие на исторический путь семантической эволюции *толк*.

⁴ Так, например, в форме родительного падежа *толк* сочетается с прилагательными, называющими характер или свойство того или иного предмета или лица: *Моя настоящая фамилия звучит как псевдоним дурного толка*. [Рычкова// Труд-7, 03.10.2003]. Форма творительного падежа *толком* образует несколько адвербиальных конструкций: *с толком* ‘с пользой’ и беспредложное *толком* ‘как следует, полностью, хорошо’. Наконец, форма множественного числа *толки* маркирует значение ‘пересуды, обсуждения’ и обычно сочетается с глаголами движения и глаголами каузации. *Речь вызвала толки, и автору, кажется, пришлось выслушать немало неприятного в свой адрес* [Забытая публикация (2004) // «Журнал Московской патриархии», 27.09.2004]. *На кухне же... в других общественных местах... люди увидят, пойдут толки* [С. И. Шуляк. Квартира номер девять. Роман с чертовщиной // «Волга», 2013].

⁵ Если поставить задачу расположить их на графике «Центр-периферия», в зависимости от степени выраженности этого признака, ядерная часть была бы заполнена конструкциями с именем *толк* с семантикой ментальной сферы, тогда как выражения типа *пошли толки, ходили толки*, эксплицирующие речевое значение, оказались бы на периферии.

⁶ Примеры для настоящего исследования были отобраны из [НКРЯ] и картотеки Института русского языка им. В. В. Виноградова РАН.

Статья построена следующим образом: в *разделе 2* мы рассмотрим процесс развития структуры многозначности исконного имени *толк*, который происходил за счет метонимических переходов (2.1), а затем лексикализации (2.2). Объектом нашего внимания в *разделе 3* станет семантический дериват *толки*: мы покажем истоки его речевой семантики (3.1), обсудим основные семантические компоненты — дифференциальные признаки этого значения (3.2) и диахронические сдвиги в конструкции с этим словом (3.3). Итоги работы подводятся в *разделе 4*.

2. Толк

2.1. Метонимические переходы

Имя *толк* обязано своим происхождением общеславянскому **tǫlkъ* [Черных 1999: 248], с которым обнаруживают родство многие современные слова: *толкование* (слова, понятия, статьи), *толковый* (словарь, человек), *бестолковый* (ребенок, поступок, шаг), *истолковать* (значение слова, этимологию). Как отмечается в [Фасмер 1996], этот славянский корень считается родственным латинскому *loquor* ‘говорить’, хотя исконно имя *толк* не имело речевой семантики.

Согласно лексикографическим [СлРЯ XI–XVII], [Дьяченко], [Срезневский URL] и корпусным данным, в древнерусском языке имя *толк* могло использоваться в двух значениях: ‘переводчик’ (1) и ‘истолкование, т. е. интерпретация’ (2).

- (1) *Здѣсе намѣ билъ целомъ Олександрѣ Труфановъ сынъ, а жалуется на вашу братью на Иньцю Зашемъбяку и на его брата на Ортимью, на местерева толка*. [Грамота Великого Новгорода Риге (1418–1421)]
- (2) *Се же тѣлкѣють гланиѣ Иаковле иже гла Иудѣ рекъ. Иудо тебе похвалатъ братья твоѧ и прочаѧ всѧ рѣчи иже сѣтъ и переди писана ѿ починка книгъ. а сѣ тѣлкы пини рѣчьмъ тѣлкъ ты же жидовине слѣшаи разѣмно. <...> [Речи к жидовину (XI–XIII вв.)]*

Первое значение ‘переводчик’ было быстро утрачено вследствие лексической конкуренции *толк* с заимствованным именем *толмач* (3), проникшим в русский язык из тюркских языков еще в XIII в. (ср. кыпч. *tylmac* тур. *dilmac*)⁷.

- (3) *Ашика Болюа. и к томоу **толмача** розоумѣюща Роускыи **тазъ**. што иметь молвити Василко приѣхавъ подѣ городъ. <...> и нача молвити горожаномъ. а Татарове слышать. послани(и) с нимъ*. [Волынская летопись (1262–1292)]

Согласно корпусным данным, *толмач* было значительно более частотно⁸ и потому смогло быстро вытеснить это значение из структуры многозначности *толк*.

⁷ Наиболее авторитетные гипотезы появления этого слова в русском, представленные в [Фасмер 1996], убеждают нас в факте лексического заимствования через тюркское посредство.

⁸ Ср. 157 вхождений *толмач* и лишь 4 вхождения *толк* ‘переводчик’ в конструкции винительного одушевленного объекта в Панхроническом корпусе.

Толк в значении ‘истолкование, интерпретация’ было более употребительным, обладало большим деривационным потенциалом и имело несколько кластеров сочетаемости. Так, в этом значении *толк* могло называть не только интерпретацию смысла какого-либо источника, но и буквальный процесс перевода между разными языками (4):

- (4) <...> *вна же, слово ѿ Ардеמידора толкомъ рѣмьскимъ азыкомъ прїимши, гла к нему <...>* [Великие Минеи-Четии, собранные всероссийским митрополитом Макарием. <...> стлб. 581–1188. XVI в.]

Неслучайно однокоренной глагол *толковать* нередко обозначал процесс лингвистического переноса понятий из одного языка в другой с помощью лексем-эквивалентов. Ср. примеры с однокоренным *толковатися* в (5) и современные употребления пассивной формы *переводиться*:

- (5) <...> *Греческимъ языкомъ хранографъ, Русскимъ [же] языкомъ тлъкуется (= ‘переводится’) временникъ <...>* [Никоновская летопись (859–1176 гг.) (1526–1530)]

В памятниках XI–XVII *толк* чаще использовался в более широком контексте интерпретации смысла фрагментов Священного Писания и агиографической литературы. Разные толкования одних и тех же источников и разные взгляды на организацию религиозно-нравственной жизни порождали новые религиозные течения. Таким образом, к XVIII в. имя *толк*, исконно обозначавшее интерпретацию текста, распространяет идею интерпретации как системы взглядов:

- (6) *И такова ради их люторскаго лжехристианскаго жития, ты, сыне мой, велми их бойся, <...> понеже их люторские толки вси богопротивны суть и правому христианству весма вредителны.* [И. Т. Посошков. Завещание отеческое к сыну своему... (1718–1725)]
- (7) *Мы вас призываем к единой святой соборной и Апостольской церкви: а вы прельщаете простых людей итти к своим толкам. Но скажите нам, к которому захочете пристать нам толку? Всякой толк говорит, что в нем содержится церковь: и один толк другой проклинает. <...>* [архиепископ Платон (Левшин). Увещание к раскольникам (1766)]
- (8) *Из сего раздору разные произошли толки: как на пример, поповщина, безпоповщина, онуфриевщина, софонтиевщина, христовщина, кадильники, перекрещиванцы и прочия многочисленныя.* [архиепископ Платон (Левшин). Увещание к раскольникам (1766)]

Так, в [САР] под *толком* понимается учение, «отступающее от общих и принятых начал веры», а само слово характеризует в первую очередь новые религиозные течения XVI–XVII вв., отошедшие от единого церковного канона.

С точки зрения лексической семантики образование этого нового значения может объяснить метонимический сдвиг — смещение фокуса с объекта интерпретации (*толк*

Жития) на субъект интерпертации (люторский толк). На поверхностном уровне проявление нового значения характеризовалось развитием валентности Посессора, которая обычно заполнялась прилагательным (6).

Ещё один семантический дериват *толк*, мотивированный лексемой ‘истолкование, интерпретация’ — это значение ‘смысл’, развившееся вследствие результативной метонимии в XVII в.

- (9) *А в обедню выдут съ евангильям и станут против царьских дверей, а поют голомя не малое; а што поют, и мы тому толку не проразумели.* [Статейный список Федота Елчина (Грузия) (1639–1640)]
- (10) *<...> а если вооружатся на меня насмешники <...> они не умеют и вместо разума имеют этот дар от природы, чтоб и худое и доброе пересмехать, не видя в обоих толку, то я скажу им сон, который я в прошедшую ночь видел.* [М. Д. Чулков. Пересмешник, или Славенские сказки (1766–1768)]

В данном случае развитие производного значения было мотивировано метонимическим переходом Действие (= ‘интерпретация X-а’) > Результат действия (= ‘понимание смысла X-а’). При этом развитие нового значения сопровождалось изменением управления за счёт расширения списка зависимых предложно-падежных форм (9)–(10).

Таким образом, к началу XVIII столетия, ознаменовавшего конец старорусского периода развития русского языка, имя *толк* характеризовалось разветвлённой структурой многозначности за счёт приращения значений ‘учение, система взглядов’ и ‘смысл’, хотя более частотно по-прежнему было ‘толкование, интерпретация’ [СлРЯ XI–XVII]. Дальнейшая семантическая эволюция всех этих лексем в XVIII–XIX вв. была связана с процессами конструционализации и лексикализации, которые привели к постепенному «стиранию» их лексического значения.

2.2. История одной конструкции

В языке XVII–XVIII вв. семантический дериват ‘учение, система взглядов’ обладал полной парадигмой и, следовательно, мог использоваться композиционно в разном лексическом окружении. Как видно из примера (7), в это время он мог оказываться как в позиции подлежащего, так и выступать в роли косвенного дополнения.

В конце XVIII века лексема *толк* ‘учение, система взглядов’ впервые появляется в генитивной конструкции, в которой выполняет атрибутивную функцию, то есть характеризует свойство субъекта с точки зрения принадлежности к тому или иному религиозному учению:

- (11) *Шекъ или Шерифъ есть великій священноначальникъ Мекскій и въ семь достоинствъ признають его всѣ Магометане, какого бы толка ни были, и онъ получаетъ отъ всѣхъ ихъ властелиновъ великіе подарки <...>* [В. Я. Баранцов. <...> (1787)]

В XIX веке конструкция вида *NP 'animate' NP-Gen толка* оказывается необыкновенно продуктивна и составляет уже большую часть вхождений *толк* в этом значении. Со временем эта конструкция расширяет сферу употребления⁹ и начинает называть не только религиозные учения, но также и политические (12), философские (13), и даже течения изобразительного искусства (14).

- (12) *Другого рода вражда, и вражда больше основательная, была между французами двух революционных толков.* [А. И. Герцен. Былое и думы. Часть шестая. Англия (1864)]
- (13) *Зато, вооружившись благочестивыми дамами и особами славянофильского толка, она набросилась на Гоголя: «Смех, благородный смех!»* [И. Е. Репин. Далекое близкое (1912–1917)]
- (14) — *Видите ли, мы — художники разного толка,* — уклончиво ответил он. — *Я понимаю вас, но в искусстве все средства хороши для выражения идеи.* [А. М. Федоров. Природа (1911)]

Вместе с расширением контекстной сочетаемости происходит постепенная генерализация исконного значения 'система взглядов X-а' до более общего 'свойство X-а'. Если в XVII–XVIII вв. эта конструкция указывала только на определенные религиозно-мировоззренческие установки, то в XX в. она начинает характеризовать субъект с точки зрения его принадлежности к группе по более общим признакам (например, внешним, психоэмоциональным, поведенческим):

- (15) *Видя лиц разных толков в Москве, я убедился, что настроение остается очень неопределенным.* [В. А. Теляковский. Дневники Директора Императорских театров (1905)]
- (16) *Группа монгольских князей «истинно-китайского» толка послали в Пекин донесение о наблюдающемся <...>* [Россия и Китай (по телеграфу от наших корреспондентов) (14.04.1911) // «Русское слово», 1911]
- (17) *Уйма памятников одного толка — тяжеловесные всадники на довольно безынтересных конях.* [И. Д. Ковальченко. Дневник (1975)]

Финальная стадия этого семантического перехода была связана с окончательным «стиранием» исконной семантики *толк* и появлением в этой конструкции качественных, а не относительных прилагательных. Как видим, в примерах типа (18) это имя выражает идею «чистого» свойства и может быть заменено лексемами *характер, свойство, тон*, обычно используемыми в этой конструкции в наши дни.

⁹ Несмотря на расширение сферы употребления в XIX веке эта конструкция отражала следы исконной контекстной сочетаемости. В примере ниже А. И. Герцен, являющийся эмигрантом со старой языковой нормой, образно называет приверженность старому политическому складу ума «старым исповеданием». *Либералы старого толка, политические эконоы старого исповедания решили, не без внутреннего удовольствия, что задача невозможная <...>* [А. И. Герцен. Русские немцы и немецкие русские (1859)]

- (18) Правда, он стыдливо скрывает эту маленькую слабость, которую Краюхин как-то назвал «отрыжкой монтекрестовщины самого дурного толка». [Аркадий Стругацкий, Борис Стругацкий. Страна багровых туч (1952–1957)]

Другая производная лексема ‘смысл’ в языке XVIII в. застывает в дискурсивных конструкциях *без толку VP-inf*, *что толку/ какой толк/ да толку-то/ и толку* (от *NP-Gen*, в *NP-Loc*), в которых имя *толк* образует синонимичный ряд с лексемами *прок*, *польза*, *резон*. Лексема *толкование* оказывается ещё более продуктивной и образует целое семейство идиоматических предикатных сочетаний *сбить с толку NP-Acc*, *NP-Inst* ‘запутать, привести в замешательство’, *знать/понимать толк в NP-Loc* ‘разбираться, хорошо знать что-либо’, *добиться толку от NP-Gen* ‘узнать, выведать, получить желаемое’, *взять NP-Acc в толк* ‘уяснить, усвоить’, *VP* ‘начала и изменения состояния’ (*выйти, получиться, быть*) *толк из NP-Gen/в NP-Loc*. Как и в случае с рассмотренным выше семантическим дериватом ‘учение, система взглядов’, каждая из этих лексем сначала встраивается в конструкцию в определенной грамматической форме и затем постепенно утрачивает исконную семантику, расширяя сферу своего использования.

Именно встраивание *толк* в конструкции разной степени лексикализированности становится причиной его экспоненциального роста употреблений в середине XIX века.

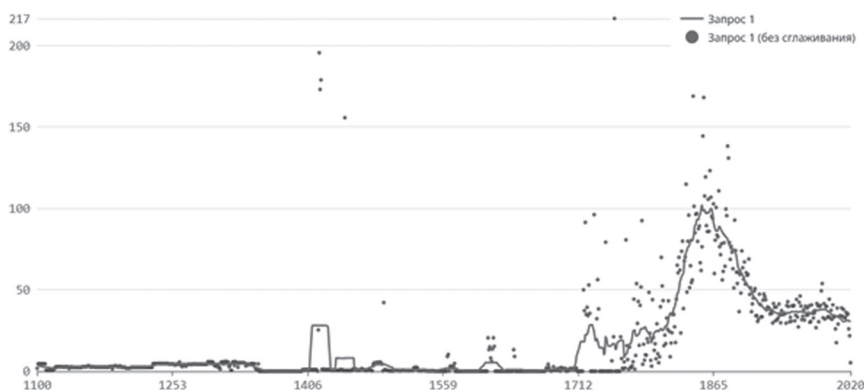


Рис. 1. Частотность лексемы *толк* в диахронии

Благодаря таким частотным конструкциям, как *без толку VP-inf* и *знать толк в NP-Loc*, это имя продолжает оставаться в узусе и в наши дни, хотя количество его употреблений уже не соответствует реальной сфере использования *толк*. Так, в газетном корпусе НКРЯ *толк* встретилось нам в 15 088 примерах. Для сравнения, частотность таких востребованных лексических единиц, как *эмоция* (25 690), *настроение* (43 961) и *мысль* (52 089), без которых невозможно представить повседневную речь, превосходит частотность имени *толк* лишь в несколько раз.

3. Толки

3.1. Истоки речевой семантики

В отличие от рассмотренных здесь семантических дериватов *толк*, лексема *толки* ‘разговоры, пересуды’, впервые встретившаяся нам в последней трети XVIII века (19), имела совершенно другую судьбу. История образования *толки* не была связана с метонимическими переходами значений, по-видимому, эта лексема была производна от самого предиката *толковать*.

(19) *Пошли в городе разные толки: одни говорили, что Фламель нашел клад; другие думали, что он знает тайну философского камня и делает золото <...>* [Н. М. Карамзин. Письма русского путешественника (1793)]

В языке старорусского периода этот глагол мог выражать как более общее значение ‘интерпретации смысла чего-либо’, так и буквальное лингвистическое ‘переводить слова между разными языками’, хотя вторая интерпретация *толковать* была впоследствии вытеснена латинской калькой *переводить*¹⁰ (см. *Но и паче же многи труды и подвиги подъях от исправления иностранных и древних пословиць, преводя на русскую рѣчь <...>* [Из Великих Миней Четых митрополита Макария. Сентябрь. Летописец (1530–1554)]).

В эту эпоху *толковать* обычно управлял беспредложным винительным падежом (*‘X толкует Y’*) или встречался в пассивной конструкции (*‘Y толкуется X-ом’*). Однако уже в XVI–XVII вв. происходит расширение модели управления этого предиката, и *толковать* постепенно встраивается в продуктивную конструкцию с аргументом в предложном падеже *VP ‘mental sphere’ o NP-Loc*, характерную для глаголов мышления и речи:

¹⁰ Традиционная версия происхождения этой лексемы, представленная в [Фасмер 1996] выглядит так: «перевод переводить. Калька из франц. traduction, traduire — то же от лат. trādūcere; см. Сандфельд, Festschrift V. Thomsen 171». В соответствии с этой гипотезой, traduire была заимствована французским из латыни (ср. *traduco* ‘вести через’; *trāns* + *dūcō*), а затем перенесена в русский из французского как семантическая калька.

По-нашему мнению, такая версия нуждается в некотором уточнении: очевидно, французскому влиянию XVIII в. предшествовало калькирование *trādūcere* непосредственно из латыни: См. *И преводят книги всѣх учителей нашихъ, елико их обрели, от еллинские бесѣды на римскую по чину и разуму грамотическому, не отмѣняюще ни малѣйше. И, преложивше на языкъ свой, дают в друкъ и размножают много, и посылают, продавающе их дешкою ценою, не точию въ Италию <...>* [Андрей Курбский. Предисловие к Новому Маргариту (1564–1583)].

Другой пример иллюстрирует тезис о том, что во второй половине XVII века языковые контакты с носителями французского ещё не были широко распространены, тогда как знание латыни было обязательно не только для переводчика, но и для большинства церковных и светских чинов этой эпохи. *Того ж числа стояли в городе Анбусе. <...> А как посланники с буржумистры говорили, и в то время речь переводил перевотчик Иван Госенц по-латине. И буржумистры посланником говорили: «Перевотчик ваш по-французски говорить умеет ли?». Стольник Петр и дьяк Семен им говорили: «Перевотчик умеет по-латине в достаток, и по-цесарски, и по-галански, и по-итальянски, а по-французски перевестъ совершенно не умеет, потому что во Французской земле по се время он не бывал».* [Статейный список П. И. Потемкина (Франция) (1668)]

- (20) <...> Саул творил, оставя пророков Божиих, ко матроне или ко фотунисе, женѣ чаровнице, притекаль, пытающе ея о сраженію настоящим, яже ему похотѣния его мечтанием бесовским Самуила пророка, аки воставаго от мертвых, в мечте показала, яко о том толкует свѣтле святыи Августын во своих книгах. [Андрей Курбский. Третье послание Курбского Ивану Грозному (1577–1583)]

Ср. также употребления других предикатов близкой семантики в этой конструкции:

- (21) Митрополитъ же **учивъ** князя о любви не токмо съ братьею, но и съ всѣми православными <...> [Московский летописный свод (1479–1492)].

Такое изменение управления *толковать* сопровождалось актантной метонимией, предполагающей смещение фокуса с Источника на Тему толкуемого источника как в (22) и (23): *толковать законы о наказаниях* => *толковать* ‘объяснять’ о наказаниях => *толковать* ‘говорить’ о наказаниях¹¹.

- (22) Четвертое следствие: судьи, судящие о преступлениях потому только, что они не законодавцы, не могут иметь права **толковать** законы о наказаниях. [Екатерина II. Наказ Комиссии о составлении проекта нового Уложения (1767)]
- (23) Таким образом продолжал он **толковать** о необходимости наказаний, потом перешел к толкам о наказаниях на том свете, о внутреннем человеке и тому подобных предметах. [Н. И. Греч. Черная женщина (1834)]

В XIX веке *толковать*, подобно многим другим предикатам речи (*говорить*, *молвить*, *судачить*), встраивается в конструкцию взаимного адресата с творительным падежом (24)–(25), которая использовалась для передачи значения совместного действия и предполагала заполнение объектной валентности актантом, обладающим агентивными свойствами¹² (ср. метафорическое прочтение пациентного аргумента в *Х говорил с океаном*).

- (24) <...> все усилия свои устремляли они сперва на то, чтоб дочка их была здорова, румяна, толста, весела <...> — потом, чуть стала она на ноги,

¹¹ См. также более ранний пример: Я слышал, как рассуждают в свете о начальствах разного рода, и готов необузданной молве всегда противоположить опыт самых событий. Но до-вольно **толковать** о сем. [И. М. Долгоруков. Повесть о рождении моем, происхождении и всей моей жизни <...> / Часть 4 / 1799–1806 (1799–1822)].

¹² В некотором смысле конструкцию такого вида можно считать реципрокальной: каждый участник называемого речевого действия является и говорящим, и слушающим одновременно. Как видится, одним из ограничений этой конструкции является использование транзитивных глаголов в конверсивном контексте. Так, глаголы типа *объяснять* и *учить*, оказываясь в этой конструкции, теряли своё конверсивное прочтение, предполагающее взаимно-разнонаправленное, а не совместное действие (ср. *Профессор читает лекцию аспиранту*. Vs *Профессор читает лекцию с аспирантом*).

начали рядить ее во всякие материи, надевать на нее бриллианты и жемчуги, возить по гостям, готовить приданое и **толковать** с нею об ее красоте и женихах. [М. П. Погодин. Суженый // «Московский вестник», 1828]

- (25) <...> Маша еще **толковала** с княжною о том, как барон часто выезжает со двора, как в это время новоприезжий сидит с баронессою, как они уговариваются ехать вместе в театр, быть вместе на бале, и проч., и проч. [В. Ф. Одоевский. Княжна Мими (1834)]

Таким образом, под давлением самой семантики этой конструкции *толковать* развивает собственно речевое значение ‘рассуждать, говорить’ (25). Подобный вид семантической деривации обычно связывают с принципом соетсiоп, описанным в [Michaelis 2004: 25]. В соответствии с этим принципом использование слова в несвойственном ему классе конструкций «вынуждает» его изменить свое лексическое значение.

Заметим, что, встраиваясь в новую конструкцию, *толковать* не просто изменяет объём своего значения за счёт добавления или вычёркивания тех или иных семантических компонентов (как в случае с уже рассмотренным здесь сдвигом в конструкции *NP ‘animate’ NP-Gen толка*), но и принимает на себя то значение, которое эксплицировалось в конструкции раньше. При этом, в соответствии с формулировкой этого принципа, такой переход становится возможен только благодаря тому, что исконное *толковать* содержало семантический компонент ‘продуцировать речь’.

В первой трети XIX века речевое значение становится доминирующим в структуре многозначности *толковать* и вытесняет лексику ‘объяснять, интерпретировать’ в рассмотренной здесь конструкции с предложным ‘темы’:

- (26) Здесь начинается ярмарка: все в движении, все гуляет и **толкует** о торговле по-прежнему. [А. И. Тургенев. Хроника русского (1825–1826)]

Неслучайно в речи современных носителей «объяснительное» значение проявляется лишь в сочетании с прямым дополнением (*толковать слово*), тогда как рассмотренные здесь конструкции предложного и творительного эксплицируют речевое значение предиката.

По-видимому, изменение управления *толковать* (встраивание в продуктивную конструкцию взаимного адресата) и утрата тех контекстов, в которых это слово называло буквальный перевод между разными языками, привели к расщеплению значения ‘интерпретировать, объяснять’ на два семантически и стилистически дифференцированных кластера сочетаемости. Так, исконное значение этого предиката ‘объяснять, интерпретировать’ терминологизируется и вместе с именем *толкование* становится атрибутом научного стиля речи. Развившееся на рубеже XVIII–XIX вв. речевое значение *толковать*, напротив, начинает характеризовать стилистически противоположный пласт разговорной лексики и постепенно переходит в периферийную зону русского просторечия.

3.2. Глагольное «наследство»: семантические компоненты

В прошлом разделе было показано, что имя *толки* — это номинализация новой речевой лексемы глагола *толковать*, которая была образована в результате перестройки его модели управления. На исконно глагольную природу деривата *толки* указывают и семантические компоненты: ‘повторяемость’, ‘неконтролируемость’, ‘интенсивность’, характеризующие процессуальный признак. Во-первых, сама форма мн. ч. *толки* выступала маркером семантики повторяемости действия (27). Ср. современные: *переговоры, дискуссии, консультации*, которые тоже сочетаются с фазовыми глаголами, глаголами движения и глаголами каузации.

- (27) *Он уже третий раз пролетает по Лене в качестве курьера, и каждый раз толки о нем долго не смолкают на пустынной реке.* [В. Г. Короленко. Ат-Даван (1892)]

Во-вторых, не менее значимым признаком того, что речевая лексема *толки* была производна от глагола, а не от имени, является семантика ‘неконтролируемости’, обычно проявляемая в примерах типа (28). По всей видимости, в данном случае неконтролируемость действия коррелировала с его массовостью — большое количество участников не позволяло Х-у контролировать процесс. Подобно пожару, *толки* могли вспыхивать и снова угасать, преодолевать обширные расстояния и наносить непоправимый урон тем, к кому они были обращены:

- (28) *Мрачные вспыхнули толки, Слез и проклятий не счесть;* [На льдине (13.03.1912) // «Русские ведомости», 1912]

В-третьих, к собственно глагольным признакам деривата *толки* можно отнести семантику интенсивности, которая указывала на градуальный характер этого процесса и проявлялась в сочетаемости с интенсификаторами:

- (29) *Везде сильные толки о «Телеграфе». Одни горько сетуют, «что единственный хороший журнал у нас уже не существует».* [А. В. Никитенко. Дневник (1834)]
 (30) *«Съ каждыѣмъ днемъ толки о бракъ усиливались и принимали почти зловерное направлѣніе <...>».* [А. В. Дружинин. Заметки петербургского туриста (1856)]

3.3. Семантический сдвиг в конструкции с глаголами движения

В первой трети XIX века *толки* встраивается в конструкцию *VP* ‘движения’ *NP-Nom* ‘речи’¹³. В одних случаях это имя проявляло идею нейтрально оценочного речевого значения предиката *толковать* и было близко лексемам *разговоры,*

¹³ Стоит отметить, что эта конструкция имеет две основных разновидности: первая и более частотная — с предикатом ‘однонаправленного движения пешком’ *пойти*, которая встроена в более общий класс конструкций с инхоативной семантикой ‘начала серии событий’. См. *пошли расспросы, жалобы, угрозы, намёки* и т. д.; вторая — с предикатом ‘разнонаправленного движения пешком’ *ходить*, указывающим на неоднократность действия.

беседы, рассказы. Ср. примеры (31) и (32), в которых *толки* и *беседы* используются взаимозаменяемо, присоединяя одну и ту же идиоматическую конструкцию.

- (31) *Я ему обрадовался от всей души. Пошли беседы о том, о сем. Дошло дело до Берлина и до Гегеля.* [Н. В. Станкевич. Письма Станкевичам (1838)]
- (32) *Она въехала к нему прямо в дом. <...> Вот, разумеется, при свидании родные обрадовались. Пошли толки о том о сем, о старине, о новизне, об урожае <...>* [В. Ф. Одоевский. История о петухе, кошке и лягушке (1834)]

В других случаях *толки* образовывало синонимический ряд с лексемами *пересуды, слухи, сплетни*, которые обычно имели негативную коннотацию и были ассоциированы с неконтролируемым стихийным распространением недостоверной информации (чаще всего пейоративного свойства). Ср. употребления *слухи* (33) и *толки* (34):

- (33) *По Дагестану ходили слухи, что он скитался между чеченцами и кой-субулинцами, утратив красоту и здоровье и даже самую отвагу; но кто же мог сказать про то утвердительно?* [А. А. Бестужев-Марлинский. Аммалат-бек (1831)]
- (34) *Об этой отпускной ходят разные толки: иные уверяют, что Клим Сидорович, увидев Блистовского, так испугался, что немедленно подписал отпускную; другие же утверждают, что Владимир за выкуп цыгана заплатил значительную сумму.* [Антоний Погорельский. Монастырка (1833)]

Как видим, в (31)–(32) речевое действие, названное конструкцией *VP* ‘движения’ *NP*-*Not* ‘речи’, сосредоточено вокруг конкретного дейктического центра и происходит в общем поле зрения локуторов, информация о которых содержится в самом контексте¹⁴. Напротив, в примерах (33)–(34) указание на конкретно-референтных участников отсутствует, при этом само действие не ограничивается временными и пространственными рамками единичного события. В данном случае распространение информации происходит не мгновенно и требует определенного времени и сменяемости участников.

Заметим, что со временем эта конструкция постепенно сужает область своего применения и употребления с лексемой *толки*, в которых последняя эксплицирует оценочно нейтральное речевое значение, исчезают из узуса.

- (35) *Пошли толки, как им будет найти заветное место? Думали, думали и положили, чтобы Кирджали сам их повел.* [А. С. Пушкин. Кирджали (1834)].

В XX веке частотность конструкции падает и *толки* вытесняется более употребительными именами *разговоры* и *слухи*.

¹⁴ См. индексальные *ему, она, к нему* в (31)–(32), указывающие на конкретно-референтного собеседника, и эгоцентризм *я*, отсылающий к фигуре самого говорящего-нарратора.



Рис. 2. Частотность лексем *толк* (1), *разговор* (2), *слух* (3)

Таким образом, в примерах типа (35), когда речь идет об обсуждении ситуации, произошедшей у говорящих на глазах, употребление конструкции уже невозможно, сегодня единственное прочтение *пошли толки* имеет негативно оценочное значение ‘стали распространяться слухи’.

4. Заключение

Как мы могли убедиться, числовые формы *толк* и *толки* имели совершенно разную судьбу, разное происхождение, и потому сегодня они отсылают к двум разным этапам семантической эволюции предиката.

Имя *толк* (‘переводчик’, ‘истолкование, интерпретация’) встречается уже в древнерусских памятниках письменности, в начальный период формирования русского языка активно развивает производные значения (‘учение’, ‘смысл’), а к концу XVIII века «застывает» в ряде идиоматических и дискурсивных конструкций.

Отглагольный дериват *толки* ‘разговоры, пересуды’ возникает на рубеже XVIII–XIX вв., быстро встраивается в частотную конструкцию с глаголами движения и уже к XX веку переходит в периферийную зону. Таким образом, эта лексема появляется к тому моменту, когда парадигма (не только числовая, но и предложно-падежная) исконного *толк* распадается, что создаёт не только формальную, но и лексическую противопоставленность двух числовых форм *толк* и *толки*.

Если развитие полисемии исконного *толк* проходило по универсальному пути именной метонимии и затем с этой готовой семантикой оно встраивалось в конструкции, то само появление лексемы *толки* ‘разговоры, пересуды’ иллюстрирует менее конвенциональный путь семантической деривации, связанный с явлением коэрсии (coersion). Имя *толки* является номинализацией новой лексемы *толковать* ‘рассуждать, говорить’, которая была образована вследствие участия этого предиката в новом классе конструкций.

Заметим, что в наши дни ни одна из лексем *толк* не является в полном смысле семантически свободной, поскольку это слово не используется за пределами его наиболее частотных конструкций. И хотя с формальной точки зрения числовая

парадигма этого имени всё ещё имеет две числовые формы сингуляриса и плюралиса, и это отражается в синхронных словарях, говорить о непосредственной семантической связи этих форм как принадлежащих одной лексеме, по-видимому, неправомерно.

Литература

Арбатский Д. И. Значения форм множественного числа имен существительных в современном русском литературном языке. Дис. ... канд. филол. наук. М., 1954

БТС — Большой толковый словарь русского языка / Гл. ред. С. А. Кузнецов. СПб., 2003.

Ляшевская О. Н. Семантика русского числа. М.: Языки славянской культуры, 2004. 390 с.

МАС — Словарь русского языка в 4-х томах: Описание ЭНИ // Фундаментальная электронная библиотека «Русская литература и фольклор» (ФЭБ). М.: 2005. [Электронный ресурс]. URL: <http://feb-web.ru/feb/mas/mas-abc/0encyc.htm>

Мельчук И. А. Поверхностный синтаксис русских числовых выражений. Вена: Венский славистический альманах, 1985. 510 с.

НКРЯ — Национальный корпус русского языка. [Электронный ресурс]. URL: <https://ruscorpora.ru/>

Ожегов, Шведова 2006 — Ожегов С. И., Шведова Н. Ю. Толковый словарь русского языка. Изд. 4-е, доп. М.: А ТЕМП, 2006. 938 с.

Поливанова А. К. Выбор числовых форм существительных в русском языке // Проблемы структурной лингвистики 1981 / отв. ред. В. П. Григорьев. М.: Наука, 1983. С. 130–145.

Полный церковнославянский словарь: (со внесением в него важнейших древнерус. слов и выражений): пособие / сост. свящ. Григорий Дьяченко. [Репр. воспр. изд. 1900 г.]. М.: Отчий дом, 2004. 1120 с.

Рахилина Е. В., Казкенова А. К. Заметки о русском числе // Вестник Российского университета дружбы народов. Серия: Лингвистика = Russian Journal of Linguistics. 2018. Т. 22. № 3. С. 605–627.

САР — Словарь Академии Российской. Санкт-Петербург: Императорская Академия наук, 1789–1794.

СЛРЯ XI–XVII — Словарь русского языка XI–XVII вв. Вып. 22. (Раскидатися-Рященко) / Ин-т рус. яз. им. В. В. Виноградова. М.: Наука, 1997. 298 с.

Срезневский И. И. Материалы для словаря древне-русского языка по письменным памятникам / Труд И. И. Срезневского. [Электронный ресурс]. URL: <http://oldrusdict.ru/dict.html>

Фасмер М. Этимологический словарь. [Электронный ресурс]. URL: <https://lexicography.online/etymology/vasmer/>

Черных П. Я. Историко-этимологический словарь современного русского языка: В 2 т. 3-е изд., стереотип. М.: Русский язык, 1999. Т. 2: Панцирь — Яшур. — 560 с.

Michaelis Laura A. Type shifting in Construction Grammar: An integrated approach to aspectual coercion // Cognitive Linguistics. 2004. Vol. 15–1. P. 1–67.

I. I. Ryzhachenkov

National Research University "Higher School of Economics"

(Russia, Moscow)

ryzhachenkov@mail.ru

TOLK AND TOLKI: DIACHRONIC PORTRAIT OF A NOUN WITH THE SEMANTICS OF THE MENTAL SPHERE

The article investigates the process of diachronic evolution of polysemy in Russian *tolk*, which is interesting by its non-trivial distribution of lexemes within the numeric paradigm. Despite the paradigm's formal completeness, in modern Russian, *tolk* and *tolki* belong to two semantically unrelated classes of constructions. In order to identify the relations of semantic derivation in the structure of polysemy of *tolk*, semantic derivations were examined from a diachronic perspective. Semantic shifts in the meaning of *tolk* and related constructions were traced using corpus data, as well as data from the card catalogs of The V. V. Vinogradov Russian Language Institute of the Russian Academy of Sciences. It was shown that the original *tolk* did not possess the semantics of speech, but received it only at the turn of the XVIII–XIX centuries as a result of the constructionalization of the predicate *tolkovat'*. At the same time, the very development of the polysemy of *tolk* was due to metonymic transitions and then lexicalization, which allowed this noun to remain in the *usus* despite the gradual weathering of the original lexical structure.

Keywords: diachronic studies, corpus linguistics, constructions, metonymic transfers, lexicalization

References

Arbatsky D. I. *Znacheniya form mnozhestvennogo chisla imen sushchestvitel'nykh v sovremennom russkom literaturnom yazyke* [The meanings of the plural forms of nouns in the modern Russian literary language]. Dissertation ... Cand. philol. nauk. Moscow, 1954 (In Russ.)

BTS — *Bol'shoi tolkovyi slovar' russkogo yazyka* [Big explanatory dictionary of the Russian language] / ed. by S. A. Kuznetsov. St. Petersburg, 2003. (In Russ.)

Chernykh P. Ya. [Historical and etymological dictionary of the modern Russian language: In 2 v.]. 3rd ed., stereotype. Moscow: Russkii yazyk Publ., 1999. Vol. 2. 560 p. (In Russ.)

Fasmer M. *Etimologicheskii slovar'* [Etymological dictionary]. Available at: <https://lexicography.online/etymology/vasmer/> (In Russ.)

Polnyi tserkovnoslavyanskii slovar': (so vneseniem v nego vazhneishikh drevnerus. slov i vyrazhenii): posobie [Full Church Slavonic Dictionary: (with the inclusion of the most important Old Russian words and expressions): manual / compiled by St. Gregory Dyachenko. Reprint of the 1900 edition]. Moscow, Father's House Publ., 2004. 1120 p. (In Russ.)

Lyashevskaya O. N. *Semantika russkogo chisla* [Semantics of Russian number]. Moscow, Yazyki slavyanskoi kul'tury Publ., 2004. 390 p. (In Russ.)

Melchuk I. A. *Poverkhnostnyi sintaksis russkikh chislovykh vyrazhenii* [Superficial syntax of Russian numerical expressions]. Vienna: Vienna Slavic Almanac Publ., 1985. 510 p. (In Russ.)

Michaelis, Laura A. Type shifting in Construction Grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics*, 2004, v. 15–1, pp. 1–67.

Natsionalnyi korpus russkogo yazyka. [The Russian National Corpus]. Available at: <https://ruscorpora.ru/>

Ozhegov S. I. Shvedova N. Y. *Tolkovyj slovar russkogo yazyka* [Explanatory Dictionary of the Russian Language]. Ed. 4-th revised edition. Moscow, 2006. 938 p. (In Russ.)

Polivanova A. K. *Vybor chislovykh form sushchestvitel'nykh v russkom yazyke* [Choice of numerical forms of nouns in Russian]. *Problems of Structural Linguistics 1981* / ed. by V. P. Grigoriev. Moscow, Nauka Publ., 1983, pp. 130–145. (In Russ.)

Rakhilina E. V., Kazkenova A. K. [Notes on Russian number]. *Bulletin of Peoples' Friendship University of Russia. Series: Linguistics = Russian Journal of Linguistics*. 2018. V. 22, no. 3, pp. 605–627. (In Russ.)

Slovar Akademii Rossiiskoi [Dictionary of the Russian Academy]. St. Petersburg: Under the Imperial Academy of Sciences, 1789–1794. (In Russ.)

Slovar russkogo yazyka v 4 tomah [Dictionary of Russian Language in 4 volumes]: Description of ENI. Fundamental electronic library “Russian Literature and Folklore” (FEB). Available at: <http://feb-web.ru/feb/mas/mas-abc/0encyc.htm> (In Russ.)

Slovar russkogo yazyka XI–XVII vv. [Dictionary of the Russian language of 11th–17th cc.] Ext. 22. (Raskidatisya–Ryashchenko) / V. V. Vinogradov Russian Language Institute. Moscow: Nauka Publ., 1997. 298 p. (In Russ.)

Sreznevsky I. I. *Materialy dlya slovary drevne-russkogo yazyka po pis'mennym pamyatnikam* [Materials for the dictionary of the Old Russian language on written monuments] / Labor I. I. Sreznevsky. Available at: <http://oldrusdict.ru/dict.html> (In Russ.)

Г. И. Кустова

Институт русского языка им. В. В. Виноградова РАН

(Россия, Москва)

galinak03@gmail.com

О НЕКОТОРЫХ СВОЙСТВАХ НЕОПРЕДЕЛЕННО-КОЛИЧЕСТВЕННЫХ ВЫРАЖЕНИЙ С МЕСТОИМЕННЫМ КОМПОНЕНТОМ

В работе обсуждаются некоторые грамматические, семантические и коммуникативные особенности неопределенно-количественных местоимений *мало кто*, *мало какой (чей)*, *много кто*, *много какой (чей)*, которые входят в обширный класс производных неопределенных местоимений (*неизвестно кто*; *кто угодно* и под.). Конструкции *мало / много кто* сравниваются с неопределенно-количественными выражениями *мало / много X-ов*. Показано, что эти выражения не полностью синонимичны и потому во многих случаях невзаимозаменяемы: *В этот поздний час в вагоне было мало людей* VS. *В этот поздний час в вагоне мало кто был*.

Внутри своей группы неопределенно-количественные местоимения также несимметричны. Статистика НКРЯ показывает, что выражений с компонентами *мало / редко (кто / какой / чей)* на порядок больше, чем с компонентом *много*. У этих выражений есть и грамматические различия (которые связаны с семантическими и коммуникативными различиями), например в выборе форм числа (для *много* какие ед. число невозможно): *Редко какая книга может вызвать такое стопроцентное отторжение* (возможный вариант: *Редко какие книги ...*) VS. *Я много в каких точках побывал* (**много в какой точке*).

Ключевые слова: неопределенные местоимения, неопределенно-количественные выражения, референция

1. Местоименные конструкции *мало / редко / много кто* среди других неопределенных местоимений

В работе будут рассмотрены некоторые свойства производных неопределенно-количественных местоимений *много / мало кто* (*что, какой, чей*). Эти выражения образуют 3 серии: *мало кто / что / какой / чей / где / куда...*; *много кто / что...*, *редко кто / что...*. Для краткости будем условно обозначать их *мало / много кто* (серия *редко*___ по употреблению похожа на *мало*___ но имеет специфические особенности, рассмотрение которых увело бы нас в сторону от основной темы. Примеры с *редко*___ будут рассматриваться, в основном, в разделе 4).

Местоимения *мало / много кто* относятся к группе производных неопределенных местоимений, ср. *неизвестно кто, черт знает кто, абы кто, кто угодно, что хочешь, что попало, где придется* и др., которая весьма многочисленна и пополняется новыми единицами. Производным неопределенным местоимениям посвящена большая литература, см., в частности, [Апресян, Иомдин 1989 / 2010], [Ермакова 1996], [Иомдин 2010], [Кустова 2016; 2017], [Падучева 2011; 2015а; 2015b], [Тестелец, Былинина 2005], [Фисун 2013; 2016], [Bylinina, Testelecs 2004], [Iomdin 2005]; о местоимениях ср. также [Булыгина, Шмелев 1997], [Татевосов 2002; 2004], [Шведова 1998], [Шмелев 2002], [Alexiadou et al. 2000], [Haspelmath 1997].

Группа *мало / много кто* занимает здесь особое место и даже не всегда включается в списки производных местоимений. Например, в работе [Былинина, Тестелец 2005], посвященной амальгамам и квазирелятивам в системе русских местоимений, группа *мало / много кто* не рассматривается.

Существуют разные взгляды на статус местоименных выражений *мало / много кто*. Поскольку они имеют гибридную природу и могут рассматриваться по-разному, мы будем называть их и местоимениями, и местоименными конструкциями, сокращенно — МК.

А. Никунласси [2008] считает эти выражения местоимениями, а элементы *мало* и *много* квалифицирует как частицы и сравнивает их с частицами *ни* и *кое-ни у кого* — *кое у кого* — *мало у кого* (в своей работе он рассматривает местоимения *мало кто / что, много кто / что* не сами по себе, а как опорные компоненты приместоименных придаточных, ср.: *Прежде были довоенный «Неизданный Хлебников», «имковский» «Неизданный Гумилев» и еще много кто, кого издавали после кончины и с большими трудностями; Даже за плату и в качестве эксперимента вы мало кого найдете, кто согласится прокатиться в метро в марлевой повязке...*).

Л. Л. Иомдин рассматривает обороты *мало / много / редко кто* как конструкции с двумя синтаксическими вершинами [Иомдин 2010: 155], т. е. *мало / много / редко* считаются самостоятельными синтаксическими элементами.

Оба автора рассматривают обороты только с субстантивными местоимениями: *мало / много кто / что* [Никунласси 2008]; *мало / много / редко кто* [Иомдин 2010]. В [Булыгина, Шмелев 1997], а позднее в [Никунласси 2008] упоминается также семантическое и коммуникативное сходство оборотов *мало какие / мало кто* с *немногие*, а *мало что* — с *немногое*.

Мы не будем рассматривать вопрос о синтаксической структуре и частеречном статусе конструкций *мало / много кто*. Отметим только, что у них действительно есть сходство с местоимениями, которые «разрываются» предлогом (*ни у кого — кое у кого — мало у кого*). При этом их можно рассматривать как не вполне сформировавшиеся местоимения, которые имеют промежуточный статус — находятся в переходной зоне между местоимениями и конструкциями. И одним из аргументов в пользу неполной сформированности является вариативность местоименного компонента в так называемых структурных падежах (ср. [Лютикова 2017: 48–75]),

т. е. в позиции подлежащего и прямого дополнения. Здесь, наряду с «правильными» конструкциями *мало (много) кто / что*, употребляются «неправильные» конструкции *мало (много) кого / чего* (этой проблеме посвящена отдельная работа [Кустова 2024]), ср.:

Но еще много что может случиться [Н. М. Амосов. Дневник (1985)].

За войну много чего случилось — все не упомнишь [Виктор Астафьев. Обертон (1995–1996)].

Я еще мало что понимала во всем происходившем вокруг, но чувствовала, что в жизнь людей вошел какой-то страх [И. К. Архипова. Музыка жизни (1996)].

А мы мало чего знаем [Наши дети (форум) (2005)].

Таким образом, конструкции *мало / много кто* занимают особое место среди других производных местоимений и отличаются от них прежде всего тем, что наряду со значением неопределенности содержат компонент количества, т. е. в семантической структуре эти МК имеют два значимых компонента, ср. с другими неопределенными местоимениями:

Об этом кое-кто знал — ‘были люди, которые об этом знали’ (бытийный компонент);

Об этом мало кто знал — ‘были люди, которые об этом знали; таких людей было немного’ (бытийный компонент + количественный компонент).

Корпусное исследование показывает, что внутри группы *мало / редко / много кто* также имеются существенные различия, которые определяют неоднородность, «несимметричность» свойств и элементов этой группы. Например, различается история формирования и функционирования серии *мало / редко кто* и серии *много кто*.

Серии *мало / редко кто* полностью освоены языком уже в XIX в. и широко представлены в текстах XIX–XX вв. Серия *много кто* сформировалась позже, в текстах XIX в. встречаются лишь единичные вхождения *много кто*, в XX в. количество *много кто* по сравнению с *мало / редко кто* также несопоставимо мало, ср. таблицу для формы Им. п. *мало кто* VS. *много кто*:

	<i>Мало кто</i>	<i>Много кто</i>
XVIII–XIX вв.	≈ 200	9
XX в.	≈ 28 500	15
XXI в.	≈ 1 500	41

В работе [Апресян, Иомдин 1989 / 2010], где рассматривались конструкции *негде спать, некого спросить*, эти конструкции сопоставлялись с конструкцией *мало кто*, поскольку они имеют свойство «разрываться» предлогом: *Не у кого спросить / Не с кем поговорить; Мало у кого спрашивал / Мало с кем говорит*.

При этом *много кто* не упоминается — видимо потому, что в 1980-е годы, когда создавалась эта работа, конструкция *много кто* ощущалась как слишком разговорная, находящаяся за пределами литературного стандарта. В более поздних работах [Июмдин 2010]; [Никунласси 2008] *много кто / что...* упоминается наряду с *мало кто / что*. Л. Л. Июмдин [2010] упоминает эти конструкции в контексте оборотов типа *неизвестно кто*.

В XXI в. происходит взрывной рост употреблений серии *много кто* в устной речи (в том числе, например, в видеоблогах). Однако в НКРЯ это отражается не в полной мере, т. к. корпус устной речи существенно уступает по объему корпусу письменных текстов.

2. Конструкции *мало / много кто* VS. *мало / много X-ов*

Мы будем рассматривать некоторые грамматические, семантические и коммуникативные свойства конструкций *мало / много кто* на фоне других количественных выражений, в первую очередь — *мало / много X-ов*.

Хотя у этих групп количественных выражений есть общие компоненты, у них есть и определенная специализация, поэтому часто они не синонимичны и не взаимозаменяемы.

Прежде всего, у *мало / много кто* и *мало / много X-ов* есть очевидные грамматические различия.

Значимой грамматической чертой выражений *мало / много X-ов* является, как известно, дефектная парадигма: они употребляются только в так называемых структурных падежах — позициях подлежащего (Им.) и прямого дополнения (Вин.). МК *мало / много кто* имеют полную парадигму.

ЗАМЕЧАНИЕ. Интересно, что если обратиться к статистике НКРЯ, то обнаружится, что и выражения *мало / много кто* с полной парадигмой употребляются преимущественно в структурных падежах — именительном и винительном. Тем не менее они могут использоваться в косвенных падежах, и это выглядит как компенсация для парадигматически дефектных *мало / много X-ов*.

Наряду с грамматическими различиями у конструкций *мало / много кто* и *мало / много X-ов* есть также семантические и коммуникативные различия.

Выражения *мало / много X-ов* тяготеют к экзистенциальным и посессивным конструкциям, обозначают оценку множества в каком-либо фрагменте пространства или личной сфере человека и часто не нуждаются ни в каких специальных глаголах (используется глагол *быть* в бытийном и посессивном значении): *В вагоне мало пассажиров; У него много книг*. Впрочем, выражение *мало X-ов* часто используется не для количественной оценки множества, а для выражения более сложной семантики.

Слова *мало* и *редко* рематичны и являются главной ремой или дополнительной ремой (дополнительным утверждением, ср. [Булыгина, Шмелев 1997]); еще одна важная особенность высказываний со словами *мало* (*У нас мало времени*), *редко* (*Он редко заходит*) и оборотами *мало кто / что*, как отмечалось в [Булыгина,

Шмелев 1997], [Иомдин 2010], состоит в том, что в них совмещаются два семантических компонента — бытийный ('существуют X-ы, которые Р') и количественный ('таких X-ов мало'). При этом количественный компонент имеет не только вариант 'количество' (*На дереве еще мало листьев*), но и вариант 'нехватка, недостаточность', ср.:

Мы не можем пойти в кино, у нас мало денег = 'деньги есть, но их мало: не хватает на Р', смысл 'деньги существуют, имеются' — в пресуппозиции, ср. [Булыгина, Шмелев 1997, Шмелев 2002: 89–96]).

Этим определяется несимметричность выражений с компонентами *много* и *мало*, в частности *много X-ов* VS. *мало X-ов*.

Много X-ов может употребляться просто как констатация: *В книге много картинок*.

Выражение *В книге мало картинок* связано с какими-то подразумеваемыми смыслами: она не очень подходит для детей, ее не стоит покупать и т. п. Разумеется, *В книге много картинок* тоже можно рассматривать как аргумент в пользу покупки, но это выражение самодостаточно в том смысле, что для такой констатации достаточно просто пролистать книгу. *В книге мало картинок* — нечто большее, чем просто констатация.

Хотя здесь *мало* не означает 'не хватает', как в случаях типа *Мы не можем это купить, у нас мало денег*; *Мы не можем пойти в кино, у нас мало времени*, тем не менее это некоторый недостаток (как в количественном, так и — часто — в оценочном смысле, т. е. скрытая отрицательная характеристика).

Собственно смысл 'небольшое количество', без дополнительных семантических (и оценочных) характеристик, обычно выражается местоимениями *немного* и *несколько*:

**Дай мне мало соли* → *Дай мне немного соли* [Шмелев 2002: 90].

В зале было мало зрителей → *несколько зрителей*.

Даже если допустимы и *мало*, и *несколько* (*мало зрителей* VS. *несколько зрителей*), это не эквивалентная замена. *Мало* содержит дополнительную информацию, в данном случае *мало зрителей* ≈ 'меньше, чем обычно; меньше, чем ожидалось; меньше, чем могло бы быть' и т. п. (ср. противопоставление: *В зале было мало зрителей. Большинство гуляло в фойе*).

Таким образом, *мало* практически никогда не сводится к смыслу 'небольшое количество' и всегда предполагает что-то еще (*мало* либо требует объяснения, либо (в случае 'не хватает') само является объяснением того, почему X не может сделать Р). Это не просто количественная оценка, это характеристика.

Хотя *мало* / *много X-ов* и *мало* / *много кто* обозначают количество и содержат один и тот же компонент *мало* (*много*), они часто невзаимозаменяемы и неэквивалентны. И контексты, где недопустимы МК (где они не могут заменить оборот *мало* / *много X-ов*), показывают семантическую специфику МК. Как мы увидим из дальнейших примеров, МК *мало* / *много кто* обычно не используются просто для

оценки количества, а содержат более сложную информацию и имеют специальные условия употребления.

Начнем с непосредственно наблюдаемых ситуаций.

Выражения *мало* / *много* *Х-ов* могут относиться к непосредственно наблюдаемой ситуации, когда внутренний наблюдатель оценивает количество *Х-ов* в некотором пространстве:

- (1) *Он зашел в местный магазинчик. Там было несколько покупателей.*
- (2) *В этот ранний час в магазине было мало покупателей / мало людей.*

Это экзистенциальные высказывания с актуальным наблюдателем.

Замена (2) на выражение:

- (3) *Мало кто был в магазине в этот ранний час — В магазине мало кто был* — невозможна, а само это выражение в данной ситуации неуместно.
- (4) *В этот поздний час в парке было мало гуляющих / Несмотря на поздний час, в парке было много гуляющих* — это оценка актуального наблюдателя, который оценивает количество гуляющих в момент наблюдения.
- (5) *В этот поздний час в парке мало кто гулял / Несмотря на поздний час, в парке много кто гулял* — это высказывание тоже относится к конкретному моменту времени, но не может быть описанием наблюдаемой картины и оценкой наблюдаемого количества ‘в парке было мало / много гуляющих’.

Оборот *мало кто* уместен в случае повторяющихся или постоянных (узуальных) ситуаций, а соответствующие высказывания содержат определенное обобщение:

- (6) *В такой поздний час мало кто гуляет в парке* (= ‘опасно, неприятно’).
- (7) *Парк находится далеко от жилого квартала, здесь мало кто гуляет.*

МК *мало кто* обычно может служить характеристикой: в данном случае это характеристика парка, причем обусловленная — временем, расположением.

Много кто также не употребляется для характеристики наблюдаемого количества:

- (8) *Поскольку сегодня выходной, много кто гуляет в парке → в парке много гуляющих.*

В описании неактуальных ситуаций МК *мало кто* (*мало какой*) также неуместны, если они просто обозначают количество. Странно сказать:

- (9) *Мало кто живет в этой деревне.*
- (10) *Мало какие крестьяне живут в этой деревне: люди предпочитают переезжать в город.*

в значении ‘в этой деревне живет мало людей / мало крестьян’. При другой интерпретации — ‘мало кто хочет жить в этой деревне’ или ‘мало кто остался в этой

деревне, большинство жителей уехало’, т. е. если есть дополнительные смыслы или если ситуация встроена в более обширный сценарий (см. ниже), — предложение допустимо.

Как видно из примеров, важное отличие *мало / много X-ов* и *мало / много кто* состоит в том, что *мало / много кто* не используются просто для оценки наблюдаемого количества или оценки количества *X-ов* в изолированной ситуации, а содержат некоторую семантическую перспективу, встраивают ситуацию в различные сценарии и имплицативные цепочки. Эти нетривиальные семантические эффекты наблюдаются и в высказываниях с глаголами СВ, которые описывают конкретные единичные ситуации (ср. *Мало кто пришел на праздник; Много кто поздравил именинника*), как мы увидим в следующем разделе.

3. Конструкции *мало кто* VS. *много кто*

Итак, мы рассмотрели некоторые различия оборотов *мало / много кто* и *мало / много X-ов*. Теперь рассмотрим различия внутри группы МК, а именно между конструкциями с компонентом *мало кто* и *много кто*.

При описании местоименных выражений неизбежно встает вопрос о квантификации и референциальном статусе соответствующих именных групп (ИГ). Тип референции в той или иной степени коррелирует с акциональным типом предиката. Высказывание с референтным субъектом может сообщать как конкретный факт (*Петя вернулся домой*), так и обобщение, свойство: *Петя отличается вспыльчивостью*; высказывание с нереферентным субъектом всегда содержит обобщение: *Холерики часто отличаются вспыльчивостью*.

Местоименные выражения могут иметь и референтный статус (ср.: *Некоторые гости опоздали*), и нереферентный статус (ср.: *Некоторые певцы не знают нот*), но, как мы увидим дальше, у них могут быть некоторые «референциальные предпочтения».

Корпусный анализ нашего материала показывает, что МК *мало* ___ тяготеют к нереферентному употреблению, а МК *много* ___ — к референтному.

Как уже отмечалось в предыдущем разделе, для *мало кто* характерно употребление в высказываниях о повторяющейся ситуации:

- (11) *В 1994-м Арафат вместе с Рабином стал лауреатом Нобелевской премии мира. С тех пор мало кто решается говорить о нем плохо* [«Коммерсантъ-Власть», 1999]

— такие предложения выражают связь между свойством и количеством участников повторяющейся ситуации: ситуация редкая (ср.: *...редко кто решается говорить о нем плохо*), участников мало;

или в высказываниях о постоянной ситуации:

- (12) *Магазин стоит на отшибе, в него мало кто заходит.*

‘Малолюдность’ (*мало кто заходит* → ‘мало покупателей’) — это не просто количественная оценка, это свойство магазина, которое обусловлено его расположением,

т. е. это импликация, которая выводится из местоположения (*Магазин стоит на отшибе*), и не надо пересчитывать покупателей, чтобы сделать такое утверждение, — достаточно общих знаний о мире.

(13) ?*Сегодня в магазин мало кто зашел.*

Для естественного употребления такого высказывания нужны специальные условия. Например, говорящий имеет в виду знакомых людей, которых он может перечислить (тогда предпочтительно общефактическое: *мало кто заходил*). Или наблюдатель периодически (или регулярно) стоит у дверей магазина и оценивает «поток»: сколько прохожих, идущих мимо магазина, заходят в него.

Высказывания с МК *много кто* имеют другие условия употребления, для них недостаточно имплицативной операции — они предполагают ту или иную степень знакомства с ситуацией и с множеством, которое обозначается выражением *много кто*:

(14) *Магазин стоит в хорошем месте, в него много кто заходит*

— из того, что магазин находится в хорошем месте, еще не следует, что там много покупателей.

Много кто заходит предполагает все-таки, что говорящий когда-то наблюдал это своими глазами, что он оценивает нечто реальное, знакомое, а не умозрительное, что говорящий имеет в виду каких-то конкретных людей либо он может перечислить и охарактеризовать, например, группы покупателей: строители с соседней стройки, школьники после школы, приезжающие на вокзал и т. п.

Вообще, это типичное различие: *мало кто* может быть импликацией, *много кто* предполагает реальное знание о количестве:

(15) *Мало кто посмотрел этот фильм, т.к. он не шел широким прокатом*

— ‘мало кто’ следует из того, что фильм не шел широким экраном.

Но если фильм шел широким экраном, то из этого еще не вытекает, что его посмотрело много зрителей. Таким образом, *много кто* требует реального знания не только объема множества, но и элементов множества:

(16) *Много кто посмотрел этот фильм*

— говорящий может указать, «предъявить» людей, которые посмотрели фильм; он говорит о каких-то известных ему людях.

В высказываниях с *много кто* часто речь идет о личном знании, и говорящий может перечислить если не всех, кто входит в множество, то многих; *много кто* используется для описания множества, с которым говорящий знаком, поэтому *много кто* — это часто реальная оценка реального множества. Ср. примеры из корпуса:

(17) *И брат покойной Линды пришел, и Ринго Старр, и много кто еще* [«Домовой», 04.12.2002]

— говорящий лично присутствовал и может перечислить тех, кто пришел.

- (18) *Сейчас в мире к фильму невероятный интерес. Его **очень много кто** хочет показать — нам каждый день звонят с предложениями из Франции, Италии, Германии, Америки, Испании* [Олег Кармунин. Алексей Герман: «Премьера «Трудно быть богом» состоится осенью» // Известия, 2013.03]

— говорящий имеет в виду конкретных людей и их конкретные предложения.

- (19) *Там вообще **много кто** учился, от Бориса Гребенщикова до Григория Перельмана* [«Новгородские ведомости», 2013]

— говорящий знает тех, кто учился, и оценивает их как ‘много’.

- (20) *Санитарки всегда требуются! У нас в больнице вечно санитарок не хватало. Да **много кто** требуется...* [В. Ю. Кунгурцева. Похождения Вани Житного, или Волшебный мел (2007)]

— говорящий знает, кто требуется, и может перечислить вакансии.

Даже если автор делает достаточно общее суждение, ср.:

- (21) ***Много кто** сначала боролся с Христом до последних сил, а потом уже, без сил, принимал Его* [«Дальний Восток», 2019]

— предполагается знание состава множества: автор высказывания изучал историю и может привести много примеров.

Таким образом, именно «знакомство с материалом» дает основание сделать вывод *много кто*.

Разумеется, это не абсолютное правило, а просто тенденция. Есть обобщенные высказывания с нереферентным *много кто* — например, в контексте модального предиката:

- (22) *Положим, пусть не в такой степени, но лохануться **много кто** из нас способен* [«Новгородские ведомости», 2013].

- (23) *То, что я делаю, **может много кто** выполнить — бесплатно или нет* [«Кот Шрёдингера», 2016].

И наоборот, встречается референтное *мало кто* — например, в контексте эпизодического предиката (глагола СВ):

- (24) ***Мало кто** вернулся из двухсот — одни расстреляны, другие накрылись деревянбушлатом, погибли от дистрофии, расстреляны при лагерных чистках* [Василий Гроссман. Все течет (1955–1963) // «Октябрь», 1989].

Впрочем, даже с глаголом СВ высказывание может иметь обобщенный смысл:

- (25) *Истина вообще кажется банальной. Бог есть Любовь. Что может быть банальней? Но эту банальность до сих пор **мало кто** понял и **про-чувствовал*** [Аркадий Мацанов. Горечь полыни на губах // «Ковчег», 2013].

Что касается будущего СВ (особенно с модальным оттенком) и узуального НСВ, то они обозначают обобщенные ситуации.

Итак, анализ корпусного материала показывает, что в высказываниях с МК *мало кто* часто содержится не только обобщение, но и обусловленность малого количества и оценка:

(26) *Мало кто* пойдет на этот фильм

— это умозаключение, основанное на качестве или других признаках фильма — ‘на такие фильмы обычно приходит мало зрителей: плохой, сложный’ (разумеется, свой семантический вклад в такую интерпретацию вносит и форма будущего времени предиката).

Типичное высказывание — с нереферентными участниками. Потому что для того, чтобы оценить количество, не надо знать участников, надо понимать статус ситуации — она редкая или нет: ситуация редкая, значит, участников мало (импликация).

Не случайно наряду с *мало* в подобных случаях употребляется *редко*. Именно на этой корреляции (‘редко’ → ‘мало’) основана синонимия *мало кто* и *редко кто*:

(27) *Мало кто* знает 3 языка / *Редко кто* знает 3 языка

— участников этой ситуации не надо знать и не надо пересчитывать, чтобы охарактеризовать их количество.

Обычно отмечается, что *редко* связано с временем, периодичностью (ср., например, [Иомдин 2010: 154], а в данном случае ситуация статическая. Периодичность, множественность здесь индуцируется за счет воображаемого субъекта оценки, который время от времени сталкивается с данной ситуацией: ‘редко встречается такая ситуация → редко встретишь такого человека, потому что мало таких людей’.

Заметим, что МК *часто кто* в русском языке (пока) не употребляется (к вопросу о несимметричности *мало* VS. *много*; *редко* VS. *часто*).

Итак, парадокс в том, что высказывания, включающие *много кто*, предполагают, что говорящий обладает реальными знаниями о множестве, обычно он может перечислить его элементы — и только в этом случае он может сделать такое утверждение, а высказывания, включающие *мало кто*, обычно предполагают лишь общие знания о мире и умение устанавливать имплицативные связи.

Разумеется, это не абсолютная корреляция, а просто тенденция, но тенденция, предопределяющая «несимметричность» МК *мало кто* и *много кто*.

4. Конструкции *мало* / *редко какой (чей)* VS. *много какой (чей)*

Аналогичное несимметричное соотношение обнаруживается между *мало* / *редко какой* и *много какой*. Эта несимметричность усугубляется еще грамматическим

отличием субстантивных и адъективных местоимений. Субстантивные местоимения *кто* и *что*, как известно, имеют фиксированное единственное число (синтаксическое), а адъективные местоимения *какой* и *чей* изменяются по числам, т. е. теоретически можно ожидать употребление форм *много / мало какой (чей)* и *много / мало какие (чи)*. Однако в употреблении числовых форм в адъективных МК обнаруживаются определенные ограничения.

Конструкции *мало / редко какой S* входят в ряд *никакой S — мало какой S — всякий / любой S*: Естественно, что за такие деньги **никакой гастарбайтер не согласится работать** [Алексей Востряков (13vainamoinen). Записи в Живом Журнале // 2013] (с формальной точки зрения сюда должна входить и МК *много какой*, но она, как и рассмотренная выше конструкция *много кто*, устроена «несимметрично»).

Редко / мало какой допускает два типа высказываний — и утвердительные, и отрицательные:

- (28) Конечно, **редко какое** жюри **удовлетворяет** ожиданиям публики и прессы, как правило, чаще бывает наоборот [«Искусство кино», 2003.06.30].
- (29) **Редко какая** мелочная бакалейная лавочка **не торгует** политурой [«Южный край» (Харьков), 1914].
- (30) **Мало какому** поколению **достается** делать революцию [Владимир Мау. В России произошла нормальная революция. Она завершилась (2001)].
- (31) Конечно, оно было не одно, такое исследование, хотя по масштабам **мало какое могло** с ним сравниться [«Знание — сила», 2003].
- (32) **Мало какой** фабрикант **не желает** гордиться звучным именем своего предприятия: им можно прикрыться и как торговой маркой, и как полно-весным брендом, и, в конце концов, как знаменем [«Мебельный бизнес», 15.10.2003]

— и обе формы числа — единственное и множественное:

- (33) **Редко какая** книга **может** вызвать такое стопроцентное отторжение, как первое издание книги «Иду в путь мой» [Л. А. Данилкин. Человек с яйцом. Жизнь и мнения Александра Проханова (2007)].
- (34) К тому же **редко какие** парни в школе одеваются со вкусом [Форум: Вы за или против школьной формы? (2007)].
- (35) **Мало какой** город мира **может** похвастаться таким количеством успешно проведенных соревнований и числом спортивных объектов [«Русский репортер», № 6 (36), 2008].
- (36) И был он столь прекрасен, что **мало какие** города **могли** соперничать с ним [Юлия Капишникова. Бретань // «Пятое измерение», 2002].

Важно также отметить, что значительное количество примеров включает модальный оператор, ср. (32), (33), (35), (36), либо допускает перифразирование с модальным словом (см. ниже).

ЗАМЕЧАНИЕ 1. При замене *мало / редко какой* на синонимичное местоимение *немногий* возможна только форма множ. числа: **Немногий город может похвастаться...* → *Немногие города...* См. ниже о *много какие*.

ЗАМЕЧАНИЕ 2. В конструкции встречаются неисчисляемые существительные, хотя такие примеры единичны:

Свои восторги от этих книг я убрал в отдельный пост, чтоб не загромождали колонку. Но «Автобиографию» Надежды Соханской (1825–1884) непременно надо обсудить здесь. Мало какая литература сильнее брала меня за жабры [«Сноб», 2018]

Кроме того, одна сцена гл. 2 вообще имеет мало какое отношение к литературе! [Алексей Шепелев. Кгыышфт Вшытундф-ТВ (2008) // «Волга», 2011]

В силу ограниченности объема статьи мы не рассматриваем контексты с неисчисляемыми существительными.

Начнем с оборота *редко какой S* и сопоставим его с прилагательным *редкий*, которое имеет в своей семантической парадигме кванторное значение и наглядно демонстрирует специфику данной серии кванторных выражений.

В словарях основное значение прилагательного *редкий*, которое можно условно назвать «периодическим», толкуется как ‘Бывающий, происходящий, повторяющийся через большие промежутки времени’ [МАС], т. е. редко, что соотносится с соответствующим значением *частый* (‘Происходящий, повторяющийся через короткие промежутки времени’): *частые удары, выстрелы, письма* — *редкие удары, выстрелы, письма* (‘часто / редко происходят, приходят’). В этом значении допустимо только множественное число, ср. **редкий / частый выстрел* (неисчисляемые существительные, не имеющие форм мн. числа, ср. *редкий / частый пульс*, употребляются, конечно, в ед. числе, но они подразумевают повторение какого-то события: *частый пульс* = ‘частые удары’). Исключения составляют личные существительные, у которых в семантике есть идея периодического повторения действия: *редкий / частый гость, посетитель* и под. (‘редко / часто приходит в гости, посещает’). Они употребляются именно в единственном числе (мы сейчас не можем отвлекаться на обсуждение того, что поскольку *редкие выстрелы* — это множество, а *редкий гость* — один человек, то, несмотря на общий семантический компонент, было бы желательно разграничить эти значения). В словаре [Кузнецов 2008] есть примеры в ед. числе типа *После боя в редкой роте не было потерь* (однако его толкование ‘Принадлежащий к числу немногих, далеко не всякий, далеко не каждый’ не кажется удачным; впрочем, нас интересуют в первую очередь грамматические свойства ИГ *редкий S*). Будем называть значение *редкий*, для которого ед. число является не только правильным, но и основным, кванторным. Кванторное значение *редкий* (в отличие от периодического) допускает единственное число с событийными существительными:

(37) *Редкое выступление этого политика обходится без скандала*

— ‘в множестве выступлений нескандальных выступлений мало’.

В таких предложениях единственное число является не просто допустимым, но и предпочтительным, т.к. у высказывания с множественным числом ИГ более вероятно другое прочтение:

(37а) *Редкие выступления этого политика обходятся без скандала*

— это предложение с большей вероятностью означает, что другие политики буйные, а этот политик выступает редко и ведет себя тихо.

Конечно, при ударе на РЕДКИЕ (т. е. при другой реме) будет выражаться значение ‘большинство выступлений скандальные’. Но для нас важно подчеркнуть, что в единственном числе (*Редкое выступление этого политика обходится без скандала*) значение ‘это тихий, нескандальный политик и он выступает редко’ вообще невозможно. Единственная интерпретация предложения *Редкое выступление этого политика обходится без скандала* — ‘большинство выступлений — скандальные, и лишь малая часть — нескандальные’. Такие высказывания содержат не один семантический компонент, как *редкие выстрелы* (‘мало выстрелов’) или *редкий гость* (‘редко приходит’), а два: ‘много выступлений со скандалом’ и ‘мало выступлений без скандала’ (ср. пример из НКРЯ: *Редкий парень уходит на военную службу, не обменявшись честным словом или не обручившись формально с какою-либо девушкой* [А. В. Амфитеатров. Марья Лусьева за границей (1911)] = ‘у большинства дома осталась девушка, мало кто не обручился перед военной службой’). Т. е. *Редкое выступление обходится без скандала* имплицитно ‘политик выступает МНОГО’. При этом есть ли в (37) импликация ‘[поскольку нескандальных выступлений мало, то] они бывают редко’, — неочевидно. Разумеется, само кванторное значение возникло на базе соотношения ‘редко’ → ‘мало’. Но главная идея говорящего в высказываниях типа (37) состоит не в том, что у политика редко бывают нескандальные выступления, а в том, что выступления данного политика, в основном, скандальные.

Аналогичное значение реализуется в обороте *редко какой*:

(38) *И все вы по вашему возрасту хорошо помните прежние до автомобильные города — города для людей, лошадей, собак; сады чуть не при каждом доме, и редко какой дом выше двух этажей — приятнейшая высота человеческого жилища* [А. И. Солженицын. Письмо вождям Советского Союза (1974)]

— ‘большинство домов — не выше двух этажей’.

Для таких высказываний можно допустить наличие двух равноценных компонентов: ‘большинство домов — невысокие; высоких домов [выше двух этажей] мало’.

В некоторых высказываниях второй компонент присутствует эксплицитно:

(39) *Только редко какую ракушку с жемчугом достанешь, пустых многонимаем* [Н. П. Колпакова. Терский берег (1936)]

— ‘преобладают ракушки без жемчуга, ракушек с жемчугом — мало’.

Однако при наличии отрицания (прямого или «непрямого») компонент ‘большинство, почти все’ выдвигается на первый план:

- (40) *От аристократов эта привычка постепенно «спустилась» к менее богатым, а там и к совсем бедным, вплоть до самых низов общества, и к XVI веку **редко какой** англичанин **не имел** собственной фамилии* [Владимир Смолицкий. Приключения фамилий // «Знание—сила», 2009]

— ‘большинство англичан имело собственную фамилию’.

- (41) ***Редко какой** праздник **обходился без** [‘не имел’] **целого апофеоза из ракет, римских свечей, бураков, шутих...*** [Ю. К. Олеша. Книга прощания (1930–1959)]

— ‘обычно на праздниках были ракеты, свечи и т. д.’

Попутно заметим, что в силу такого акцентирования компонента ‘большинство Х-ов — Р’ идея редкости событий отходит на задний план или даже исчезает (в [Иомдин 2010: 154] отмечается, что «фраземы, образованные наречием *редко*, сочетают идею малочисленности некоторых объектов с идеей временной редкости», или «редкости ситуаций», из-за чего «неуместно сказать что-л. вроде **редко кто сидел в этом зале*». Для статической ситуации типа (39) ‘большинство домов не выше двух этажей’ идея редкости событий, по-видимому, не релевантна; на первый план выступает идея малочисленности объектов с заданным признаком).

Аналогичным образом устроены обороты с *мало какой / чей*.

Интересно, что утвердительные высказывания не симметричны отрицательным. Утвердительные высказывания (не любые, но многие) часто предполагают сравнение группы объектов Х с объектом Y и имплицитно сообщают, что Y — особенный, уникальный. При этом объект Y может быть эксплицитно обозначен в высказывании:

- (42) *Падуб обыкновенный — вечнозеленый кустарник или, реже, небольшое дерево с пирамидальной вытянутой формой кроны. **Мало какие растения** могут сравниться с ним по теневыносливости* [«Сад своими руками», 15.11.2002].
- (43) *Давным-давно на берегу моря у самой воды стоял город Ис. И был он столь прекрасен, что **мало какие города** могли соперничать с ним* [Юлия Капишникова. Бретань // «Пятое измерение», 2002].

Но даже при отсутствии Y-а в предложении он может подразумеваться (и обнаруживается в контексте):

- (44) — *Из современных русских композиторов **мало чей юбилей** отмечается на Западе так громко — Год Щедрина* [Комсомольская правда, 2002]

— ‘так громко, как юбилей Щедрина’.

Типичное отрицательное высказывание содержит сообщение о группе объектов *X* и на первый план выдвигается смысл ‘большинство *X*-ов — *P*’:

- (45) *Сейчас мало какое ведомство не обзавелось пресс-службами и пресс-секретарями* [«Известия», 17.01.2002]

— ‘большинство ведомств имеют пресс-службу, почти нет ведомств без своей пресс-службы’.

Поскольку во всех высказываниях с *редко какой X / мало какой X* речь идет о множестве, допустимо множественное число (*мало какие растения*), однако абсолютное большинство примеров в основном корпусе НКРЯ со счетными существительными — в единственном числе и лишь несколько примеров — во множественном числе:

мало какой	32
мало какие	3
редко какой	78
редко какие	5

Конструкция *много какие S* со счетными существительными допускает только форму множ. числа:

- (46) *Много какие города подавали заявку на проведение Олимпиады*

— **много какой город подавал заявку на проведение Олимпиады*.

- (47) *Михалыч, я много в каких точках побывал-послужил* [Виктор Ремизов. Воля вольная // «Новый мир», 2013]

— **много в какой точке побывал*.

- (48) *Гагарину доведется оказываться много на каких обложках, о нем будут писать на первых полосах «Нью-Йорк таймс», «Дейли телеграф» и «Фигаро» — но первой им заинтересовалась «Заводская правда», многотиражка завода имени Ухтомского* [Л. А. Данилкин. Юрий Гагарин (2011)]

— **много на какой обложке оказывался*.

ЗАМЕЧАНИЕ 1. С несчетными существительными, конечно, употребляется форма ед. ч.: *Много какой литературы после знакомства с вами мне удалось перечитать; читал в том числе и китайскую литературу, в оригинале; но не нашёл удовлетворяющий меня ответ* [Эдгар Крейс. Любовь и Вера. Вспомнить Бога, 2021]; *Много какой информацией обогатился* [Фёдор Ошевнев. Билет на тот свет // «Ковчег», 2015].

ЗАМЕЧАНИЕ 2. Что касается утвердительности-отрицательности, теоретически МК *много какие* может встречаться в отрицательных высказываниях (ср. *Я много в каких городах еще не был*), однако реальные примеры в НКРЯ пока не встретились.

Естественным объяснением запрета на единственное число *много какой* могло бы быть то, что *много* семантически предполагает множество и потому требует множественного числа. Однако *редко* и *мало* тоже семантически предполагают множество. Поэтому объяснение «несимметричности» *мало какой* и *много какой* более сложное и связано с параметром референтность-нереферентность.

В большинстве примеров из НКРЯ с ИГ *мало / редко какой S* в ед. числе ИГ нереферентная. Этот статус может «размываться» за счет того, что само утверждение нередко делается по поводу какого-то конкретного объекта, введенного в рассмотрение, ср. (33) *Редко какая книга может вызвать такое стопроцентное отторжение, как первое издание книги «Иду в путь мой»*. Показательно, что в таких высказываниях часто присутствуют модальные операторы (*может, не хочет, смог бы* и под.). Нереферентные ИГ, как показано в работах Е. В. Падучевой, типичны в контексте модальных операторов, а также в других контекстах со снятой утвердительностью (см., например, [Падучева 1985; 2011; 2015a; 2015b]). При нереферентной интерпретации ИГ единственное число — законная форма (впрочем, поскольку при нереферентной ИГ противопоставление счетных существительных по числу нейтрализуется, возможно и множ. ч.: *Никакое лекарство не поможет от старости — Никакие лекарства не помогут от старости; Мало какой город сравнится с нашим — Мало какие города сравнятся с нашим; Всякий человек смертен — Все люди смертны*, — но нам важно, что форма ед. ч. допустима). Нереферентность ИГ коррелирует с соответствующим статусом предиката: он может иметь модальное значение, форму сослагательного наклонения, форму будущего времени или форму настоящего в значении абстрактного настоящего, но не значение единичной конкретной ситуации, поэтому форма СОВ. вида прош. времени в таких высказываниях ненормативна:

- (49) *Очень много людей погибло из стрелков. Редко кто остался из тех дней*
[В. Н. Гельфанд. Дневники 1941–1943 гг. (1941–1943)]

(в данном примере фигурирует оборот *редко кто*, однако у него такие же свойства, как у *редко какой S*, ср.: *Редко какой боец остался в живых*. Здесь предпочтительны формы абстрактного настоящего или модальные предикаты: *Редко какой боец / редко кто остается в живых / может / мог бы...* и т. п.).

- (35a) *Мало какой город мира может похвастаться таким количеством спортивных объектов.*

— это нереферентное значение ‘почти никакой’ в контексте модального предиката.

При описании единичной конкретной ситуации единственное число неуместно:

- (46a) *Мало какой город подал заявку на Олимпиаду.*

Ср.: (46b) *В этот раз мало какие города подали заявку на Олимпиаду.*

На приведенной выше шкале *никакой* — *мало какой* — *всякий / любой*, — МК *мало какой* коррелирует с обобщенными высказываниями с нереферентными ИГ.

Именно единственное число характерно для высказываний с нереферентными субъектами (*Всякий человек смертен*), поскольку использование формы единственного числа для обозначения открытого множества подчеркивает обобщенный характер высказывания. Обобщение характерно и для *мало какой* (а также для *мало кто*, как показано выше). Наша гипотеза состоит в том, что *мало какой S*, в силу характерной для этой МК нереферентности и благодаря значению ‘почти никакой’, «тяготеет» к «отрицательному» к полюсу — к нереферентной МК *никакой S* — и, по аналогии с ней, предпочитает единственное число, хотя допускает и множественное.

ЗАМЕЧАНИЕ. Впрочем, *много какие* — не единственная МК, которая требует множественного числа. Множественного числа требуют и ИГ с *некоторый*. В значении неопределенного местоимения (‘некий, какой-то, кое-какой, небольшой’) *некоторый* допускает ед. число, ср.:

Атрибут, характеризующий некоторый объект, принимает определенное значение [«Информационные технологии», 2003]

В некоторый момент времени велосипедиста догоняет поезд // Брусок под действием пружины придет в движение и, пройдя некоторый путь l, остановится // С помощью дощечки AC веревка отведена от стены и образует некоторый угол [Владимир Лукашик, Елена Иванова. Сборник задач по физике. 7–9 кл. (2003)]

Некоторый оптимизм в отечественных инвесторов вселяет заявление нефтяного министра одной из стран — участниц ОПЕК [«Финансовая Россия», 2002.09.19]

Есть даже некоторый смысл в том, что вся эта тягомотина будет длиться какое-то время [Известия», 2002.10.25]

Однако в функции квантора (в составе экзистенциальных нереферентных ИГ) возможно только множественное число: *Некоторые лекарства вредят печени* (‘существует хотя бы одно лекарство, такое, что оно Р [вредит печени]’), ср.: *Некоторое лекарство вредит печени*.

Попутно отметим, что высказывания с *некоторые* содержат только бытийный компонент (‘существуют’), высказывания с *мало / редко какой*, как уже упоминалось, имеют сложную семантическую и коммуникативную структуру и акцентируют именно количественный компонент: *Редко какие лекарства не вредят печени // мало какие...* — ‘лекарств, которые не вредят печени, мало’.

Итак, высказывания с *редко / мало какой S*, за счет бытийного компонента, тяготеют к нереферентной интерпретации и предпочитают форму единственного числа МК.

Высказывания с оборотами *много какие / чьи* относятся преимущественно к референтным множествам и отражают реальное количество объектов. Кроме того, не исключено, что на выбор формы множественного числа оказывает системное давление семантика *много*.

Литература

Апресян Ю. Д., Иомдин Л. Л. Конструкции типа *негде спать* в русском языке: синтаксис, семантика, лексикография // Апресян Ю. Д. и др. Теоретические проблемы русского синтаксиса: Взаимодействие грамматики и словаря. М.: Языки славянских культур, 2010. С. 59–113.

Булыгина Т. В., Шмелев А. Д. Механизмы квантификации в естественном языке и семантика количественной оценки // Булыгина Т. В., Шмелев А. Д. Языковая концептуализация мира (на материале русской грамматики). М.: Школа «Языки русской культуры», 1997. С. 193–207.

Ермакова О. П. Составные местоимения в русском языке // Словарь. Грамматика. Текст. Отв. ред. Ю. Н. Караулов & М. В. Ляпон. М.: Институт русского языка им. В. В. Виноградова РАН, 1996. С. 195–204.

Иомдин Л. Л. Синтаксические фраземы: между лексикой и синтаксисом // Апресян Ю. Д. и др. Теоретические проблемы русского синтаксиса: Взаимодействие грамматики и словаря. М.: Языки славянских культур, 2010. С. 141–190.

Кустова Г. И. Свободные релятивы как редуцированные конструкции: значение и употребление // Язык: поиски, факты, гипотезы. К 100-летию со дня рождения академика Н. Ю. Шведовой / Отв. ред. д. ф. н. М. В. Ляпон. М.: ЛЕКСРУС, 2016. С. 302–319.

Кустова Г. И. Семантика и типы употребления наречных квазирелятивов // Труды Института русского языка им. В. В. Виноградова. Вып. XIII. Культура русской речи. Гл. ред. А. М. Молдован. 2017. С. 257–264.

Кустова Г. И. Грамматическая вариативность в местоименных конструкциях: *много кто / что* VS. *много кого / чего* // Труды Института русского языка им. В. В. Виноградова. 2024. № 2. С. 174–195.

Лютикова Е. А. Формальные модели падежа: теории и приложения. М.: ЯСК, 2017.

МАС — Словарь русского языка: В 4-х тт. / Под ред. А. П. Евгеньевой. М., 1981–1984.

Никунласси А. Приместоименно-относительные конструкции в современном русском языке. Helsinki, 2008.

Падучева Е. В. Высказывание и его соотносительность с действительностью (Референциальные аспекты семантики местоимений). М.: Наука, 1985.

Падучева Е. В. ИмPLICITное отрицание и местоимения с отрицательной поляризацией // Вопросы языкознания. 2011. № 1. С. 3–18.

Падучева Е. В. Снятая утвердительность и неверидиктальность (на примере русских местоимений отрицательной полярности) // Russian Linguistics. 2015a. Vol. 39. № 2. Р. 129–162.

Падучева Е. В. Нереферентные местоимения на *-нибудь* (2015b) // Русская корпусная грамматика. <http://rusgram.ru>

Татевосов С. Г. Семантика составляющих именной группы: квантовые слова. М.: ИМЛИ РАН, 2002.

Татевосов С. Г. Семантическое картирование: теория и методика // Вестник Московского университета. Серия 9. Филология. 2004. № 1. С. 123–139.

Тестелец Я. Г., Былинина Е. Г. О некоторых конструкциях со значением неопределенных местоимений в русском языке: амальгамы и квазирелятивы (2005) [Электронный ресурс]. URL: http://www.rsuh.ru/binary/1787534_99.1322270635_82662.pdf

Фисун Р. С. Референциальный тип деривации в системе словообразования русских местоимений // Язык, литература, культура / Под ред. Клобуковой Л. П. и др. М.: МАКС Пресс, 2013. С. 122–126.

Фисун Р. С. Об использовании семантического картирования в описании местоименного компонента значения русских составных местоимений // Филологические науки. Вопросы теории и практики. 2016. № 11 (65): в 3-х ч. Ч. 2. Тамбов: Грамота. С. 148–158.

Шведова Н. Ю. Местоимение и смысл. Класс русских местоимений и открываемые ими смысловые пространства. М.: Азбуковник, 1998.

Шмелев А. Д. Русский язык и внеязыковая действительность. М.: Языки славянской культуры, 2002. (In Russ.)

Alexiadou A., Law P., Meinunger A., Wilder C. Introduction. In: The Syntax of Relative Clauses. Artemis Alexiadou, Paul Law, André Meinunger & Chris Wilder (eds.). Amsterdam / Philadelphia: John Benjamins, 2000. P. 1–51.

Bylinina E. G., Testelelets Ya. G. Sluicing-Based Indefinites in Russian. *Formal Approaches to Slavic Linguistics 13: The South Carolina Meeting*. Steven Franks, Frank Y. Gladney and Mila Tasseva-Kurkchieva (eds.). Ann Arbor, MI: Michigan Slavic Publications. 2004. P. 355–365.

Haspelmath M. Indefinite pronouns (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press, 1997.

Iomdin L. L. Hypothesis of two Syntactic Starts // Восток-Запад: Вторая международная конференция по модели «Смысл ↔ Текст». М.: Языки славянской культуры, 2005. С. 165–175.

G. I. Kustova

*V. V. Vinogradov Russian Language Institute,
Russian Academy of Sciences
(Russia, Moscow)
galinak03@gmail.com*

ON SOME PROPERTIES OF INDEFINITE-QUANTITATIVE EXPRESSIONS WITH A PRONOUN COMPONENT

The paper discusses grammatical, semantic and communicative features of indefinite-quantitative pronouns *malo kto* ('few people'), *malo kakoi / chei* ('few'), *mnogo kto* ('many people'), *mnogo kakoi / chei* ('many'), that are included in a large class of derived

indefinite pronouns (*neizvestno kto* ‘unknown who’; *kto ugodno* ‘anyone’, etc.). The constructions *malo / mnogo kto* (‘few / many people’) are compared with the indefinite-quantitative expressions *malo / mnogo X* (‘few / many X’). It is shown that these expressions are not completely synonymous and therefore in many cases are not interchangeable: *V etot pozdnii chas v vagone bylo malo lyudei* VS. **V etot pozdnii chas v vagone malo kto byl* ‘At this late hour there were few people in the carriage’.

Within their group, indefinite-quantitative pronouns are also asymmetrical. NRC statistics shows that there are more expressions with components *malo / redko (kto / kakoi / chei)* ‘few / rare (who / what / which)’ than with component *mnogo* (‘many’). These expressions also have grammatical differences (which are determined by semantic and communicative differences), for example, in the choice of number forms (for *mnogo* ‘many’ singular numbers are impossible): *It’s rare that a book can cause such 100% rejection* (possible option: *some books*) VS. *I have visited many places* (*many place).

Keywords: indefinite pronouns, indefinite-quantitative expressions, reference

References

Alexiadou A., Law P., Meinunger A., Wilder C. Introduction. In: *The Syntax of Relative Clauses*. Artemis Alexiadou, Paul Law, André Meinunger & Chris Wilder (eds.). Amsterdam / Philadelphia, John Benjamins, 2000, pp. 1–51.

Апресян Ю. Д., Иомдин Л. Л. [Construction *negde spat’*: syntax, semantics, lexicography]. Apresjan Yu. D. et al. *Teoreticheskie problemy russkogo sintaksisa. Vzaimodeistvie grammatiki i slovaryia* [Theoretical problems of the Russian syntax: Interaction of grammar and vocabulary]. Moscow, “Yazyki slavyanskoi kul’tury” Publ., 2010, pp. 59–113. (In Russ.)

Бulygina T. V., Shmelev A. D. [Mechanisms of quantification in natural language and semantics of quantitative evaluation]. Bulygina T. V., Shmelev A. D. *Yazykovaya konceptualizaciya mira (na materiale russkoj grammatiki)*. [Language conceptualization of the world (based on Russian grammar)]. Moscow, “Yazyki russkoj kul’tury” Publ., 1997, pp. 193–207. (In Russ.)

Bylinina E. G., Testelet Ya. G. Sluicing-Based Indefinites in Russian. *Formal Approaches to Slavic Linguistics 13: The South Carolina Meeting*. Steven Franks, Frank Y. Gladney and Mila Tasseva-Kurkchieva (eds.). Ann Arbor, MI: Michigan Slavic Publications, 2004, pp. 355–365.

Ермакова О. П. [Compound pronouns in the Russian]. Karaulov Yu. N., M. V. Lya-pon (eds.), *Slovar’ Grammatika. Tekst*. [Dictionary. Grammar. Text]. Moscow, Institut russkogo yazyka im. V. V. Vinogradova RAN, 1996, pp. 195–204. (In Russ.)

Фисун Р. С. [Referential type of derivation in the word formation system of Russian pronouns]. *Yazyk, literatura, kul’tura* [Language, literature, culture]. Klobukova L. P. et al. (eds.). Moscow, “MAKS Press” Publ., 2013, pp. 122–126. (In Russ.)

Фисун Р. С. [On the use of semantic mapping in the description of the pronominal component of the meaning of Russian compound pronouns]. *Filologicheskie nauki. Voprosy teorii i praktiki* [Philology. Theory & Practice]. 2016, no. 11 (65), part 2, Tambov, “Gramota” Publ., pp. 148–158. (In Russ.)

Haspelmath M. *Indefinite pronouns* (Oxford Studies in Typology and Linguistic Theory). Oxford, Oxford University Press, 1997.

Iomdin L. L. Hypothesis of two Syntactic Starts. *Vostok-Zapad: Vtoraja mezhdunarodnaja konferencija po modeli «Smysl ↔ Tekst»* [East-West: Second international conference on the “Meaning ↔ Text” model]. Moscow, “Jazyki slavyanskoy kul’tury” Publ., 2005, pp. 165–175.

Iomdin L. L. [Syntactic phrasemes: between vocabulary and syntax]. Apresjan Yu. D. et al. *Teoreticheskie problemy russkogo sintaksisa. Vzaimodejstvie grammatiki i slovary* [Theoretical problems of the Russian syntax: Interaction of grammar and vocabulary]. Moscow, “Jazyki slavyanskoi kul’tury” Publ., 2010, pp. 141–190. (In Russ.)

Kustova G. I. [Free relatives as reduced constructions: meaning and use]. *Jazyk: poiski, fakty, gipotezy* [Language: searches, facts, hypotheses]. Lyapon M. V. (ed.). Moscow, LEKSUS Publ., 2016, pp. 302–319. (In Russ.)

Kustova G. I. [Adverbial free relative: the semantics and the types of using]. *Trudy Instituta russkogo jazyka im. V. V. Vinogradova* [Proceedings of the V. V. Vinogradov Russian Language Institute], A. M. Moldovan (ed.), 2017, no. XIII, pp. 257–264. (In Russ.)

Kustova G. I. Grammatical variability in pronominal constructions: *mnogo kto / chto* VS. *mnogo kogo / chego* [Proceedings of the V. V. Vinogradov Russian Language Institute], A. M. Moldovan (ed.), 2024, no. 2, pp. 174–195. (In Russ.)

Kuznetsov S. A. (resp. ed.). *Bol’shoj tolkovyj slovar’ russkogo jazyka* [Great Dictionary of Russian language]. St. Petersburg, “Norint” Publ.; Moscow, “Ripol klassik” Publ., 2008. (In Russ.)

Lyutikova E. A. *Formal’nye modeli padezha: teorii i prilozheniya* [Formal case models: theories and applications]. Moscow, YaSK Publ., 2017. (In Russ.)

Nikunlassi A. *Primestoimenno-otnositel’nye konstruksii v sovremennom russkom jazyke* [Pronominal-relative constructions in modern Russian]. Helsinki, 2008. (In Russ.)

Paducheva E. V. *Vyskazyvanie i ego sootnesennost’ s dejstvitel’nost’yu* [Utterance and its Correlation with Reality]. Moscow, “Nauka” Publ., 1985. (In Russ.)

Paducheva E. V. [Implicit negation and pronouns of negative polarity]. *Voprosy jazykoznanija*, 2011, no. 1, pp. 3–18. (In Russ.)

Paducheva E. V. [Subtracted affirmativeness and non-veridicality (using the example of Russian pronouns of negative polarity)]. *Russian Linguistics*, 2015a, vol. 39, no. 2, pp. 129–162. (In Russ.)

Paducheva E. V. Nereferentnye mestoimeniya na *-nibud’* (2015b) [Non-referential pronouns *-nibud’*]. *Russkaya korpusnaya grammatika* [Russian corpus grammar]. <http://rusgram.ru>. (In Russ.)

Shvedova N. Yu. *Mestoimienie i smysl. Klass russkikh mestoimenij i otkryvaemye imi smyslovye prostranstva* [Pronoun and meaning. The class of Russian pronouns and the semantic spaces they open]. Moscow, “Azbukovnik” Publ., 1998. (In Russ.)

Shmelev A. D. *Russkiy jazyk i vneyazykovaya dejstvitel’nost’* [Russian language and extra-linguistic reality]. Moscow, “Jazyki slavjanskoj kul’tury” Publ., 2002. (In Russ.)

Slovar' russkogo yazyka [Russian Dictionary]. Vol. 1–4. Ed. by A. P. Evgen'eva. Moscow, 1981–1984. (In Russ.)

Tatevosov S. G. *Semantika sostavlyayushchikh imennoi gruppy: kvantovye slova* [Semantics of noun phrase components: quantum words]. Moscow, IMLI RAN Publ., 2002. (In Russ.)

Tatevosov S. G. Semanticheskoe kartirovanie: teoriya i metodika [Semantic mapping: theory and methodology]. *Vestnik Moskovskogo universiteta* [Bulletin of Moscow University], seriya 9. Filologiya, 2004, no. 1, pp. 123–139. (In Russ.)

Testeleys Ya. G., Bylinina E. G. O nekotorykh konstruktivnykh so znacheniem neo-predelennykh mestoimenii v russkom yazyke: amal'gamy i kvazirelyativy [On some constructions with the meaning of indefinite pronouns in Russian: amalgams and quasi-relatives] (2005) [Electronic resource]. URL: http://www.rsuh.ru/binary/1787534_99.1322270635.82662.pdf (In Russ.)

**И. М. Богуславский, П. В. Дяченко, Е. С. Иншакова, Л. Л. Иомдин,
А. В. Лазурский, Л. Г. Митюшин, А. А. Мовсисян, И. П. Рыгаев,
В. Г. Сизов, С. П. Тимошенко, Т. И. Фролова, А. В. Чага**

**ИППИ РАН им. А. А. Харкевича
(Москва)**

*igor.m.boguslavsky@gmail.com, pavel.v.djachenko@gmail.com,
e.s.inshakova@gmail.com, iomdin@gmail.com, lazursky@mail.ru,
lmityushin@gmail.com, derise@iitp.ru, irygaev@jent.ru, victor.sizov@gmail.com,
nyrestein@gmail.com, tfrolova@gmail.com, chagachaga@gmail.com*

СОВРЕМЕННОЕ СОСТОЯНИЕ КОРПУСА СИНТАГРУС¹

Предлагается описание основных особенностей и опций многосторонне размеченного корпуса русских текстов СинТагРус. Корпус был разработан в ИППИ РАН им. А. А. Харкевича и в настоящее время выступает как один из подкорпусов НКРЯ, где он называется «Синтаксическим корпусом». Излагаются основные подходы к выбору текстов для корпуса и к их метаразметке, лингвистические принципы, лежащие в основе разметки разных типов — морфологической, синтаксической, лексико-семантической, лексико-функциональной, эллиптической, микросинтаксической, кореферентной и темпоральной. Приводятся статистические данные, характеризующие различные аспекты СинТагРуса и его фрагментов. СинТагРус является корпусом со стопроцентно дизамбигуированной на всех уровнях разметкой: в статье описываются очевидные достоинства такой разметки и в то же время отмечаются трудности, связанные с необходимостью всегда принимать определенные решения и выбирать единственные варианты разметки даже в тех случаях, когда языковой материал принципиально допускает неединственное лингвистическое описание. Значительное внимание уделяется описанию некоторых различий, существующих между СинТагРусом и основными подкорпусами НКРЯ — разделению материала по частям речи, различным морфологическим решениям, принятыми в СинТагРусе и НКРЯ (таким, как трактовка отдельных морфологических категорий — вида и залога глагола, некоторых падежей существительных и др.).

¹ Данная работа выполнена в рамках гранта Министерства науки и высшего образования РФ № 075-15-2020-793. Авторы выражают благодарность Министерству за поддержку.

Ключевые слова: СинТагРус, синтаксический корпус, морфосинтаксическая разметка, лексическая разметка, эллиптическая разметка, микросинтаксис, кореллированная разметка, темпоральная разметка

1. Общая информация о корпусе

В 2023 году синтаксически размеченный корпус русских текстов, известный как СинТагРус (или SynTagRus, Syntactically Tagged Russian text corpus), отмечает 25-летие своего существования. Работа над корпусом началась в 1998 году в Лаборатории компьютерной лингвистики ИППИ РАН им. А. А. Харкевича и поддерживалась грантами РФФИ, РНФ и РГНФ. Ход работы над корпусом и различные теоретические и практические аспекты его функционирования и использования отражены в ряде публикаций его авторов (см., в частности, [Апресян и др. 2005], [Boguslavsky et al. 2000], [Богуславский и др. 2008], [Iomdin, Sizov 2009], [Шеманаева, Фролова 2010], [Boguslavsky 2014], [Дяченко и др. 2015], [Iomdin 2016], [Маракасова, Иомдин 2016], [Iomdin 2017], [Тимошенко и др. 2021], [Chaga 2021]).

Весьма интенсивная работа по развитию корпуса была проведена в 2020–2022 гг. в рамках мегапроекта Corpus 2.0, выполняемого по гранту Министерства науки и высшего образования № 075-15-2020-793. В этот период корпус увеличился на 35 % и достиг объема, превышающего 1,5 млн слов (1308 текстов, свыше 107 тыс. предложений). Благодаря продлению гранта на 2023 год, к концу года объем корпуса еще увеличился и к началу 2024 года составил около 1 570 000 слов (1364 текста, около 111 тыс. предложений)

Корпус включает следующие типы текстов: художественная проза XX–XXI веков; современная научно-популярная литература; публицистика; биографии; новостные ленты (тексты общественно-политического, культурного, экономического и научно-технического характера).

В настоящее время СинТагРус входит в состав Национального корпуса русского языка и представляет собой отдельный его фрагмент — Синтаксический корпус (<https://ruscorpora.ru/new/search-syntax.html>).

Разметка СинТагРус осуществляется в Институте проблем передачи информации РАН и по мере готовности очередного пополнения передается администраторам НКРЯ для внесения в текущую версию последнего. Информация, содержащаяся в новейшей версии Синтаксического корпуса, идентична информации в соответствующей версии СинТагРус, однако возможности поиска по сложным запросам в Синтаксическом корпусе несколько ограничены по сравнению с поиском по СинТагРусу. Работа над расширением таких возможностей ведется, однако в силу разнообразия нетривиальной лингвистической информации в СинТагРусе, она еще не завершена. В дальнейшем изложении мы сосредоточимся именно на особенностях СинТагРус как исходного ресурса (отмечая, впрочем, некоторые расхождения с Синтаксическим корпусом, если они являются существенными).

На данный момент в СинТагРусе имеются следующие виды разметки: 1) морфологическая; 2) синтаксическая; 3) лексико-семантическая; 4) лексико-функ-

циональная; 5) эллиптическая; 6) микросинтаксическая; 7) кореферентная; 8) темпоральная; 9) метатекстовая.

Виды разметки 1–6 покрывают весь корпус. Работа над кореферентной и темпоральной разметкой была начата относительно недавно, и в настоящее время эти типы разметки выполнены для части текстов.

Разметка текстов СинТагРуса выполняется в полуавтоматическом режиме. Сначала исходный текст обрабатывается парсером многофункционального лингвистического процессора ЭТАП-4, созданного в Лаборатории компьютерной лингвистики ИППИ РАН, в результате чего для каждого предложения строится его морфологическая и синтаксическая структура, а также идентифицируются лексические функции. Затем эти данные проверяются и корректируются экспертами-лингвистами с помощью программного комплекса «Редактор структур», разработанного специально для работы с корпусом². На следующем шаге к построенным структурам применяются специализированные модули процессора, маркирующие микросинтаксические единицы, кореферентные связи и темпоральные выражения. Эти данные также проверяются лингвистами.

Морфологическая разметка состоит в том, что каждому слову текста ставится в соответствие его основная форма, часть речи и набор морфологических характеристик (значений морфологических категорий числа, падежа, вида, наклонения, времени, лица и т. д.).

Существенно подчеркнуть, что морфологическая разметка корпуса характеризуется стопроцентно снятой неоднозначностью. Она соответствует синтаксической разметке корпуса, которая также полностью однозначна.

В целом морфологическая разметка СинТагРуса соответствует принципам и конвенциям морфологической разметки ряда корпусов НКРЯ, в том числе основного корпуса НКРЯ, однако существуют некоторые расхождения. Основные расхождения состоят в следующем.

1) В СинТагРусе вид глагола рассматривается как словоизменительная категория. Какие морфологические средства используются для образования форм совершенного и несовершенного вида глагола, не имеет значения: это могут быть суффиксальные средства (*начинать* — *начать*), префиксальные средства (*писать* — *написать*), ударение (*просыпать* — *просы́пать*), смешанные средства (*становиться* — *стать*). В случае, если у соответствующего глагола вообще существует форма несовершенного вида, именно она выбирается в качестве лексемы. Например, словоформа *стал* в выражении *Я стал инженером* представляется лексемой СТАНОВИТЬСЯ1, в выражении *Я стал в очередь* — лексемой СТАНОВИТЬСЯ2, а в выражении *Я стал хуже слышать* — лексемой СТАТЬ1 (поскольку глагол *стать*, употребляемый в качестве вспомогательного при инфинитиве, не

² «Редактор структур», или Structure Editor (см., в частности, [Iomdin, Sizov 2009]) — основная программная система, разработанная в Лаборатории компьютерной лингвистики ИППИ РАН специально для построения, хранения и поддержки СинТагРуса и ряда других современных компьютерно-лингвистических ресурсов.

имеет несовершенного вида: в таких случаях именем лексемы считается форма совершенного вида)³.

2) Категория залога в СинТагРусе имеет два значения: действительный залог и страдательный залог. Медиальный залог, как в НКРЯ, не используется. В выражении *Старик с трудом поднимается со стула* словоформа *поднимается* представлена как форма действительного залога, лексема при этом — ПОДНИМАТЬСЯ, а в выражении *Этот вес штангистом еще не поднимался* словоформа *поднимался* представлена как форма страдательного залога глагольной лексемы ПОДНИМАТЬ.

3) В СинТагРусе не используется часть речи «предикатив». В большинстве случаев формы, отражаемые в НКРЯ как принадлежащие этой части речи, представлены как краткие прилагательные (*грустно* в *Маше было грустно*) или наречия (*жаль* в *Жаль, что ты не придешь*). Сведения о предикативности единицы отражаются не в морфологической, а в синтаксической структуре: в синтаксическом дереве конструкции типа *Мне было холодно* существенно отличаются от конструкций типа *Я был холоден*⁴.

4) В СинТагРусе не используются «составные» части речи типа «числительное-прилагательное», «существительное-местоимение», «прилагательное-местоимение», местоименное наречие. Порядковые слова типа *четвертый* считаются прилагательными, слова типа *я, он, кто-то, что-нибудь, себя* — существительными, слова *мой, такой, любой, чей-либо* — прилагательными, слова *там, кое-где, туда* — наречиями. (Свойства местоименности таких слов, разумеется, не утрачиваются, но используются за пределами морфологии.)

5) Словоформы *его, ее, их* в притяжательной конструкции типа *его дом* представляются лексемами существительных ОН, ОНО, ОНА, ОНИ в родительном падеже, в то время как в НКРЯ такие словоформы трактуются как местоименные прилагательные без падежа с лексемами ЕГО, ЕЕ, ИХ.

6) Словоформы типа *него, нее, ним, них* представлены особыми лексемами — существительными НЕГО1 (мужского рода), НЕГО2 (среднего рода), НЕЕ, НИХ и т. д. в нужном падеже (родительном, дательном, винительном, творительном, предложном: именительного падежа у этих существительных нет). В НКРЯ здесь используются стандартные местоименные существительные ОН, ОНО, ОНА, ОНИ.

7) Вариативные формы творительного падежа (*землей* — *землею*) и сравнительной степени (*веселее* — *веселей*) представлены одинаковыми наборами

³ Естественно, такое решение может влиять на тактику поиска по СинТагРусу: пользователь, удивившись, что в корпусе не обнаруживается ни одного вхождения слова *написать*, заподозрит неладное и постарается поискать слова с лексемой несовершенного вида — *писать*.

⁴ Решение отказаться от использования такой части речи было принято задолго до начала разметки СинТагРус, в ходе разработки русского синтаксического парсера ЭТАП. Важная мотивировка этого отказа состояла в следующем: авторы ЭТАПа сочли, что в выражениях типа *Казалось полезным выучить новый маршрут* и *Было полезно выучить новый маршрут* словоформы *полезным* и *полезно* относятся к одной и той же лексеме, а первая из них может быть только прилагательным.

морфологических характеристик, однако менее стандартные из них (*землю, веселей*) получают дополнительную вспомогательную характеристику *alt*.

8) В СинТагРусе не используется принятая в НКРЯ характеристика «2-й винительный» для представления словоформ типа *депутаты, солдаты* и т. д. в конструкциях типа *кандидат в депутаты* или *пойти в солдаты*. Эти словоформы трактуются как существительные именительного падежа множественного числа, а их синтаксический статус дополнения отражается в синтаксической структуре⁵.

9) Страдательные причастия настоящего времени, образованные с помощью суффиксов типа «-ем-» и суффиксов типа «-ющ-» и частицы «-ся» (*рассматриваемый — рассматривающийся*) получают тождественные морфологические разборы.

10) Формы «смягченной» сравнительной степени с приставкой *по-* получают специальную морфологическую характеристику «смяг» в дополнение к характеристике «срав», тогда как в НКРЯ используется единая характеристика «сравнительная 2».

Синтаксическая разметка означает, что каждому предложению приписывается синтаксическая структура в виде дерева зависимостей, узлами которого являются слова предложения, а дуги помечены именами синтаксических отношений в соответствии с моделью «Смысл ↔ Текст» И. А. Мельчука.

Всего в корпусе используется 68 различных синтаксических отношений. Полный комментированный список частей речи, морфологических категорий и характеристик, а также синтаксических отношений можно найти в инструкции к Синтаксическому корпусу [Инструкция 2023].

Узлами синтаксических структур являются слова текста. В предложении естественно выделяются графические слова: цепочки алфавитных символов, разделенные пробелами и знаками препинания (внутри цепочек допускается дефис). Чаще всего один узел синтаксической структуры соответствует одному графическому слову предложения, однако в некоторых случаях СинТагРус рассматривает несколько следующих друг за другом графических слов как одну лексическую единицу — например, сочетания *как бы то ни было, ни много ни мало, до поры до времени, один на один* (в лингвистическом процессоре ЭТАП-4 такие сочетания именуются безусловными оборотами). Каждому такому сочетанию приписывается одна словарная статья, которая считается цельным словом, невзирая на пробелы. Словарь процессора ЭТАП-4 включает более 200 таких статей, и им соответствует около 11 тыс. вхождений в СинТагРусе⁶.

⁵ Стоит добавить, что в основном корпусе НКРЯ эта характеристика применяется весьма непоследовательно.

⁶ В последнее время в Синтаксический корпус была введена опция, при которой элементы безусловных оборотов СинТагРуса могут по запросу пользователя отражаться в синтаксической структуре предложения порознь и представляться независимо. При этом фрагмент синтаксической структуры, соответствующий безусловному обороту, строится заранее по специально сформулированным разработчиками СинТагРуса правилам.

Лексико-семантическая разметка означает, что для многозначных (или омонимичных) слов фиксируется их лексическое значение, реализованное в данном предложении. Это значение обычно представляется именем лексемы, состоящим из леммы и следующей за ней цифры. Например, форма *толковали* в предложении *Два старика сидели за столом, пили чай и толковали о жизни* (Ю. Домбровский) представлена лексемой ТОЛКОВАТЬ1 ‘беседовать’, а такая же форма в предложении *Органы власти по-разному толковали понятие равноправия* — лексемой ТОЛКОВАТЬ2 ‘интерпретировать’. Имена лексем соответствуют именам словарных статей комбинаторного словаря лингвистического процессора ЭТАП-4. В «Редакторе структур», обслуживающем СинТагРус, предусмотрена удобная возможность заглянуть в соответствующие статьи комбинаторного словаря.

Хотя этот словарь достаточно объёмен (он насчитывает свыше 110 тысяч вхождений), разумеется, не все встречающиеся в СинТагРусе слова в нем представлены. В случае отсутствия слова в словаре ЭТАП-4 лексико-семантическая разметка ограничивается леммой. Значительную часть таких слов составляют сложные слова типа *деревообработка*, *Минобрнауки* или *двадцатипятиэтажный*. Несмотря на то, что подобные единицы вполне адекватно обрабатываются блоком композитной морфологии лингвистического процессора, в СинТагРусе было принято решение не ставить им в соответствие конкретных слов или групп слов комбинаторного словаря.

Особо следует отметить случаи, когда в корпусе присутствует слово, значение которого не отражено в комбинаторном словаре, хотя другие лексические значения этого слова в нем представлены (и, в случае неоднозначности, помечены цифрами при леммах). В таких ситуациях слово не получает никакого цифрового индекса и трактуется как не имеющее отражения в комбинаторном словаре. Заметная доля таких ситуаций приходится на случаи, когда в СинТагРусе некоторое слово представляет название, совпадающее с именем реального слова, но не вполне соответствующее ему по набору морфологических характеристик. Так, названия изданий «Коммерсант» или «Наш современник», кораблей «Викинг» или «Варяг» используют реальные слова русского языка, но выступают, в отличие от последних, как неодушевленные существительные. С другой стороны, личное имя *Урал* выступает как одушевленное существительное, в отличие от присутствующего в словаре слова *Урал* ‘горный массив’. По этой причине такие лексические единицы не наследуют свойств исходных слов и считаются отсутствующими в комбинаторном словаре.

Лексико-функциональная разметка состоит в том, что в тексте маркируются пары (а в особых случаях и тройки) слов, допускающие интерпретацию в терминах лексических функций (ЛФ) лингвистической теории «Смысл ↔ Текст». Эту разметку можно представлять себе как установление направленных связей: лексико-функциональное отношение связывает слово-аргумент лексической функции (главный член отношения) со словом, представляющим значение этой лексической функции (зависимый член отношения). В общем случае направление лексико-

функциональной связи между двумя словами не обязано совпадать с направлением синтаксического отношения между этими словами или даже предполагать наличие такого отношения.

Например, в предложении *Хлынул проливной дождь* слово *хлынул* представляет значение ЛФ INCEPFUNC0 от аргумента *дождь*, а слово *проливной* представляет значение ЛФ MAGN от *дождь*. В древовидной синтаксической структуре слово *дождь* синтаксически зависит от слова *хлынул*, тогда как лексико-функциональное отношение INCEPFUNC0 идет в обратном направлении — от аргумента *дождь* к значению *хлынул*. В то же время в этой синтаксической структуре слово *дождь* синтаксически подчиняет слово *проливной*, но и лексико-функциональное отношение MAGN имеет то же направление — от аргумента *дождь* к значению *проливной*. Если же рассмотреть предложение *Дождь был проливной*, то направление лексико-функционального отношения MAGN между *дождь* и *проливной* сохранится, а непосредственной синтаксической связи между этими словами не будет вовсе.

Эллиптическая разметка вводит в синтаксическую структуру предложения СинТагРус слова, не представленные в графической записи предложения. Эта разметка затрагивает далеко не все типы и не все случаи эллипсиса и других разновидностей синтаксического отсутствия, существующие в языке и рассматриваемые в специальных исследованиях, и применяется достаточно избирательно. В первую очередь восстанавливаются такие типы эллипсиса, которые естественно нормализуют синтаксическую структуру присутствующей в СинТагРусе фразы (например, при опущении одного из предикатных слов в сочинительной цепочке).

В частности, при представлении предложения *Он пошел налево, а она направо* разумно принять, что между *она* и *направо* содержится лексическая единица, соответствующая слову *пошла*, но физически отсутствующая в тексте. В результате введения такого слова в синтаксическую структуру предложения оказывается естественным связать слово *пошла* с подлежащим *она* и с направлятельным словом *налево*. С другой стороны, в предложениях типа *Я предпочитаю светлое пиво, а мой друг любит темное* вводить для нормализации синтаксической структуры опущенное слово *пиво* в конец предложения нет необходимости, поскольку слово *темное* достаточно для того, чтобы отразить наличие у глагола *любит* прямого дополнения. Подробнее об эллиптической разметке см. ниже в разделе 2.

Микросинтаксическая разметка идентифицирует синтаксически своеобразные фразеологические и полуфразеологические выражения самых разных типов (именуемые микросинтаксическими единицами), содержащиеся в текстах СинТагРуса и отличающиеся своеобразием синтаксического поведения. Примерами таких единиц могут быть обороты типа *всё равно* (в разных значениях), составные союзы типа *как будто бы, разве что*, парные союзы типа *не только ... но и, не то ... не то*,

составные предлоги типа *со стороны, в качестве*, адverbials типа *с виду, на виду, про запас* и многие другие. Число разных микросинтаксических единиц, которые встречаются в корпусе СинТагРус, превышает 3200.

Этот тип разметки применяется относительно недавно, однако в настоящее время уровень покрытия достиг всего объема СинТагРус, а текущая разметка осуществляется одновременно с пополнением корпуса новыми текстами. Подробнее микросинтаксическая разметка будет рассмотрена в разделе 3.

Кореферентная разметка предполагает установление особых связей между словами текста, референты которых совпадают, включая анафорические связи между местоимениями и их антецедентами. В отличие от охарактеризованных выше типов разметки, кореферентная разметка не ограничивается отдельными предложениями, а покрывает весь текст целиком. Благодаря этому обстоятельству кореферентные цепочки в одном тексте могут быть весьма длинными и достигать сотен элементов.

Создатели СинТагРус приступили к кореферентной разметке относительно недавно, причем в первом варианте разметка сводилась только к установлению местоименной анафоры и выполнялась в экспериментальном режиме для текстов корпуса, размеченных в 2017–2019 гг. [Иншакова и др. 2019]. С развитием проекта Corpus 2.0 и разработкой механизма установления неместоименной кореферентности было принято решение существенно увеличить объем подлежащих кореферентной разметке текстов. В соответствии с этим решением объем текстов СинТагРус с кореферентной разметкой к концу 2023 года составил более 120 тыс. слов, или около 9200 предложений.

Данный тип разметки будет подробнее рассмотрен в разделе 4.

Темпоральная разметка предназначена для идентификации в тексте выражений, описывающих протекание событий во времени, а также связей между ними. Принципы темпоральной разметки СинТагРус были подробно описаны в работе [Тимошенко и др. 2021]. В 2020–2022 годах темпоральная разметка СинТагРус была выполнена для нескольких текстов корпуса. Однако в текущей версии СинТагРус развитие этого типа разметки и пополнение ею новых текстов не производится; к ней предполагается вернуться через некоторое время.

В рамках проекта Corpus 2.0 была проведена большая работа по совершенствованию **метаразметки** СинТагРус. Формат метатекстовой разметки был приведен в соответствие со стандартом, принятым в основном корпусе НКРЯ и описанным в [Савчук 2005]. Исключением являются, в частности, поля «возраст аудитории» и «размер аудитории», которые не были использованы нами из-за неоднозначности их трактовки и их непоследовательного использования в НКРЯ. Поля «хронотоп» и «тематика текста» в СинТагРусе объединены в одно поле, поскольку первое указывается только для художественных текстов, а второе — только для нехудожественных. Кроме того, есть некоторые поля, присутствующие в СинТагРусе, но отсутствующие в НКРЯ, включая год разметки текста, часть текста

(если размечен фрагмент), имена аннотаторов текста и дополнительные комментарии.

Общее число полей метатекстовой разметки возросло с 7 до 17. Заполнение новых полей и правка старых для всех текстов корпуса были проведены вручную с использованием «Редактора структур», в который был добавлен функционал, обеспечивающий единообразие разметки между разными текстами и частями одного текста. В том числе обеспечивается одинаковость написания имени одного и того же автора в разных текстах.

На рис. 1 приведен пример метаразметки текста СинТагРус в варианте Синтаксического корпуса НКРЯ. Отличие от разметки в самом СинТагРусе состоит в отсутствии некоторых дополнительных полей и включении пола и года рождения автора (в СинТагРусе эта информация является частью сведений об авторе, хранящихся отдельно от самих текстов).

Автор	Михаил Афанасьевич Булгаков
Пол	мужской
Год рождения	1891
Название	Мастер и Маргарита
Год создания	1928-1940
Сфера функционирования	художественная
Жанр текста	фантастика
Тематика текста	ирреальный мир
Тип текста	роман
Носитель	электронный текст
Источник	http://bibliotekar.ru/bulgakov/2.htm
Дата публикации	-
Издание	Библиотекарь.Ру
Издательство	-
Часть	глава 2
Год разметки	2018

Рис. 1. Метаразметка текста 2-й главы «Мастера и Маргариты» (в варианте Синтаксического корпуса НКРЯ)

2. Несколько иллюстративных примеров

Для иллюстрации основных видов разметки СинТагРус (морфологической, синтаксической, лексико-семантической, лексико-функциональной и эллиптической) мы рассмотрим подробнее представление нескольких реальных примеров предложений, содержащихся в корпусе.

Для начала посмотрим, как в СинТагРусе отражено предложение

- (1) *Тут прокуратор поднялся с кресла, сжал голову руками, и на желтоватом его бритом лице выразился ужас.* (М. Булгаков, «Мастер и Маргарита»).

В «Редакторе структур» предложение (1) выглядит следующим образом:

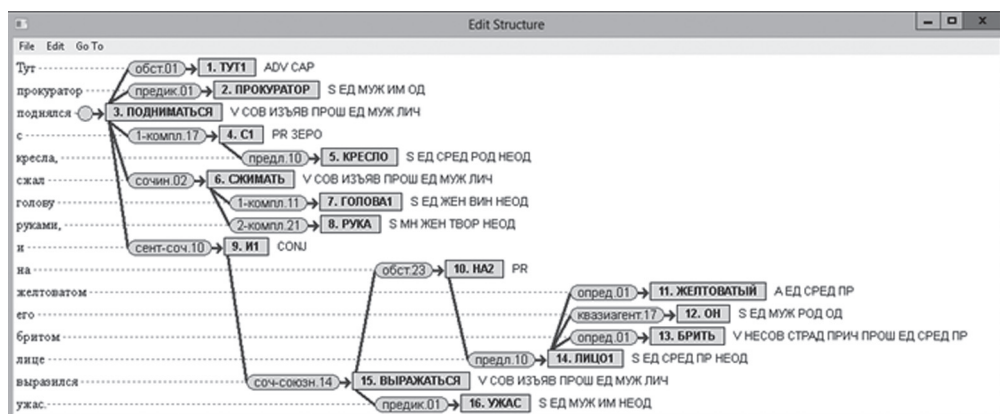


Рис. 2. Структура предложения (1)

Для каждого слова предложения, записываемого слева, в прямоугольной рамке указано соответствующее ему имя лексемы, связывающее его со статьей комбинаторного словаря лингвистического процессора ЭТАП-4, и представляющее собой лемму с добавленным цифровым индексом в случае неоднозначности (все равно, омонимии или полисемии). Например, ТУТ1 — статья наречия (а не существительного), С1 — предлог, управляющий родительным падежом, ГОЛОВА1 — часть тела (а не начальник), И1 — союз (а не частица), НА2 — предлог, управляющий предложным падежом; ЛИЦО1 — неодушевленное существительное со значением «часть тела», а не одушевленное существительное со значением «персона». Это и есть лексико-семантическая разметка корпуса.

Если разметчики СинТагРусе или его пользователи захотят удостовериться в том, что в разметке выбрано нужное значение слова, они могут воспользоваться особой опцией «Редактора структур», чтобы в реальном времени заглянуть в комбинаторный словарь и увидеть, каким лексическим значением представлено соответствующее слово. В частности, воспользовавшись данной опцией, разметчик или пользователь сможет увидеть представительные фрагменты словарных статей слов, основные формы которых совпадают, и подтвердить или изменить сделанный в СинТагРусе выбор.

Так, для неоднозначного слова ТУТ «Редактор структур» покажет по запросу два фрагмента комбинаторного словаря: для наречия ТУТ1 (рис. 3) и для существительного ТУТ2 «тутовое дерево» (рис. 4).

Аналогичным образом, при просмотре лексических значений слова ГОЛОВА мы сможем увидеть фрагменты словарных статей, соответствующих лексеме ГОЛОВА1 («часть тела») и ГОЛОВА2 («начальник», как в *городской голова*). При этом демонстрация разных лексических значений слова не зависит от того, насколько реально появление этих значений в рассматриваемой фразе: система показывает все значения.



Рис. 3. Фрагмент словарной статьи ТУТ1, доступный при работе с СинТагРусом над предложением, содержащем слово ТУТ

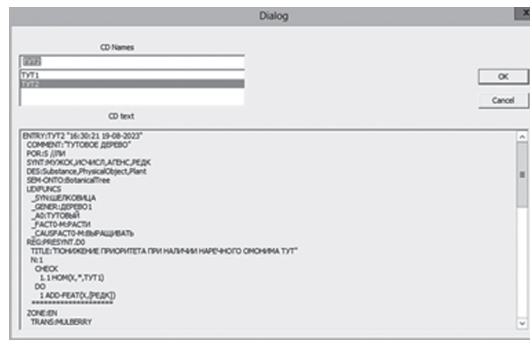


Рис. 4. Фрагмент словарной статьи ТУТ2, доступный при работе с СинТагРусом над предложением, содержащим слово ТУТ

Справа от каждой прямоугольной рамки на рис. 2 указывается часть речи и список морфологических характеристик слова: напомним, что как морфологическая, так и синтаксическая структуры в СинТагРусе являются стопроцентно дизамбигуированными.

Основные морфологические характеристики, представленные на рис. 2, кажутся самоочевидными.

В овалах записаны стандартно сокращенные имена синтаксических отношений (обстоятельственное, предикативное, 1-е комплетивное и т. п.). Как нетрудно увидеть, при каждом слове записано имя единственного синтаксического отношения, входящего в это слово; существует, однако, ровно одно слово (вершина дерева зависимостей), в которое не входит никакое синтаксическое отношение. В примере (1) такой вершиной является третье по порядку слово — глагол *поднялся*.

Цифры в овалах после имени синтаксического отношения указывают на номер синтаксического правила (синтагмы), пользуясь которым синтаксический парсер процессора ЭТАП-4 установил данное отношение. Данный тип информации пред-

назначается для разработчиков СинТагРуса, а не для его пользователей. В Синтаксическом корпусе эти номера не воспроизводятся.

Рассмотрим еще одно предложение СинТагРуса, чтобы подробнее проиллюстрировать его лексико-функциональную разметку:

- (2) *Присуждение Нобелевской премии этого года в области физиологии и медицины прошло под знаком точных наук.* (Ф. Смирнов, «Нобелевские премии: физики снова в почете»)

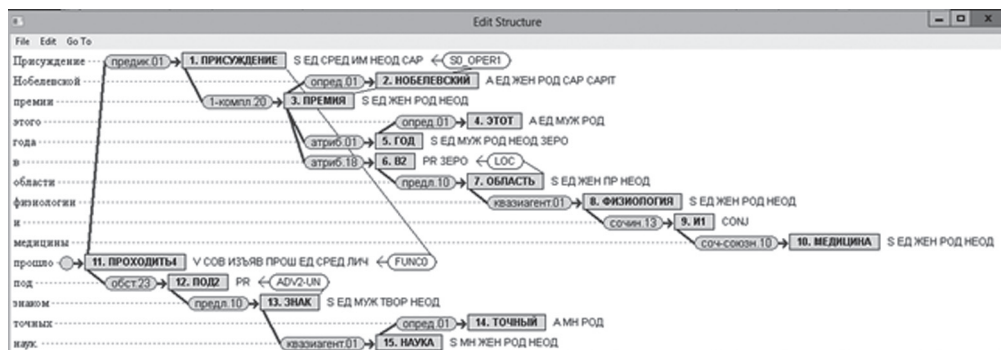


Рис. 5. Структура предложения (2)

Как уже отмечалось, лексико-функциональная разметка фиксирует пары слов в предложении, между которыми можно установить связи в терминах лексических функций модели «Смысл ↔ Текст». В предложении (2) таких пар слов (лексических коррелятов) четыре: (а) *присуждение премии*, (б) *прошло присуждение*; (в) *в области* и (г) *под знаком*. Во всех четырех парах первое слово является значением некоторой лексической функции, а второе — аргументом (ключевым словом) этой функции. В паре (а) имеет место ЛФ S0_OPER1; это по существу субстантивация ЛФ OPER1 (‘делать то, что обозначено ключевым словом’) — *присуждать премию* → *присуждение премии*, в (б) — ЛФ FUNC0 (‘иметь место’), в (в) — ЛФ LOC (типичный предлог, характеризующий ключевое слово в качестве места) и в (г) — ЛФ ADV2-UN (типичный предлог, образующий адвербиал, который характеризует ситуацию, описываемую ключевым словом, с точки зрения его второго участника). Во всех четырех случаях аргумент и значение ЛФ синтаксически связаны между собой (что, как мы видели в разделе 1, бывает не всегда), причем в трех случаях (а, в и г) связь направлена от значения ЛФ к аргументу, а в одном случае (б) — от аргумента к значению.

На рис. 5 направление лексико-функциональной связи графически представлено стрелками, внешний вид которых отличается от представления синтаксических отношений.

Добавим в заключение, что лексико-функциональные связи в СинТагРусе представлены одним из двух широких классов ЛФ — так называемыми ЛФ-параметрами. Другой широкий класс ЛФ — ЛФ-замены, такие как ЛФ SYN (сино-

ним), ANTI (антоним), CONV (конверсив) и ряд других (в том числе синтаксические дериваты), в корпусе не отражается. Такое решение было принято потому, что, как правило, в тексте практически не встречаются одновременно аргументы и значения таких функций. Между тем в некоторых ситуациях значения ЛФ данного класса все-таки появляются в предложении вместе с их аргументами; например, в случае ЛФ GENER (родовое слово). Помимо хорошо известных цитат типа *Спит животное Собака, Дремлет птица Воробей* (Н. Заболоцкий), где слова *животное* и *птица* выступают как значения ЛФ GENER при словах *собака* и *воробей* соответственно, аналогичные примеры встречаются и во вполне нейтральных текстах, в том числе и в СинТагРусе; ср., например

- (3) *Появился новый препарат эпоэтин, близкий к гормону эритропоэтину, способному повышать выработку эритроцитов и тем самым улучшать снабжение организма кислородом и справляться с такими видами малокровия, которые ранее были неизлечимы* (В. Прозоровский, «Лекарства от усталости»),

где слова *препарат* и *гормон* выступают в качестве значения ЛФ GENER к словам *эпоэтин* и *эритропоэтин* соответственно. Сейчас такие ЛФ не фиксируются, однако обсуждается возможность дополнить лексико-функциональную разметку и некоторыми функциями, относящимися к классу замен.

Комментированный список всех используемых в СинТагРусе лексических функций приводится в инструкции к Синтаксическому корпусу [Инструкция 2023]. В настоящее время в текстах корпуса встречается 143 различных лексических функции, они реализованы приблизительно в 44 тыс. пар слов. Количество предложений, содержащих хотя бы одно вхождение ЛФ, составляет около 32 тыс., или 30% всех предложений корпуса.

Обратимся теперь к одному из предложений СинТагРуса, содержащих эллиптическую разметку:

- (4) — *Взять бы этого Канта, да за такие доказательства года на три в Соловки!* — *совершенно неожиданно бухнул Иван Николаевич* (М. Булгаков, «Мастер и Маргарита»).

Представляется очевидным, что в предложении (4) опущено некоторое глагольное слово, иначе весьма затруднительно понять, к чему могут относиться предложно-именные сочетания *за такие доказательства, года на три* и *в Соловки*. Разработчики СинТагРуса приняли решение в подобных случаях восстанавливать такие опущенные предикаты, как и прототипический эллипсис, помещая в определенное место фразы соответствующую лексическую единицу. Технически такое решение осуществляется следующим образом: (1) выбирается конкретная позиция в предложении; (2) на эту позицию помещается опущенное слово, которому приписывается определенное словарное имя, нулевая словоформа и нужный набор морфологических характеристик; (3) это опущенное слово встраивается в синтаксическую структуру предложения. Само опущенное слово получает поряд-

ковый номер в предложении и дополнительную помету Phantom. Результат такой обработки предложения (4) представлен на рис. 6:

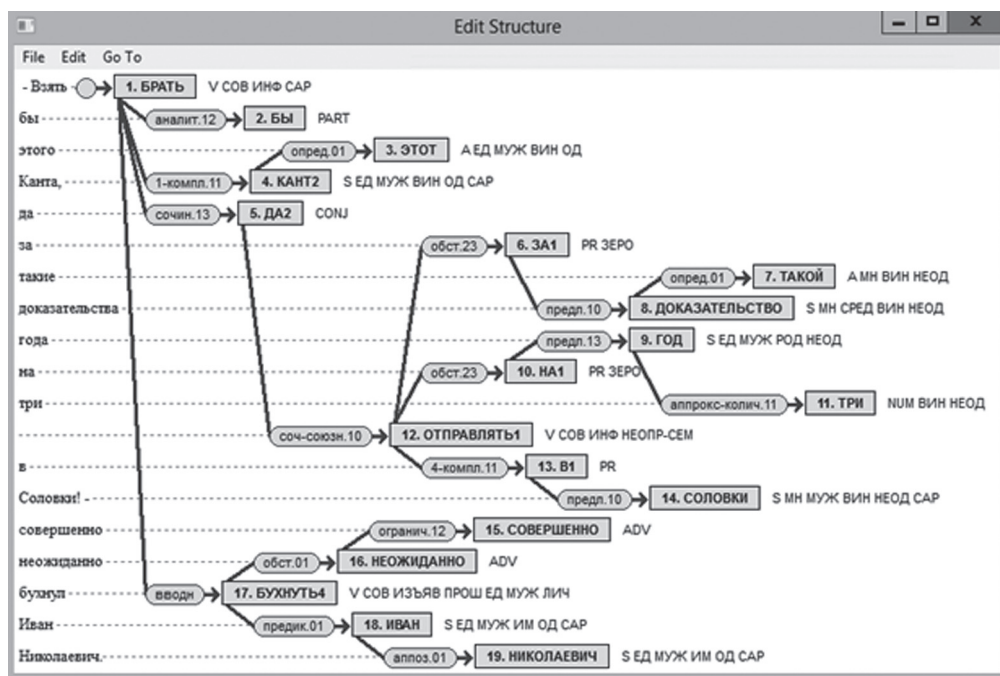


Рис. 6. Структура предложения (4)

Как и в случаях со стандартными узлами структуры, информацию о лексических свойствах элидированного слова можно получить посредством описанной выше опции обращения к комбинаторному словарю (см. рис. 7).

Содержание словарной статьи ОТПРАВЛЯТЬ1 можно посмотреть, нажав на кнопку «. . .» справа от поля «КС Name».

Поскольку лингвистический процессор ЭТАП-4 не располагает механизмом автоматического восстановления эллипсиса, узлы с пустыми текстовыми элементами («фантомы») добавляются в синтаксическую структуру предложения вручную в ходе ее проверки и коррекции лингвистами (после работы этапа автоматического синтаксического анализа, конструирующего морфосинтаксическую структуру предложения).

В большинстве случаев добавление «фантомов» в структуру не вызывает затруднений, так как они повторяют другие лексические единицы текста (разумеется, с возможной морфологической модификацией). Если же параллелей в окружающем тексте нет, ситуация становится более неопределенной. Именно так обстоит дело в примере (4): подходящим кандидатом в «фантомы» будут многие глаголы: *отправить*, *выслать*, *депортировать*, *загнать* и др. (причем, учитывая личность персонажа текста, глагол *отправить* окажется чересчур мягким). Тем не менее в подобных случаях лингвист, осуществляющий разметку, выбирает по своему

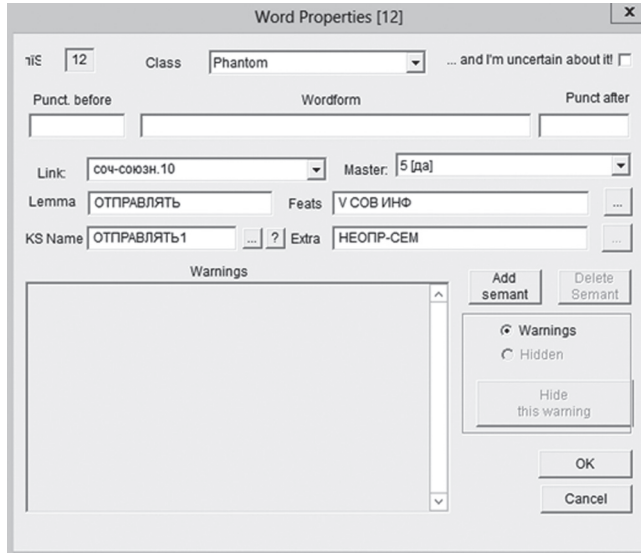


Рис. 7. Карточка узла фантомной словоформы *отправить*, доступная при работе с СинТагРусом над предложением, содержащем слово ОТПРАВЛЯТЬ1 с признаком Phantom

усмотрению один из семантически наиболее нейтральных вариантов и приписывает новому узлу специальный признак «неопр-сем» (неопределенная семантика).

Аналогичным образом разметчики поступают и в других подобных ситуациях. Скажем, в предложении

(5) *Татьяна в лес; медведь за нею...* (А. Пушкин, «Евгений Онегин»)

подходящими кандидатами в «фантомы» будут многие глаголы, означающие быстрое движение: *Татьяна ... бросилась, помчалась, побежала...*; *медведь ... погнался, побежал...*

В настоящее время в корпусе насчитывается около 4 тыс. «фантомов» эллиптической разметки, в том числе 450 с признаком «неопр-сем».

В отдельных случаях при эллиптической разметке корпуса оказывается затруднительным не только зафиксировать конкретное опущенное слово, но даже сколько-нибудь надежно определить семантический класс таких слов. В подобных ситуациях мы прибегаем к лексическому заполнению «фантома» символом неопознанного слова, по возможности фиксируя полный или частичный набор его морфологических характеристик. Примерами могут служить незаконченные предложения СинТагРуса типа

(6) *Поскольку изоляция большой страны с ядерным оружием...*

или

(7) *Между прочим, здоровые люди в это время пьют водку, а я...*

В предложение (6) после слова *оружием* добавляется «фантом», обозначаемый как фиктивный глагол настоящего времени; он синтаксически подчинен подчинительному союзу *поскольку*, а слово *изоляция* подчинено этому «фантому», выступая в роли подлежащего при нем. В предложении (7) «фантом» также помещается в конец предложения и также обозначается как фиктивный глагол настоящего времени. Он синтаксически подчинен сочинительному союзу *а* и, в свою очередь, подчиняет слово *я* в качестве подлежащего.

3. Микросинтаксическая разметка СинТагРуса

Микросинтаксическая разметка СинТагРуса осуществляется разработчиками корпуса начиная с 2016 года. Решение пополнить корпус этой разметкой было обусловлено появлением теории микросинтаксиса (см., в частности, [Иомдин 2008] и [Iomdin 2016]), предлагающей принципы описания широкого круга языковых явлений, лежащих на границе словаря и грамматики. Эта теория имеет много общего с Грамматикой конструкций Ч. Филлмора и его коллег (начиная с ранних работ, таких как [Fillmore et al. 1988]), однако развивалась совершенно независимо от Грамматики конструкций, первоначально на материале русского языка. Следует также добавить, что теория микросинтаксиса и создаваемые на ее основе ресурсы (микросинтаксическая разметка СинТагРуса, Микросинтаксический словарь русского языка) имеют ряд общих черт с разрабатываемым с недавнего времени лингвистическим ресурсом «Русский конструктикон» [Endresen et al. 2020], однако и в данном случае соответствующие подходы и ресурсы развиваются независимо.

Основные принципы и подходы к микросинтаксической разметке СинТагРуса были подробно изложены в [Иншакова и др. 2019] и здесь не повторяются. Мы ограничимся здесь сведениями о современном состоянии этой разметки и о возможностях ее использования в лингвистических исследованиях.

В настоящее время микросинтаксическая разметка распространяется на весь СинТагРус целиком. Число разных микросинтаксических единиц, отраженных в разметке СинТагРуса, составляет 3220. Меньшую часть этих единиц (не более ста) представляют нестандартные синтаксические конструкции (главным свойством которых является синтаксическая специфика и минимальное число конкретных лексических единиц).

Это, например, сформированные дательным падежом и инфинитивом инфинитивно-модальные конструкции — с отрицанием; ср.

- (8) *Ведь никому не понять, что с ней происходит и почему она в троллейбусе* (Ю. Трифонов) ≈ ‘Отсутствует перспектива, что кто-нибудь сможет понять...’

или без него; ср.

- (9) *Мне еще на поезд успеть...* (В. Шукшин) ≈ ‘Я должен успеть на поезд’.

Среди инфинитивно-модальных конструкций особо выделяются конструкции с вопросительными словами, причем вопросительные слова могут быть за преде-

лами самой конструкции, как в (10) или (12), или занимать позицию дательного субъекта в ней, как в (11):

- (10) — *А для чего тебе владеть звездами?* (А. де Сент-Экзюпери, «Маленький принц»);
- (11) *Какому быть главнее?* (М. Шиманский, «Василь Быков: язык — душа народа»);
- (12) — *Проблема двуязычия легче разрешима, скажем, в Латвии или Армении, а как быть уйгуру или дунганину в Киргизии?* (Ibid.)⁷

Другой подкласс нестандартных синтаксических конструкций представлен единицами, содержащими два вопросительных местоимения. Главным компонентом значения этой конструкции является распределение элементов двух множественных объектов, участвующих в ситуации; ср.

- (13) *Важны были другие знания: кто что спрашивает, кому как отвечать (один любит сразу, другой — подумавши), как легче заучить наизусть формулы или формулировку* (И. Грекова, «Кафедра»);
- (14) *Можно отказаться от обсуждения проблем: мало ли что где происходит — кому надо, те разберутся* (Г. Онищенко, «Как закупать дорогостоящее оборудование»).

Целый ряд нестандартных синтаксических конструкций представляют собой единицы с лексическими повторами и с весьма нетривиальным значением; например, $X_{нов}$ не $X_{нов}$:

- (15) *Работай не работай, а денежки идут* (Л. Радзиховский, «Вызов и ответ»);

$X_{нов}$ да $X_{нов}$:

- (16) ... *это все Пашка, пристал ко мне: напиши да напиши* (А. Мамедов, «Дзэн в городе N»).

В некоторых случаях лексические повторы требуют присутствия в контексте других конкретных слов. Таково, например, слово *рознь* в единице XX -у *рознь*:

- (17) *Ошибки практически неизбежно сопровождают любое научное исследование, но ошибки ошибкам рознь* (Е. Александров, Ю. Ефремов, «Как воевать с лженаукой»).

⁷ Не вдаваясь в сложные подробности, отметим, что предложение (12) представляет собой пример весьма частой ситуации, когда две микросинтаксические единицы частично пересекаются: в данном случае это инфинитивно-модальная конструкция рассмотренного только что типа и синтаксическая фраза *как быть*. Аналогичная ситуация имеет место и в предложении (14), где частично пересекается синтаксическая фраза «мало ли + ВОПРОС» и нестандартная синтаксическая конструкция с двумя вопросительными словами.

Подавляющее большинство микросинтаксических единиц, размеченных в Син-TagРусе, относится к классу синтаксических фразем. Среди разрядов таких единиц можно выделить, в частности, 1) отдельные неоднословные выражения, более или менее эквивалентные словам, в том числе многозначные, типа *все равно* (в некоторых значениях), *через раз*, *на ночь глядя* и десятки других; 2) составные союзы и предлоги типа *как только*, *потому что*, *пока что*, *тогда как*; *со стороны*, *в качестве*, *по поводу*, 3) предложно-именные слабокомпозиционные или некомпозиционные словосочетания типа *с виду*, *на вид*, *на виду*, *под видом*, *в бытность*, *в возрасте*, *в миру*, *под рукой*, *в духе* (в разных значениях: (а) 'на принципах', как в *воспитывать в духе толерантности* и (б) 'в определенном настроении', как в *Он сегодня совершенно не в духе*) и т. д.; 4) некомпозиционные словосочетания, принципиально не сводимые к единому слову, такие как *то и дело*, *то ли дело*, *в том числе*, *в упор не видеть* и сотни других).

Микросинтаксическая разметка корпуса осуществляется в «Редакторе структур». На рис. 8 показано, как выглядит эта разметка для фрагмента текста корпуса, а на рис. 9 представлена разметка одного предложения.

В левой части окна, представленного на рис. 8, помещены отдельные предложения текста, а в правой его части отмечены микросинтаксические единицы, встречающиеся в соответствующих предложениях. Например, единица *по тем временам* присутствует в предложении 3 и занимает отрезок от слова 2 до слова 4, а единицы *как только* в предложении 21 и *в одиночку* в предложении 26 занимают

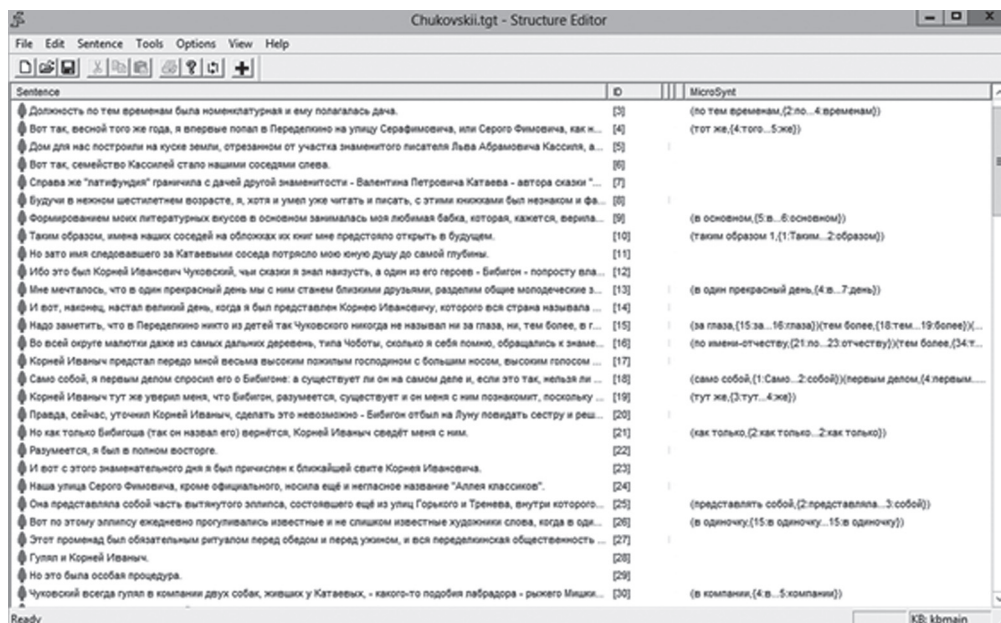


Рис. 8. Микросинтаксическая разметка фрагмента эссе К. Смирнова «Улица Серого Фимовича»

отрезок длиной в одно-единственное слово (2 в предложении 21 и 15 в предложении 26). Последние два случая объясняются тем, что в число микросинтаксических единиц входит большинство безусловных оборотов, которые лингвистический процессор ЭТАП-4 трактует как единые слова.

В нижней части рисунка перечисляются присутствующие в предложении четыре микросинтаксические единицы: составные союзы *да и* и *так что* (последний занимает отрезок в одно слово), временная частица *только что* и местоименное прилагательное тождества *тот же*.

Мы уже отмечали, что микросинтаксическая разметка осуществляется на пространстве всего корпуса СинТагРус. Результаты этой разметки, осуществляемой почти исключительно вручную, можно резюмировать следующими цифрами.

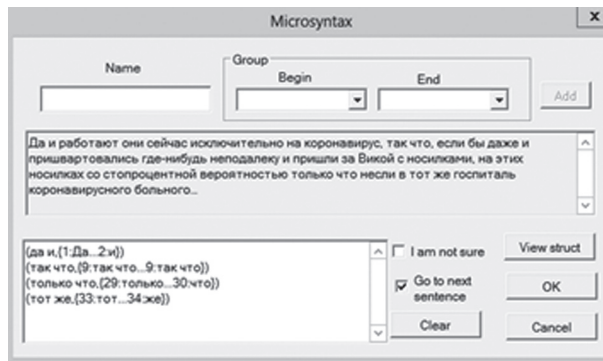


Рис. 9. Микросинтаксическая разметка предложения из очерка Н. Раппопорт «Набережная исцелимых»

Число микросинтаксических единиц, входящих в микросинтаксическую разметку, составляет более 45 600. При этом число предложений, содержащих хотя бы одну микросинтаксическую единицу, составляет около 34 300. Разница между двумя последними цифрами естественно объясняется тем, что многие предложения содержат по несколько микросинтаксических единиц. Доля предложений СинТагРуса, в которых присутствует хотя бы один микросинтаксический элемент, составляет почти 32 %. Этот факт, на наш взгляд, убедительно свидетельствует о высокой степени базовой (нефразеологической) идиоматичности русского текста.

Важно при этом подчеркнуть следующее. Несмотря на то, что объем микросинтаксической разметки СинТагРуса достаточно велик (насколько известно авторам, в мире не существует корпусов текстов со сравнимым количеством детектированных и тщательно атрибутированных в тексте идиоматичных выражений), его все еще недостаточно для полномасштабных исследований функционирования таких единиц в языке. В частности, многие микросинтаксические единицы присутствуют в корпусе в весьма ограниченном количестве: из 3220 таких единиц 1235 (почти 40 %) встречаются в текстах корпуса всего один раз, в то время как количество микросинтаксических единиц, присутствующих в корпусе 50 раз и более, составляет всего 200.

Кроме того, ввиду ограниченности объема всего СинТагРуса, определенные микросинтаксические единицы, в том числе весьма интересные как с синтаксической, так и семантической точек зрения, вообще туда не попали (таковы, например, единицы *под шумок*, *дать дёру* или *вдруг да и*: *Бывают в жизни тупики, говорила я, которые только кажутся тупиками, а вдруг да и расступятся*. (Л. Чуковская, «Предсмертие»)⁸.

Тем не менее, несмотря на недостаточную представительность микросинтаксической разметки корпуса, даже при этом объеме удается провести ряд интересных лингвистических исследований. Завершая раздел о микросинтаксической разметке, мы кратко изложим в качестве примера результаты одного из таких исследований.

При анализе материалов микросинтаксической разметки СинТагРуса обращает на себя внимание следующий факт: среди микросинтаксических единиц обнаруживается неожиданно много выражений, сформированных существительными во втором родительном (будем для краткости именовать его партитивным, независимо от того, передает он значение партитивности или нет) и во втором предложном (местном) падежах. В первую очередь это единицы, состоящие из предлога и управляемого им существительного и в целом синтаксически более или менее эквивалентные наречию (приглагольному или приименному).

Единицы с партитивным падежом включают, например, такие выражения, как *с виду*, *с ходу*, *с лёту*, *с разбегу*, *с налёту*, *с маху*, *с размаху*, *с краю*, *с краешку*, *с перепугу*, *с перепой*, *с глузду*, *без толку*, *без спросу*, *без умолку*, *без удержу*, *из дому*, *от роду*, *от веку*, *до упаду*, *до зарезу*, *до свету* 'до рассвета'.

Легко заметить, что многие существительные, образующие эти выражения, за пределами микросинтаксических единиц не встречаются или почти не встречаются: существительные *перепуг*, *перепой*, *глузд*, *удерж*, *умолк*, *упад*, *зарез* сами по себе в современных текстах не фигурируют.

Кроме того, в число конструкций с партитивом входят и другие единицы, с обязательным или факультативным участием других слов (в частности, с прилагательным — определением к существительному в партитиве, ср. *со всего маху* <*размаху*>, *с какого боку*, *с одного боку*, *с другого боку*, *без всякого толку*, разнообразные конструкции с повторением либо существительного, либо предлога (*с глаза на глаз*, *с боку на бок*, *час от часу*, *с пылу с жару*) и отдельные полилексемные конструкции с предлогом (*баба с возу*, *без году неделя*, *без роду и племени*, *с миру по нитке*, *с бору по сосенке*, *сбивать с толку*, *сбивать с панталыку*, *терять из виду*, *не до жиру*, *беситься с жиру*, *с боку припёка*) или без него (*что толку*, *толку чуть*, *спору нет*, *сносу нет*, *нашего полку прибыло*) и т. д.

Доля таких единиц в общем объеме выражений современного русского языка, содержащих партитив, непропорционально высока. В частности, при поиске в основном корпусе НКРЯ выражений, содержащих **S**, **gen2**, из 100 первых фрагментов текстов, в общей сложности содержащих 159 вхождений партитива, нефразео-

⁸ Следует, впрочем, оговориться, что в процессе многолетней работы над микросинтаксическими единицами в поле зрения авторов оказалось весьма небольшое количество таких единиц, не представленных в корпусе.

логических сочетаний с существительным в партитиве обнаружилось 29 (18 %); во всех остальных случаях мы имеем дело с бесспорными фразеологическими выражениями (почти исключительно с микросинтаксическими конструкциями).

Еще красноречивее оказывается распределение фразеологических и нефразеологических выражений среди сочетаний «предлог + партитив»: 100 первых фрагментов текстов основного корпуса НКРЯ, удовлетворяющих запросу *Pr+gen2*, содержат 248 таких сочетаний, из которых только 8 (3,2 %) являются нефразеологическими (*вместо табаку, [поднял] с полу, [умер] с голоду, после чаю* и некоторые другие).

Единицы со вторым предложным падежом представлены, в частности, выражениями *на ходу, на бегу, на лету, в ходу, в строю, в краю* и т. д. В этих конструкциях вырисовывается похожая картина: доля микросинтаксических единиц среди таких конструкций тоже весьма высока, хотя и не столь впечатляющая (в процентном отношении их несколько меньше).

Так, в первых ста фрагментах текстов, полученных по запросу к основному корпусу НКРЯ *Pr+loc2*, встречается 191 такое словосочетание, из которых не менее 120 (63 %) следует расценивать как микросинтаксические единицы (правда, некоторые из них, например, *в виду* или *в связи с*, встречаются многократно).

По существу, каждая микросинтаксическая единица рассматриваемого здесь типа требует индивидуального исследования — даже с точки зрения взаимодействия предлога с конкретным падежом существительного. Ср., например, микросинтаксическую единицу *на дому* с весьма специфическим значением, указывающим на то, что какая-то профессиональная деятельность осуществляется в месте, где живет человек, занимающийся этой деятельностью (в качестве активного участника или клиента), как в примере

- (18) *Домашние сборища запрещали, но Ляпунов продолжал читать на дому лекции по теории программирования.* (Д. Гранин, «Зубр»)

и нефразеологическое выражение *в дому* в песне А. Галича:

- (19) *Я живу теперь в дому — чаша полная, Даже брюки у меня и те на «молнии».*

И в том, и в другом случае мы встречаем форму второго предложного *дому*, но *на дому* — это микросинтаксическая единица, а *в дому* — это просторечие (может быть, стилизованное).

Конечно, обнаруженные нами закономерности в употреблении второго родительного и второго предложного падежей носят вероятностный, а не абсолютный характер. Ср., например, нефразеологическое выражение *в поту* со вторым предложным, как в примере

- (20) *Весь пол был заляпан белыми пятнами, все мы были в поту.* (М. Булгаков, «Полотенце с петухом»)

и фразеологическую единицу *в поте лица* в выражении типа *трудиться в поте лица*.

Тем не менее в совокупности приведенные факты убедительно свидетельствуют о консервативности фразеологических единиц, массово сохраняющих в языке уходящие формы, и о морфологической устойчивости таких единиц.

Верно и обратное утверждение: уходящие из языка реликты находят своеобразное убежище во фразеологии и тем самым обретают достаточно надежную нишу, в которой продлевают свое присутствие в языке в целом.

4. Разметка кореферентных связей в СинТагРусе

В конце 2023 г. объем размеченной такими связями части СинТагРусе составил более 120 тыс. слов, или около 9200 предложений в 33 текстах.

Особенность текстов с кореферентной разметкой из СинТагРусе заключается в том, что почти все они имеют достаточно большую длину по сравнению с текстами из других корпусов с кореферентной разметкой. На данный момент длина таких текстов в нашем корпусе составляет от 30 до 685 предложений (для сравнения, в корпусе RuCor максимальная длина текста — 170 предложений [Toldova et al. 2014], в тестовом корпусе AnCor [Budnikov et al. 2019] — 336 предложений, в RuCoCo [Dobrovolskii et al. 2022] — 202 предложения, в ARRAU (подкорпус GNOME) [Uryupina et al. 2020] — около 320 предложений, в Пражском корпусе зависимостей [Hajič et al. 2018] — 231 предложение). Это обусловлено выбором для СинТагРусе текстов определенных жанров — публицистических и научно-популярных статей, интервью, глав из художественной литературы и мемуаров (многие из них входят и в Уппсальский корпус русских текстов). Одним из следствий этого являются большие расстояния между некоторыми элементами кореферентных цепочек, которые могут составлять сотни предложений.

Кореферентные связи в СинТагРусе проводятся поверх синтаксической разметки. Поскольку она представляет собой деревья зависимостей без указания границ составляющих, кореферентные связи соединяют словоформы — вершины именных групп или некоторых других составляющих и не содержат информации об их границах (однако существует возможность вычислить эти границы по соответствующим поддеревьям с помощью разработанной в лаборатории программы).

Одна кореферентная связь соединяет два упоминания одной сущности (в направлении от правого к левому, за исключением случаев катафоры) и содержит их координаты в тексте вида $(-y;)x:$, где x — номер словоформы в предложении (обязательная координата), а y — количество предложений между первым и вторым упоминаниями (указывается, если они находятся в разных предложениях). Анафорические связи от возвратных и взаимно-возвратных местоимений проводятся к их ближайшим неместоименным antecedентам, а кореферентные связи от неместоименных существительных и местоимений 3-го лица — к ближайшим неместоименным упоминаниям той же сущности. Если antecedент является сочиненным или расщепленным, то от местоимения проводятся связи к каждому из его элементов. Из отдельных пар упоминаний с помощью специальной программы можно получить упорядоченные множества упоминаний, соответствующие сущностям. Вид

корреферентной разметки фрагмента текста СинТагРуса в окне «Редактора структур» приведен на рис. 10.

Цементные столбики выкрошились, обнажив ржавый каркас, загогулины об...	[280]	
Спускаясь по этим щербатым ступеням, особенно зимой в гололед, Энн ка...	[281]	(19:который,16:архитектора)(5:ступеням,-1:9:ступен
Впрочем, виноват ли был зодчий?	[282]	(5:зодчий,-1:16:архитектора)
Вряд ли он, проектируя дом, входил в психологию старика, которому, спуск...	[283]	(9:которому,8:старика)(2:он,-1:5:зодчий)(4:дом,-4:10;
Город вообще жесток к старикам.	[284]	
Время от времени в связи с какой-нибудь датой ЖЭК срочно проводил "ко...	[285]	
В обычное время, между датами, дом стоял страшноватый, как престарел...	[286]	(6:дом,-3:4:дом)
К проживанию он был мало приспособлен, главным образом из-за шума.	[287]	(3:он,-1:6:дом)
Обращенный всей своей парадностью на крупную магистраль, он день и но...	[288]	(3:своей,8:он)(8:он,-2:6:дом)
Оконные рамы дрожали, посуда подпрыгивала, с потолка сыпались хлопья...	[289]	
Зимой заклеенные окна чуть-чуть умеряли шум; летом он становился невы...	[290]	(8:он,6:шум)(6:шум,-3:10:шума)
Из четырех комнат квартиры обитаемыми были, в сущности, только две, са...	[291]	(29:Энн,-8:8:старика)
Две большие комнаты, окнами на проспект, как говорил Энн, были заняты...	[292]	(12:шумом,-2:6:шум)(9:Энн,-1:29:Энн)(6:проспект,-4
Он, впрочем, привык к своему дому и даже в каком-то смысле его любил.	[293]	(5:своему,1:Он)(1:Он,-1:9:Энн)(12:его,6:дому)

Рис. 10. Фрагмент СинТагРуса с корреферентной разметкой

В корпусе размечаются следующие виды связей:

1) анафорические связи:

- от возвратных (*себя, свой*) и взаимно-возвратных (*друг друга / друг дружку*) местоимений — в любых случаях, включая связанную анафору с кванторными antecedентами (...*никто_i из них, даже кляня свою_i судьбу, не поменял бы ее ни на чью другую*);
- от местоимений 3-го лица, включая «местоимение переключения референции» *тот*, — везде, за исключением случаев, когда местоимение и его antecedент отсылают к разным референтам (*Во-вторых, создать условия, вынуждающие руководителей предприятий и организаций не заказывать лишних специалистов, а эффективнее их_i использовать*);
- от всех видов относительных местоимений, включая наречные (*который, кто, что, чей, какой, где, куда, откуда, когда*);
- от местоимений-наречий *здесь, тут, там, туда, оттуда, тогда*. Если их antecedентом является не существительное или наречие, а предложная группа, то корреферентная связь проводится к предлогу (*в_i доме / у_i дома / за_i домом ↔ там_i*);
- от местоимения *это*, если у него есть именной antecedент (...*они_i сами, и вырытые ими окопы_j — все это_{i+j} было уже в тылу у немцев*).

2) корреферентные связи:

- а. между неместоименными существительными, а также прилагательными и причастиями, которые выполняют функции существительных (*Тот черненький_i, интеллигентненький, что примечал его при регистрации как своего, прямо направился теперь к Монахову как к своему. <...> Он искал глазами того_i, черненького, и не нашел*) или остаются после эллипсиса существительного-вершины в именной группе (*Одна пушка, брошенная всеми, завалилась набок ... <...> Около другой_i суеился расчет. Они на руках скатывали ее_i вниз, надеясь открыть огонь...*). Они могут быть предметными, событийными или обозначениями отрезков времени и иметь

любой референциальный статус, кроме предикативного, при котором именная группа не отсылает ни к какой сущности или классу сущностей (*Лева рос в так называемой академической среде и с детства мечтал стать ученым; Он ринулся в самую гущу толпы. И вышел победителем в борьбе за трап*). В частности, в нашем корпусе проводятся кореферентные связи между родовыми именными группами, отсылающими к одному и тому же классу (*Поэтому очень важно снизить металлоемкость полиграфической техники, тем самым добиваясь наибольшей экономии металла. А нельзя ли придумать такой способ печати, чтобы обходиться вообще без металла?₁?; — Лес₁, оказывается, — с надеждой продолжил Монахов, — не просто много деревьев... И не каждое дерево, а лес₁ в целом существует, как единый организм...*);

- b. между существительными и притяжательными прилагательными или некоторыми видами относительных прилагательных (*Люда ↔ Людин; Сибирь ↔ сибирский*);
- c. между местоимениями 1-го и 2-го лица и другими упоминаниями (*Тогда Иценко₁ закричал: — Ты что, проверяешь меня₁?*).

Разметка кореферентных связей в СинТагРусе проводится одним лингвистом в полуавтоматическом режиме: для каждого предложения сначала запускаются правила разрешения анафоры и кореферентности в системе ЭТАП-4, а если они устанавливают неправильные связи, разметчик исправляет их вручную (также вручную размечаются длинные связи, которые правила ЭТАП-4 не могут установить). После этого проводится автоматическая проверка ошибок разметки, а затем размеченный текст проверяется другим лингвистом.

На данный момент подкорпус СинТагРуса с кореферентной разметкой может использоваться в исследовательских целях, а в перспективе, по мере его увеличения, — и для обучения систем разрешения анафоры и кореферентности. О режиме доступа к данному подкорпусу см. <https://ruscorpora.ru/page/corpora-datasets/>.

Литература

Апресян Ю. Д., Богуславский И. М., Иомдин Б. Л., Иомдин Л. Л., Санников А. В., Санников В. З., Сизов В. Г., Цинман Л. Л. Синтаксически и семантически аннотированный корпус русского языка: современное состояние и перспективы // Национальный корпус русского языка: 2003–2005. Результаты и перспективы. М.: Индрик, 2005. С. 193–214.

Богуславский И. М., Иомдин Л. Л., Митюшин Л. Г., Сизов В. Г. Длина синтаксических связей в русском аннотированном корпусе // Международная конференция «Корпусная лингвистика — 2008», Санкт-Петербург, 2008. С. 75–82.

Дяченко П. В., Иомдин Л. Л., Лазурский А. В., Митюшин Л. Г., Подлесская О. Ю., Сизов В. Г., Фролова Т. И., Цинман Л. Л. Современное состояние глубоко аннотированного корпуса текстов русского языка (СинТагРус) // Труды Института русского языка им. В. В. Виноградова / Национальный корпус русского языка: 10 лет проекту. 2015. Вып. 6. С. 272–299.

Инструкция 2023: Синтаксическая разметка [Электронный ресурс]. URL: <https://ruscorpora.ru/page/instruction-syntax>

Иниакова Е. С., Иомдин Л. Л., Митюшин Л. Г., Сизов В. Г., Фролова Т. И., Цинман Л. Л. СинТагРус сегодня. // Труды Института русского языка им. В. В. Виноградова. 2019. № 21. С. 14–41. DOI: 10.31912/pvrl-2019.21.1

Маракасова А. А., Иомдин Л. Л. Микросинтаксическая разметка в корпусе русских текстов СинТагРус // Информационные технологии и системы 2016 (ИТиС'2016). Труды 40-й междисциплинарной школы-конференции ИППИ РАН. СПб., 2016. С. 445–449.

Савчук С. О. Метатекстовая разметка в Национальном корпусе русского языка: базовые принципы и основные функции // Национальный корпус русского языка: 2003–2005. Результаты и перспективы. М.: Индрик, 2005. С. 62–88.

Тимошенко С. П., Иомдин Л. Л., Гладили С. А., Иниакова Е. С. СинТагРус в составе НКРЯ: новые возможности // Труды международной конференции «Корпусная лингвистика-2021» (г. Санкт-Петербург, 01–03 июля 2021 г.). СПб.: Изд-во Санкт-Петербургского университета, 2021. С. 31–43.

Шеманаева О. Ю., Фролова Т. И. Лексико-функциональная разметка текстов в СинТагРус // Информационные технологии и системы 2010 (ИТиС'10). Труды 33-й конференции молодых ученых и специалистов ИППИ РАН. М.: ИППИ, 2010. С. 320–324.

Boguslavsky I. SynTagRus — a Deeply Annotated Corpus of Russian // Les émotions dans le discours. Emotions in Discourse / ed. by P. Blumenthal, I. Novakova, and D. Siepmann. Peter Lang Edition, 2014. pp. 367–381.

Boguslavsky I., Grigorieva S., Grigoriev N., Kreidlin L., Frid N. Dependency Treebank for Russian: Concept, Tools, Types of Information // Proc. of the 18th International Conference on Computational Linguistics (COLING 2000). San Francisco, Kaufmann, 2000. pp. 987–991.

Budnikov A., Toldova S., Zvereva D., Maximova D., Ionov M. Ru-Eval-2019: Evaluating anaphora and coreference resolution for Russian // Computational Linguistics and Intellectual Technologies. Supplementary Volume, 2019. pp. 2–13.

Chaga A. On a specific Russian construction with saturative verbs and negation // Annual International Conference “Dialogue” 2021, Student Session, Moscow, 2021.

Dobrovolskii V., Michurina M., Ivoylova A. RuCoCo: a new Russian corpus with coreference annotation // Computational Linguistics and Intellectual Technologies, 2022. pp. 141–149.

Endresen A. A., Zhukova V. A., Mordashova D. D., Rakhilina E. V., Lyashevskaya O. N. The Russian Constructicon: A New Linguistic Resource, Its Design and Key Characteristics // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue” (2020). Issue 19 (26). pp. 241–255.

Fillmore Charles J., Kay Paul, O'Connor Mary Catherine. Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone // Language. Vol. 64. No. 3 (Sep., 1988). P. 501–538.

Hajič J., Bejček E., Bémová A., Buráňová E., Hajičová E., Havelka J., Homola P., Kárník J., Kettnerová V., Klyueva N., Kolářová V., Kučová L., Lopatková M., Mikulová M., Mirovský J., Nedoluzhko A., Pajas P., Panevová J., Poláková L., Rysová M., Sgall P., Spoustová J., Straňák P., Synková P., Ševčíková M., Štěpánek J., Urešová Z., Vidová Hladká B., Zeman D., Zikánová Š., Žabokrtský Z. Prague Dependency Treebank 3.5 [Электронный ресурс]. URL: <http://hdl.handle.net/11234/1-2621>.

Iomdin L. Microsyntactic Phenomena as a Computational Linguistics Issue // Grammar and Lexicon: Interactions and Interfaces. Proc. of the Workshop. Osaka, 2016, pp. 8–18. Available at: <http://aclweb.org/anthology/W/W16/W16-38.pdf>.

Iomdin L. Microsyntactic Annotation of Corpora and its Use in Computational Linguistics Tasks // Jazykovedný časopis, ročník 86, číslo 2, 2017. pp. 169–178.

Iomdin L., Sizov V. Structure Editor: a Powerful Environment for Tagged Corpora // MONDILEX Fifth Open Workshop. Ljubljana, 2009. pp. 1–12.

Toldova S., Roytberg A., Nedoluzhko A., Kurzukov M., Ladygina A., Vasilyeva M., Azerkovich I., Grishina Y., Sim G., Ivanova A., Gorshkov D. Evaluating Anaphora and Coreference Resolution for Russian // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue” (2014). Issue 13 (20). pp. 681–695.

Uryupina O., Artstein R., Bristot A., Cavicchio F., Delogu F., Rodriguez K., Poesio M. Annotating a broad range of anaphoric phenomena, in a variety of genres: The ARRAU Corpus // Natural Language Engineering, 26 (1), 2020. pp. 95–128.

**Igor M. Boguslavsky, Alexandra V. Chaga, Pavel V. Djachenko,
Tatyana I. Frolova, Evgenia S. Inshakova, Leonid L. Iomdin,
Alexander V. Lazurski, Leonid G. Mityushin,
Andrey A. Movsesyan, Ivan P. Rygaev, Victor G. Sizov,
Svetlana P. Timoshenko**

*A. A. Kharkevich Institute for Information Transmission Problems, RAS
(Moscow)*

*igor.m.boguslavsky@gmail.com, chagachaga@gmail.com,
pavel.v.djachenko@gmail.com, tfrolova@gmail.com, e.s.inshakova@gmail.com,
iomdin@gmail.com, lazurski@mail.ru, lmityushin@gmail.com, derise@iitp.ru,
irygaev@jent.ru, victor.sizov@gmail.com, nyrestein@gmail.com*

THE CURRENT STATE OF THE SYNTAGRUS CORPUS

The paper presents a description of the main features and options of a diversely tagged corpus of Russian texts called SynTagRus. The corpus has been developed by the A. A. Kharkevich Institute for Information Transmission Problems, RAS, and is currently considered to be a subcorpus of RNC, where it is referred to as the “Syntactic Corpus”.

Much attention is given to the linguistic principles underlying the different annotation types: morphological, syntactic, lexical semantic, lexical functional, elliptical, microsyntactic, coreferential, and temporal. Statistical data are given which characterize a variety of aspects of SynTagRus and its fragments. SynTagRus is a corpus with a 100-percent disambiguation at all levels of annotation. The paper outlines the obvious advantages of this approach but at the same time notes the difficulties associated with the need to always make definite decisions and choose single annotation options even in cases when the linguistic material undeniably allows for multiple linguistic description. Much attention is given to certain differences that exist between the SynTagRus and the main RNC subcorpora, such as distribution of words by parts of speech or specific morphological solutions that are accepted in SynTagRus in contrast to RNC (e.g. individual morphological categories, like verbal aspect and voice, certain cases of nouns etc.).

Keywords: SynTagRus, Syntactic Corpus, Morpho-syntactic Tagging, Lexical Tagging, Elliptical Tagging, Microsyntax, Coreference Tagging, Temporal Tagging

References

- Apresjan Ju. D., Boguslavsky I. M., Iomdin B. L., Iomdin L. L., Sannikov A. V., Sannikov V. Z., Sizov V. G., Cinman L. L. [Syntactically and semantically tagged corpus of Russian: state of the art and prospects]. *Natsional'nyi korpus russkogo yazyka: 2003–2005* [The Russian National Corpus: 2003–2005. Results and Prospects]. Moscow, Indrik Publ., 2005, pp. 193–214. (In Russ.)
- Boguslavsky I. SynTagRus — a Deeply Annotated Corpus of Russian. *Les émotions dans le discours. Emotions in Discourse* / ed. by P. Blumenthal, I. Novakova, and D. Siepmann. Peter Lang Edition, 2014. P. 367–381.
- Boguslavsky I., Grigorieva S., Grigoriev N., Kreidlin L., Frid N. Dependency Treebank for Russian: Concept, Tools, Types of Information. *Proc. of the 18th International Conference on Computational Linguistics (COLING 2000)*. San Francisco, Kaufmann, 2000, pp. 987–991.
- Boguslavsky I. M., Iomdin L. L., Mitjushin L. G., Sizov V. G. [The length of syntactic links in the Russian tagged corpus]. *Mezhdunarodnaya konferentsiya "Korpusnaya lingvistika — 2008"* [Proc. of the International Conference "Corpus Linguistics — 2008"]. St. Petersburg, 2008b, pp. 75–82. (In Russ.)
- Budnikov A., Toldova S., Zvereva D., Maximova D., Ionov M. Ru-Eval-2019: Evaluating anaphora and coreference resolution for Russian. *Computational Linguistics and Intellectual Technologies. Supplementary Volume*. Moscow, RSUH Publ., 2019, pp. 2–13.
- Chaga A. On a specific Russian construction with saturative verbs and negation. *Annual International Conference "Dialogue" 2021*, Student Session, Moscow, 2021.
- Dobrovolskii V., Michurina M., Ivoylova A. RuCoCo: a new Russian corpus with coreference annotation. *Computational Linguistics and Intellectual Technologies*. Moscow, RSUH Publ., 2022, pp. 141–149.
- Dyachenko P. V., Iomdin L. L., Lazursky A. V., Mityushin L. G., Podlesskaya O. Yu., Sizov V. G., Frolova T. I., Tsinman L. L. [A deeply annotated corpus of Russian texts (SynTagRus): contemporary state of affairs]. *Natsional'nyi korpus russkogo yazyka:*

10 let proektu. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova. Vyp. 6* [The Russian National Corpus: 10 Years of the Project. Proc. of the V. V. Vinogradov Russian Language Institute. Iss. 6]. Moscow, 2015, pp. 272–299. (In Russ.)

Endresen A. A., Zhukova V. A., Mordashova D. D., Rakhilina E. V., Lyashevskaya O. N. The Russian Constructicon: A New Linguistic Resource, Its Design and Key Characteristics. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue"* (2020). Issue 19 (26). Moscow, RSUH Publ., 2020, pp. 241–255.

Fillmore Charles J., Kay Paul, O'Connor Mary Catherine. Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone. *Language*, vol. 64, no. 3 (Sep., 1988), pp. 501–538.

Hajič, J., Bejček, E., Bémová, A., Buráňová, E., Hajičová, E., Havelka, J., Homola, P., Kárník, J., Kettnerová, V., Klyueva, N., Kolářová, V., Kučová, L., Lopatková, M., Mikulová, M., Mírovský, J., Nedoluzhko, A., Pajas, P., Panevová, J., Poláková, L., Rysová, M., Sgall, P., Spoustová, J., Straňák, P., Synková, P., Ševčíková, M., Štěpánek, J., Urešová, Z., Vidová Hladká, B., Zeman, D., Zikánová, Š., Žabokrtský, Z. *Prague Dependency Treebank 3.5*. Available at: <http://hdl.handle.net/11234/1-2621> (accessed 10.10.2023)

Inshakova E. S., Iomdin L. L., Mitjushin L. G., Sizov V. G., Frolova T. I., Cinman L. L. [SynTagRus today]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova* [Proc. of V. V. Vinogradov Russian Language Institute]. Vol. 21, Moscow, 2019, pp. 14–41. (In Russ.)

Instrukcija 2023: Sintaksicheskaja razmetka [Instruction 2023: Syntactic annotation]. Available at: <https://ruscorpora.ru/page/instruction-syntax> (accessed 10.10.2023)

Iomdin L. Microsyntactic Phenomena as a Computational Linguistics Issue // *Grammar and Lexicon: Interactions and Interfaces. Proc. of the Workshop*. Osaka, 2016, pp. 8–18. Available at: <http://aclweb.org/anthology/W/W16/W16-38.pdf>.

Iomdin L. Microsyntactic Annotation of Corpora and its Use in Computational Linguistics Tasks. *Jazykovedný časopis, ročník 86, číslo 2*, 2017, pp. 169–178.

Iomdin L., Sizov V. Structure Editor: a Powerful Environment for Tagged Corpora // *MONDILEX Fifth Open Workshop*. Ljubljana, 2009, pp. 1–12.

Marakasova A. A., Iomdin L. L. [Microsyntactic tagging in the SynTagRus corpus of Russian texts]. *Informatsionnye tekhnologii i sistemy 2016 (ITiS'2016)*. *Trudy 40-i mezhdistsiplinarnoi shkoly-konferentsii IPPI RAN* [Information Technologies and Systems 2016 (ITiS'2016). Proc. of the 40th Interdisciplinary School-Conference of IITP RAS]. St. Petersburg, 2016, pp. 445–449. (In Russ.)

Savchuk S. O. [Metatext markup in the National Corpus of the Russian Language: basic principles and main functions]. *Nacional'nyj korpus russkogo yazyka: 2003–2005. Rezul'taty i perspektivy*. [Russian National Corpus: 2003–2005. Results and prospects.]. Moscow, Indrik Publ., 2005, pp. 62–88. (In Russ.)

Shemanaeva O. Yu., Frolova T. I. [Tagging with lexical functions in SynTagRus]. *Informacionnye tekhnologii i sistemy 2010 (ITiS'10)*. *Trudy 33-i Konferencii molodykh uchenykh i spetsialistov IPPI RAN* [Information Technologies and Systems 2010. Proc.

of the 33rd Conference of Young Scientists and Specialists of IITP RAS]. Moscow, IITP, 2010, pp. 320–324. (In Russ.)

Timoshenko S. P., Iomdin L. L., Gladilin S. A., Inshakova E. S. [SynTagRus as a part of RNC: new perspectives]. *Trudy mezhdunarodnoi konferentsii “Korpusnaya lingvistika-2021”* [Proc. Int. Conf. “Corpus Linguistics-2021”]. Saint Petersburg, 2021, pp. 31–43. (In Russ.)

Toldova S., Roytberg A., Nedoluzhko A., Kurzukov M., Ladygina A., Vasilyeva M., Azerkovich I., Grishina Y., Sim G., Ivanova A., Gorshkov D. Evaluating Anaphora and Coreference Resolution for Russian. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”* (2014). Issue 13 (20). Moscow, RSUH Publ., 2014, pp. 681–695.

Uryupina O., Artstein R., Bristot A., Cavicchio F., Delogu F., Rodriguez K., Poesio M. Annotating a broad range of anaphoric phenomena, in a variety of genres: The ARRAU Corpus. *Natural Language Engineering*, 26(1), 2020, pp. 95–128.

С. О. Савчук

Институт русского языка им. В. В. Виноградова РАН

(Россия, Москва)

savsvetlana@mail.ru

ОБУЧАЮЩИЙ КОРПУС В СОСТАВЕ НКРЯ КАК ИНСТРУМЕНТ УЧИТЕЛЯ-СЛОВЕСНИКА

Обучающий корпус предназначен для преподавания русского языка и литературы в школе, для подготовки учителей-словесников в педагогических институтах, а также для преподавания русского языка иностранцам. Обучающий корпус построен на иных принципах, отличающих его от основного корпуса письменных текстов. Отличия касаются сбалансированности текстов в составе корпуса, соответствия их школьной программе и современным орфографическим нормам.

Объем корпуса превышает 13 млн словоупотреблений. В соответствии со школьным курсом русского языка в корпусе представлены тексты, относящиеся к различным сферам употребления языка: художественная литература, публицистика, деловая, научная, разговорная речь. Большую часть корпуса составляют произведения, входящие в программу по литературе для основной и старшей школы, в том числе рекомендованные для внеклассного чтения.

Морфологическая разметка в Обучающем корпусе адаптирована к школьной программе по русскому языку и обеспечивает именно традиционный, упрощенный грамматический разбор. В целях соответствия задачам школьного преподавания в разметку введены дополнительные морфологические признаки — словоизменяемые типы имен и глаголов (склонения, спряжения) и лексико-грамматические разряды существительных, прилагательных, местоимений, наречий. В корпусе использована как морфологическая разметка с ручным снятием омонимии, так и морфологическая и синтаксическая разметка с автоматическим снятием неоднозначности на основе нейросетевых алгоритмов.

Новая разметка делает доступными все новейшие функциональные возможности: «Поиск коллокаций», «Частотный словарь», «Частотность выдачи», «Портрет слова», включающий «Скетчи» и «Похожие слова» и др., которые призваны повысить привлекательность корпуса для пользователя и превратить его в необходимый инструмент учителя-словесника.

Ключевые слова: Обучающий корпус, состав и структура корпуса, морфологическая разметка, синтаксическая разметка, новые функциональные возможности, использование Обучающего корпуса в преподавании

1. Создание Обучающего (педагогического) корпуса

Одной из основных сфер практического применения лингвистических корпусов является практика преподавания языка: практически во всех странах, где имеются национальные корпуса, существуют специальные учебники и программы, ориентированные на корпус. Преподаватель испытывает необходимость искать наглядные и реальные примеры на конкретную лексику или грамматические конструкции, в том числе и для того, чтобы разрабатывать на основе этих примеров соответствующий тренировочный и контрольный материал. Для преподавателя русского языка источником таких примеров является корпус русских текстов, прежде всего, Национальный корпус русского языка (НКРЯ), в котором собраны тексты, представляющие русский язык во всём многообразии его функциональных стилей, жанров и эпох. Действительно, с самого начала своего создания Национальный корпус русского языка стал широко применяться в обучении русскому языку в России и за рубежом. Свидетельством этого стало большое количество публикаций, конференций, мастер-классов и других мероприятий по использованию Корпуса в образовании [Герасимова, Грудева 2014; Добрушина 2006; Добрушина, Левинзон 2006; Левинзон 2014; Крылова 2022; Крылова 2023; Ольховская 2019; Ольховская, Парамонова 2019; Рычкова, Киеня 2010; Фомичева 2015 и др.].

В 2007 году в составе НКРЯ был создан Обучающий корпус русского языка — небольшой корпус со снятой омонимией, ориентированный на преподавание русского языка в школе [Савчук, Сичинава 2009]. Если очерчивать область практического использования этого корпуса, то он предназначен для школьного и вузовского (на непрофильных факультетах, например, факультете журналистики) преподавания русского языка и литературы носителям языка, а также для преподавания языка иностранцам и подготовки преподавателей русского языка в педагогических институтах.

Со временем практика использования Обучающего корпуса показала, что при всех своих достоинствах он нуждается в развитии и совершенствовании, поэтому была поставлена задача модернизации Обучающего корпуса с целью превращения его в важный и необходимый инструмент учителя-словесника.

Формулируя концепцию модернизации, исходили из того, что основные принципы построения корпуса, проверенные временем, останутся неизменными, при этом будут учтены замечания и пожелания преподавателей и расширены полезные свойства и функциональные возможности корпуса. В качестве основных целей модернизации Обучающего корпуса были намечены следующие: увеличение объема корпуса, регулярное пополнение состава текстов, расширение сферы его исполь-

зования. Для достижения этих целей были решены следующие задачи, о которых пойдет речь в настоящей статье:

- обновление и расширение состава текстов;
- совершенствование лингвистической разметки текстов;
- создание новых функциональных возможностей и дополнительных ресурсоспутников.

2. Обновление и расширение состава текстов

В процессе подготовки Обучающего корпуса русского языка определились следующие принципы отбора материала, которые отличают его от других корпусов в составе НКРЯ [Савчук, Сичинава 2009, 320–322].

Сбалансированность корпуса. Поскольку школьный курс русского языка предполагает обучение учащегося всем функциональным разновидностям русской речи, корпус включает в себя тексты, принадлежащие к различным сферам употребления языка: в нем представлены художественная литература, публицистика, деловая, научная, разговорная речь. Однако принцип пропорциональной балансировки текстов различных функциональных сфер здесь иной, чем в Основном корпусе в составе Национального корпуса русского языка. «Литературоцентричность» обучения в российской школе связана с тем, что учебные предметы «Русский язык» и «Литература» составляют единую предметную область «Русский язык и литература», а литература рассматривается как «высшая форма существования российской духовности и языка»¹. Поэтому объём художественных текстов в Обучающем корпусе кратно превышает объём других текстов. Публицистика, учебно-научные, деловые тексты, образцы разговорного стиля представлены в корпусе довольно широко и разнообразно, в том объёме, который предусмотрен школьной программой.

Подбор текстов и школьная программа. Выбор образцовых художественных текстов для включения в корпус применительно к школьному преподаванию определяется существующей школьной программой, которая задаёт хрестоматийный канон как имён, так и текстов. Подкорпус художественной литературы в Обучающем корпусе включает в себя преимущественно те произведения, которые входят в школьный курс по русской литературе XVIII–XXI вв., а также в программу по литературному чтению для начальной школы. Круг произведений максимально расширен, учтен вариант программ углубленного уровня для средней и старшей школы, а также рекомендации по внеклассному чтению.

Орфографические особенности текстов. Тексты, входящие в Обучающий корпус, с точки зрения орфографических и пунктуационных характеристик должны соответствовать требованиям современной русской орфографии и пунктуации. Однако реальные тексты, изданные до 1956 года, от этого идеала отклоняются, они

¹ Концепции преподавания русского языка и литературы в РФ (утв. распоряжением Правительства РФ от 9 апреля 2016 г. N 637-р).

отражают орфографические нормы своего времени. Это же касается академических изданий классиков художественной литературы XIX-го — начала XX в., поскольку в них воссоздается «авторская орфография», отчего возникает связанная с ней проблема, хорошо известная каждому школьному учителю. Поэтому в качестве письменных источников электронных версий текстов Обучающего корпуса предпочтительнее выбирать школьные издания (серии «Школьная библиотека», «Классики и современники» и др.), в большинстве случаев нормализующие орфографию, а в ряде случаев и пунктуацию классиков. Что касается бытовых текстов, текстов публицистики (восходящих к электронным версиям газет) и иных нехудожественных жанров, то здесь составители корпуса по мере обнаружения недочётов выполняют функцию корректора (там, где публикации не было или же она оказалась почему-либо неудовлетворительной с точки зрения целей Обучающего корпуса).

Расширение состава Обучающего корпуса. Пополнение Обучающего корпуса в целом опиралось на сформулированные принципы. Однако, в отличие от ограниченного по объёму первоначального ядра корпуса, включавшего большие по объёму произведения в отрывках, при расширении корпуса были сняты количественные ограничения: все тексты школьной программы включаются в корпус целиком, нет ограничений и на количество произведений одного и того же автора в корпусе. Такое решение открывает возможность использовать корпус для исследований языка и стиля русских классиков и отдельных произведений (пример см. ниже в разделе 4.3).

Отбор текстов для пополнения Обучающего корпуса начался с изучения учебных программ по русскому языку и литературе для основной, начальной и старшей школы. Были рассмотрены также варианты рабочих программ по литературе, принятых в отдельных учебных заведениях, что помогло конкретизировать те положения стандартной программы по литературе, которые сформулированы как рекомендации, оставляющие выбор авторов или произведений при изучении каких-то тем на усмотрение учителя.

В результате изучения программ по литературе были составлены списки авторов и произведений, в том числе и для внеклассного чтения, с указанием классов, в которых эти произведения изучаются или рекомендованы к прочтению. На основании этих списков был осуществлён отбор текстов для Обучающего корпуса. Часть из них была представлена в основном корпусе письменных текстов, поэтому им требовалась только проверка на соответствие орфографическим и пунктуационным нормам и дополнительное редактирование. Значительная часть текстов (более 2 млн словоупотреблений), отсутствующая в НКРЯ, подготовлена для Обучающего корпуса с нуля: была произведена оцифровка изданий, редактирование электронных изданий, метатекстовая разметка структуры текстов, разметка xml версий.

В Программах по русскому языку было проанализировано содержание по разделам «Язык и речь», «Система языка», «Текст», «Функциональные разновидности языка», «Культура речи» с целью выделения разновидностей речи, функциональных

стилей и речевых жанров, изучаемых в школе. В соответствии с требованиями школьной программы был определен состав стилей и жанров, изучаемых в школе, и производился отбор текстов для наполнения корпуса.

Метатекстовая разметка. Каждый текст Обучающего корпуса характеризуется с помощью системы признаков, которые называются метатекстовой разметкой. Метаразметка текста позволяет пользователю (при помощи интерфейса на русском языке) отобрать подкорпус по релевантным параметрам. Среди этих параметров дата рождения автора, дата создания текста, функциональная сфера и жанр, тип текста, место и время описываемых событий, источник текста. Характеристика аудитории включает размер, возраст и уровень подготовки аудитории. В Обучающем корпусе для характеристики возраста аудитории используются признаки «начальная школа» (1–4 классы), «средняя школа» (5–9 классы), «старшая школа» (10–11 классы), то есть указывается период обучения, который определен школьной программой для изучения соответствующего текста. Произведения многих авторов изучаются в школе дважды: первое знакомство в отрывках происходит в начальном или среднем звене, а полностью произведение изучается в курсе литературы в 9–11 классах. В таком случае указывается период первого знакомства с целым произведением. В частности, рассказы И. С. Тургенева из цикла «Записки охотника» изучаются в 7 классе, роман в стихах А. С. Пушкина «Евгений Онегин» — в 9 классе, поэтому имеют помету «средняя школа».

Метатекстовая разметка позволяет также разработчикам корпуса контролировать его состав. Состав текстов Обучающего корпуса представлен в Таблице 1.

Таблица 1. Состав текстов Обучающего корпуса

Разделы Обучающего корпуса	Кол-во текстов	Кол-во словоупотреблений	Доля в корпусе
Художественные тексты	470	12 120 540	88,1 %
Публицистические тексты (включая мемуары)	506	870 880	6,3%
Учебно-научные тексты	140	401 179	2,9%
Официально-деловые тексты	57	227 258	1,7%
Разговорные тексты (устная публичная, непубличная речь, дневники, бытовая переписка)	71	141 397	1,0%
Всего	1244	13 761 254	100%

1. Художественные произведения

Как видно из Таблицы 1, художественные тексты, рекомендуемые школьной программой по литературе, составляет 88% состава корпуса. По периодам тексты распределяются следующим образом.

XVIII век — 4 текста, 87 595 словоупотреблений. Представлены произведения Д. И. Фонвизина, Н. М. Карамзина, А. Погорельского, А. Н. Радищева.

XIX век — 159 текстов, 3 585 568 словоупотреблений. Среди авторов С. Т. Аксаков, В. М. Гаршин, А. И. Герцен, Н. В. Гоголь, И. А. Гончаров, Ф. М. Достоевский,

М. Ю. Лермонтов, Н. С. Лесков, Д. Н. Мамин-Сибиряк, А. Н. Островский, А. С. Пушкин, М. Е. Салтыков-Щедрин, Л. Н. Толстой, И. С. Тургенев, А. П. Чехов и др.

XX век — 455 текстов, 11 782 440 словоупотреблений, при этом тексты первой и второй половины XX века представлены примерно поровну: 1-я пол. XX в. — 187 текстов 4 191 164 слова, 2-я пол. XX в. — 339 текстов 5 053 094 слова. Среди авторов В. П. Аксенов, Л. Н. Андреев, В. П. Астафьев, Г. Я. Бакланов, М. А. Булгаков, И. А. Бунин, Г. Газданов, Ю. О. Домбровский, В. Ю. Драгунский, М. М. Зощенко, Э. Г. Казакевич, В. Г. Короленко, А. И. Куприн, К. Г. Паустовский, А. П. Платонов, М. М. Пришвин, В. Г. Распутин, Т. Н. Толстая, А. Н. Толстой, В. Т. Шаламов, В. М. Шукшин и др.

XXI век — 14 текстов, 315 179 словоупотреблений. Включены тексты Л. Волковой, Д. Данилова, Н. Дашевской, В. Крапивина, Т. Михеевой, З. Прилепина и др.

Такая широкая представленность в корпусе художественных текстов позволяет использовать Обучающий корпус при изучении языка художественной литературы, стиля писателя, в том числе в проведении сравнительных исследований, например, выразительно-изобразительных средств художественной прозы и поэзии и под.

2. Нехудожественные тексты

Публицистические тексты составляют значительную часть корпуса. Помимо публикаций из газет и журналов XX — начала XXI в., в корпус включены очерки и мемуарно-биографическая проза А. Ф. Кони, Даниила Гранина, Ирины Архиповой, Игоря Кио и др., а также критико-филологические тексты как классиков критического жанра, таких, как В. Г. Белинский и И. А. Гончаров, так и современных филологов. В разделе представлены основные публицистические жанры, изучаемые в курсе русского языка: статья, рецензия, интервью, комментарий, заметка, фельетон, репортаж, очерк. Тематика включаемых текстов охватывает практически весь представленный в «большом» корпусе репертуар, в том числе это такие темы, как искусство и культура, образование, экономика и бизнес, домашнее хозяйство, политика и общественная жизнь, право, здоровье и медицина, досуг и развлечения, а также некоторые научно-популярные заметки.

Учебно-научные тексты включают как предназначенные для квалифицированного читателя статьи из научных журналов, так и (в большинстве) — научно-популярные статьи и публикации из таких журналов, как «Знание — сила», «Первое сентября», «Наука и жизнь», «Химия и жизнь», «Техника — молодежи», «Кот Шредингера» и др. В раздел включены научно-популярные тексты о русском языке и культуре речи (Н. А. Еськовой, М. Королевой, Ю. Сафоновой и др.). Среди авторов учебно-научных статей и интервью — филологи М. М. Бахтин, Ю. М. Лотман, Д. С. Лихачев, психолог А. Н. Леонтьев, математик В. А. Успенский. Представлены следующие типы текстов: аннотация к научной статье, рецензия, статья, лекция, конспект. Научные дисциплины включают филологию, историю, культурологию, математику, физику, психологию, медицину.

Официально-деловые тексты относятся к жанрам, с которыми знакомятся школьники в курсе русского языка: законы, договоры, характеристики, деловые письма, доверенности и некоторые другие типы текстов, характерные для официально-деловой речи. Документы относятся к разным сторонам жизни — общественно-политической, бизнесу, праву, науке.

Разговорные тексты представлены записями устной речи (домашние разговоры и публичные выступления), а также непринуждёнными письменными текстами, относящимися к бытовой сфере (личные дневники, письма, записки, смс-сообщения, форумы).

Сбалансированная часть Обучающего корпуса. Для занятий по стилистике русского языка и культуре речи удобно выделить часть текстов одного периода — 2-й пол. XX в. — начала XXI в., относящихся к разным функциональным стилям, в отдельный подкорпус объемом около 2 млн словоупотреблений. Соотношение текстов разных стилей в нем представлено в Таблице 2. На материале подобного сбалансированного подкорпуса возможна постановка такой задачи, как анализ употребления тех или иных слов, словосочетаний, грамматических форм в том или ином функциональном стиле русского языка, применение корпусных инструментов и основ статистических методов исследования.

Таблица 2. Соотношение текстов в сбалансированной части Обучающего корпуса

Тексты, относящиеся к определённому стилю	Кол-во с/у	Доля, %
Язык художественной литературы	715 000	35,7%
Публицистический стиль (включая мемуары)	582 000	29,1%
Учебно-научный стиль	372 000	18,6%
Официально-деловой стиль	223 000	11,1%
Разговорный стиль (устная публичная, непубличная речь, дневники, бытовая переписка)	110 000	5,5%
Всего	2 002 000	100%

В планы ближайшего пополнения входит включение в Обучающий корпус справочно-энциклопедических текстов (они увеличат долю учебно-научных и публицистических текстов), нехудожественной прозы (мемуарно-биографической, исторической и очерковой литературы), художественных произведений современных авторов для детей и подростков. Предполагается проанализировать учебные программы по предметам «Родной язык (русский)», «Родная литература (русская)» для основного общего образования, а также программы по изучению русского языка как иностранного с рекомендуемыми списками литературы.

3. Лингвистическая разметка текстов

Морфологическая разметка указывает значения всех основных грамматических категорий русских слов — то есть, в школьных терминах, дает «грамматический разбор» каждого слова. Поскольку Обучающий корпус ориентирован на

задачи школьного преподавания и практическую помощь учителю-словеснику, в нем должен присутствовать именно традиционный, упрощенный грамматический разбор. Поэтому в Обучающем корпусе используется морфологическая аннотация, адаптированная к программам общеобразовательной школы и педагогических вузов.

С одной стороны, в Обучающем корпусе не используются некоторые дополнительные категории, принятые в современном научном описании русского языка: второй компаратив (на *по-*), счётная форма (*два часа*), местоимение-предикатив (*некого, нечего*); отдельные термины заменены на традиционные: звательный падеж — звательная форма), предикатив — категория состояния (предикативное наречие), родительный 2, винительный 2, предложный 2 — родительный на *-у* (чаю), винительный с утратой одушевлённости (в гости), предложный на *-у* (в лесу).

С другой стороны, в Обучающий корпус внесены дополнительные параметры описания морфологических категорий, используемые в школе. Это прежде всего информация о типе словоизменения, включающая в себя:

- Инвентарь трёх склонений (I склонение *голова* — *головы*, II склонение *стол* — *стола* и III склонение *тетрадь* — *тетради*) пополнен также изучаемыми в школе особыми подтипами II склонения на *-ий, -ие* и I склонения на *-ия* (*санаторий, призвание, история*), несклоняемыми существительными (*пальто, кафе, кенгуру, пари* и др.), разносклоняемыми словами (*имя, семья, стремя, племя* и др.), а также особыми морфологическими типами, такими, как адъективное склонение («склоняется как прилагательное»: *военный, дежурная*), склонение фамилий.
- Типы спряжений: первое (*болеть*), второе (*гасить*), глаголы разноспрягаемые (*хотеть, бежать*) и неправильные (*есть, дать*).

Во вторых, это информация о лексико-грамматических разрядах частей речи, изучаемых в школе:

- существительных (конкретные, абстрактные, вещественные, собирательные),
- прилагательных (относительные, притяжательные, качественные),
- местоимений (личные, возвратное, вопросительно-относительные, неопределённые и т. п.),
- наречий (образа действия, меры и степени, места, времени, причины, цели).

В Обучающем корпусе сведена к минимуму грамматическая неоднозначность вида «краткие прилагательные — наречия — слова категории состояния (предикативные наречия)»; «союзы — союзные слова (местоимения, местоименные наречия) — частицы». Все такие места проверяются экспертом.

Адаптация системы лингвистической разметки к стандартным нормативным пособиям по русскому языку произведена путем пополнения электронного [грамматического] словаря корпуса, в который была внесена информация о словоизменительных классах существительных и глаголов в виде двух дополнительных баз данных. Одна из них представляет собой список существительных с приписанным

каждой лемме типом склонения, вторая — список глаголов, распределенных по типам спряжений. Распределение по типам склонений и спряжений произведено программно при помощи эвристики, ориентирующейся на представленные в корпусе со снятой омонимией словоформы, с последующей ручной коррекцией получившегося грамматического словаря. В словарь внесена также информация о лексико-грамматических разрядах частей речи (существительных, прилагательных, местоимений, наречий), извлеченная из семантического словаря, который используется для семантической разметки в Национальном корпусе русского языка. В процессе морфологической аннотации тексты, отобранные для Обучающего корпуса, подвергаются морфологической разметке программными средствами, с последующим редактированием результатов морфологической аннотации.

Часть текстов Обучающего корпуса имеет морфологическую разметку со снятой вручную омонимией; двойной разбор допускается лишь в случае реальной неоднозначности трактовки той или иной словоформы. Другая часть текстов, дополнивших корпус, имеет *морфологическую разметку с автоматическим разрешением омонимии* (а также допускает поиск и по неснятой омонимии). Это новый вид разметки, который выполнен с использованием нейросетевых методов. Для аннотации данных НКРЯ была выбрана модель, основанная на qbic [Anastasyev 2020], в версии, получившей название Rubic [Lyashevskaya et al. 2023; Ляшевская и др. 2022а, 2022б]. В процессе обработки выполняется классификация частей речи, лемматизация и синтаксический анализ зависимостей соответственно. Разметка лемм и морфологии воспроизводит схему, применяемую в подкорпусе со снятой вручную грамматической омонимией в составе основного корпуса НКРЯ, разметка синтаксиса — схему UDSynTagRus. Благодаря автоматическому снятию неоднозначности удастся развести грамматические омонимы: словоформы с разным набором признаков получают разные разборы. Как показали эксперименты, точность разборов применяемой версии модели Rubic достаточно высока и для отдельных категорий текстов превышает 99% [Lyashevskaya et al. 2023]. Возможные ошибки в выборе грамматических помет, а также в выборе (и облике) лемм устраняются при последующем тестировании.

Синтаксическая разметка является еще одним новым типом лингвистической разметки в Обучающем корпусе. Она производится одновременно с автоматическим разрешением грамматической неоднозначности. Она построена на синтаксисе составляющих и выделяет такие синтаксические параметры как синтаксическая роль (вершина предложения, аргументы, модификаторы, грамматические маркеры структуры), тип клаузы (глагольная, именная, главная, подчиненная, клауза с эллипсисом и др.). В текстах корпуса размечены синтаксические группы, например, типы клауз (предикативных групп), группы подлежащего и сказуемого и другие параметры. Состав синтаксических помет был адаптирован к школьному преподаванию синтаксиса, понятийный аппарат которого значительно отличается от того, который принят в научных синтаксических теориях. Были исключены некоторые категории, используемые в современном научном описании русского

языка: *вершина предложения, аргументы, квази-аргументы, комплемент, модификатор, клауза, клаузальный комплемент* и под. В то же время предусмотрена система дополнительного поиска тех форм, которые охватываются данными пометами. В поисковом интерфейсе использована синтаксическая терминология, которая соответствует школьному курсу русского языка.

Благодаря синтаксической разметке можно строить запросы относительно синтаксических конструкций и получать гораздо более точные ответы, чем с использованием только морфологической аннотации. Например, на запрос о двух существительных, из которых второе является приложением к первому, получаем контексты употребления со следующими сочетаниями: *князь Андрей, граф Нулин, товарищ командир, господа офицеры, протопоп Аввакум, город Москва, село Ясное, фрегат «Аврора», художники-абстракционисты, женщина-космонавт, телефон-автомат, глаза-маслишки* и мн. др.

Разметка морфемной структуры слова доступна в окне «Морфемный разбор» в разделе «Портрет слова». Разбор слов *по составу* (в их словарной форме) в Обучающем корпусе имеет экспериментальный характер и отличается от разметки Основного корпуса. В Основном корпусе морфемный разбор строится для слов, присутствующих в базовом словаре морфемного анализа [Ляшевская и др. 2009], разработанного на основе словаря морфем [Кузнецова, Ефремова 1986]. Для тех слов, которые отсутствуют в этом словаре, морфемная структура строится автоматически с помощью нейросетевой модели НейроКРЯ (см. об этом в статье [Савчук и др. 2023: 12]). Такие разборы снабжены специальным признаком «сгенерировано НейроКРЯ».

Обучающий корпус должен учитывать требования, предъявляемые к морфемному анализу в школе. Школьный морфемный разбор тесно связан со словообразовательным анализом, то есть выделение морфем в членимой основе должно сопровождаться доказательством наличия семантических и словообразовательных отношений между производящей и производной основами на каждой ступени анализа.

Если вариантов разбора несколько, школьная практика опирается на один из авторитетных морфемных и словообразовательных словарей, к которым относятся [Тихонов 2002], [Николина 2013], [Потиха 1999], при этом в ряде случаев трактовки морфемного членения в этих словарях могут отличаться.

В практике школьного морфемного анализа не принято определять морфемный состав имен собственных и многих высокочастотных слов — местоимений, служебных слов. В ряде случаев сложилась своя традиция обозначения морфем: постфикс *-ся* обозначается как суффикс; *-то, -либо, -нибудь* принято называть суффиксами, окончания у несклоняемых существительных не выделяются и не обозначаются.

Выбор варианта разметки морфемной структуры слова для Обучающего корпуса опирается на экспертный лингвистический анализ результатов морфемного членения материала с помощью словаря морфемного анализа и с помощью нейросетевой модели НейроКРЯ (анализ был проведен М. М. Литвиновой при участии

Г. И. Кустовой и С. О. Савчук). Результаты морфемного анализа, выполненные нейросетевыми методами, содержали много ошибок и отличий от стандартов школьного разбора, поэтому было принято решение использовать морфемную разметку только на основе словаря. В качестве словаря выбран «Морфемно-орфографический словарь» А. Н. Тихонова, поскольку он самый большой (около 100 тыс. слов) и считается авторитетным источником в школьном преподавании. Слова, не входящие в словарь, пока не получают разметку. Исключены из разборов также имена собственные, местоимения, числительные, служебные слова. В дальнейшем предполагается продолжить обучение модели, с тем чтобы улучшить результаты разметки слов, отсутствующих в словаре.

4. Новые функциональные возможности

Современный облик Обучающего корпуса определяется, несомненно, новейшими функциональными возможностями. Это прежде всего новые виды выдачи результатов поиска («График по годам», «Статистика», «Частотность», «Н-граммы»), новые виды поиска («Поиск коллокаций»), а также новые инструменты анализа — «Портрет слова», включающий «Скетчи» и «Похожие слова»; «Частотный словарь», «Портрет корпуса» и «Портрет подкорпуса», позволяющие пользователю анализировать особенности корпуса или подкорпуса и сравнивать их. Ниже будут показаны возможности применения новых инструментов при выполнении заданий и упражнений.

4.1. Поиск коллокаций

Возможности поиска коллокаций продемонстрируем на примере выполнения упражнения, в котором предлагается проанализировать плеонастические словосочетания.

Задание 1. Какие из приведенных сочетаний закрепились в языке и стали допустимыми, а какие воспринимаются как избыточные и не соответствующие языковой норме? Подберите пример из корпуса.

Возвращаться обратно, отступить назад, совместное сотрудничество, информационное сообщение, ведущий лидер, конечный итог, короткое мгновение, будущие потомки, неожиданный сюрприз, внутренний интерьер, реальная действительность, памятный сувенир, преискурант цен.

Плеоназмом (от гр. *pleonasmos* — излишество) называется употребление в речи близких по смыслу и потому излишних слов (*главная суть, повседневная обыденность* и т. д.). Разновидностью плеоназма является тавтология (из гр. *tauto* — то же самое, *logos* — слово), которая может возникать при повторении однокоренных слов (*умножить во много раз, спросить вопрос*), а также при соединении иноязычного и русского слова, дублирующего его значение (*памятный сувенир, впервые дебютировал*), в последнем случае говорят о скрытой тавтологии [Голуб 1997, 20–21]. Пособия по стилистике и культуре речи рекомендуют их избегать. Однако некоторые плеонастические и тавтологические сочетания по разным причинам

закрепились в речи и уже не воспринимаются как ошибочные, несмотря на рекомендации справочных пособий, например, *период времени, экспонаты выставки* и под. [Голуб 1997, 21–22].

Можно выполнить задание привычным способом, задав конкретное словосочетание в лексико-грамматическом поиске, а затем просмотреть выданные контексты и найти или не найти среди них анализируемое сочетание. Ускорить просмотр контекстов поможет режим выдачи KWIC или режим n-грамм, в последнем случае будет подсчитано количество вхождений и относительная частота каждой формы словосочетания. Например, для сочетания *возвращаться обратно* выдача в режиме биграмм представлена в таблице 3. По ссылке в последнем столбце можно перейти к полной выдаче и посмотреть контексты употребления данной формы словосочетания.

Таблица 3. Формат выдачи результатов лексико-грамматического поиска в режиме биграмм

№	Тексты	Вхождения	ipm	Фрагмент
1	5	5	0.4	<i>возвращался обратно</i>
2	4	4	0.32	<i>возвращаться обратно</i>
3	2	2	0.16	<i>возвращается обратно</i>
4	2	2	0.16	<i>возвращались обратно</i>
5	2	2	0.16	<i>возвращаясь обратно</i>
6	2	2	0.16	<i>возвращаюсь обратно</i>
7	2	2	0.16	<i>возвращаются обратно</i>
8	1	1	0.08	<i>возвращалась обратно</i>

Примеры использования: *Но так же, как и сотни лет назад, рыбаки уходят в море за салакой. И так же, как и сотни лет назад, не все **возвращаются обратно**. [К. Г. Паустовский. Золотая роза (1955)]. Через месяц Юшка обыкновенно **возвращался обратно** в город и опять работал с утра до вечера в кузнице. [Андрей Платонов. Юшка (1935)].*

Таким образом, результаты проверки сочетания *возвращаться обратно* показывают, что, несмотря на речевую избыточность (*возвращаться* — идти обратно, поэтому добавление наречия в данном сочетании излишне), оно давно и достаточно широко употребляется в речи.

Еще быстрее справиться с задачей поможет использование поиска коллокаций, поскольку результаты поиска обрабатываются статистическими инструментами и выдаются в упорядоченном (ранжированном) виде, удобном для анализа. Коллокации — это статистически устойчивые словосочетания, в которых один компонент (коллокат) понимается как единица, встречающаяся с заданным словом (ключом) чаще, чем случайно. Выдача коллокаций ранжирована: в верхнюю часть списка попадают слова, наиболее тесно связанные друг с другом. Мера неслучайности вычисляется с помощью статистических метрик t-score, MI3, LogDice и Loglikelihood,

которые основаны на частотности каждого из элементов коллокации и размере корпуса, однако дают несколько разное ранжирование. Выполнение упражнения сводится к тому, чтобы проверить, входит ли исследуемое сочетание в список коллокаций. Если мы найдем его в результатах поиска, следовательно, оно достаточно частотно и допустимо в речи. Если не найдем, будем делать вывод о неприемлемости данного сочетания. При этом мы исходим из того, что тексты, собранные в Обучающем корпусе, отличаются достаточно высоким уровнем речевой культуры.

Поиск устроен следующим образом. Ключ, так же как и коллокат, можно задать с помощью конкретной леммы или словоформы или при помощи набора грамматических и/или семантических признаков. В поле «ключ» вводим первое слово анализируемого словосочетания, в поле «коллокат» — часть речи второго слова. Тем самым мы ожидаем получить наиболее частотные сочетания определенной структуры с первым словом. Если указать в качестве ключа часть речи первого слова, а в качестве коллоката — конкретную лексику, то получим наиболее частотные сочетания со вторым словом. Указываем расстояние $-1 +1$, это означает, что коллокат может находиться справа и слева от ключевого слова, и запускаем поиск. Результаты выдаются в виде таблицы (см. таблицу 4). Для того чтобы упорядочить список по частоте совместной встречаемости, когда вверху списка находятся самые частотные сочетания, лучше выбрать метрику *t-score*, она дает ранжирование, близкое простой частоте.

Таблица 4. Формат выдачи результатов поиска коллокаций

ruscorp.pora.ru

Национальный корпус русского языка: поиск

Коллокации

Скачать

Ключ	Коллокат	Совместная частота	Частота ключа	Частота коллоката	LogDice	Loglikelihood	MI ³	t-score	Ап. мера	Конкорданс
отступать	медленно	8	524	2912	8.63	52.81	8.44	2.79	4.85	Примеры
отступать	назад	8	524	4677	8.22	45.35	7.96	2.77	4.69	Примеры
отступать	уже	8	524	27764	6.52	18.58	6.18	2.45	3.90	Примеры
отступать	далеко	5	524	4578	7.77	23.95	6.57	2.16	3.77	Примеры
отступать	далее	4	524	1682	8.38	25.24	6.91	1.97	3.69	Примеры
отступать	вместе	4	524	6004	7.30	15.38	5.63	1.89	3.26	Примеры

Как видим, сочетание *отступать назад*, которое приводится в пособиях по стилистике и культуре речи как пример плеоназма (*отступать* — ‘ступая, передвигаться назад’, следовательно наречие *назад* избыточно), занимает довольно высокую позицию в списке коллокаций, что позволяет считать его допустимым в речи так же, как и рассмотренное выше сочетание *возвращаться обратно*. Примеры из текстов С. Т. Аксакова, А. К. Толстого, Л. Н. Толстого, Ч. Айтматова, А. Рыбакова, Б. Васильева и др. подтверждают этот вывод: ***Отступать назад, туда, где прятался безоружный и напуганный Волков, было невозможно, и Плужников бросился в другую сторону.*** [Борис Васильев. В списках не значился (1974)].

Проверив подобным образом все заданные словосочетания, получаем следующие результаты.

Закрепились в языке и стали допустимыми сочетания	Примеры из Обучающего корпуса
информационное сообщение	При разработке соответствующего комплекса программ его авторами была построена информационная модель научной конференции, описывающая последовательность действий как оргкомитета (начиная от написания первого <i>информационного сообщения</i> и заканчивая формированием программы), так и участника конференции. [В. Барахнин. Будущее компьютерных технологий закладывается сегодня // «Наука в Сибири» (Новосибирск), 2001]
конечный итог	В <i>конечном итоге</i> , целью изучения любой знаковой системы является определение ее содержания. [Ю. М. Лотман. Структура художественного текста (1998)] Как всегда в истории, исход сражения в <i>конечном итоге</i> решал народ. [Анатолий Приставкин. Ночевала тучка золотая (1981)]
короткое мгновение	Личное человеческое чувство на <i>короткое мгновение</i> взяло верх над тем искусственным призраком жизни, которому он служил так долго. [Л. Н. Толстой. Война и мир. Том третий (1867–1869)]
неожиданный сюрприз	В действительности нас ожидал ряд <i>неожиданных сюрпризов</i> ... [О. Н. Писаржевский. Как было изготовлено звездное вещество // «Юный Техник», 1956 № 02, 1956]
реальная действительность	Он создавал вокруг себя М-поле, преобразующее, говоря просто, <i>реальную действительность</i> в действительность сказочную. [Аркадий Стругацкий, Борис Стругацкий. Понедельник начинается в субботу (1964)]
прейскурант цен	Узел имеет право в одностороннем порядке вносить изменения и дополнения в Регламент и <i>Прейскурант цен</i> . [Договор о доступе в интернет (2003)]
отступать назад	— Какое по сердцу, батюшка! — простонал Михеич, <i>отступая назад</i> . [А. К. Толстой. Князь Серебряный (1861–1863)]
возвращаться обратно	Но какой ему смысл отгонять лодку на такое расстояние и затем <i>возвращаться обратно</i> ? [Анатолий Рыбаков. Бронзовая птица (1955–1956)]

Остальные сочетания воспринимаются как избыточные и не соответствуют языковой норме: *ведущий лидер, внутренний интерьер, совместное сотрудничество, памятный сувенир, будущие потомки*.

4.2. Портрет слова

На странице Портрета слова собраны общие параметры слова, соответствующие его грамматической и семантической разметке в корпусе, а также в компактном виде представлена самая разная информация о функционировании слова в заданном корпусе.

Во-первых, в Портрет слова входят скетчи, каждый из которых представляет собой список из 10 самых частотных коллокаций, распределенных по типу синтаксической связи. Для каждой части речи определен свой список синтаксических отношений. Например, для прилагательных скетчи включают в себя определяемые

существительные к заданному прилагательному, сочетания с наречиями, другие прилагательные, связанные с заданным сочинительной связью. Скетчи позволяют пользователю быстро оценить сочетаемость выбранной лексики, прежде чем приступить к более детальному исследованию.

Во-вторых, дистрибуция форм слова в корпусе отображена с помощью таблицы «формы слова» (см. рис. 1). Можно получить информацию о том, какие формы слова употребляются в корпусе чаще (выделены более интенсивным цветом), а какие не встречаются в корпусе вовсе. Распределение форм слова *воробей* показывает, что наиболее частотными в корпусе являются формы им. п. ед. и мн. ч., а примеры употребления форм тв. п. ед. ч. и предл. п. обоих чисел в корпусе отсутствуют.

В третьих, относительная частотность слова показана насыщенностью цветового фона на шкале частотности, состоящей из 6 диапазонов, при этом количественный показатель *ipm* слова выдается во всплывающем окне.

В четвертых, в Портрете слова присутствует разбор морфемной структуры слова (подробнее о разметке морфемной структуры слова см. на с. 9). Морфемный разбор в Обучающем корпусе строится только для тех слов, которые присутствуют в «Морфемно-орфографическом словаре» [Тихонов 2002] и отражает морфемное членение слова, представленное в этом словаре. Если слово отсутствует в словаре, появляется сообщение «По этому запросу ничего не найдено».

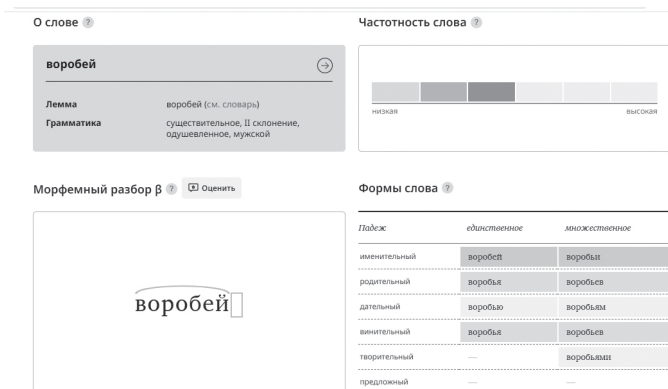


Рис. 1. Фрагмент интерфейса «Портрет слова»

В-пятых, Портрет слова позволяет увидеть так называемые ассоциаты слова или «похожие слова» — слова, которые наиболее часто встречаются в корпусе в одном и том же контексте. В основе представляемого облака слов лежат модели дистрибутивной семантики — матрицы взаимной встречаемости слов (см. рис. 2). Характерно, что разные корпуса, отличающиеся объемом и составом текстов, дают разные списки «похожих слов» к одной и той же лексеме, потому что наиболее частотные контексты одного и того же слова в них будут разными (подробный анализ дистрибутивных семантических моделей см. в статье А. А. Бонч-Осмоловской в настоящем номере).

В-шестых, функционирование слова в данном корпусе представлено на диаграмме «Статистика слова», которая показывает частотное распределение слова в текстах, которые группируются по метатекстовым признакам: сферам функционирования, тематике текстов, полу автора. На диаграмме можно сразу увидеть, что, например, слова *когнитивный* и *виртуальный* чаще встречается в нехудожественных текстах, относящихся к учебно-научной и публицистической сферам, а слово *лазоревый* — исключительно в художественной литературе. Функция полезна для стилистических упражнений.

В-седьмых, график распределения результатов поиска по датам показывает относительную частоту (частоту на миллион словоформ) встречаемости слов в текстах в зависимости от даты их написания. На графике можно увидеть, когда слово впервые зафиксировано в корпусе, в какие периоды употреблялось более активно. От количественных показателей, представленных на графике, можно перейти к контекстам употребления слов в корпусе. Этот виджет полезен при изучении истории слов.

В конце страницы приводится 5 случайных примеров из текстов корпуса, иллюстрирующих употребление слова. От этих примеров по ссылке можно перейти к полной выдаче по запросу на анализируемое слово.

Как представляется, Портрет слова может служить удобной средой для различного рода упражнений и заданий по морфемному, синтаксическому, стилистическому анализу слова (см., например, [Крылова 2023]). Приведем несколько примеров.

Задание 2. Рассмотрите похожие слова в портрете слова *смелый* (рис. 2). Найдите среди них синонимы. Выпишите их и подберите примеры из текстов корпуса.

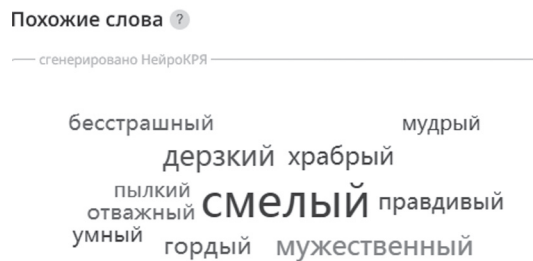


Рис. 2. Фрагмент интерфейса «Портрет слова»

Синонимы к слову **смелый** — отважный, храбрый, мужественный, бесстрашный, дерзкий.

Примеры: В тот же вечер был большой праздник, и под бой барабанов племя *смелых* охотников за акулами приняло нас в почетные вожди. [Владислав Крапивин. Тень Каравеллы (1970)]. Вы — человек *мужественный* и ради других пренебрегаете опасностями. [Анатолий Рыбаков. Прах и пепел (1991–1994)]. Я думаю об этих

временах и завидую прекрасным женщинам и *отважным* мужчинам, чья любовь расцветет под небом веселой и вольной страны. [Константин Паустовский. Северная повесть (1938)]. А он должен быть упорный и очень *храбрый* — такой, что много не говорит, а делает лучше других. [Л. Н. Толстой. Севастопольские рассказы/ Севастополь в августе 1855 года (1855)]. Были люди и добрые, и человеческие, и *бесстрашные*. [Д. С. Лихачев. Заметки о русском (1984)]. Печорину очень легко было обольстить ее: стоило только казаться непонятым и таинственным и быть *дерзким*. [В. Г. Белинский. Герой нашего времени (1988)]

Задание 3. В скетче слова *смелый* (рис. 3) представлено 10 самых частотных сочетаний прилагательного с определяемым словом. Одинаково ли значение слова *смелый* в этих сочетаниях? Как их можно сгруппировать? Определите значение слова *смелый* в каждой группе.

Скетчи ?

смелый Прилагательное

Определяемые слова		Словосочетания с наречием		Сочетанные прилагательные	
1. предположение	8,22	1. отчаянно	7,73	1. честный	8,73
2. гипотеза	7,96	2. слишком	7,03	2. настойчивый	8,7
3. воин	7,95	3. чрезвычайно	6,64	3. решительный	8,6
4. нападение	7,8	4. очень	2,99	4. благородный	8,39
5. надежда	7,74			5. насмешливый	8,38
6. юноша	7,7			6. смелый	8,05
7. поступок	7,65			7. гордый	8,01
8. поход	7,56			8. сильный	7,92
9. мнение	7,15			9. умный	7,25
10. купец	7,06			10. веселый	6,69

Рис. 3. Фрагмент интерфейса «Портрет слова»

Словосочетания можно разбить на две группы:

1. Смелый воин, юноша, купец, нападение, поход, поступок. Значение прилагательного *смелый* в этой группе 'не знающий страха, решительный'.
2. Смелое предположение, гипотеза, мнение, надежда. Значение прилагательного — 'отличающийся новизной и оригинальностью, новаторский'.

Задание 4. Разберите по составу слова *правдивый*, *благоразумный*, *рассудительный*. Сравните свой ответ с морфемный разбор в портретах этих слов в Обучающем корпусе.

Задание 5. Найдите разбор по составу слова *усовершенствование* в портрете этого слова в Обучающем корпусе. Подберите слово, которое отличается от него одним аффиксом, определите значение этого аффикса.

Разбор по составу слова *усовершенствование* представлен на рис. 4.

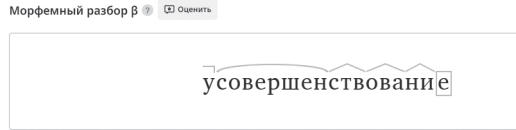


Рис. 4. Фрагмент интерфейса «Портрет слова»

Усовершенствовать → усовершенствование. Суффикс *-ниѳ-*, образует отглагольное существительное со значением действия, представленного как предмет.

Пользуясь Портретом слова в Обучающем корпусе, нужно иметь в виду следующее. Во-первых, поскольку портрет представляет свойства отдельной лексемы в отдельном корпусе, эти свойства зависят от размера корпуса и состава его текстов. Поэтому портреты одного и того же слова могут отличаться в разных корпусах. Во-вторых, чем больше корпус, тем детальнее портрет, тем точнее он отражает семантические и синтаксические связи слова в тексте. Портреты строятся на основе частотных показателей, поэтому для редких слов, встретившихся в корпусе 1–2 раза, трудно сделать вывод, является ли низкочастотное сочетание слов типичным или случайным. Обучающий корпус — небольшой, по меркам статистики, поэтому для многих слов, особенно редких, портрет в Обучающем корпусе будет неполным. Наполненность Портрета слова надо предварительно проверить, прежде чем формулировать задание. В случае отсутствия какой-то информации можно обращаться к соответствующему функционалу в Основном корпусе.

4.3. Новые виды выдачи результатов: частотность

Объем Обучающего корпуса позволяет проводить на его основе небольшие самостоятельные исследования. Покажем это на примере исследования эмоциональной лексики в произведениях Ф. М. Достоевского и Л. Н. Толстого, при сборе материалов в корпусе продемонстрируем новый вид выдачи результатов «Частотность».

Принято считать, что мировосприятие Ф. М. Достоевского и Л. Н. Толстого разительно отличается, так, например, у В. В. Вересаева встречаем: «Трудно во всемирной литературе найти двух художников, у которых отношение к жизни было бы до такой степени противоположно, как у Толстого и у Достоевского...» [В. В. Вересаев. «Да здравствует весь мир!» (о Льве Толстом) (1909–1910)]. Можно предположить, что в текстах Достоевского будут преобладать слова с названием отрицательных эмоций, а в текстах Толстого — положительных.

На первом этапе определим предмет исследования: круг базовых эмоций (радость, грусть, гнев, презрение, отвращение) и состав слов, которые эти эмоции называют. С положительными эмоциями связаны слова *веселый, радостный* и однокоренные *веселье, весело, веселиться, радость, радоваться* и др., с отрицательными — слова *грустный, печальный, тоскливый, гневный, злой, сердитый, презрительный, отвратительный* и однокоренные слова других частей

речи. Если в запросе по лемме указать часть слова со звездочкой, без уточнения части речи, получаем контексты для всех однокоренных слов с заданной начальной частью независимо от их частеречной принадлежности (иногда для таких единиц используется пока еще не общепринятый термин гиперлексема (см., например, [Максименко 2019]). Так, на запрос, начинающейся с *груст**, получаем 330 примеров из произведений Ф. М. Достоевского и Л. Н. Толстого, содержащих существительные, прилагательные, наречия, глаголы с корнем *груст-*: *грусть*, *грустный*, *грустно*, *грустить*.

Для исследования распределения эмоциональной лексики в произведениях Ф. М. Достоевского и Л. Н. Толстого создадим два подкорпуса текстов. Подкорпус текстов Достоевского составляет 496 284 словоупотреблений (7 текстов), подкорпус текстов Толстого — 682 677 словоупотреблений (15 текстов). Наша задача — изучить, используя инструменты Обучающего корпуса, особенности употребления эмоциональной лексики в произведениях каждого из авторов.

На втором этапе собираем материал путем построения запросов к корпусу. Запросы можно строить по-разному, используя лексико-грамматический поиск, — все зависит от желаемого результата поиска. Если указать в поле леммы конкретное слово, получим все контексты для отдельной лексемы. Для пилотного исследования (если не ставить задачу исчерпывающего изучения всех однокоренных слов, связанных с названием изучаемых эмоций) можно построить запрос, содержащий начальные части всех исследуемых лексем (*груст*|печал*|тоск*|гнев** и т. д.). Тогда в результатах выдачи получаем контексты употребления всех интересующих нас слов в подкорпусе Достоевского, а затем, проделав то же самое, в подкорпусе Толстого.

Для предварительного количественного анализа и сохранения результатов удобно использовать инструмент выдачи результатов «Частотность» (рис. 5). В этом случае результаты предварительно обработаны и представлены в таблице, которая имеет ряд настроек. Меняя параметры настроек, можно рассмотреть результаты в разных аспектах, наиболее соответствующих целям анализа. В первом столбце содержатся слова, которые соответствуют словам из запроса, их можно отобразить в режиме леммы, словоформы или грамматических признаков. Столбцы с количественными данными содержат сведения о количестве вхождений слова и относительной частоте на 1 млн словоформ (ipm). Последний столбец «конкорданс» позволяет по ссылке на примеры выйти в корпус и посмотреть контексты употребления слова, выбрать и сохранить нужные примеры.

Вид выдачи «Частотность» позволяет в режиме онлайн оценить предварительные результаты запроса и уточнить дальнейшие шаги по сбору материала. Если исследование предполагает анализ и дальнейшую обработку всего собранного материала, результаты нужно сохранить, нажав на кнопку «скачать».

Из таблицы видно, что самое частотное эмоциональное слово у Достоевского *тоска* (с частотой 96 вхождений, ipm 193), с небольшим отрывом (93, ipm 187) идет слово *веселый*. У Толстого в вершине списка — *веселый* (с частотой 302 вхож-

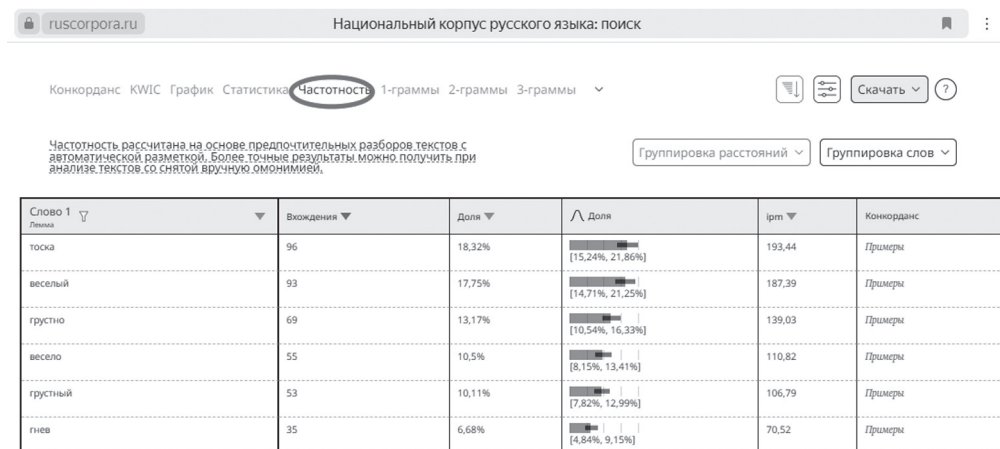


Рис. 5. Способ представления результатов выдачи с помощью инструмента «Частотность»

дения, ipm 442). Однако суммарный анализ результатов собранного материала может внести коррективы в эти первоначальные наблюдения.

Завершающий этап работы — анализ собранных количественных данных. Сведения о суммарном количестве вхождений лексем, отобранных для анализа эмоциональной лексики, собираются отдельно по корпусам Толстого и Достоевского, вычисляется относительная частота для каждой группы однокоренных слов, после чего результаты сравниваются по этому показателю. Наиболее наглядно различие между распределением эмоциональной лексики в произведениях Достоевского и Толстого представлено на рис. 6.

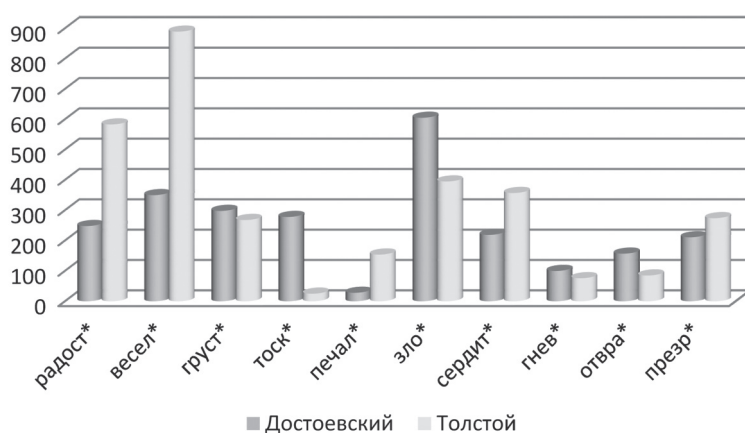


Рис. 6. Частотное распределение лексики с названием эмоций в произведениях Ф. М. Достоевского и Л. Н. Толстого

Диаграмма показывает, что слов, связанных с положительными эмоциями (*радость, радостный, радостно, веселый, весело, веселье* и др.), в произведениях Л. Н. Толстого в два раза больше, чем у Ф. М. Достоевского. Слов, связанных с эмоцией грусти (*грустный, тоскливый, печальный* и др.) в сумме больше у Достоевского, чем у Толстого (в процентном соотношении 57% к 43%), однако если в текстах Достоевского разлита тоска, то в текстах Толстого чаще, чем у Достоевского говорится о печали. К эмоции гнева мы отнесли многочисленные слова с корнем *зл-* (*злой, злость, злобный, злиться* и др.), а также слова *сердитый, сердито, сердить, сердиться, гнев, гневный, гневно, гневаться, гневить*. Как видим, слова с корнем *гнев-* имеют невысокую частоту в подкорпусах текстов обоих писателей, что касается слов с корнями *зл-* и *серд-*, то у Достоевского больше злятся, у Толстого — сердятся. Наконец слова *отвращение, отвратительный, отвратительно* преобладают в текстах Достоевского, а частота слов *презирать, презрение, презрительный* и т. д. чуть выше у Толстого. Таким образом, наше пилотное исследование лексики, называющей эмоции, в текстах Достоевского и Толстого в целом подтверждает представление о различном эмоциональном фоне их произведений, особенности же употребления конкретных слов определенной семантики требуют более тщательного изучения.

5. Проекты-спутники

Важным направлением работы по модернизации Обучающего корпуса было повышение его привлекательности для пользователя и превращение корпуса в необходимый инструмент учителя-словесника. С этой целью были созданы проекты-спутники, предоставляющие информацию об использовании корпуса, демонстрирующие его полезные свойства, а также удобная и наглядная навигация по сайту ruscorpora.ru.

К ресурсам, ориентированным на преподавателя, относятся:

- Раздел «Упражнения на основе корпуса» <https://ruscorpora.ru/exercise>
- Образовательный портал Национального корпуса русского языка Studiorum <https://studiorum.ruscorpora.ru>
- Национальный корпус русского языка в школе <http://ruscorpora.ru/page/corpora-school/>

В разделе «Упражнения на основе корпуса» представлены упражнения, составленные на основе Обучающего корпуса и других корпусов НКРЯ. Упражнения относятся к разным типам: тренировочные, повторительные, развивающие, творческие, контрольные. Их можно использовать для самостоятельной работы в школе и дома, а также для контроля знаний.

Каждое упражнение оформлено в отдельный файл, содержащий формулировку задания и текст упражнения и снабженный метаданными, к которым относится заголовок, называющий тему упражнения, и раздел курса русского языка (лексикология, морфология, синтаксис, орфография, текст) или учебный предмет (литературоведение, русский язык как иностранный). Упражнения объединены в базу

данных, имеющую следующие поля: заголовок, автор, дата создания, учебный предмет или раздел лингвистики, язык (в настоящее время в разделе содержатся задания только по русскому языку), школьный класс, источник упражнения (для упражнений, ранее опубликованных в учебно-методических изданиях), уровень сложности. В настоящее время раздел содержит более 100 упражнений. Их авторы — лингвисты и преподаватели: Н. Р. Добрушина, Е. Р. Добрушина, Е. С. Абелюк, А. И. Левинзон, Г. И. Кустова, С. О. Савчук.

Для работы с упражнениями организован удобный интерфейс на странице <https://ruscorpora.ru/exercise>. Слева приводится список всех упражнений по названиям, справа — инструменты поиска и отбора. Для выбора упражнений, относящихся к определенному разделу курса русского языка, предназначена панель с кнопками, называющими тему. Для выбора уровня сложности и возраста учащихся используются кнопки с номерами классов. Сочетание двух кнопок (темы и класса) сужает поиск и позволяет отобрать упражнения по темам, изучаемым в определенном классе, например, в разделе морфологии русского языка в 7-м классе. Можно скачать файл с текстом упражнения, нажав на иконку справа от названия темы, распечатать его и предложить учащимся для выполнения работы в классе. Можно поделиться списком отобранных упражнений с коллегами и учениками, нажав кнопку «Скопировать ссылку». По образцу представленных упражнений преподаватели могут составлять свои уникальные задания и упражнения на материале Обучающего корпуса и делиться опытом с другими преподавателями.

Таким образом, с помощью раздела «Упражнения на основе корпуса» решается несколько задач: во-первых, предлагаются готовые методические ресурсы, которыми учитель может воспользоваться, скачав и распечатав файл с заданиями; во-вторых, даются образцы, подсказывающие, как можно самостоятельно строить свои задания; и в-третьих, в некоторых упражнениях демонстрируется путь решения задачи, алгоритм работы с корпусом.

Остановимся кратко на других ресурсах, адресованных преподавателю.

Образовательный портал Национального корпуса русского языка знакомит преподавателей с корпусными методами обучения, помогает освоить использование Национального корпуса русского языка в их повседневной работе и служит для создания новых методик преподавания, новых словарей, справочников и пособий.

Страница «*Национальный корпус русского языка в школе*» раскрывает педагогический аспект использования корпуса: «Работа с Национальным корпусом русского языка при обучении гуманитарным предметам, например, литературе, русскому языку, истории, может наглядно продемонстрировать, что язык — это живая развивающаяся система, у которой есть вчера, сегодня и завтра. А ученики могут попробовать себя в роли исследователей: вывести определение из наблюдения за контекстами, увидеть, как значение слова меняется на протяжении нескольких лет, спрогнозировать правописание новых слов» (<https://ruscorpora.ru/page/corpora-school/>).

Страница открывает доступ к материалам, которые полезны в практике преподавания и организации учебной и исследовательской деятельности учащихся.

Помимо *Обучающего корпуса*, *Портрета слова* и раздела упражнений, описанного выше, в помощь учителю создано еще четыре раздела.

В разделе *«Примеры творческих заданий с использованием Корпуса»* содержатся примеры заданий творческого характера, предназначенных для учащихся старших классов, с образцами их выполнения. Темы заданий: «Жизнь слова (изменение значения и грамматических характеристик слова со временем)», «Оттенки значения (работа с паронимами)», «Частотный словарь (слово в произведениях писателя, за выбранный период)». Примеры заданий:

Когда в русском языке появилось слово «прикол» и как менялось его значение? Всегда ли слово «положительно» значило то же самое, что сейчас? (9–10 класс); Как корпус русского языка показывает, что бывает рыбьим, а что рыбным? (9–10 класс); Какие прилагательные со значением высокой оценки использовались в произведениях начала XIX века, а какие — в произведениях начала XXI века? (9–10 класс) и др.

В разделе *«Примеры школьных исследовательских работ с использованием Корпуса»* собраны конкретные работы, в которых представлены небольшие исследования, выполненные школьниками с использованием корпусных подходов и инструментов, например, «Процесс диахронических изменений в моделях управления некоторых глаголов в XVIII–XXI веках», «Лексикографическое описание тематической группы «ткани» в современных словарях русского языка» и др.

Раздел *«Руководство пользователя»* содержит инструкции для исследователей и преподавателей, в которых подробно рассматриваются функциональные возможности корпуса и на конкретных примерах демонстрируются работа корпусного инструментария, способы формулирования запросов, необходимых для решения разных типов задач. Пользователи могут выбрать удобный для себя формат инструкции.

В разделе *«Методические материалы»* представлены учебные пособия, записи лекций, мастер-классов и др., обобщающих опыт использования корпуса в учебном процессе или открывающих новые возможности применения корпусного инструментария в практике преподавания, а также содержатся примеры учебных материалов.

6. Перспективы развития Обучающего корпуса русского языка

Обновление Обучающего корпуса значительно расширило возможности его использования в практике преподавания. Многократно увеличилось разнообразие текстов для составления заданий для контрольных и самостоятельных работ при изучении различных разделов русского языка. Корпус предоставляет материал и аналитические инструменты для небольших исследований, которые могут проводиться школьниками в качестве зачетных работ, написания рефератов и пр.

Работа с Обучающим корпусом формирует и развивает исследовательские навыки, функциональную грамотность, интерес к языку, представление об истории

языка, культуру использования справочных ресурсов, умение анализировать контекст, понимание особенностей жизни слова в художественном тексте.

В ближайших планах развития корпуса — включение в его состав русской классической и детской поэзии, изучаемой в школе, с целью создания в Обучающем корпусе раздела поэзии. Во-первых, это сделает корпус более репрезентативным для использования на уроках литературы, а во-вторых, при изучении русского языка образцы поэзии используются в качестве иллюстративных примеров столь же часто, как и прозаические примеры. Включение поэзии в корпус расширит тематику самостоятельных исследований русской литературы, даст возможность сравнивать творчество поэтов и писателей, особенности поэзии и прозы и т. д.

Расширение состава корпуса будет осуществляться также за счет включения в него произведений современной литературы, в том числе книг для детей и подростков. Это потребует тщательного отбора текстов, изучения рекомендаций по детскому чтению, консультаций со специалистами по образованию. Думается, что работа в этом направлении приведет к более тесной интеграции Обучающего корпуса и корпуса «От 2 до 15».

Другое направление развития Обучающего корпуса связано с разработкой лингвистической разметки текстов и на ее основе — различных инструментов поиска и представления данных. Развитие этого технологического направления предполагает как совершенствование уже существующих видов разметки (о морфемной и синтаксической разметке говорилось выше), так и введение новых видов разметки — лексической, стилистической и др.

Третье направление предполагает развитие ресурсов, адресованных преподавателю: наполнение материалами методических разделов, прежде всего раздела упражнений, в том числе и с применением новых инструментов поиска и обработки результатов, описанных в статье.

Все это будет способствовать превращению Обучающего корпуса русского языка из средства для подбора примеров в рабочий инструмент учителя-словесника, незаменимый помощник в подготовке и организации учебной и исследовательской работы школьников и студентов.

Литература

Герасимова М. Н., Грудева Е. В. Использование лингвистически аннотированного корпуса текстов на занятиях по дисциплинам «Русский язык» и «Русский язык и культура речи»: из опыта работы // Вестник Череповецкого государственного университета. 2014. № 2. С. 65–68.

Голуб И. Б. Стилистика русского языка. М. Айрис-пресс, 1997. 443 с.

Добрушина Н. Р. Как использовать Национальный корпус русского языка в образовании? // Национальный корпус русского языка: 2003–2005. М.: Индрик, 2005. С. 308–329.

Добрушина Н. Р. Что такое Национальный корпус русского языка и зачем он нужен школьному учителю // Русский язык. 2006. № 7. URL: <https://rus.1sept.ru/article.php?ID=200600712>

Добрушина Н. Р., Левинзон А. И. Национальный корпус русского языка в образовании // Cuadernos de Rusística Española. 2012. № 2. С. 41–60.

Добрушина Н. Р. Корпусные методики обучения русскому языку // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. СПб: Нестор-История, 2009. С. 335–352.

Концепции преподавания русского языка и литературы в РФ (утв. распоряжением Правительства РФ от 9 апреля 2016 г. N 637-п) URL: <https://docs.edu.gov.ru/document/216174b3a3e742ed4198233094d948ac/download/2745/>

Крылова Л. В. Ресурс, который поможет учителю составлять уникальные упражнения для уроков русского языка // Учительская мастерская Людмилы Крыловой URL: <https://dzen.ru/a/YjOAo9qiRi-3tWuL>

Крылова Л. В. Национальный корпус русского языка в помощь учителю-словеснику: как использовать функционал “Портрет слова” // Учительская Мастерская Людмилы Крыловой. URL: <https://dzen.ru/a/ZHdRJyf0xmTq-oi9>

Кузнецова А. И., Ефремова Т. Ф. Словарь морфем русского языка: Ок. 52 000 слов. М.: Русский язык, 1986. 1132 с.

Левинзон А. И. Корпусное преподавание в российской школе // Труды Института русского языка им. В. В. Виноградова. 2015. № 6. С. 641–658.

Ляшевская О. Н. и др. (2022а), 6-й Колмогоровский семинар по компьютерной лингвистике и наукам о языке (Москва). Доклад: К задаче разработки версии корпусов НКРЯ с разрешенной неоднозначностью морфологической и синтаксической разметки.

Ляшевская О. Н. и др. (2022б), 46-я школа-конференция ИППИ РАН «Информационные технологии и системы» (ИТиС-2022) (Огниково Московской области). Доклад: Опыт применения моделей-трансформеров для лемматизации современных и исторических текстов русского языка.

Ляшевская О., Гришина Е., Тагабилева М., Иткин И. О задачах и методах словообразовательной разметки в корпусе текста // Полярный вестник. 2009. № 12. С. 5–25.

Максименко О. И. Автоматизированный дистрибутивно-статистический анализ как системная обработка текста // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2019. Т. 10. № 1. С. 92–100. doi: 10.22363/2313-2299-2019-10-1-92-100.

Николина Н. А. Школьный морфемный словарь русского языка. М.: Русское слово, 2013. 368 с.

Образовательный портал Национального корпуса русского языка. Studiorum. Электронный ресурс. URL: <https://studiorum.ruscorpora.ru/education/exercises/>

Ольховская А. И., Парамонова М. К. Корпус в преподавании русского языка и литературы // Rhema. Рема. 2019. № 1. С. 93–123.

Ольховская А. И. Корпусное преподавание русского языка // Вестник Томского государственного педагогического университета (TSPU Bulletin). 2019. № 2 (199). С. 98–107.

Потиха З. А. Школьный словарь строения слов русского языка. М.: Просвещение, 1999. 319 с.

Рычкова Л. В., Киеня С. Н. Корпусные технологии в преподавании РКИ // Этнокультурный и социолингвистический аспекты в теории и практике преподавания языков в негуманитарных вузах: сборник научных статей. Минск, 2010. С. 32–43.

Савчук С. О., Сичинава Д. В. Обучающий корпус русского языка и его использование в преподавательской практике // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009. С. 317–334.

Савчук С. О. Национальный корпус русского языка в зеркале статистики // Труды международной конференции «Корпусная лингвистика-2021». СПб.: Скифия-принт, 2021. С. 18–30.

Савчук С. О., Архангельский Т. А., Бонч-Осмоловская А. А., Доница О. В., Кузнецова Ю. Н., Ляшевская О. Н., Орехов Б. В., Подрядчикова М. В. Национальный корпус русского языка 2.0: новые возможности и перспективы развития. Вопросы языкознания. 2024. № 2. С. 7–34.

Тихонов А. Н. Морфемно-орфографический словарь: Около 100 000 слов. М.: АСТ: Астрель. 2002. 704 с.

Федеральная рабочая программа начального общего образования. Литературное чтение (для 1–4 классов образовательных организаций). М., 2022. 49 с.

Федеральная рабочая программа основного общего образования. Русский язык (для 5–9 классов образовательных организаций). М., 2022. 130 с.

Федеральная рабочая программа среднего общего образования Русский язык (для 10–11 классов образовательных организаций). М., 2023. 62 с.

Федеральная рабочая программа среднего общего образования Литература углублённый уровень (для 10–11 классов общеобразовательных организаций). М.: Институт развития стратегии образования, 2023. 111 с.

Фомичева Н. В. От умения взаимодействовать с информационными ресурсами — к образовательной самостоятельности // Вестник Пермского национального исследовательского политехнического университета. Проблемы языкознания и педагогики. 2015. С. 127–134.

Anastasyev D. Exploring pretrained models for joint morphosyntactic parsing of Russian // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной международной конференции «Диалог». Вып. 19. М.: РГГУ, 2020. С. 1–12.

Lyashevskaya O., Afanasev I., Rebrikov S., Shishkina Y., Suleymanova E., Trofimov I., Vlasova N. Disambiguation in context in the Russian National Corpus: 20 years later. Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной международной конференции «Диалог». Вып. 22. М.: РГГУ, 2023. С. 307–318.

S. O. Savchuk

*V. V. Vinogradov Russian Language Institute,
Russian Academy of Sciences
(Russia, Moscow)
savsvetlana@mail.ru*

THE EDUCATIONAL CORPUS OF THE RNC AS VALUABLE RESOURCE FOR EDUCATORS

The Russian educational corpus is primarily intended for school lessons of the Russian language and literature. It can also be used in university teaching as well as for teaching Russian as a foreign language and for training language teachers. The Educational corpus is based on other principles that distinguish it from the Main corpus of written texts. The differences relate to the balance of texts in the corpus, their compliance with the school curriculum and modern spelling standards.

The volume of the corpus exceeds 13 million word usage. The main part of the corpus is made up of works included in the literature curriculum for secondary and high school, including those recommended for extracurricular reading. Non-fiction texts included in the corpus belong to the functional styles that are studied in the course of the Russian language (journalistic, official, educational, academic and colloquial styles).

The morphological markup in the Educational corpus is adapted to the standard Russian language manuals and provides the traditional, simplified grammatical analysis. In addition, in order to comply with the purposes of school teaching, additional morphological features were introduced into the annotation scheme: inflectional types of nouns and verbs (declension, conjugation) and lexical and grammatical categories of nouns, adjectives, pronouns, adverbs. Morphological markup has been performed automatically in the texts of the Educational corpus using a special program, while grammatical homonyms have been disambiguated. A small part of the texts of the Educational corpus are manually disambiguated.

The new markup makes available all the latest functionality. First of all, these are new types of search results (Graph by year, Statistics, Frequency, N-grams), new types of search query (Collocation Search), as well as a new analytical tool — Word at a Glance which includes Word Sketches and Similar Words widgets, etc. These tools are designed to increase the attractiveness of the corpus for the user and turn it into a necessary tool for a language and literature teacher.

Keywords: Educational corpus, composition and structure of the corpus, morphological markup, syntactic markup, new functionality, use of the Educational corpus in teaching

References

Anastasyev D. Exploring pretrained models for joint morphosyntactic parsing of Russian. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: po materialam ezhegodnoi mezhdunarodnoi konferentsii "Dialog"* [Computational Linguistics and Intellectual

Technologies: Papers from the Annual Conference “Dialogue”]. Vol. 19. Moscow, RSUH Publ., 2020, pp. 1–12.

Dobrushina N. R. [How to use the Russian National Corpus in education?]. *Natsional'nyi korpus russkogo yazyka: 2003–2005* [The Russian National Corpus: 2003–2005]. Moscow, Indrik Publ., 2005, pp. 308–329. (In Russ)

Dobrushina N. R., Levinzon A. I. The Russian National corpus in education. *Cuadernos de Rusística Española*, 2012, no. 2, pp. 41–60. (In Russ)

Dobrushina N. R. [Corpus methods of teaching Russian]. *Natsional'nyi korpus russkogo yazyka: 2006–2008. Novye rezul'taty i perspektivy* [The Russian National Corpus: 2006–2008. New results and perspectives]. St. Petersburg, Nestor-Istoriya Publ., 2009, pp. 335–352. (In Russ)

Federal'naya rabochaya programma nachal'nogo obshchego obrazovaniya. Literaturnoe chtenie (dlya 1–4 klassov obrazovatel'nykh organizatsii) [Federal working programme of primary general education. Literary Reading (for grades 1–4 of educational organisations)]. Moscow, 2022. 49 p. (In Russ)

Federal'naya rabochaya programma osnovnogo obshchego obrazovaniya. Russkii yazyk (dlya 5–9 klassov obrazovatel'nykh organizatsii) [Federal working programme of basic general education. Russian language (for grades 5–9 of educational organisations)]. Moscow, 2022. 130 p. (In Russ)

Federal'naya rabochaya programma srednego obshchego obrazovaniya. Russkii yazyk (dlya 10–11 klassov obrazovatel'nykh organizatsii) [Federal working programme of basic general education. Russian language (for grades 5–9 of educational organisations)]. Moscow, 2023. 62 p. (In Russ)

Federal'naya rabochaya programma srednego obshchego obrazovaniya. Literatura uglublennyy uroven' (dlya 10–11 klassov obrazovatel'nykh organizatsii) [Federal working programme of secondary general education. Literature (advanced level) (for grades 10–11 of educational organisations)]. Moscow, Institute for Educational Strategy Development Publ., 2023. 111 p. (In Russ)

Fomicheva N. V. [From the ability to interact with information resources to educational independence]. *Vestnik Permskogo natsional'nogo issledovatel'skogo politekhnicheskogo universiteta. Problemy yazykoznaniya i pedagogiki* [Bulletin of Perm National Research Polytechnic University. Problems of linguistics and pedagogy.], 2015, p. 127–134. (In Russ)

Gerasimova M. N., Grudeva E. V. [Use of linguistically annotated corpus of texts in classes on “Russian language” and “Russian language and speech culture”: from the experience of work]. *Vestnik Cherepovetskogo gosudarstvennogo universiteta*, 2014, no. 2, pp. 65–68. (In Russ)

Kontseptsii prepodavaniya russkogo yazyka i literatury v RF [Concept of Teaching the Russian Language and Literature in the Russian Federation (approved by the Russian Government Order N 637-r of 9 April 2016)] (In Russ)

Krylova L. V. Resurs, kotoryi pomozhet uchitelyu sostavlyat' unikal'nye uprazhneniya dlya urokov russkogo yazyka [A resource to help teachers create unique exercises for Russian language lessons]. *Uchitel'skaya masterskaya Lyudmily Krylovoi* [Ludmila

Krylova's Teacher's Workshop]. Available at: <https://dzen.ru/a/YjOAo9qiRi-3tWuL> (In Russ)

Krylova L. V. [The Russian National Corpus of to help the teacher of the Russian language: how to use the "Word Portrait" functionality]. *Uchitel'skaya masterskaya Lyudmily Krylovoi* [Ludmila Krylova's Teacher's Workshop]. Available at: <https://dzen.ru/a/ZHdRJyf0xmTq-oi9> (In Russ)

Kuznetsova A. I., Efremova T. F. *Slovar' morfem russkogo yazyka* : Ok. 52 000 slov [Dictionary of Russian morphemes : Circa 52 000 words]. Moscow, Russkii yazyk Publ., 1986. 1132 p. (In Russ)

Levinzon A. I. [Corpus-based teaching in Russian school]. *Trudy Instituta russkogo yazyka im. V. V. Vinogradova*, 2015, no. 6, pp. 641–658. (In Russ)

Lyashevskaja O. N. et al. (2022a), *6-j Kolmogorovskii seminar po komp'yuternoj lingvistike i naukam o yazyke* [6th Kolmogorov Seminar on Computational Linguistics and Language Sciences (Moscow)]. Report: [On the developing a version of the RNC corpora with resolved ambiguity of morphological and syntactic markup]. (In Russ)

Lyashevskaja O. N. et al. (2022b), *46-ja shkola-konferentsiya IPPI RAN "Informatsionnye tehnologii i sistemy"* (ITiS-2022) [The 46th School-conference of IPPI RAS "Information Technologies and Systems" (ITiS-2022)] (Ognikovo, Moscow region). Report: [Application of transformer models for lemmatization of modern and historical texts of the Russian language]. (In Russ)

Lyashevskaya O., Grishina E., Tagabileva M., Itkin I. [On the tasks and methods of word formation markup in a text corpus]. *Polyarnyi vestnik*, 2009, no. 12, pp. 5–25. (In Russ)

Lyashevskaya O., Afanasev I., Rebrikov S., Shishkina Y., Suleymanova E., Trofimov I., Vlasova N. Disambiguation in context in the Russian National Corpus: 20 years later. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: po materialam ezhegodnoi mezhdunarodnoi konferentsii "Dialog"* [Computational Linguistics and Intellectual Technologies: Papers from the Annual Conference "Dialogue"], vol. 22. Moscow, RSUH Publ., 2023, pp. 307–318.

Maksimenko O. I. [Automatic distributive-statistic analysis as system text processing]. *RUDN Journal of Language Studies, Semiotics and Semantics*, 2019, no. 10 (1), pp. 92–100. doi: 10.22363/2313-2299-2019-10-1-92-100. (In Russ)

Nikolina N. A. *Shkol'nyi morfemnyi slovar' russkogo yazyka* [School morpheme dictionary of the Russian language]. Moscow, Russkoe slovo Publ., 2013. 368 p. (In Russ)

Ol'khovskaya A. I., Paramonova M. K. [Corpus in the teaching of Russian language and literature]. *Rhema*, 2019, no. 1, pp. 93–123. (In Russ)

Ol'khovskaya A. I. [Corpus-based teaching of the Russian language]. *Vestnik Tomskogo gosudarstvennogo pedagogicheskogo universiteta* (TSPU Bulletin), 2019, no. 2 (199), pp. 98–107. (In Russ)

Potikha Z. A. *Shkol'nyi slovar' stroeniya slov russkogo yazyka* [School Dictionary of Russian Word Structure]. Moscow, Prosveshchenie publ., 1999. 319 p. (In Russ)

Rychkova L. V., Kienya S. N. [Corpus technologies in teaching Russian as a second language]. *Etnokul'turnyi i sotsiolingvisticheskiy aspekty v teorii i praktike prepodavaniya*

yazykov v negumanitarnykh vuzakh: sbornik nauchnykh statei [Ethnocultural and sociolinguistic aspects in the theory and practice of language teaching in non-humanitarian higher education institutions: A collection of articles]. Minsk, 2010, pp. 32–43. (In Russ)

Savchuk S. O., Sichinava D. V. [Educational corpus of the Russian language and its use in teaching practice]. *Natsional'nyi korpus russkogo yazyka: 2006–2008. Nove rezul'taty i perspektivy* [The Russian National Corpus: 2006–2008. New results and perspectives]. St. Petersburg, Nestor-Istoriya Publ., 2009, pp. 317–334. (In Russ)

Savchuk S. O. [The Russian National Corpus through the prism of statistics]. *Trudy mezhdunarodnoi konferentsii "Korpusnaya lingvistika-2021"* [Proceedings of the International conference "Corpus Linguistics 2021"]. St. Petersburg, Skifiya-print Publ., 2021, pp. 18–30. (In Russ)

Savchuk S. O., Arkhangel'skii T. A., Bonch-Osmolovskaya A. A., Donina O. V., Kuznetsova Yu. N., Lyashevskaya O. N., Orekhov B. V., Podryadchikova M. V. Russian National Corpus 2.0: New opportunities and development prospects. *Voprosy Jazykoznanija*, 2024, 2: 7–34. (In Russ)

Studiorum. Educational portal of the Russian National Corpus. Electronic resource. Available at: <https://studiorum.ruscorpora.ru/education/exercises/> (In Russ)

Tikhonov A. N. *Morfemno-orfograficheskii slovar': Okolo 100 000 slov* [Morpheme and Spelling Dictionary: Circa 100,000 words]. Moscow, AST; Astrel' Publ., 2002. 704 p. (In Russ)

А. А. Бонч-Осмоловская

Институт русского языка им. В. В. Виноградова РАН

(Россия, Москва)

abonch@gmail.com

ДИСТРИБУТИВНЫЕ МОДЕЛИ В ПРОЕКЦИИ ТЕОРЕТИЧЕСКИХ ПОДХОДОВ ЛЕКСИЧЕСКОЙ СЕМАНТИКИ (ПО ДАННЫМ НКРЯ)

В статье анализируются данные семантической векторной модели Основного корпуса НКРЯ. Рассматриваются группы ассоциатов существительных, контекстная близость которых превышает порог 0,5. К корпусным данным групп ассоциатов применяются теоретические методы анализа лексико-семантических отношений. Более абстрактно основной вопрос статьи можно сформулировать следующим образом: насколько имеющиеся теоретические представления о лексико-семантических отношениях между лексемами в языке — такие как гиперонимия (родо-видовые отношения), меронимия (отношения часть-целое), синонимия — соотносимы с корпусными данными о близости контекстов употребления этих лексем. В статье показывается, что привлечение данных близости семантических векторов дает возможность увидеть более сложно организованную систему лексико-семантических отношений. Эта система, с одной стороны, позволяет сравнивать пары лексем внутри одного типа отношений по шкале ближе-дальше и, таким образом, выделить аналог семантических прототипов на основании корпусных данных. С другой стороны, выявляется зависимость реализации лексико-семантических отношений от экстралингвистических параметров, таких как дистрибуция по сфере употребления текстов, связанность употребления лексем с определенными текстовыми жанрами, а также социокультурные факторы, влияющие на доминацию определенных контекстов.

Ключевые слова: дистрибутивная семантика, векторные модели, Национальный корпус русского языка, лексические отношения, гиперонимия, синонимия, меронимия

1. Постановка задачи

Открытие технологии word2vec стало важнейшим поворотом в развитии инструментов автоматической обработки естественного языка, предопределившим, в частности, достижения больших языковых моделей (LLM), лежащих в основе

современных систем искусственного интеллекта. Впервые эта технология была описана в 2013 году в статье [Mikolov et al. 2013], где была предложена идея представления корпуса текстов в виде многомерного векторного пространства. В рамках этого подхода каждому слову в корпусе соответствует вектор, состоящий из вероятностей его «контекстных репрезентаций», т. е. вероятностей встретить вместе с этим словом те или иные слова на заданном ограниченном расстоянии. Собственно, ключевой инновацией стала сама возможность математического вычисления расстояния между векторами: вектора, которые близки по расстоянию в получаемом многомерном пространстве, соответствуют словам, которые встречаются в похожем лексическом окружении, а значит, они тем или иным образом семантически связаны. Идея математически вычисляемой семантической близости открыла возможности для создания семантических векторных моделей языка, основанных на корпусных данных. Применение этой технологии сразу дало существенный качественный прирост в решении классов задач, связанных с автоматическим извлечением текстовых данных, тематической классификацией, анализом тональности и пр. Вместо затратных в реализации электронных словарей, тезаурусов, онтологий, лексических сетей типа Wordnet, которые широко использовались как компоненты компьютерных лингвистических технологий, стали применяться векторные многомерные пространства, автоматически получаемые из корпусов текстов. Такой подход получил название «дистрибутивной семантики» [Baroni et al. 2014], в его основе лежит достаточно давно сформулированная лингвистами идея о том, что значение слова во многом определяется его контекстом. Исследования свойств дистрибутивных семантических моделей мотивировались прежде всего прикладной задачей оценки качества получаемых векторов, т. е. тем, насколько близость векторных репрезентаций соответствует близости семантической. Надо отметить, что задача понять, насколько хорошо работает дистрибутивная семантическая модель, оказалась достаточно сложной [Bakarov 2018; Faruqui et al. 2016]. В целом сформировалось три основных подхода: экспертный подход — разметчики вручную оценивают то, насколько близки по значению слова, обнаруживающие близкое расстояние между их векторами, автоматический словарный подход — состав векторов сравнивается с лексическими базами тезаурусного типа, автоматический ассоциативный подход — состав векторов сравнивается с ассоциативными базами, собранными с помощью психолингвистических экспериментов [Панченко и др. 2015]. Однако каждый из этих подходов имеет свои недостатки: экспертный подход достаточно субъективный (само понятие близости или связанности очень расплывчато), сравнение с лексической и ассоциативной базой может давать результаты, которые не всегда легко интерпретировать, поскольку векторная репрезентация слова в принципе отличается от информации о слове, зафиксированной в лексических базах или собранной с помощью ассоциативных экспериментов. Можно ли утверждать, что вектор контекстуальной репрезентации слова, не совпадающий ни с группами слов в тезаурусе, ни с набором ассоциаций к слову, заведомо плохой? Возможно,

он отражает какой-то иной аспект лексического содержания слова точно так же, как исследования корпусной дистрибуции грамматических категорий обнаруживают взаимосвязи и предпочтения, которые не отражаются в дескриптивных грамматиках. Проблема интерпретации дистрибутивных семантических моделей подробно обсуждается в работе [Günter, Rinaldi, Morelli 2019], в частности, большое внимание уделяется тому, насколько соотносимы дистрибутивные семантические модели с психологическими моделями, описывающими усвоение и хранение лексических значений в ментальном лексиконе. В целом, следует отметить, что исследования дистрибутивных семантических моделей в контексте когнитивных моделей хранения семантики в ментальном лексиконе является одним из основных направлений, теоретически осмысляющих природу векторного семантического пространства и его интерпретируемость [Bhatia 2017; Rodriguez, Merlo 2020]. В некотором смысле, это неудивительно: и статистически получаемые векторные репрезентации, и ментальный лексикон являются своего рода «черными ящиками», мы не можем исследовать их свойства иначе как экспериментально. Сопоставление результатов независимых экспериментов может пролить свет на то, насколько эти две модели изоморфны друг другу [Nematzadeh, Meylan, Griffiths 2017].

Возможно, именно по этой причине теоретические исследования лингвистических оснований дистрибутивной семантической модели встречаются гораздо реже. Фактически перед исследователями встает сложная задача выбрать некое теоретическое утверждение о свойствах языка и проверить экспериментально, воспроизводится ли оно на уровне векторов слов. Это достаточно непросто, учитывая то, что экспериментально полученные данные о контекстуальном поведении слов необходимо перевести на определенный уровень абстракции, чтобы соотнести их с теми или иными теоретическими представлениями. Тем не менее, в этом направлении достигнуты интересные и многообещающие результаты. Так, в [Boledo 2020] представлен обзор того, как дистрибутивные семантические модели могут быть использованы для исследования самых разных лингвистических проблем: семантических примитивов, диахронических семантических сдвигов, гранулярности полисемии и семантики словообразовательных аффиксов. В [Рыжова 2020] рассматриваются методы дистрибутивной семантики в применении к задачам лексической типологии, показывается, что существует корреляция между типологическими и дистрибутивными пространствами.

Настоящее исследование будет также посвящено проблемам выявления взаимосвязи дистрибутивной семантической модели и теоретическим знаниям о языке, а именно лексической семантике. Задачи исследования ставятся следующим образом:

- Возможно ли описать векторную близость слов с помощью инструментов теоретического лексико-семантического анализа?
- Что нового мы узнаем о системных лексических отношениях в языке с помощью дистрибутивной семантической модели?

Следует отметить, что близкие по задачам исследования были представлены в работах [Тимофеева 2018a] и [Тимофеева 2018b]. В первой работе автор рассматривает 520 пар близких по векторному представлению слов и дифференцирует их семантические отношения по восьми типам, включающим в себя собственно лексические отношения синонимии и антонимии, экстралингвистические, но важные для организации лексической системы отношения род-вид (гиперонимы/гипонимы) и часть-целое (меронимы/холонимы), а также четыре отношения, определяемые не в терминах лексической семантики, а в целом характеризующие примерные типы «близости» векторных репрезентаций — признак, операция, ситуационная связанность, словообразовательный вариант. Во второй работе семантические отношения между векторными представлениями глаголов рассматриваются в рамках теории лексических функций модели «Смысл — Текст» [Мельчук 1999]. Как будет показано ниже, лексические функции оказываются очень мощным инструментом для характеристики семантических отношений дистрибутивной модели не только для глаголов, но и для существительных.

В настоящей статье будет предпринята попытка описать различные типы семантических отношений между векторными репрезентациями существительных, которые были получены с помощью дистрибутивной семантической модели Основного корпуса Национального корпуса русского языка и представлены в интерфейсе портрета слова Основного корпуса в качестве «похожих слов». В отличие от упомянутых выше работ [Тимофеева 2018a и 2018b] в фокусе внимание будет не столько анализ частотности разных лексико-семантических отношений, сколько изучение того, как именно они реализуются на векторном пространстве, как они могут быть связаны, как взаимодействуют собственно лексические компоненты значения слов с социокультурными коннотациями, задающимися тематическими и стилистическими особенностями текстов, на которых обучалась модель. В первую очередь нас будет интересовать не дистрибутивная семантическая модель в целом, а исследование возможностей реализации лексико-семантических отношений и тех системных усложнений (и обогащений), которые привносятся в них благодаря корпусным данным, отраженным в векторном пространстве модели. Будут рассмотрены три основных отношения, которые традиционно используются для моделирования лексики: родо-видовые отношения, отношения часть-целое и синонимия. Каждое из этих отношений будет проанализировано на примере векторных репрезентаций слов разной семантики. Значимыми параметрами анализа станут, во-первых, сравнение косинусных мер близости векторов (коэффициентов близости), а во-вторых гомогенность векторной репрезентации — т. е. то, насколько семантическая дистрибутивная модель связывает ближайшие слова с рассматриваемым словом единым лексическим отношением, или же их отношения определяются на разных основаниях. Слово, векторная репрезентация которого будет анализироваться, будет далее называться «основным словом», а наиболее близкие к нему по семантическим векторам слова — «группой ассоциатов».

Следует отдельно подчеркнуть, что сам состав группы ассоциатов полностью зависит от исходных корпусных данных, одно и то же слово может иметь

совершенно разные ассоциативные ряды, если векторные модели построены на разных корпусах. Так, например, ассоциатами слова *звезда* в Основном корпусе НКРЯ будут такие слова как *небосклон*, *созвездие*, *луна*, *светило*, а в Газетном корпусе — *знаменитость*, *шлягер*, *эстрада*, *хит* [Савчук и др. 2024]. Поэтому смысл представленного анализа будет не в том, чтобы зафиксировать конкретный набор ассоциатов основного слова, но на примере их анализа показать лексические связи и противопоставления, заложенные в модели. В то же время, как уже было сказано, векторная близость слов может основываться не только на лексическом сходстве, но еще и задаваться текстовыми дискурсивными факторами, определяющими выбор слов в зависимости от тематических, стилистических или жанровых характеристик текста. Такие социокультурные сдвиги, наблюдаемые в векторном пространстве, являются еще одним очень популярным направлением исследования семантических дистрибутивных моделей [Caliskan, Bryson, Narayanan 2017; Kozlowski, Taddy, Evans 2019; Литвинова, Котлярова, Заварзина 2022]. Поскольку в статье рассматривается векторная модель, построенная на данных Основного корпуса Национального корпуса русского языка, представляется, что наблюдения о социокультурных сдвигах, проявляющихся в векторных репрезентациях его дистрибутивной семантической модели, могут представлять особый интерес, во-первых, как инструмент анализа корпусных коллекций, а во-вторых, как проекция культурных стереотипов, отраженных в Основном корпусе как наиболее полном и репрезентативном собрании текстов на русском языке с XVIII-го по XXI-й век.

Статья будет построена следующим образом: в разделе 2 будут кратко охарактеризованы исходные данные векторной модели, в разделе 3 будут рассмотрены основные лексические отношения (род-вид, часть-целое, синонимия) и их реализации в пространстве семантических векторов, а в разделе 4 будут предложены выводы на основе представленных наблюдений.

2. Исходные данные

Источником данных для проведенного анализа стала дистрибутивная семантическая модель, построенная на данных Основного корпуса Национального корпуса русского языка. Основной корпус на сегодняшний момент имеет объем более 370 миллионов токенов и включает в себя более 130 тысячи текстов. Модель была построена с помощью алгоритма CBOW, размер векторов модели равен 300, контекстуальное окно было выбрано в объеме пяти токенов. Слова, встретившиеся в корпусе реже пяти раз, при построении модели не учитывались. Смысл алгоритма CBOW (Continuous bag of words) состоит в том, что этот алгоритм предсказывает слово, опираясь на слова его контекста, в нашем случае количество слов составляет пять слов (токенов) слева и пять слов справа. В начале обучения каждому слову присваивается случайное значение вектора, а в процессе обучения его значение оптимизируется так, чтобы минимизировать ошибку предсказания слова. В результате слова, которые встречаются в одних и тех же контекстах, будут иметь близкое значение векторов. Для исследования рассматривались только леммы

существительных, группы ассоциатов включали в себя десять наиболее близких существительных к заданному слову, при этом порог коэффициента близости должен был быть выше 0,5, при общей шкале измерения от 0 до 1, где 1 — это вектор исходного слова. Таким образом, если не все из десяти ближайших слов преодолевали установленный порог близости, в группу ассоциатов включалось меньшее количество слов. Из рассмотрения также исключались слова, полученные в результате ошибочной лемматизации.

3. Типы ассоциативных связей существительных

В целом, отношения между ассоциатами и основным словом можно разделить на условно регулярные, т. е. те, которые легко классифицировать как синонимию, родо-видовые отношения или отношения часть-целое, и неопределенные, которые, на первый взгляд, не являются системными, хотя и подразумевают какую-то связь. Отношения такого типа обычно называются в дистрибутивной семантике «семантической связанностью» (*semantic relatedness*), и более четкая их классификация выглядит затруднительной. На самом деле, как будет показано ниже, неопределенность семантических отношений во многом определяется тем, что ассоциаты связаны с основным словом не через конкретную лексему, а через более сложный семантический комплекс, имеющий свое семантическое ядро и ряд модифицирующих его параметров. Этот комплекс не имеет собственной лексической реализации, но выражен в языке через ряд лексем, встречающихся в близком контексте, где каждая лексема содержит в своем значении, кроме общего семантического ядра, дополнительное значение параметра. Напротив, регулярные лексические отношения реализованы на уровне полноценных пар лексем, выражающих заданное лексическое отношение. В анализе, представленном ниже, мы будем сначала опираться на регулярные лексические отношения, рассматривая слова, которые ожидаемо должны быть ими связаны (например, у лексемы *дерево* имеется родо-видовое отношение с лексемой *береза*). Однако в том же ряду мы будем рассматривать и такие случаи, когда ассоциативный ряд связывается через основное слово с лексически нереализованным смысловым комплексом, который может быть определен только описательно через противопоставление собственного лексического наполнения ассоциатов и основного слова.

3.1. Родо-видовые отношения

Родо-видовые отношения (или отношения гиперонимии/гипонимии) являются ключевым типом отношений для формализованных лексических описаний. Так, на гиперонимии/гипонимии построена лексическая сеть Wordnet (Fellbaum 1998). Прочие типы отношений, очевидно, являются дополнительными, именно гиперонимия/гипонимия задает иерархичную древесную структуру сети. Это экстралингвистическое отношение в том смысле, что за ним стоят экстралингвистические таксономии, категоризирующие многочисленные области знаний, как научных, так и бытовых. В то же время, очевидно, что за его распространенностью в языке стоят

глубинные когнитивные процессы, которые позволяют человеку структурировать окружающую действительность, понимая связь объектов друг с другом, их общие (родовые) свойства и собственные (видовые) характеристики. Именно поэтому эти отношения так глубоко и системно укоренены в языке и регулярным образом воспроизводятся в дистрибутивной семантической модели.

Наиболее предсказуемыми ассоциативными группами являются лексемы, входящие в один таксономический ряд с основным словом и являющиеся вместе с ним представителями одного класса объектов, или когипонимами. Чем более четко определена экстралингвистическая таксономия, тем выше будет коэффициент векторной близости между лексемами, обозначающими ее таксоны. Закрытые списки элементов одного класса обычно имеют настолько близкое контекстуальное употребление, что они полностью заполняют ассоциативную группу. Так, любое название месяцев имеет контекстный вектор, наиболее близкими к которому являются названия других месяцев, см. Таблицу 1.

Таблица 1. Ассоциаты названий месяцев, когипонимы общего гиперонима *месяц*

Основное слово	Ассоциаты
январь	когипонимы: ноябрь (0.97), декабрь (0.97), февраль (0.97), март (0.97), апрель (0.96), май (0.95), июнь (0.94), июль (0.94), сентябрь (0.95), октябрь (0.95)
июнь	когипонимы: июль (0.98), август (0.98), сентябрь (0.98), апрель (0.97), май (0.97), октябрь (0.96), ноябрь (0.96), декабрь (0.95), февраль (0.95), март (0.95)
октябрь	когипонимы: ноябрь (0.97), декабрь (0.97), сентябрь (0.96), февраль (0.96), март (0.96), апрель (0.96), май (0.96), июнь (0.96), январь (0.95)

Хотя ассоциаты, показанные в таблице, выглядят совершенно одинаково, внимательное наблюдение над тем, как распределяются сотые доли значений, позволяет заметить, что языковые «времена года» текстов НКРЯ не вполне совпадают с фактическими. Так, мы видим, что и ноябрь, и март так же близки к январю, как декабрь и февраль, все они имеют один и тот же коэффициент близости (0.97), т. е. фактически они ведут себя как «зимние месяцы». Напротив, сентябрь близок к июню так же, как июль, с коэффициентом (0.98), т. е. контекстуально это «летний», а не «осенний» месяц. Заметим, что общий гипероним *месяц* ни разу не попадает в ассоциаты, в ассоциативные группы названия месяца собираются именно когипонимы. Еще одной особенностью функционирования лексем, обозначающих названия месяцев, возможно, во многом объясняющей высокий коэффициент сходства их векторов, т. е. повторяемость их контекстов, является то, что в Основном корпусе они чаще всего встречаются в коллекции нехудожественных текстов. Распределение по сфере употребления для названий месяцев составляет около 95% для нехудожественных текстов и около 5% для художественной литературы. Можно предположить, что в нехудожественных текстах для названий месяцев встречаются одни и те же контексты, указывающие на время действия

событий, и именно они дают такие высокие значения близости между семантическими векторами.

Связь между сферой функционирования текстов и гомогенностью лексических отношений в группах ассоциатов хорошо видна на примере другого таксономического списка — дней недели. В отличие от двенадцати названий месяцев, дней недели только семь, таким образом, если рассматривать группу ассоциатов из десяти слов с уровнем близости выше 0.5, часть ассоциатов неизбежно будет заполнена словами не из основного таксономического ряда. При этом в ассоциативные группы дней недели никогда не входит слово *среда*. Это многозначное слово, и оно часто используется в других значениях (вещество, пространство, условия), поэтому его семантический вектор сильно отличается от векторов дней недели. Таким образом, в группу ассоциатов к дням недели входят всегда только пять слов из таксономического списка, исключая собственно основное слово и слово *среда*. В ассоциативной группе остается еще несколько потенциальных слотов для слов, контекстно близких к основному слову (дню недели), но не связанных с ним отношениями таксономической когипонимии. Сравнение ассоциативных групп дней недели очень хорошо демонстрирует зависимость разнообразия лексических отношений от частотности употребления основного слова в текстах художественной литературы, см. Таблицу 2.

Таблица 2. Группы ассоциатов дней недели с учетом количества употреблений основного слова в художественных и нехудожественных текстах

Основное слово	Нехудож. тексты	Худож. тексты	Ассоциаты
понедельник	86%	14%	<u>когипонимы (таксономия)</u> : четверг (0.91), пятница (0.9), воскресенье (0.82), суббота (0.81), вторник (0.67) <u>варианты когипонимов</u> : среда (0.68), выходной (0.64) <u>гипероним</u> : неделя (0.54) <u>меронимы</u> : утро (0.61), вечер (0.59)
вторник	90%	10%	<u>когипонимы (таксономия)</u> : пятница (0.73), четверг (0.73), понедельник (0.7), суббота (0.67), воскресенье (0.63), <u>варианты когипонимов</u> : среда (0.62), выходной (0.53)
четверг	87%	13%	<u>когипонимы (таксономия)</u> : пятница (0.94), суббота (0.89), воскресенье (0.83), понедельник (0.81), вторник (0.67) <u>варианты когипонимов</u> : среда (0.72), выходной (0.65) <u>однокоренное слово</u> : четверток (0.64) <u>гипероним</u> : неделя (0.56) <u>мероним</u> : вечер (0.58)
пятница	85%	15%	<u>когипонимы (таксономия)</u> : понедельник (0.9), четверг (0.94), суббота (0.92), воскресенье (0.82), вторник (0.73) <u>варианты когипонимов</u> : среда (0.74), выходной (0.68), четверток (0.64) <u>гипероним</u> : неделя (0.57) <u>мероним</u> : вечер (0.59)

Основное слово	Нехудож. тексты	Худож. тексты	Ассоциаты
суббота	80%	20%	<u>когипонимы (таксономия)</u> : воскресенье (0.92), пятница (0.92), четверг (0.88), понедельник (0.87), вторник (0.67) <u>варианты когипонимов</u> : среда (0.72), выходной (0.73), четверг (0.62) <u>гипероним</u> : неделя (0.54) <u>мероним</u> : вечер (0.63) <u>квазисиноним</u> : сочельник (0.59)
воскресенье	76%	24%	<u>когипонимы (таксономия)</u> : суббота (0.92), пятница (0.86), четверг (0.83), понедельник (0.82), вторник (0.63) <u>варианты когипонимов</u> : среда (0.68), выходной (0.64) <u>гипероним</u> : неделя (0.70) <u>квазисиноним</u> : пасха (0.63), праздник (0.63) <u>мероним</u> : вечер (0.67)

Из Таблицы 2 мы видим, что, чем чаще встречается слово в художественных текстах, тем разнообразнее лексические отношения у ассоциатов с основным словом. Показательно сравнение групп ассоциатов *вторника* и *воскресенья*. *Вторник*, 90% употреблений которого приходится на нехудожественные тексты, употребляется только как член таксономического ряда названий дней недели, кроме этого ряда только два слова имеют семантические вектора, чья близость к основному слову оказывается выше барьера косинусной близости 0.5. Семантический вектор *вторника* оказывается близким векторам пяти дней недели, а также устаревшего варианта слова *среда* — *среда* и синониму его когипонимов, слову *выходной*. Интересно также, что *вторник* — единственное слово из этого таксономического ряда, ассоциаты которого не включают общий гипероним — *неделя*. Употребления слова *воскресенье*, напротив, почти на четверть состоят из контекстов, встречающихся в художественной литературе. И мы видим, что группа ассоциатов *воскресенья* включает в себя, кроме собственно таксономии дней недели, слова *пасха* и *праздник*, которые можно отнести к квазисинонимам. Более подробно лексическое отношение квазисинонимии будет обсуждаться в разделе (3.3). Пока же отметим, что слова *воскресенье* и *праздник* объединяет общее лексическое значение «дни, когда не работают», более того, это значение может быть эксплицитно выражено с помощью сочинительной конструкции с этими двумя лексемами. Эта устойчивая конструкция входит в список десяти скетчей сочиненных существительных, отраженных и в портрете слова *воскресенье*, и в портрете слова *праздник* Основного корпуса, ср. примеры (1–3):

- (1) Например, в *воскресенья* и *праздники* грешно было работать, понедельник — тяжелый день, и таким образом в году набиралось около двухсот дней, когда поневоле приходилось сидеть сложа руки. [А. П. Чехов. Скрипка Ротшильда (1894)]

- (2) Домой он приходил только в субботу, чтобы отмыться в бане заводскую сажу да пообедать в **воскресенье** или **праздник**. [Д. Н. Мамин-Сибиряк. Под домной (1891)]
- (3) По **воскресеньям** и **праздникам** он весь день оставался один, и тогда единственным утешением его было чтение книг, которые учитель позволял ему брать из небольшой своей библиотеки. [Антоний Погорельский. Черная курица (1829)]

Список скетчей слова *воскресенье* в портрете слова Основного корпуса для десяти устойчивых конструкций с атрибутивным синтаксическим отношением включает в себя семь определений, которые вместе с *воскресеньем* образуют названия церковных праздников: *вербный, прощенный, фомин, христов, светлый, пасхальный, крестопоклонный*.

Тесная лексическая связь дней недели и календаря христианских праздников обнаруживается и у слова *суббота*, в группе ассоциатов которого присутствует слово *сочельник*. Нетрудно достроить семантическую пропорцию: страстная суббота относится к дню Христова воскресенья так же, как сочельник к дню рождения Христова, т. е. *субботу* и *сочельник* связывает то, что это дни накануне самых главных христианских праздников. Таким образом, внутри ассоциативного ряда мы видим стройные парадигматические отношения четырех лексем — *суббота : пасха / (Христово, светлое) воскресенье, сочельник : Рождество (Христово)*. Однако само это соотношение определяется не столько внутренним лексическим потенциалом слова *суббота*, сколько экстралингвистическим историко-культурным контекстом, заданным текстами, входящими в коллекции Основного корпуса. Заметим, что группа ассоциатов слова *суббота* в Газетном корпусе слово *сочельник* уже не включает.

Другие таксономические группы когипонимов, например, названия овощей, ягод, птиц, животных, растений, ведут себя так же, как и описанные выше названия месяцев и дней недели: ассоциатами основного слова становятся преимущественно слова того же таксономического класса. Общий гипероним (родовое слово) может тоже входить в ассоциативную группу, однако связь между когипонимами оказывается теснее, чем у гиперонима и гипонима. Это выражается в том, что коэффициенты сходства векторов с основным словом у когипонимов выше, чем у гиперонима. Выше в Таблице 2 мы видели, что слово *неделя* входит в ассоциаты всех дней недели, кроме слова *вторник*. В названиях ягод мы встречаем слово *ягода* в ассоциативных группах слов *брусника, клюква, малина, земляника, черника*, но это слово не входит в десятку ассоциатов слов *голубика, костяника, клубника, облепиха, ежевика, крыжовник*. В названиях деревьев слово *дерево* является ассоциатом слов *дуб* и *сосна*, но не входит в группу ассоциатов слов *ольха, клен, тополь, липа, лиственница*. Можно предположить, что связь гипоним — гипероним в своем чистом виде оказывается наиболее сильной для частотных и культурно значимых представителей класса. Таким образом, векторное пространство Основного корпуса дает интересный материал для уточнения теории семантических прототипов

на материале корпусных данных. *Вторник* не является прототипическим днем недели, *ольха* — не прототипическое дерево, *костяника* и садовые ягоды *клубника* и *крыжовник* — не прототипические ягоды, поэтому их семантические вектора дальше от семантического вектора их гиперонимов, чем вектора более частотных представителей этих таксономических классов.

Еще одно важное свойство когипонимов-ассоциатов, которое мы будем более подробно рассматривать ниже на других примерах, это случаи, когда ряд когипонимов формируется не устоявшейся внеязыковой таксономией, а моделируется дополнительным фреймом, не имеющим собственной лексической реализации. Так когипонимами слова *жираф* становятся названия совершенно разных животных, объединенных тем, что место их обитания — не средняя полоса, а экзотические страны, своего рода «животные в зоопарке»: *антилопа* (0.72), *носорог* (0.71), *кенгуру* (0.7), *бегемот* (0.7), *гиппопотам* (0.7), *зебра* (0.7), *тапир* (0.69), *леопард* (0.68), *павлин* (0.67), *слон* (0.6). В ассоциативную группу слова *шампиньон* входят не грибы, но растения, которые используются в пищу, этот фрейм можно было бы обозначить как «растительный кулинарный ингредиент»: *баклажан* (0.78), *спаржа* (0.81), *морковка* (0.76), *редиска* (0.76), *укроп* (0.76), *оливка* (0.76), *артишок* (0.76), *фасоль* (0.76), *клубника* (0.75). Иначе говоря, кроме классических таксонов, зафиксированных в словарях и онтологиях и воспроизводящихся в текстах, векторные модели отражают фреймовые «usage-based» таксоны. В этом случае родовым понятием становится типическая ситуация, определяемая социокультурно мотивированным семантическим комплексом: дни накануне христианских праздников, животные, обитающие в дальних странах, растения, используемые в рецептах и т. д.

3.2. Отношение часть-целое

Отношение часть-целое представляет собой еще один тип иерархической организации лексики. Однако в отличие от родо-видовых отношений, отношение меронимов (часть) и холонимов (целое) устроены гораздо менее регулярным образом. Если гиперонимия выстраивает системные отношения в лексике и определяет возможность ответа на вопрос «что это? что это за объект?» через референцию к общему классу, то отношения часть-целое существенно большим образом зависят от того, насколько в целом объект мыслится как имеющий значимые части. Как отмечается в [Рахилина 2008: 39] «(..) для понимания природы этих отношений чрезвычайно существенно, что они не присущи действительности как таковой, а вносятся человеческим разумом. Это человек «видит» объекты как объединенные в множества или расчлененные на части и соответствующим образом называет их. Самую причудливую абстрактную фигуру, нарисованную на бумаге, мы, как носители языка, будем считать «целым», нечленимым объектом, но у нарисованного дерева безошибочно выделим ствол, ветки и корни. Разгадка такой «непоследовательности» поведения носителей языка в том, что как полноценные части объекта воспринимаются только те его фрагменты, которые для данного объекта функционально значимы.»

Отношения часть-целое в целом гораздо слабее предопределены и лексически, и соответственно, иерархически. Может ли часть существовать как самостоятельный объект отдельно от целого? Например, части тела человека или животных — нос, рука, нога, клюв — то, что в типологической литературе называется классом неотъемлемой принадлежности, дефолтно нераздельно связаны со своим целым — телом. При этом уже части артефактов могут быть связаны со своим целым иначе. Например, *багажник* или *фара* являются неотъемлемой частью автомобиля как средства передвижения, а *колесо* является самостоятельным объектом. Кроме этого, лексические значения слов, которые используются для обозначения части, могут содержать в себе очень слабые указания на лексические ограничения тех сущностей, которые выступают в качестве их целых, или не содержать их вовсе. Например, *стекло* является частью *окна*, но при этом *стекло* может также являться частью множества других предметов, относящихся к семантическому классу артефактов. То же самое можно сказать и про холонимы (целое): например, существительное *коллекция* обозначает множество артефактов определенного вида, таким образом, частью коллекции могут быть разные лексически не связанные между собой объекты — *картины*, *марки*, *растения*. Наконец, существуют такие слова как *кусок*, *деталь* или, собственно, *часть*, лексическое значение которых состоит именно в том, чтобы выражать отношение меронимии в сочетании с принципиальной неопределенностью или слабой определенностью их холонима. Соответственно, анализ лексической связи типа часть-целое в группах ассоциатов не может не учитывать разнообразие возможностей реализации этой связи. Обычно, кроме собственно части и целого, выделяют еще несколько типов меронимии, или, точнее сказать, отношений подобных меронимии. Это, во-первых, отношение типа «объект-материал объекта», а во-вторых, отношение типа «элемент-совокупность». Так, например, в датасете EVALution, собранного для тренировки и оценки дистрибутивных моделей и состоящего из пар слов, связанных разными лексическими отношениями (синонимия, антонимия, гиперонимия, меронимия), выделяются три подтипа меронимов: “PartOf”, “MemberOf”, и “MadeOf” [Santus et al. 2015]. Именно эти три разновидности меронимических отношений были взяты за основу для анализа, представленного ниже. Во-первых, анализировалась базовая реализация отношения часть-целое — *часть тела* : *тело*. В рамках этого отношения X (мероним) является непосредственной и неотторжимой частью Y (холонима) [Рахилина 2008: 46], а функциональные свойства меронимов (*частей тела*) являются неотъемлемым условием нормального функционирования холонима (*тела*). Во-вторых, рассматривались такие случаи, когда Y (холоним) состоит из X (меронима), который может быть материалом или веществом, например, как в случаях *окно* : *стекло*, *дом* : *кирпич*. В-третьих, в фокус рассмотрения были включены две лексические функции модели «Смысл — Текст» [Мельчук 1999] — функция Sing: X (мероним) обозначает один экземпляр, элемент, порцию, штуку или единицу Y-а (холонима), например, *кусок* : *пирог* и функция Mult: X (холоним) обозначает естественную совокупность Y-ов, например, *флот* : *корабль*. Перечисленные типы меронимических отношений различаются тем, насколько

лексически устанавливаются симметричные и взаимно-однозначные соответствия между меронимом и холонимом. Ниже рассмотрим последовательно каждое из этих отношений.

3.2.1. Часть тела как неотторжимое целое

Части тела являются прототипическим примером отношения частей, неотторжимых от целого. Это их свойство, которое называется денотативной фиксированностью (Рахилина 2008: 42), проявляется на уровне грамматической сочетаемости. Например, они не сочетаются в генитивных конструкциях с именем целого — **нос лица*, **нога тела*, а также не могут употребляться с предлогом отторжения от — **нос от лица*, **нога от тела*, кроме специальных ситуативных контекстов, поддерживающих особую интерпретацию конструкции — *рука от куклы*. Денотативная фиксированность, по-видимому, обеспечивает близость векторов комеронимов — т. е. других частей того же целого.

Посмотрим на Таблицу 3, в которой приведена ассоциативная группа слова *нос*.

Таблица 3. Группа ассоциатов слова *нос*

Основное слово	Ассоциаты
нос	<u>собственные меронимы</u>: ноздря (0.6), переносица (0.6) <u>комеронимы</u>: подбородок (0.67), кадык (0.65), ус (0.6), лоб (0.6) <u>холонимы</u>: морда (0.62), мордочка (0.6) <u>квазисиноним</u>: клюв (0.6) <u>однокоренное слово</u>: носик (0.75)

Несмотря на то, что лексические связи группы ассоциатов и основного слова кажутся, на первый взгляд, очень разнообразными, все они, кроме уменьшительного деривата, так или иначе связаны с отношением часть-целое и, более того, весьма любопытным образом его структурируют. Во-первых, выделяется ряд слов, которые являются частью того же целого — *лица* (за исключением *кадыка*, об этом будет сказано чуть ниже). Во-вторых, о том, что целым для носа с точки зрения векторного пространства является именно лицо, а не голова и не тело, мы можем судить по присутствию холонимов *морда* и *мордочка*. Мы не видим в этом ряду основной лексики *лицо*, но это, скорее всего, объясняется полисемией этого слова и множеством контекстов, в котором слово *лицо* значит не переднюю часть головы, а человека, индивидуума или вообще переднюю часть предмета. Вектор *лица* с учетом многообразия контекстов для разных значений оказывается очень далек от слова *нос*, для которого значение части тела является доминирующим. Заметим также, что базовое лексическое значение слова *морда* как передней части головы животного оказывается контекстуально объединено с *мордой* как стилистически сниженным синонимом *лица*. То есть с точки зрения векторной семантики *нос собаки* и *нос неприятного человека* оказываются неразличимы. Векторная близость общих контекстов человека и животных поддерживается квазисинонимом *клюв*.

Как будет видно дальше из Таблицы 5, такое сближение номинаций человеческих частей тела и частей тела животных проявляется в семантическом пространстве Основного корпуса для частей, обозначающих крупные элементы — *спина, рука, нога, туловище*. Наконец, в группу ассоциатов слова *нос* попали его собственные меронимы — *переносица* и *ноздра*. Они также являются денотативно фиксированными, причем не только по отношению к общему целому — *лицу*, но и к своему непосредственному холониму — *носу*. Возвращаясь к группе комеронимов (*подбородок, кадык, ус, лоб*), отметим, что они объединены топологически, — все они расположены на одной прямой и являются выступающими частями. Иначе говоря — это те части лица, которые образуют его профиль, если мы представляем его включающим шею: *лоб — нос* (основное слово) — *ус — подбородок — кадык*. Интересно, что это ярко выраженный мужской профиль. Можно предположить, что близость соответствующих контекстов предопределена культурно-текстовыми параметрами словесного портрета в русской литературе, — слово *нос* в 71% случаях своего употребления встречается в художественной литературе и в 29% в нехудожественной. Иначе говоря, векторная близость ассоциатов *кадык* и *ус* указывает на то, что *нос* является элементом именно мужского литературного портрета. Характерный комплекс женского портрета можно воссоздать с помощью группы ассоциатов слово *щека* (Таблица 4):

Таблица 4. Группа ассоциатов слова *щека*

Основное слово	Ассоциаты
щека	комеронимы: подбородок (0.71), висок (0.71), скула (0.7), губа (0.7), лоб (0.7), грудь (0.69), затылок (0.66), шея (0.65) синоним: ланита (0.68) однокоренное слово: щечка (0.81)

На женскость этого портрета нам указывает комероним *грудь* и устаревший синоним *ланита*, который чаще всего использовался для поэтических описаний женского облика. Это, кстати, хорошо видно и по собственной группе ассоциатов слова *ланита*, куда входят четыре дескриптивных адъективных существительных *розовость* (0.6), *румянец* (0.59), *смуглота* (0.56) и *багрянец* (0.56). Интересно, что женский портрет — это скорее не профиль, а бюст, он трехмерен (в ассоциатах присутствуют *висок* и *затылок*), и на него можно посмотреть в фас — мы видим *скулы* и *губы*.

Изучение групп ассоциатов разных частей тела показывает очень интересную особенность: ассоциаты группируются локально по нескольким основным зонам, иначе говоря, дополнительным основанием для сближения меронимов являются метонимические отношения смежности. Однако границы зон не определены жестко и как бы меняются в зависимости от ракурса, который задается основным словом. Можно утверждать, что семантическое векторное пространство Основного корпуса задает в целом идею вертикали тела от затылка до пяток. При этом каждое слово определяет свою ближайшую зону на этой вертикали. Отметим сразу, что

в вертикаль могут входить или не входить *руки*, но при этом они в любом случае входят не полностью, а до *запястья*. Зона *ладоней* является максимально удаленной от всех частей тела и образует собственный «ракурс» тела от руки до лица: *ладонь* — *локоть* — *рука* — *грудь* — *щека*. Кроме того, совершенно особенно ведет себя слово *глаз* — его зона минимальна и определяется только его собственными меронимами — словами *веко* и *зрачок*, остальными его ассоциатами являются дериваты и синонимы. Семантический вектор слова *глаз* оказывается так мало похож на другие части тела (лица), что *глаз* исключается из меронимических отношений, которые связывают другие названия частей тела или лица. В Таблице 5 приведены группы ассоциатов частей тела, комеронимы разделены по четырем категориям соответственно четырем потенциальным холонимам — *голова*, *туловище*, *рука*, *нога*, которые, в свою очередь, выступают как целое для своих функциональных частей. Кроме того, мы видим, что в группах ассоциатов слов *туловище*, *спина*, *рука* и *нога* появляются части тела животного: *плавник*, *загрявок*, *лапа*. Таких слов не очень много, и это, вероятно, связано с тем, что тело животного, по-видимому, содержит в себе меньшее число функционально значимых лексикализованных элементов, отличных от тела человека. В частности, чрезвычайно лексически разработанными для тела человека оказываются части *рук* и *ног* в отличие от *лап* животных, которые труднее разложить на части из отдельных лексем.

В Таблице 5 размечены все типы лексических отношений, которые мы наблюдаем в группах ассоциатов слов со значением частей тела.

Таблица 5. Группы ассоциатов слов со значением частей тела

Основное слово	Ассоциаты с ключевыми зонами (голова, руки, ноги, туловище)
голова	<u>собственные меронимы</u> : затылок (0.61), темя (0.58), чуб (0.57) <u>однокоренное слово</u> : головенка (0.73), головка (0.68), головушка (0.59) <u>синоним</u> : башка (0.71)
лоб	<u>комеронимы</u> (голова): затылок (0.76), висок (0.74), переносица (0.73), лобик (0.71), щека (0.7), темя (0.69), подбородок (0.69), скула (0.69), переносье (0.67), лысина (0.66)
глаз	<u>собственные меронимы</u> : зрачок (0.75), веко (0.69) <u>однокоренное слово</u> : глазенки (0.81), глазок (0.8), глазик (0.68) <u>синоним</u> : око (0.69) <u>квазисиноним</u> : взор (0.69) (<i>ср. обвести глазами/взором</i>)
подбородок	<u>комеронимы</u> (голова): скула (0.72), переносица (0.71), щека (0.71), кадык (0.71), затылок (0.7), лоб (0.68), висок (0.67), переносье (0.67) (туловище): брюшко (0.68), ключица (0.67)
туловище	<u>собственные меронимы</u> торс (0.69), конечность (0.65), брюшко (0.63), ягодица (0.62), позвоночник (0.61) <u>комеронимы</u> (нога): ступня (0.67), голень (0.66), бедро (0.65) <u>синоним</u> : тело (0.64) <u>альтернативный мероним</u> : плавник (0.62)

Продолжение табл. 5

Основное слово	Ассоциаты с ключевыми зонами (голова, руки, ноги, туловище)
грудь	<u>комеронимы</u> (голова): висок (0.69), щека (0.69), затылок (0.68) (туловище): живот (0.8), шея (0.71), плечо (0.68), ягодица (0.68) (рука): ладонь (0.68), запястье (0.66) (нога): бедро (0.73)
шея	<u>комеронимы</u> (голова): щека (0.65), затылок (0.69) (туловище): грудь (0.7), плечо (0.66), спина (0.65) (рука): запястье (0.68) (нога): бедро (0.66), лодыжка (0.65), ляжка (0.64), нога (0.63)
плечо	<u>комеронимы</u> (голова): затылок (0.65), (туловище): спина (0.76), грудь (0.68), шея (0.66), ягодица (0.63) (рука): рука (0.64), локоть (0.67) (нога): бедро (0.68), коленка (0.64) <u>однокоренное слово</u> : плечико (0.75)
локоть	<u>комеронимы</u> (туловище): плечо (0.67), ягодица (0.67), лопатка (0.65) (рука): ладонь (0.67) (нога): коленка (0.75), нога (0.67), колено (0.67), лодыжка (0.66), бедро (0.66) <u>однокоренное слово</u> : локоток (0.73)
спина	<u>комеронимы</u> (голова): затылок (0.69) (туловище): плечо (0.76), шея (0.65), бок (0.65), грудь (0.65), зад (0.64), живот (0.64) (нога): нога (0.65), коленка (0.62) <u>альтернативный мероним</u> : загривок (0.63)
живот	<u>комеронимы</u> (туловище): грудь (0.8), поясница (0.69), ягодица (0.69), пах (0.66) (нога): бедро (0.7), ляжка (0.68), коленка (0.67) <u>однокоренное слово</u> : животик (0.73) <u>синоним</u> : пузо (0.73), брюхо (0.77)
бедро	<u>комеронимы</u> (туловище): ягодица (0.84), ключица (0.76), предплечье (0.74), грудь (0.73), живот (0.7) (рука): запястье (0.71) (нога): голень (0.77), лодыжка (0.77), ступня (0.72) <u>синоним</u> : ляжка (0.77)
колено	<u>гипероним</u> : нога (0.71) <u>комеронимы</u> (туловище): грудь (0.65), плечо (0.63), шея (0.62), живот (0.61) (рука): локоть (0.67), ладонь (0.61) (нога): щиколотка (0.63) <u>однокоренное слово</u> : коленка (0.8) <u>квазисиноним</u> : корточки (0.65) (<i>ср: сидеть на корточках/на коленях</i>)

Основное слово	Ассоциаты с ключевыми зонами (голова, руки, ноги, туловище)
нога	<u>собственные меронимы</u> : коленка (0.77), ступня (0.75), колено (0.72), пятка (0.71), лодыжка (0.68) <u>комеронимы</u> (рука): локоть (0.67) <u>однокоренное слово</u> : ножонка (0.73), ножка (0.73) <u>квазисиноним</u> : лапа (0.71), лапка (0.66)
пятка	<u>комеронимы</u> (туловище): задница (0.66), зад (0.64), ягодица (0.62), пах (0.62) (нога): коленка (0.71), ступня (0.69), ляжка (0.65), носок (0.64), лодыжка (0.62) <u>холоном</u> : нога (0.71)
рука	<u>собственные меронимы</u> : ладонь (0.73), ладошка (0.68), палец (0.67), пятерня (0.66) <u>комероним</u> : подмышка (0.65) <u>квазисиноним</u> : лапа (0.66) <u>однокоренное слово</u> : ручонка (0.74), ручка (0.71), ручища (0.65)
ладонь	<u>собственные меронимы</u> : палец (0.76) <u>комеронимы</u> (голова): щека (0.65) (туловище): грудь (0.68) (рука): кулачок (0.68), локоть (0.66), (нога): коленка (0.65) <u>холоном</u> : рука (0.73), ручонка (0.64) <u>синоним</u> : пятерня (0.73) <u>однокоренное слово</u> : ладошка (0.85),

В Таблице 6 представлено обобщение того, как меронимы и комеронимы, входящие в ассоциативные группы разных слов, покрывают четыре ключевых семантических зоны тела — *голову, туловище, руку, ногу*, а также указано наличие в этих группах лексем с другим типом отношений (лексических или словообразовательных).

Обобщение, приведенное в Таблице 6, демонстрирует внутреннюю сложность семантических отношений между векторными представлениями частей тела. Мы видим систему, состоящую из четырех противопоставленных зон, которые условно можно локализовать как «голова», «верх туловища — плечевой пояс, живот, спина», «низ туловища — живот, ноги» и отдельно «руки». Все лексеммы со значением частей тела распределяются относительно этих четырех зон через группы своих комеронимов или собственных меронимов, соотнесенных либо с той же самой зоной, либо еще и с другими зонами, либо со всеми четырьмя зонами сразу. Во-первых, есть слова, которые имеют ассоциативные ряды, сконцентрированные вокруг одной зоны — это зона головы, к ней относятся само слово *голова*, *лоб* и *глаз*. Интересно, что *лоб* и *глаз*, хотя и относятся к одной и той же зоне, ведут себя совершенно по-разному. Как уже говорилось выше, *глаз* — единственная часть тела (лица) из рассмотренных, которая не связана никакими

Таблица 6. Меронимы и комеронимы в группах ассоциатов и их связь с «зонами» тела

Векторное зонирование тела	Основное слово	Голова	Рука	Туловище	Нога	Другие типы связей
голова	голова	3				однокоренные слова — 3, синоним — 1
	лоб	10				
	глаз	2				однокоренные слова — 3, синоним — 1, квазисиноним — 1
руки	рука		5			однокоренные слова — 3, квазисиноним — 1
руки и ноги	нога		1		5	однокоренные слова — 2, квазисиноним — 2
голова и туловище без рук и ног	подбородок	8		2		
голова и туловище без рук	спина	1		6	2	квазисиноним — 1
голова, руки, верхняя часть туловища	грудь	3	2	5		
туловище без головы с руками и ногами	локоть		1	3	4	однокоренное слово — 1
	бедро		1	4	4	синоним — 1
	колени		2	4	1	гипероним — 1 однокоренное слово — 1, квазисиноним — 1
туловище без головы с ногами	туловище			5	3	синоним — 1, мероним животного — 1
	живот			4	3	однокоренное слово — 1, синоним — 2
	пятка			4	6	
все тело	шея	2	1	3	4	
	плечо	1	2	4	2	однокоренное слово — 1
	ладонь	1	5	1	1	однокоренное слово — 1, синоним — 1

комеронимическими отношениями, кроме собственных меронимов. Напротив, *лоб* — единственное из рассмотренных слов, чья группа ассоциатов связана отношением комеронимии в пределах одной зоны. Наблюдаемое различие показывает то, как контексты употребления слова модернизируют его внутреннее лексическое содержание. Контекстуально *глаза* (в отличие от *лба*) теряют свою денотативную фиксированность и как бы «отрываются» от лица. *Глаза* и *лицо* могут быть сопоставлены как два отдельных объекта, например, они могут выражать

противоположные эмоции (глаза улыбаются, а лицо серьезно), такое сопоставление оказывается значимым и повторяющимся способом репрезентации портрета человека в тексте. Корпусные примеры (4–6) ниже демонстрируют три синтаксические конструкции, в которых *глаза* и *лицо* выступают как два равноправных объекта. В примере (4) приведена редкая, но встречающаяся в корпусе конструкция с родительным падежом *глаза лица*, недопустимая для других частей тела/лица [Рахилина 2008: 46].

- (4) В живых черных *глазах* этого *лица* видно много уцелевшего огня и нежности, а характерные заломы в углах тонких губ говорят о силе воли и сдержанности. [Н. С. Лесков. Некуда (1864)].

В примере (5) мы видим конструкцию с эллипсисом при общей вершине — *выражение глаз и лица*

- (5) Там Мостовской в последний раз видел Софью Осиповну, — она стояла посреди пыльного двора, без пилотки, с сорванными знаками различия, и восхитила Мостовского угрюмым, злобным выражением *глаз* и *лица*. [Василий Гроссман. Жизнь и судьба, ч. 1 (1960)].

Наконец, в примере (6) использована конструкция глагола со значением перемещения от исходной точки, в которой исходная точка выражена именной группой с предлогом *с*, ср: *слезть с дивана, свисать с крыши, сорваться со скалы*, и в том числе и *смотреть с балкона*, где *смотреть* описывает перемещение взгляда.

- (6) Острые, как шпильки, *глаза* смотрели прямо на него с *лица*, очень худого, но не болезненного; могучий и кроткий сарказм змеился в углу тонких губ, обведенных длинной золотистой бородой, завивавшейся наподобие штопора и висевшей ниже второй пуговицы цветного жилета. [А. С. Грин. Ива (1923)].

В то же время, *голова* является наиболее меронимически дифференцированной зоной тела, ее части максимально функционально значимы, и именно поэтому *лоб*, как наиболее денотативно фиксированная часть *головы*, собирает в своей группе ассоциатов максимальное количество комеронимов.

Кроме слов, сконцентрированных вокруг одной зоны (помимо *головы*, к ним относится еще и *рука*, чьи собственные меронимы принадлежат только к зоне своего холонима), выделяется группа слов, ассоциаты которых группируются по зонам, противопоставленным условно верх-низ, при этом конечности (руки и ноги) могут входить или не входить в них. Это слова «верха»: *подбородок, спина, грудь*, слова «средней зоны»: *локоть, бедро, колено*, и слова «низа»: *туловище, живот, пятка*. Как видно из этих списков, векторное семантическое «зонирование» не всегда соответствует зонированию биологическому. Наконец, есть группа слов, которые имеют в группе ассоциатов меронимы и комеронимы из всех зон. Это слова, обозначающие подвижные суставы верхней части тела — *шея, плечо, ладонь*.

Важным признаком также является «плотность» меронимического ряда: насколько вектора комеронимов доминируют в группах ассоциатов и полностью заполняют собой список из наиболее близких по векторному представлению слов. Мы видим четыре функционально значимых части тела: голова, глаз, нога, рука, контекстные вектора которых в гораздо меньшей степени связаны с собственно меронимическими отношениями, эти части тела в своих контекстных репрезентациях ведут себя более самостоятельно, чем другие. Их самостоятельность очень хорошо видна на уровне способности образовывать лексико-грамматические конструкции, устойчивые сочетания и входить в идиомы, т. е. включаться в совершенно иные семантические поля и контексты; ср. такие выражения, как *выбросить из головы*, *прийти в голову*, *вешать голову*, *вскружить голову*, *глаза на мокром месте*, *глаза разбежались*, *мозолить глаза*, *как рукой сняло*, *встать не с той ноги*, *выбить почву из-под ног*, *легкая рука*, *широкая рука*, *попасть под горячую руку*, и т. д. Возможно, идиоматические обороты также влияют на уникальность их контекстов и, соответственно, на отдаление их векторов от комеронимов частей тела.

Подводя итог рассмотрению семантических отношений векторов слов со значением частей тела, заметим, что в семантическом векторном пространстве Основного корпуса тело человека представлено в разных перспективах — как портрет или бюст, как вертикаль фигуры, как отдельная часть фигуры, или же как фигура полностью. Перспектива задается семантическим вектором основного слова. Источник этих перспектив, их реализация в языке и, в частности, в текстах русской литературы, может стать весьма интересным вопросом для отдельного исследования.

3.2.2. Часть как материал

Для отбора слов в значении материала были использованы инструменты Основного поиска: в лексико-семантическом поиске была задан запрос, состоящий из существительного с семантикой предмета, предлога *из* и существительного с семантикой материала или вещества. Полученные результаты были проанализированы в таблице частотности. Для анализа были выбраны лексемы, различающиеся по своим семантическим свойствам — естественность или искусственность материала, наличие родственного прилагательного для передачи значений «сделано из, состоит из» (типа *камень — каменный*), наличие прототипического холонима — лексемы определенного семантического класса:

- *папье-маше* — искусственный материал, употребляется в основном в значении материала, нет родственного прилагательного для выражения «сделано из», может быть меронимом разных объектов;
- *кирпич* — отдельный объект, использующийся как материал, значение «сделано из» может быть заменено прилагательным *кирпичный*, прототипический холоним — лексемы со значением строения;
- *стекло* — артефакт и материал, употребляется не только в значении материала, может быть не только материалом, но именно частью предмета

(окно, очки), прилагательное *стеклянный* может выражать значение «сделано из», есть несколько прототипических холонимов разных классов — *окно, очки, посуда*;

- *камень* — естественный материал, может быть самостоятельным объектом и употребляться в разных значениях, прилагательное *каменный* может выражать значение «состоит из, сделано из», нет прототипического холонима;
- *изумруд* — естественный материал, прилагательное *изумрудный* может выражать значение «сделано из», прототипическим холонимом являются лексемы, имеющие общий семантический компонент со значением «украшение».

Группы ассоциатов для рассматриваемых лексем приведены в Таблице 7. В группу включались до 15 ассоциатов существительных с уровнем близости выше 0,5. Псевдослова, получившиеся в результате неправильной лемматизации, игнорировались.

Таблица 7. Группы ассоциатов слов со значением материала

Основное слово	Ассоциаты
папье-маше	<u>когипонимы</u> : терракота (0.73), оргстекло (0.72), пластилин (0.7), плексиглас (0.7), нержавейка (0.69), пластмасса (0.67), пенопласт (0.7), стекловолокно (0.65), кожаменитель (0.65), пластик (0.64), латунь (0.64), кожзам (0.64), стеклопластик (0.64)
кирпич	<u>комеронимы/когипонимы</u> : черепица (0.74), шифер (0.73), глина (0.72), бетон (0.71), плитняк (0.7), бревно (0.69), цемент (0.67), тес (0.67), известка (0.67), камень (0.67), алебастр (0.66), щебенка (0.66), булыжник (0.65) <u>лексическая функция</u> : кладка (0.66)
стекло	<u>когипоним</u> : слюда (0.65), оргстекло (0.6), жель (0.6) <u>комероним</u> : рама (0.67), замазка (0.61) <u>холоним</u> : оконница (0.73), окно (0.67), оконце (0.62), иллюминатор (0.6) <u>квазисиноним</u> : стенка (0.6) <u>однокоренное слово</u> : стеклышко (0.68), стекляшка (0.61)
камень	<u>квазисиноним/гипоним</u> : валун (0.74), булыжник (0.72), гранит (0.68), галька (0.68), плитняк (0.67), голыш (0.65) <u>квазисиноним/холоним</u> : глыба (0.72), скала (0.71) <u>когипоним</u> : кирпич (0.67), песок (0.66) <u>однокоренное слово</u> : камешек (0.72), камня (0.71), камушек (0.7) <u>холоним</u> : камео (0.67)
изумруд	<u>когипоним</u> : рубин (0.9), сапфир (0.88), бирюза (0.85), аметист (0.84), яхонт (0.84), топаз (0.80), смарагд (0.80), сердолик (0.8), кабошон (0.79), брильянт (0.78), хризолит (0.78), аквамарин (0.77), агат (0.77), алмаз (0.77), бриллиант (0.76)

Из Таблицы 7 мы видим, что слова, содержащие в своем значении семантический компонент материала или вещества, обнаруживают разное поведение. Слова, у которых это значение является основным, собирают в группе ассоциатов ряд когипонимов, объединенных более или менее эксплицитно выражаемым общим

понятием. Так, для слова *панье-маше* такой ряд определен понятием «материалы, используемые для создания артефактов» — *оргстекло, пластилин, кожзаменитель* и др. Ассоциатами слова *изумруд* являются названия других драгоценных камней. И в том, и в другом случае мы видим, что вся группа наиболее близких ассоциатов связана единым родо-видовым отношением.

Немного менее очевидным представляется ряд лексических отношений слова *кирпич*. Очевидно, что основным родовым понятием для этого слова является «строительный материал», таким образом, его представители — *бетон, черепица, шифер, тес, бревно*, одновременно являясь когипонимами, могут обозначать и составные части строения, то есть находиться в отношении комеронимии — «частей одного целого». В отличие от частей тела, где каждая такая часть является функционально и структурно единственно возможным компонентом целого, материалы и вещества вариативны, дополнительные друг к другу и необязательны. Дом может быть построен из кирпича или из бревен, а может быть и из того и другого. В результате меронимические отношения ассоциатов материалов и частей тела принципиально различаются. Части тела, как было показано выше, задают четкую систему локальных отношений, когда контекстуальные отношения семантической близости соотнесены с определенной зоной или профилем, выделенными из общего целого. Ассоциаты материалов, напротив, образуют своего рода «серую зону», в которой отношения гипонимии, меронимии и неточной синонимии оказываются сложно различимы и не противопоставлены. Это особенно хорошо видно на примере группы ассоциатов слова *камень*, в отличие от рукотворных артефактов природные объекты не существуют для предзаданной, пусть даже и очень широкой, функции. Выделенные лексические отношения в этой группе очень условны. Является ли ряд *валун, булыжник, голыш, галька, гранит* видами (гипонимами) *камня*, или же их можно рассматривать как ряд квазисинонимов [Апресян 1995: 235], имеющих общее семантическое ядро, но различающихся спецификацией формы, размера, жесткости? Является ли *камень* меронимом (веществом, материалом) *скалы*, или же они выступают как контекстно заменимые синонимы?

Группы ассоциатов лексем со значением материалов могут, таким образом, быть связаны общей семантикой функционально заданной, фреймовой когипонимии (то, с помощью чего можно делать артефакты; то, с помощью чего можно строить; драгоценные камни, которыми украшают; ср. выше ассоциативный ряд слова *шампиньон*, связанный не с классом грибов, а с функциональным классом кулинарных растительных ингредиентов); члены этого ряда будут противопоставлены по своим функциональным свойствам в рамках одного фрейма (строительство, украшение камнями, производство артефактов) и будут специфицировать определенные параметры общего семантического ядра. Но если такая семантика не выводится ясным образом из ряда ассоциатов, т. е. мы не можем определить фундамент для противопоставления, то такие отношения скорей всего будут квалифицированы как квазисинонимические (или как неточные синонимы). Важно понимать,

что ассоциаты-когипонимы, связанные разницей функциональных характеристик своих свойств в рамках одного фрейма (*кирпич, бетон, тес*), и ассоциаты-квазисинонимы, задающие спецификацию параметров объекта (*валун, булыжник, галька*), связаны, на самом деле, семантически подобными отношениями. Можно представить себе ситуацию, в которой спецификация параметров объекта оказывается функционально противопоставленной (например, камни разного размера, которые используются в некой воображаемой игре) и квазисинонимы будут переосмыслены как когипонимы. Важно также подчеркнуть, что чем менее параметрически и функционально определено слово в базовой семантике (ср. *камень*), тем более неоднородным и «лабильным» оказывается его ассоциативный ряд. Представляется, что эти наблюдения имеют базовый характер и сохраняются независимо от конкретных реализаций ассоциативных рядов, которые, как уже было сказано, напрямую определяются исходными корпусными данными.

Отдельный интерес представляет группа ассоциатов слова *стекло*. Здесь мы действительно видим «следы» меронимических отношений типа “MadeOf”. Во-первых, в ассоциативном ряду присутствуют сразу несколько холонимов, обозначающих целое, составной частью или материалом которого может быть *стекло*: это словообразовательные варианты лексемы *окно*, а также слово *иллюминатор*. Вариативность лексем с общим корнем *окно* указывает на то, что это слово (корень) является здесь прототипическим холонимом. Во-вторых, в группе ассоциатов имеются комеронимы к слову *окно* и его дериватам — *рама* и менее очевидное *замазка*. Последнее — хороший пример того, как экстралингвистические знания отличаются от лексической системы, строящейся на данных контекстов употребления. Едва ли можно сказать, что конструктивно *замазка* является частью *окна*, подобно *раме* или *стеклу*. Однако корпусные примеры показывают, что *замазка* именно принадлежит *окну* как целому и соотносится с его частями:

- (7) *Рамы*, составленные из обломков *стекол*, грубо скрепленных *замазкой*. [В. А. Каверин. Открытая книга (1949–1956)].
- (8) Затем, еще раз оглядевшись, Куренцов достал перочинный ножик, отколупнул сухую *замазку*, отогнул гвоздики и вынул двойное *стекло*. [Григорий Горин. Обнаженный Куренцов (1960–1985)].
- (9) В окошечках не было ни одного цельного *стекла*: в переплетах — множество осколков, сплетенных *замазкой*, скрепленных лучинками. [Ф. В. Гладков. Повесть о детстве (1948)].

Наконец, *стекло* имеет и ряд когипонимов разной степени близости (*слюда, оргстекло, жемчуг*). Таким образом, *стекло* отчасти выступает структурной частью целого (подобно частям тела), и связывается со своими холонимами и комеронимами, но одновременно выступает и как материал, собирая вокруг себя функциональный ряд когипонимов.

Квазисиноним *стенка* имеет с основным словом общий функциональный семантический компонент — так же, как и *стекло*, *стенка* разделяет два пространства, является преградой.

Таким образом, дистрибутивная семантическая модель задает новую смысловую проекцию, выделяя семантический компонент, который формируется в результате употребления слов (*usage-based*), но не содержится в их собственном лексическом ядре.

Завершая рассмотрение групп ассоциатов лексем со значением материала, обратим внимания на слово *кладка*, ассоциат слова *кирпич*. В Таблице 7 это отношение определяется общим термином «лексическая функция». Как будет показано ниже, ассоциативные отношения часто соотносятся не с базовыми словарными отношениями, такими как гиперонимия, синонимия, антонимия, а с параметрическими, чрезвычайно похожими на лексические функции модели «Смысл — Текст» [Мельчук 1999]. Так, в толково-комбинаторном словаре в толкование слова *кирпич* входит глагол *класть*, связанный с ним с помощью лексической функции *Real₁* — ‘выполнять, делать, что положено’ [Мельчук, Жолковский 1984: 369]. Очевидно, что субстантивация глагола *класть* и неизбежный семантический сдвиг образованной лексемы модифицирует отношения пары *кирпич:кладка*, однако они все равно остаются параметрическими, определяемыми в духе лексикографических подходов московской семантической школы.

Далее мы рассмотрим две лексические функции, связывающие именно существительные с существительными и так или иначе относящиеся к семантическому типу отношений часть-целое. Это лексические функции **Sing** и **Mult**, противоположные по своему значению. В первой *X*-ом обозначается элемент, порция, штука *Y*-а, вторая, наоборот, определяет *X* как естественную совокупность объектов *Y*-а.

3.2.3. Лексические функции *Sing* и *Mult*

Для анализа групп ассоциатов, которые потенциально могут быть связаны с помощью функции *Sing*, были выбраны четыре слова. Три из них — *кусок*, *лосок*, *ломоть* — содержат в своей семантике значение части целого, последнее — *нота* — скорее может быть определено как минимальный элемент, необходимый для получения целого (*музыкального произведения, мелодии* и т. п.). Ряды ассоциатов к этим лексемам представлены в Таблице 8.

Как видно из Таблицы 8, слова *кусок* и *ломоть* оказываются очень похожи по тому, как выстраиваются лексические отношения в их группах ассоциатов. Оба собирают ряды квазисинонимов, объединенных функцией **Sing** по отношению к холониму, который сам не входит в ассоциативный ряд. Заметим, что, по-видимому, под влиянием социокультурных факторов, проявившихся в текстах Основного корпуса, ясно виден семантический прототип и целого, и его частей: целое — это хлеб (*коврига, каравай, батон*), части — названия, специфически употребляемые для частей хлеба (*краюха, горбушка, полбуханки*). Очевидно, что для хлеба, как элемента текстового нарратива, само соотношение (и противопоставление) части-целого оказывается особенно хорошо лексически разработанным и контекстуально выявленным. Отметим также, что в ассоциативных рядах обоих слов присутствуют

Таблица 8. Ассоциаты слов с лексической функцией Sing

Основное слово	Ассоциаты
кусок	<u>холоним</u> : коврига (0.65) <u>квазисиноним</u> : ломоть (0.67), шматок (0.65), лоскут (0.64), краюха (0.63), обрезок (0.63), лоскуток (0.62), шмат (0.62), ломтик (0.61) <u>однокоренное слово</u> : кусочек (0.85), кус (0.68)
ломоть	<u>холоним</u> : коврига (0.75), каравай (0.7), батон (0.67) <u>квазисиноним</u> : краюха (0.79), горбушка (0.76), краюшка (0.72), шмат (0.7), шматок (0.66), кусок (0.67), скибка (0.67), , полбуханки (0.68) <u>однокоренное слово</u> : ломтик (0.71),
лоскут	<u>комероним</u> : тесьма (0.69), тесемочка (0.68), тесемка (0.67) <u>мероним</u> : сатин (0.67) (в значении материала лоскута) <u>квазисиноним</u> : холстина (0.74), мешковина (0.7), коленкор (0.7), обрезок (0.68), тряпица (0.72), дерюга (0.7), <u>однокоренное слово</u> : лоскуток (0.86)
нота	<u>холоним</u> : аккорд (0.64), октава (0.61), мелодия (0.60), трезвучие (0.56), камертон (0.56), партитура (0.56), рулада (0.56) <u>лексическая функция</u> : регистр (0.57), струна (0.56), фальцет (0.56), альт (0.55) <u>однокоренное слово</u> : нотка (0.56)

слова *шмат* и *шматок*, связанные функцией **Sing** с другим прототипическим объектом — *сало*. Такого рода семантические пропорции чрезвычайно характерны для ассоциативных групп в целом (ср. описанную выше связь *субботы* и *сочельника*): к основному слову X_1 имеется ассоциат X_2 , при этом X_1 связан с Y_1 с помощью той же лексической функции, что X_2 с Y_2 .

Ассоциативный ряд слова *лоскут* выглядит иначе, собственно к функции **Sing** может быть отнесен квазисиноним *обрезок*, а также холоним *сатин*, при этом в группу ассоциатов попадает целый ряд лексем, объединенных значением «кусок ткани для функционального использования». Эта функциональность может быть выражена лексически — *дерюга*, *коленкор*, ср. примеры из корпуса (10–13):

- (10) Давид стал молиться о младенце, постился, ходил одетый в *дерюгу* и спал прямо на земле. [Библия. Современный русский перевод Книги Священного Писания Ветхого и Нового Завета Канонические. Ветхий завет. Вторая книга Царств (2011)].
- (11) Вернее — обернуть кусок безобразной *дерюги* вокруг пояса. [Андрей Пермяков. Темная сторона света // «Волга», 2013].
- (12) Тусклая лампочка в проволочной сетке. Обитая *коленкором* дверь. Коммунальные соседи — Гендлины, Моргулисы, Шарашенидзе. [Сергей Довлатов. Иная жизнь (1984)].
- (13) Гослитовский том 1955 года издания в *коленкоре* мышинового, под стать фамилии, цвета. [С. Г. Боровиков. Небунин, или Россия, которую мы не потеряли // «Волга», 2015].

Функциональность ткани также может быть эксплицитно определена с помощью корня и суффикса *-ин*: *мешковина* — ткань для мешков или *холстина* — ткань, которую можно использовать для разных нужд, ср. примеры (14–16):

- (14) Были остатки солдатской гимнастерки, летние штаны из посконной *холстины*, имелось пять рубах. [Даниил Гранин. Зубр (1987)].
- (15) Тут же продают ящики, свежую стружку, *холстину* для обшивки. [Ю. О. Домбровский. Факультет ненужных вещей, часть 1 (1978)].
- (16) По воспоминаниям современников, вся площадь была заполнена людьми, охваченными восторгом, в момент снятия с монумента *холстины* толпа прокричала «ура». [Наталья Бесхлебная. Пампуш на Твербуле. Краткая история памятника Пушкину: от Достоевского до митинга (06.06.2017)].

Наконец, *лоскут* так же, как и *стекло*, может выступать в качестве «конструктивной части», при котором целое образуется благодаря соединению с другой функционально отличающейся частью. Такая часть, вернее словообразовательный ряд — *тесьма*, *тесемка*, *тесемочка* попадают в отношения комеронимии со словом *лоскут*, однако их потенциальный общий холоним не попадает в топ списка группы ассоциатов.

Наконец, принципиально отличным от всех слов, рассматриваемых в рамках анализа отношений часть-целое, выглядит ассоциативная группа лексемы *нота*. Мы видим очень богатый ряд холонимов, связанных с *нотой* с помощью отношения **Sing**: *нота* является элементом, экземпляром таких сущностей как *аккорд*, *октава*, *трезвучие*, *камертон*, *партитура*, *рулада*. Действительно, в данном случае лексически противопоставленными оказываются не отдельные меронимы (разные по тональности и длительности музыкальные звуки), а именно их комбинации (*аккорд*, *октава*, *трезвучие*). Таким образом, мы наблюдаем случай, когда у нас ряд когипонимов общего понятия «музыкальные фразы» объединяется в одной ассоциативной группе их общего меронима, их элементарного компонента — *ноте*. Отметим, впрочем, что мы имеем дело с лемматизированными словами, однако иногда во множественном числе они приобретают дополнительную семантику. Так, можно предложить, что если рассматривать не слово *нота*, а его множественное число *ноты*, то скорее всего, *партитура* является его синонимом. Наконец, отметим, что кроме ряда холонимов / когипонимов в группе ассоциатов к слову *нота* выделяется также ряд несомненно связанных с ним слов, лексическая связь которых с главным словом не укладываются ни в родо-видовые, ни в партитивные, ни в синонимические отношения — *регистр*, *струна*, *фальцет*, *альт*. Такие слова мы условно относим к общей категории других лексических функций, более подробное исследование которых остается пока за рамками настоящей статьи.

Для анализа функции **Mult** были выбраны два слова, которые иллюстрируют использование этой функции в толковом комбинаторном словаре — *флот* и *студенчество* [Мельчук, Жолковский 1984: 83]. В первом случае мы ожидаем увидеть многообразие возможных меронимов — единиц, которые входят во *флот*,

во втором случае таких меронимов может быть только два — *студент* и *студентка*. Данные ассоциативных групп этих двух лексем приводятся в Таблице 9.

Таблица 9. Ассоциативные группы слов с функцией Mult

Основное слово	Ассоциаты
флот	<u>синоним</u>: эскадра (0.77) <u>холоним/комероним</u>: армия (0.68) <u>мероним</u>: корабль (0.66), судно (0.66), крейсер (0.64), фрегат (0.62) <u>комероним</u>: авиация (0.63), войско (0.59) <u>лексическая функция</u> десант (0.61), порт (0.59) <u>однокоренное слово</u>: флотилия (0.73)
студенчество	<u>квазисиноним/когипоним</u>: интеллигенция (0.62), молодежь (0.62), профессура (0.6), офицерство (0.58), эмиграция (0.52), <u>мероним</u>: студент (0.51) <u>гипероним</u>: общественность (0.5) <u>неинтрепретируемое</u>: журналистика (0.55),

Ассоциативная группа лексемы *флот* включает в себя достаточно разнообразные и ясно дифференцируемые лексические связи. Это, во-первых, как и ожидалось, собственно, меронимы (единицы, объединение которых в одну совокупность образует *флот*): более общие *корабль*, *судно* и более специфические *крейсер*, *фрегат*. Во-вторых, лексема *флот* сама выступает как элемент множества, будучи связана с помощью функции **Sing** с еще более общей совокупностью — *армией* в значении «вооруженные силы» в целом. Однако, поскольку существует выражение *армия и флот*, можно понимать *армию* и как комероним *флота*, включенного в ряд элементов общего множества *вооруженных сил*. В ассоциативной группе мы видим и другие комеронимы к этому общему холониму. Частью вооруженных сил, кроме *флота*, могут быть также *авиация* и *войско*. Наконец, *флот* имеет среди ассоциатов сравнительно близкий и по лексическому наполнению, и по значению контекстуального сходства (0.77) синоним *эскадра*, ср. определение в Толковом словаре Кузнецова: «Крупное соединение военных судов или самолётов» [БТС 1998], а также несколько слов, связанных с помощью лексических функций: S_{instr} — *десант* (*флот* как средство передвижения *десанта*) и S_c — *порт* (место размещения *флота*).

Гораздо менее структурированной (но не менее интересной) является ассоциативная группа лексемы *студенчество*. Как и ожидалось, в группу ассоциатов входит ее основной и единственный (если не считать однокоренного феминитива) мероним — *студент*. При этом уровень его близости к основному слову составляет 0.51, что немногим выше заградительного барьера в 0.5. В основном группу заполняют квазисинонимы, объединенные тем же самым значением лексической функции **Mult**, но по отношению к другим меронимам — *интеллигенция*, *молодежь*, *профессура*, *офицерство*. Наиболее интересным представляется наличие в группе ассоциатов общего слова к этому ряду — лексемы *общественность*. В Толковом словаре Кузнецова это слово толкуется как «Часть общества, активно

участвующая в общественной жизни, выражающая общественное мнение» [БТС 1998]. Семантика «активной части общества» дополняет приведенный выше квазисинонимический ряд лексическим отношением когипонимии: ассоциаты выстраиваются не только как альтернативные названия похожих совокупностей, но как и представители классов страт, выражающих общественное мнение. Эта семантика не собственно словарная, но приобретенная, контекстуальная. Здесь нельзя не обратить внимание на слово *эмиграция*, имеющее сравнительно низкий коэффициент близости (0.52). Очевидно, что это слово означает не действие, но совокупность *эмигрантов* (ср. «белая эмиграция»). В этом смысле слово *эмиграция* встает в один ряд со своими когипонимами, такими как *студенчество* или *молодежь*.

Итак, мы рассмотрели разные примеры отношений типа часть-целое, в отличие от родо-видовых связей такие связи гораздо менее регулярны. Части тела образуют совершенно особую лексическую систему, практически непроницаемую и непохожую ни на какое другое семантическое поле. Группы ассоциатов к словам со значением материала оказываются очень плохо структурированными и, как правило, тяготеют в сторону функционально заданных гипонимических рядов или неточной синонимии. Наконец, слова, связанные с лексическими функциями **Sing** и **Mult**, собирают в группах своих ассоциатов квазисинонимы — слова, характеризующиеся той же самой лексической функцией, которая может быть связана с общей лексемой (*кусок, ломоть*) или с другими лексемами (*студенчество, интеллигенция*). Более подробно то, как работает квазисинонимия и шире синонимия в векторном семантическом пространстве, будет рассмотрено в следующем разделе статьи.

3.3. Синонимические отношения

Лексическому отношению синонимии посвящено огромное множество работ по теоретической лексикологии. Одновременно, синонимия оказывается ключевым понятием прикладной лексикографии, для толкования слова лексикографы подбирают синонимы, отдельно синонимические ряды публикуются в многочисленных словарях синонимов.

Сразу оговоримся, что в задачи настоящего исследования не входит обсуждение разных возможных подтипов синонимических отношений, нас интересует в целом то, как и насколько синонимические отношения проявляются в дистрибутивной семантической модели Основного корпуса. Тем не менее, чтобы как-то разделить синонимию и квазисинонимию, обратимся к определению, представленному в [Апресян 1995]. Ю. Д. Апресян противопоставляет точные синонимы и квазисинонимы. Для первых характерными признаками является 1) одинаковое толкование с помощью выражений семантического языка, 2) одинаковое число валентностей и совпадение актантов, заполняющих эти валентности, по семантическим ролям и 3) принадлежность к одной и той же части речи. Для квазисинонимов первый признак может выполняться не полностью [Апресян 1995: 223]. Отмечается, что общая часть толкований квазисинонимов должна быть больше их различий, однако собственно то, каким образом можно сравнить объемы общей

и различающихся частей, остается непроявленным [Апресян 1995: 158]. В этом смысле интересным является вопрос — если векторная семантическая модель отражает синонимические отношения, является ли более высокий индекс схожести слов отражением большего объема их общего значения относительно имеющих различий. Ответ на этот вопрос не очевиден. С одной стороны, известно, что одним из популярных методов выделения синонимов является дистрибутивный критерий синонимичности — полная или частичная взаимозаменяемость слов в одном и том же контексте, с другой стороны, как отмечается в [Апресян 1995], очень часто синонимы (или квазисинонимы) не встречаются в одних и тех же контекстах: они имеют общее толкование, но различаются ограничениями по лексической сочетаемости. При этом нужно понимать, что сам метод дистрибутивной семантики не предполагает абсолютное равенство контекстов при построении векторной модели: сравниваемые контексты преобразуются в так называемые «мешки слов» — списки лемм, не учитывающих порядок слов и их синтактико-грамматические признаки. Таким образом, близкие по расстоянию вектора могут иметь слова, которые в принципе часто встречаются рядом в окружении одних и тех же слов, находясь при этом в разных синтаксических позициях (именно это, по-видимому, имеет место в обсуждаемой выше паре *кирпич* — *кладка*). Кроме того, из того, что слова встречаются в идентичных контекстах, не следует идентичность их значения. Это очень хорошо показывают векторные репрезентации представителей закрытых списков названий, как например, в случае названий месяцев, рассмотренных выше.

Говоря об источниках точной синонимии, Ю. Д. Апресян указывает, во-первых, на экспрессивную лексику, во-вторых, на метафорическое употребление и идиомы, кроме того, источниками точной синонимии могут быть заимствования, архаизмы и дериваты [Апресян 1995: 282]. Поскольку в группах ассоциатов невозможно отделить метафорическое или идиоматическое употребление от основного, а рассмотрение лексических отношений дериватов не входит в задачи исследования, ограничимся рассмотрением экспрессивно маркированных слов, заимствований и архаизмов, чтобы оценить насколько в группы их ассоциатов попадают слова близкого или идентичного значения. Кроме того, будет рассмотрено существительное, выражающее эмоции — *радость*, потенциально богатое синонимическими и антонимическими отношениями, а также два источника квазисинонимов в векторных представлениях — отглагольные существительные и существительные с диминутивными суффиксами.

3.3.1. Экспрессивные слова, заимствования, архаизмы, эмоции

Для анализа ассоциативных рядов экспрессивных слов были выбраны два слова с пейоративным значением — *идиот* и *подлец*. Следует сразу отметить, что при анализе лексических отношений ассоциатов и основного слова разделение между синонимами и квазисинонимами производилось достаточно грубо. С точки зрения теории лексико-семантических отношений, изложенной в [Апресян 1995], большинство ассоциатов, отнесенных к синонимам, являются неточными синонимами

или квазисинонимами. Однако, представляется важным развести слова, которые, выражая общее значение, различаются логическим ударением [Апресян 1995: 245], акцентом на один из семантических компонентов или наложением на него дополнительных ограничений и т. д. от слов, которые также имеют общий семантический компонент, но на очень абстрактном уровне, а на уровне лексической реализации их значения сильно различаются. В последнем случае мы имеем дело с синонимией на уровне абстрактного семантического ядра, усложненной лексико-параметрическими отношениями, в которых различные спецификации параметров этого ядра дают разные реализации лексем. Первый класс случаев будет условно называться синонимами, а второй класс случаев — квазисинонимами. Продемонстрируем это различие на данных существительных с экспрессивным значением, представленных в Таблице 10.

Таблица 10. Группы ассоциатов экспрессивных существительных

Основное слово	Ассоциаты
идиот	<u>синоним</u> : кретин (0.82), дебил (0.74), болван (0.73), дурак (0.7), придурак (0.7), тупица (0.69), недоумок (0.68) <u>квазисиноним</u> : псих (0.68), негодяй (0.67) <u>однокоренное слово</u> : идиотка (0.74)
подлец	<u>синоним</u> : мерзавец (0.86), негодяй (0.8), сволочь (0.77), прохвост (0.76), каналья (0.71) <u>квазисиноним</u> : хам (0.71) дурак (0.7), мошенник (0.7), жулик (0.69), трус (0.68)

Итак, в Таблице 10 мы видим синонимический ряд к слову *идиот*, сюда входят экспрессивные слова, обозначающие человека с ограниченными умственными способностями. Можно предположить, что они различаются, во-первых, в оценке недостаточности умственных способностей по шкале от минимального (*недоумок*, *придурак*) до максимального (*кретин*, *дебил*), а во-вторых, в силе выражаемой экспрессии. Однако так или иначе, в целом, семантика этих слов совпадает. К квазисинонимам были отнесены другие пейоративы, связанные с отсутствием иных базовых качеств, которые ожидаются от человека в обществе. Таким образом, абстрактным значением для квазисинонимов этого ряда становится «осуждение человека за несоответствие норме поведения в результате отсутствия некоторого качества X», а сопутствующая лексическая функция специфицирует это качество и соотносит получившееся значение с лексической реализацией:

X = ум >> *идиот*

X = контроль за поведением, самоконтроль >> *псих*

X = моральные ориентиры >> *негодяй*

Ту же самую картину мы наблюдаем в группе ассоциатов слова *подлец*. В этом случае, синонимический ряд собирается вокруг значения «аморальности», а квазисинонимический ряд вокруг отсутствия других качеств: вежливости — *хам*, ума — *дурак*, честности — *мошенник*, *жулик*, смелости — *трус*. Таким образом, синони-

мические и квазисинонимические ряды составляют на самом деле «двухъярусные» родо-видовые отношения: ряд синонимов является когипонимами к общему специфицированному признаку — недостаток ума для ассоциатов слова *идиот*, аморальность для ассоциатов слова *подлец*, а квазисинонимы в свою очередь являются когипонимами на еще более высоком абстрактном уровне «отсутствие социально значимого качества». Структура лексических отношений, проявляющаяся в ассоциативных группах экспрессивных слов, очень похожа на иерархическую структуру лексикографической сети Wordnet, в которой синонимические ряды упорядочиваются именно в рамках отношений гипероним-гипоним от наиболее абстрактных смыслов к более специфицированным.

В Таблице 11 приведены группы ассоциатов двух заимствованных слов *эффективность* и *абракадабра*. В первом случае мы имеем дело с рядом условных (неточных) синонимов, их объединяет значение «положительной оценки экономической деятельности предприятия». Квазисинонимами являются лексемы, в которых сохраняется заданный фрейм «положительная оценка деятельности», но специфицируется тип предприятия и соответствующий параметр оценки: оценка деятельности банка — *надежность*, оценка деятельности СМИ — *информативность*. Заметим, что *эффективность* по сути является гиперонимом ко всем словам группы.

В группе ассоциатов к слову *абракадабра* девять слов из десяти могут быть объединены в синонимический ряд, при этом, хотя значение «непонятности, непрозрачности смысла» сохраняется, усиливается значение негативной оценки. В квазисинонимы попадает слово *пошлятина* с достаточно низким по сравнению с другими членами группы уровнем близости (0.62), общим абстрактным семантическим ядром здесь является, по-видимому, именно значение негативной оценки, а также, возможно, объект, на который оно может быть направлено и который может быть определен как «продукт креативной деятельности человека». *Абракадабра* может быть употреблена как негативная оценка речевой/языковой деятельности, текста, *пошлятина* — скорее как негативная оценка любого художественного произведения.

Таблица 11. Группы ассоциатов заимствованных слов *эффективность* и *абракадабра*

Основное слово	Ассоциаты
эффективность	синоним: результативность (0.75), прибыльность (0.7), экономичность (0.66), продуктивность (0.64), оперативность (0.62), перспективность (0.62), конкурентоспособность (0.61) квазисиноним: надежность (0.7), информативность (0.6) однокоренное слово: энергоэффективность (0.62)
абракадабра	синоним: тарабарщина (0.72), несурязица (0.69), галиматья (0.68), белиберда (0.67), бессмыслица (0.64), нелепица (0.64), чепуха (0.63), невнятица (0.62), чушь (0.61) квазисиноним синоним: пошлятина (0.62)

Таким образом, действительно, экспрессивная и заимствованная лексика существует в зоне слабой «семантической дифференциации» [Апресян 1995: 224–225] и в группы их ассоциатов попадают в первую очередь слова, выражающие близкие значения. Близость их контекстов на уровне векторной репрезентации соответствует близости их толкования.

Иначе обстоит дело с архаизмами. Мы рассмотрим два слова *алчность* и *агнец* (Таблица 12).

Таблица 12. Группы ассоциатов устаревших слов *алчность* и *агнец*

Основное слово	Ассоциаты
алчность	<u>синоним</u> : корыстолюбие (0.77), стяжательство (0.74), сребролюбие (0.71) <u>квазисиноним</u> : властолюбие (0.74), кровожадность (0.7), хищничество (0.69), сластолюбие (0.68), похоть (0.67), честолюбие (0.66)
агнец	<u>синоним</u> : ягненок (0.68), овечка (0.61) <u>квазисиноним</u> : серафим (0.58), голубица (0.58), херувим (0.56), богочеловек (0.56) <u>лексическая функция</u> : заклание (0.58), всеожжение (0.55)

В группу ассоциатов, связанных синонимическим отношениями, основного слова *алчность* попадают такие же, если не более архаичные слова *корыстолюбие*, *стяжательство*, *сребролюбие*. В группе квазисинонимов, объединенной общим абстрактным значением «значимая отрицательная черта характера человека, порок», слова *властолюбие*, *хищничество*, *сластолюбие* также являются устаревшими. Таким образом, векторная модель сохраняет стилистический компонент архаичности, а контекстные вектора современных слов соответствующих значений оказываются слишком далеки по косинусной близости: из контекстов, свойственных слову *алчность*, плохо предсказывается слово *жадность*, но хорошо *сребролюбие*. Этот факт отличает архаизмы и от заимствований, и от экспрессивной лексики и важен для понимания внутренней системности дистрибутивной семантической модели.

Наиболее интересной является группа ассоциатов слова *агнец*: кроме собственно синонимов *ягненок*, *овечка*, в группу ассоциатов включается целый ряд лексем, общее значение которых может быть определено не на семантическом уровне, а на уровне нарративного источника: «символические фигуры Евангелия, посланники Бога» — *серафим*, *голубица*, *херувим*, *богочеловек*. Само наличие этого ряда много говорит нам о внутренней структуре Евангельского текста — мы видим, что символические связанные фигуры устойчиво появляются в одних и тех же контекстах, и это выражается в сближении их семантических векторов. Кроме того, можно предположить, что эти контексты оказываются устойчивыми не только в самом тексте Евангелия, но воспроизводятся в дальнейшем при его цитировании и отсылках в других текстах Основного корпуса. Еще более интересным выглядит наличие в группе ассоциатов двух отглагольных существительных *заклание* и *всеожжение*. Связь основного слова *агнец* со словом *заклание* мы можем определить с помощью субстантивации лексической функции **S Oper** (*агнец*) = *заклание*.

Лексема *всесожжение* (козла, быка, барана, ягненка) используется в Старом завете для обозначения акта жертвоприношения животных. Таким образом, если для *агнца* функция **SO_{per}** оказывается общеязыковой — основное действие, которое производят с *агнцем* — это заклание, то для группы домашних животных (потенциальных когипонимов *агнца*, не входящих в его ассоциативную группу) такая функция существует только в пространстве текста Старого завета. В результате образуется сложная конструкция, где общим значением для квазисинонимов является значение «символического жертвоприношения в пространстве Библейского текста», специфицирующим лексическим параметром является противопоставление текста Старого завета, где говорится о всесожжении разных животных (домашнего скота) и метафорического выражения *агнец на заклание* в Новом завете и многочисленных последующих цитатах этой метафоры.

Таким образом, мы видим, что векторные репрезентации архаизмов оказываются очень сильно связаны именно с их употреблением в старых текстах, в отличие от заимствованной лексики, они не размывают семантическую дифференциацию, но, напротив, сохраняют более стройную систему семантических противопоставлений «общий признак vs. параметрическое различие». Это выражается и в разнообразии квазисинонимов слова *алчность*, и особенно в подчиненности семантической структуры текстовому символизму, которую мы наблюдаем в группе ассоциатов слова *агнец*.

Завершив анализ синонимических связей ассоциатов и основного слова рассмотрением существительного со значением эмоций — *радость* (Таблица 13).

Таблица 13. Группа ассоциатов слова *радость*

радость	синоним: счастье (0.68), восторг (0.68), удовольствие (0.67), наслаждение (0.67), веселие (0.66), блаженство (0.66) антоним: печаль (0.7), горесть (0.69), горе (0.67), огорчение (0.66)
---------	---

В ассоциативную группу слова *радость* входят не только синонимы, но и антонимы, причем наиболее близким по векторному значению к основному слову является именно его антоним *печаль*. Таким образом, группа ассоциатов слова *радость* можно определить в терминах Л. Г. Бабенко как синонимико-антонимический комплекс, «в котором ядром является антонимическая оппозиция, а приядерную зону заполняют синонимы, конкретизирующие ее опорные компоненты в левой и правой частях оппозиции» (Бабенко 2020: 12). В Новом объяснительном словаре синонимов слово *печаль* рассматривается как один из представителей синонимического ряда с доминантным словом *тоска*. Общее значение этого ряда объясняется как «неприятное чувство, какое бывает, когда нет того, что человек хочет, и когда он думает, что желаемое невозможно». Слово *радость* приведено в этой же статье в качестве антонима к слову *печаль* (НОСС 2004: 1165). В введении к «Новому объяснительному словарю синонимов» антоним определен как лексема, противопоставленная другой лексеме по одному из нескольких параметров. В рассматриваемом случае релевантны два параметра: ‘Р’ — ‘не Р’ и ‘хороший’ — ‘плохой’.

Таким образом, вся группа ассоциатов слова *радость* может быть связана с помощью одного общего абстрактного значения, в котором определенные компоненты оказываются неопределенными в рамках бинарных значений хороший-плохой и истинность/ложность события: «плохое/хорошее чувство, какое бывает, когда нет/есть то, что человек хочет, и когда он думает, что желаемое невозможно/реально». Спецификация этих трех компонентов семантики относительно их бинарных шкал в сочетании со значением интенсивности чувства так или иначе реализуется во всех ассоциатах рассматриваемой группы. Таким образом, единое семантическое ядро и противопоставленность лексических единиц вокруг этого ядра воспроизводится и на уровне векторной контекстной семантики. Отметим также близость всех ассоциатов группы друг к другу, показатели сходства их векторов колеблются в пределах от 0.7 до 0.66.

3.3.2. Квазисинонимы

Для анализа ассоциативных рядов, связанных отношением квазисинонимии, рассмотрим лексические отношения в группах ассоциатов четырех отглагольных существительных: *крик*, *стук*, *взлет* и *чтение*. Все эти слова имеют ассоциативные группы, которые преимущественно состоят именно из квазисинонимов — рядов лексем с общим компонентом значения, противопоставленных параметрически (Таблица 14).

Таблица 14. Группы ассоциатов отглагольных существительных

Основное слово	Ассоциаты
крик	<u>синоним</u> : вопль (0.89) <u>квазисиноним</u> : визг (0.81), рев (0.8), вой (0.75), возглас (0.74), гогот (0.71), рычание (0.7), стон (0.7), взвизгивание (0.7) <u>однокоренное слово</u> : вскрик (0.77), выкрик (0.73)
стук	<u>квазисиноним</u> : скрип (0.76), скрип (0.69), дребезжание (0.7), звяканье (0.67), звон (0.67), позвякивание (0.66), звяк (0.66), скрипение (0.66), цоканье (0.65) <u>однокоренное слово</u> : постукивание (0.74), перестук (0.73), постукиваньё (0.69)
взлет	<u>квазисиноним</u> : подъем (0.64), падение (0.59), приземление (0.58), парение (0.56), паденье (0.55), вираж (0.55), посадка (0.52), прыжок (0.52), разворот (0.52), старт (0.5) <u>однокоренное слово</u> : полет (0.63), перелет (0.51)
чтение	<u>квазисиноним</u> : переписывание (0.6), просмотр (0.59), написание (0.59), редактирование (0.58), слушание (0.57), декламация (0.57), печатание (0.56) <u>однокоренное слово</u> : чтение (0.69), перечитывание (0.65), читка (0.64), прочтение (0.61), чтение (0.55)

Ассоциативные группы отглагольных существительных со значением звука *крик* и *стук* представляют собой два стройных ряда, из которых видно, что их векторные репрезентации включают в себя кроме общего значения собственно звука устойчиво сохраняющиеся семантические характеристики. Так для слова *крик* такими характеристиками являются, во-первых, резкость (или внезапность, как

в слове *возглас*) звука, а во-вторых, то, что звук производится живым существом (хотя и не обязательно человеком — ср. *вой, рычание*). Параметрическое противопоставление происходит здесь по качеству звучания звука. Для слова *стук* такой устойчивой характеристикой является то, что звук возникает благодаря взаимодействию поверхностей двух материалов. В данном случае, спецификация материалов определяет выбор лексемы.

Квазисинонимы ассоциативных групп слов *взлет* и *чтение* устроены сложнее. Слово *взлет* соотносится со своими ассоциатами через общее значение «движение в пространстве». В рамках этого значения ассоциаты делятся на подгруппы по общим семантическим параметрам. Часть слов может быть описана как движение между твердой поверхностью и воздухом (*падение, посадка, приземление, прыжок*), часть слов объединяет значение движения именно в воздухе (*парение, вираж*), для части слов это отношение остается неопределенным (*подъем, разворот, старт*). При этом противопоставляющей параметрической характеристикой является, во-первых, направление движения:

подъем — вверх, *падение, приземление, посадка* — вниз, *прыжок* — в любую сторону,
старт — вперед, *разворот* — вперед и назад, *вираж* — в любую сторону,
парение — неопределенное направление,

а во-вторых, осуществляется ли движение на транспортном средстве (*приземление, посадка, вираж*) или самостоятельно живым существом (*прыжок*) или этот параметр не определен (*подъем, падение, разворот, старт, парение*).

Рассматривая ассоциативную группу слова *чтение*, отметим, во-первых, очень большое количество однокоренных слов и вариантов, выражающих разные «логические ударения» действия глагола *читать* [Апресян 1995: 245]. Сам ряд квазисинонимов легче всего описать не через семантическое ядро, а через фрейм: группа ассоциатов связана со сложной когнитивной деятельностью, в которой человек фокусируется на взаимодействии с текстом. Внутри этого фрейма мы видим две группы слов — одна из них касается взаимодействия с письменным текстом — *переписывание, написание, редактирование, чтение, печатание*, другая связана с устным воспроизведением текста — *декламация, слушание*. Наконец, слово *просмотр* заменяет семантический компонент текста в этом фрейме на объект синтетического искусства (фильм, театральное представление, и т. д.).

Следует отметить, что такая дифференцированная параметризация неслучайно наблюдается именно у отглагольных существительных, это след их «происхождения» — глаголов, которые с помощью зависящих от них актантов и сирконстантов задают и специфицируют характеристики описываемой ситуации. Эти характеристики сохраняются на уровне векторной контекстной репрезентации и воспроизводятся в ассоциативных группах существительных-derivатов.

В завершение рассмотрим еще два примера ассоциативных групп, состоящих из квазисинонимов слов с диминутивным суффиком. Интересной особенностью, важной для понимания системности векторного семантического пространства,

является то, что диминутивный суффикс основных слов (*коробочка*, *вдовушка*) сохраняется и у ассоциатов. Состав ассоциативных групп этих слов представлен в Таблице 15.

Таблица 15. Ассоциативные группы слов с диминутивными суффиксами

Основное слово	Ассоциаты
коробочка	<u>квазисиноним</u> : ящичек (0.84), футлярчик (0.79), баночка (0.79), шкатулочка (0.79), шкатулка (0.78), пакетик (0.78), флакончик (0.75), мешочек (0.74), вазочка (0.74) <u>однокоренное слово</u> : коробка (0.82)
вдовушка	<u>квазисиноним</u> : немочка (0.74), бабенка (0.72), дамочка (0.71), мешаночка (0.69), барынька (0.69), девушка (0.69), барышня (0.69), блондиночка (0.68), девица (0.67), толстушка (0.66)

Группа ассоциатов слова *коробочка* содержит в себе, во-первых, однокоренное слово *коробка*, от которого, собственно, и образованно уменьшительное *коробочка*. Это слово имеет очень близкий семантический вектор к основному слову (0.82), однако вектор квазисинонима *ящичек* оказывается еще ближе к основному слову (0.84). Весь ряд квазисинонимов объединяет общее значение «небольшой контейнер», и только одно слово *шкатулка* не имеет дополнительной диминутивной семантики, выражаемой суффиксально. Иначе говоря, словообразовательная семантика сохраняется в векторном представлении: контексты, при которых выбираются именно уменьшительные формы названий объектов-контейнеров, оказываются чрезвычайно близки. Сам набор ассоциатов (*шкатулка*, *флакончик*, *вазочка*) позволяет предположить, что скорей всего описываются вещи, принадлежащие женщинам. Таким образом, векторная модель основного корпуса воспроизводит не только лексические отношения, но и комплекс социокультурных, стилистических и дискурсивных факторов, определяющих в том числе и выбор словообразовательной структуры существительных. Отметим также, что современная дистрибутивная семантика позволяет вычислить вектор не только слова, но и вектор семантически нагруженного аффикса, см. например, определение семантического вектора английского суффикса *-er* со значением деятеля в [Boleda 2020]. Семантическую емкость диминутивных суффиксов демонстрирует группа ассоциатов существительного *вдовушка*. То, что *вдовушка* и *вдова* очень различны по своей сочетаемости, хорошо видно из скетчей портрета слова. В списке атрибутивных скетчей слова *вдова* пять самых устойчиво встречающихся с ним определений составляют прилагательные *безутешный*, *неутешный*, *бездетный*, *унтер-офицерский* (как героиня Гоголя), *соломенный* (женщина, живущая при отсутствующем муже), пять главных сказуемых включают такие глаголы, как *пожертвовать*, *проживать*, *жить*, *скончаться*, *рыдать*. Список скетчей слова *вдовушка* представляет практически противоположную картину: в пять самых устойчиво связанных определений входят такие слова, как *молоденький*, *соломенный*, *хорошенький*, *преlestный*, *богатый*, а типичное сказуемое представлено

только одним глаголом *жить*. Безусловно, такие различия в сочетаемости отражают разницу в социокультурных коннотациях, а не собственно разницу лексических значений. Группа ассоциатов слова *вдовушка* выводит этот комплекс коннотаций на поверхность. График диахронической дистрибуции этого слова показывает, что особенно часто оно упоминалось в литературе середины XIX-го — начала XX-го века (Рисунок 1).

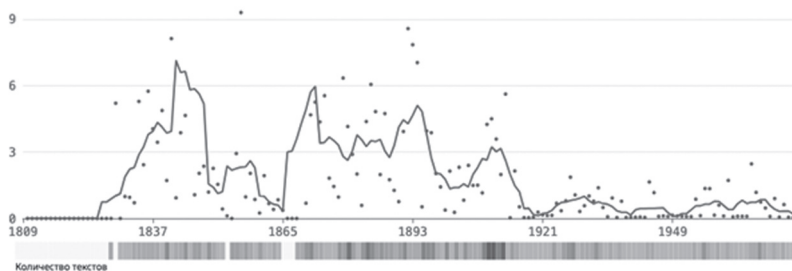


Рис. 1. График распределения по годам слова *вдовушка*

Квазисинонимы слова *вдовушка* как будто складываются в определенный женский образ, который воспроизводился в русской литературе на протяжении XIX-го века вплоть до революции. У этого образа есть несколько характеристик, которые, в отличие, например, от квазисинонимов отглагольных существительных, не противопоставлены параметрически, но выделяют его определенные характеристики: социальный статус — *мещаночка*, *барынька*, *барышня*, *бабенка*, *дамочка*, сюда же можно отнести и национальность — *немочка*, внешность — *блондиночка*, *толстушка*, возраст — *девушка*, *девица*. Следует подчеркнуть, что в отличие от приведенных выше коллокаций, ассоциаты совершенно не обязательно входят в устойчивые сочетания с основным словом, *вдовушка* едва ли может описываться как *девица*, но при этом вся группа квазисинонимов *вдовушки* определяет повторяющиеся черты женского образа, отраженного в текстах XIX-го и начала XX-го вв. коллекции Основного корпуса.

4. Выводы

Выше был представлен обзор того, какими лексическими отношениями могут быть связаны ассоциаты существительных, полученные в результате сравнения расстояния между векторами в дистрибутивной семантической модели Основного корпуса НКРЯ. Цель статьи состояла в том, чтобы очертить возможности лексикографических исследований на материале корпусных данных, представленных в виде контекстных векторных репрезентаций. За пределами описания остались множество наблюдений над группами ассоциатов существительных разных лексических классов и семантических полей, отдельного изучения требуют лексические отношения ассоциатов глаголов, прилагательных и наречий. При этом ключевым вопросом для дальнейшего развития исследований в области векторной

лексикографии является то, до какой степени типы лексических отношений ассоциатов и разнообразие их представленности зависит от корпусных данных, а в какой степени может быть предсказано из семантической структуры основной лексемы. Представляется, что несмотря на то, что конкретное заполнение ассоциативных групп обусловлено спецификой коллекции корпуса, лежащей в основе построенной векторной модели (в нашем случае Основного корпуса НКРЯ), общие наблюдения, высказанные выше в статье, могут тем не менее оставаться валидными. Резюмируем их в виде нескольких ключевых выводов:

1. Слова имеют разный «ассоциативный потенциал», он определяется уровнем лексической разработанности семантического поля. Лексемы из лексически «богатых» полей собирают гомогенные ассоциативные группы, связанные единым лексическим отношением когипонимии, синонимии или квазисинонимии. К таким полям относятся экстралингвистические классификации, части тела, зоны экспрессивной и эмоциональной лексики и др. В будущем было бы интересно описать список таких полей, опираясь именно на данные векторной семантической модели.
2. Системность лексики в языке проявляется на уровне дистрибуции контекстов, и это отражается в системности лексических отношений основного слова и его ассоциатов. Под системностью имеется в виду возможность выделить общее семантическое ядро, связывающее ассоциаты и основное слово, и параметрические характеристики общего значения, с помощью которых ассоциаты противопоставляются друг другу внутри группы. Именно поэтому инструмент лексикографического описания, разработанный в рамках теории «Модель Смысл — Текст», оказывается чрезвычайно мощным средством описания внутренней структуры лексических отношений. Возможно, корпусные данные векторной модели могли бы, во-первых, расширить список лексических функций, а во-вторых, оценить их распространенность, частотность реализаций и связанность с определенными семантическими полями.
3. Еще один признак системности — это глубинные родо-видовые отношения, которые могут быть устроены более сложно, чем базовые отношения «класс — представитель класса». Понятие «класса» далеко не всегда имеет лексическое выражение, более того, базовые семантические категории, такие как «материал», «контейнер», «звук», «движение» усложняются дополнительными семантическими характеристиками, задающими лексические функции и параметрически организующими лексические группы ассоциатов. Рассмотренное выше отношение квазисинонимии по сути является родо-видовым отношением, где родом является абстрактный семантический комплекс, а видами его конкретные лексические реализации.
4. Лексические отношения групп ассоциатов и основного слова, а также ранжирование ассоциатов по близости векторов, определяются не только сходством или различием семантической структуры, но и комплексом

социокультурных коннотаций, воспроизводящихся в текстах корпуса. Таким образом, исследование лексических отношений в векторных моделях открывает доступ к парадигматическим лексическим отношениям внутри определенного жанра, нарратива, тематики или стилистики текста. При этом коннотативную нагрузку могут получать не только сами лексические единицы, но и их структурные элементы, например, суффиксы-диминытивы. Выявление таких лексических парадигм, зависимых от типа текста, помогает не только лучше понять природу лексических отношений, но и выявить мотивы, стереотипы, понятийные установки, воспроизводимые в разных социокультурных дискурсах.

Литература

Апресян Ю. Д. Избранные труды, том I. Лексическая семантика. М.: Школа «Языки русской культуры», 1995. 481 с.

Бабенко Л. Г. Антонимия в контексте синонимии: принципы лексикографического описания в новом идеографическом словаре // Филологический класс. 2020. № 25 (1). С. 7–17.

БТС — Большой толковый словарь русского языка / гл. ред. С. А. Кузнецов. Санкт-Петербург: Норинт, 2008. 1534 с.

Литвинова Т. А., Котлярова Е. С., Заварзина В. А. Фактор гендера в ассоциативных связях слов: данные словаря и дистрибутивно-семантической модели. // Научный диалог. 2022. № 11 (5). С. 136–156.

Мельчук И. А. Опыт теории лингвистических моделей СМЫСЛ⇔ТЕКСТ. М.: Школа «Языки русской культуры», 1999. 345 с.

Мельчук И. А., Жолковский А. К. Толково-комбинаторный словарь русского языка: Опыт семантико-синтаксического описания русской лексики. Вена. Wiener Slavistische Almanach, 1984. 992 с.

НОСС — Новый объяснительный словарь синонимов русского языка. Второе издание, исправленное и дополненное / Авторы словарных статей: В. Ю. Апресян, Ю. Д. Апресян, Е. Э. Бабаева, О. Ю. Богуславская, И. В. Галактионова, М. Я. Гловинская, С. А. Григорьева, Б. Л. Иомдин, Т. В. Крылова, И. Б. Левонтина, А. В. Птенцова, А. В. Санников, Е. В. Урысон. Под общим руководством акад. Ю. Д. Апресяна. М.; Вена: Языки славянской культуры: Венский славистический альманах, 2004. 1488 с. — (Studia philologica).

Панченко А., Лукашевич Н. В., Усталов Д., Паперно Д., Мейер К. М., Константинова Н. RUSSE: семинар по оценке семантической близости для русского языка. Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27–30 мая 2015). Вып. 14. Т. 2. М.: Изд-во РГГУ, 2015. С. 89–105.

Рахилина Е. В. Когнитивный анализ предметных имен: семантика и сочетаемость. М.: Русские словари, 2000. 415 с.

Рыжова Д. А. Типология лексики. Компьютерные методы и инструменты. СПб. Алетейя, 2020. 202 с.

Савчук С. О., Архангельский Т. А., Бонч-Осмоловская А. А., Донина О. В., Кузнецова Ю. Н., Ляшевская О. Н., Орехов Б. В., Подрядчикова М. В. Национальный корпус русского языка 2.0: новые возможности и перспективы развития // Вопросы языкознания. 2024. № 2. С. 7–34.

Тимофеева М. К. Возможности использования сервиса RusVectōrēs для выявления семантических ассоциатов глаголов русского языка // Научный диалог. 2018a. № 9. С. 117–131.

Тимофеева М. К. Типология семантических отношений, выявляемых посредством инструмента RusVectōrēs. Научный диалог. 2018b. № 8. С. 74–87.

Bakarov A. A survey of word embeddings evaluation methods // arXiv preprint arXiv:1801.09536. 2018. 26 p.

Baroni M., Bernardi R., Zamparelli R. Frege in Space: A Program for Compositional Distributional Semantics // Linguistic Issues in Language Technologies. 2014. № 9. P. 241–346.

Bhatia S. The semantic representation of prejudice and stereotypes // Cognition. 2017. № 164. P. 46–60.

Boleda G. Distributional semantics and linguistic theory // Annual Review of Linguistics. № 2020. № 6. С. 213–234.

Caliskan A., Bryson J. J., Narayanan A. Semantics derived automatically from language corpora contain human-like biases // Science. 2017. № 356 (6334). С. 183–186.

Faruqui M., Tsvetkov Y., Rastogi P., Dyer C. Problems with evaluation of word embeddings using word similarity tasks // arXiv preprint arXiv:1605.02276. 2016. 6 c.

Günther F., Rinaldi L., Marelli M. Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions // Perspectives on Psychological Science. 2019. № 14 (6). P. 1006–1033.

Kozłowski A. C., Taddy M., Evans J. A. The geometry of culture: Analyzing the meanings of class through word embeddings // American Sociological Review. 2019. № 84 (5). P. 905–949.

Mikolov T., Chen K., Corrado G., Dean J. Efficient estimation of word representations in vector space // arXiv preprint arXiv:1301.3781. 2013. 12 p.

Nematzadeh A., Meylan S. C., Griffiths T. L. Evaluating vector-space models of word representation, or, the unreasonable effectiveness of counting words near other words // Proceedings of the 39th Annual Meeting of the Cognitive Science Society. 2017. P. 859–864.

Rodriguez M. A., Merlo P. Word associations and the distance properties of context-aware word embeddings // Proceedings of the 24th Conference on Computational Natural Language Learning (2020, November). P. 376–385.

Santus E., Yung F., Lenci A., Huang C. R. Evaluation 1.0: an evolving semantic dataset for training and evaluation of distributional semantic models // Proceedings of the 4th Workshop on Linked Data in Linguistics: Resources and Applications (2015, July). P. 64–69.

WordNet: An Electronic Lexical Database / ed. by Christian Fellbaum. Cambridge, MA: MIT Press. 1998. 422 p.

A. A. Bonch-Osmolovskaya
*V. V. Vinogradov Russian Language Institute,
Russian Academy of Sciences
(Russia, Moscow)
abonch@gmail.com*

DISTRIBUTIVE MODELS IN THE PROJECTION OF THEORETICAL APPROACHES OF LEXICAL SEMANTICS (ACCORDING TO THE RNC)

The article analyses the data of the semantic vector model of the main corpus of RNC. Groups of noun associates whose contextual proximity exceeds the threshold of 0.5 are considered. Theoretical methods of analysing lexical-semantic relations are applied to the corpus data of the associate groups. More abstractly, the main question of the article can be formulated as follows: to what extent do existing theoretical ideas about lexical-semantic relations between lexemes in a language, such as hyperonymy (generic-species relations), meronymy (part-to-part relations), synonymy, correlate with corpus data on the proximity of the contexts of use of these lexemes? The article shows that the attraction of data on the proximity of semantic vectors makes it possible to see a more complexly organised system of lexical-semantic relations. This system allows us, on the one hand, to compare pairs of lexemes within a type of relation on the closer-farther scale and thus to identify the analogue of semantic prototypes on the basis of corpus data. On the other hand, we show the dependence of lexical-semantic relations on extra-linguistic parameters, such as their distribution in the sphere of textual use, the association of lexemes' use with certain textual genres, as well as socio-cultural factors influencing the dominance of certain contexts.

Keywords: distributive semantics, vector models, the Russian National Corpus, lexical relations, hyperonymy, synonymy

References

- Apresyan Yu. D. *Izbrannye trudy, tom I. Leksicheskaya semantika* [Selected Works, V. I. Lexical semantics]. Moscow, Shkola "Yazyki russkoi kul'tury" Publ., 1995. 481 p. (In Russ.)
- Babenko L. G. [Antonymy in the context of synonymy: principles of lexicographic description in a new ideographic dictionary]. *Filologicheskii klass*, 2020, no. 25 (1), pp. 7–17. (In Russ.)
- Bakarov A. A survey of word embeddings evaluation methods. *arXiv preprint arXiv: 1801.09536*. 2018. 26 p.
- Bhatia S. The semantic representation of prejudice and stereotypes. *Cognition*, 2017, no. 164, pp. 46–60.
- Boleda G. Distributional semantics and linguistic theory. *Annual Review of Linguistics*, 2020, no. 6, pp. 213–234.

Bol'shoi tolkovyi slovar' russkogo yazyka [Big explanatory dictionary of the Russian language] / ed. by S. A. Kuznetsov. St. Petersburg, Norint Publ., 2008. 1534 p. (In Russ.)

Caliskan A., Bryson J. J., Narayanan A. Semantics derived automatically from language corpora contain human-like biases. *Science*, 2017, no. 356 (6334), pp. 183–186.

Faruqui M., Tsvetkov Y., Rastogi P., Dyer C. Problems with evaluation of word embeddings using word similarity tasks. *arXiv preprint arXiv:1605.02276*. 2016, 6 p.

WordNet: An Electronic Lexical Database / Fellbaum Christian (ed.). Cambridge, MA: MIT Press, 1998, 422 p.

Günther F., Rinaldi L., Marelli M. Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, 2019, no. 14 (6), pp. 1006–1033.

Kozlowski A. C., Taddy M., Evans J. A. The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, 2019, no. 84 (5), pp. 905–949.

Litvinova T. A., Kotlyarova E. S., Zavarzina V. A. [The gender factor in word association: evidence from a dictionary and distributional semantic model]. *Nauchnyi dialog*, 2022, no. 11 (5), pp. 136–156. (In Russ.)

Mel'chuk I. A. *Opyt teorii lingvisticheskikh modelei SMYSL ⇔ TEKST* [The Meaning ⇔ Text Theory of linguistic models]. Moscow, Shkola "Yazyki russkoi kul'tury" Publ., 1999, 345 p. (In Russ.)

Mel'chuk I. A., Zholkovskii A. K. *Tolkovo-kombinatornyi slovar' russkogo yazyka: Opyty semantiko-sintaksicheskogo opisaniya russkoi leksiki* [Explanatory Combinatory dictionary of the Russian language: Experience on semantic-syntactic description of the Russian lexicon]. Vienna. Wiener Slavistische Almanach, 1984, 992 p. (In Russ.)

Mikolov T., Chen K., Corrado G., Dean J. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. 2013. 12 p.

Nematzadeh A., Meylan S. C., Griffiths T. L. Evaluating vector-space models of word representation, or, the unreasonable effectiveness of counting words near other words. *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, 2017, pp. 859–864.

Novyi ob'yasnitel'nyi slovar' sinonimov russkogo yazyka [A new explanatory dictionary of synonyms of the Russian language]. 2nd ed., revised and enlarged / Under the general supervision of akad. Yu. D. Apresyan. Moscow, Vienna: Yazyki slavyanskoi kul'tury Publ., Venskii slavisticheskii al'manakh, 2004. 1488 p. (Studia philologica). (In Russ.)

Panchenko A., Lukashevich N. V., Ustalov D., Paperno D., Meier K. M., Konstantinova N. RUSSE: The First Workshop on Russian Semantic Similarity. *Komp'yuternaya lingvistika i intellektual'nye tekhnologii: Po materialam ezhegodnoi Mezhdunarodnoi konferentsii "Dialog"* (Moscow, May 27–30, 2015) [Computational Linguistics and Intellectual Technologies: Papers from the Annual Conference "Dialogue"]. Vol. 14 (2). Moscow, RSUH Publ., 2015, pp. 89–105.

Rakhilina E. V. *Kognitivnyi analiz predmetnykh imen: semantika i sochetaemost'* [Cognitive analysis of subject names: semantics and compatibility]. Moscow, Russkie slovari Publ., 2000, 415 p. (In Russ.)

Rodriguez, M. A., & Merlo, P. Word associations and the distance properties of context-aware word embeddings. *Proceedings of the 24th Conference on Computational Natural Language Learning*. 2020, November, p. 376–385.

Ryzhova D. A. *Tipologiya leksiki. Komp'yuternye metody i instrumenty* [The typology of vocabulary. Computer methods and tools]. St. Petersburg, Aleteiya Publ., 2020, 202 p. (In Russ.)

Santus, E., Yung, F., Lenci, A., & Huang, C. R. Evaluation 1.0: an evolving semantic dataset for training and evaluation of distributional semantic models. *Proceedings of the 4th Workshop on Linked Data in Linguistics: Resources and Applications*. Beijing, Association for Computational Linguistics, 2015, pp. 64–69.

Savchuk S. O., Arkhangel'skii T. A., Bonch-Osmolovskaya A. A., Donina O. V., Kuznetsova Yu. N., Lyashevskaya O. N., Orekhov B. V., Podryadchikova M. V. Russian National Corpus 2.0: New opportunities and development prospects. *Voprosy Jazykoznanija*, 2024, no. 2, pp. 7–34. (In Russ)

Timofeeva M. K. [Possibilities of using RusVectōrēs service to identify semantic associates of Russian verbs]. *Nauchnyi dialog*, 2018a, no. 9, pp. 117–131. (In Russ.)

Timofeeva M. K. [Typology of semantic relations revealed through the RusVectōrēs tool]. *Nauchnyi dialog*, 2018b, no. 8, pp. 74–87. (In Russ.)

II. СЛОВО И ЖЕСТ: ВОПРОСЫ МУЛЬТИМОДАЛЬНОЙ КОММУНИКАЦИИ

Т. Е. Янко

Институт языкознания РАН

(Moscow, Russia)

tanya_yanko@list.ru

ПРОСОДИЯ ВОПРОСА ПРИ ОПУЩЕНИИ ПОКАЗАТЕЛЯ ИЛЛОКУТИВНОЙ СИЛЫ¹

Анализируется проблема расчленения речевого акта на два иллокутивных компонента: собственно иллокутивный (например, рему сообщения и неизвестное вопроса) и несобственно иллокутивный (например, тему и известное вопроса). Показано, что речевые акты подразделяются на имеющие четкую границу между иллокутивными компонентами и не имеющие очевидной границы. Вводится понятие иллокутивного шва, который может разделять компоненты речевого акта и который служит результатом специальных линейно-просодических преобразований, примененных к условно исходным предложениям. Понятие иллокутивного шва демонстрируется на примере русского вопросительного предложения. Рассмотрены механизмы опущения собственно иллокутивного компонента по иллокутивному шву, если лексико-синтаксическая структура собственно иллокутивного компонента может быть реконструирована из контекста.

Ключевые слова: просодия, речевой акт, собственно иллокутивный компонент, несобственно иллокутивный компонент, иллокутивный шов, иллокутивное опущение, просодическое кодирование

1. Введение

Речевой акт в теориях актуального членения предложения и коммуникативной (информационной) структуры предложения понимается как расчлененный на два иллокутивных компонента: тему и рему, известное и неизвестное вопроса, см., например, обзорную работу по коммуникативной структуре русского предложения, включая вопрос [Падучева 2015]. Между тем четко определенная граница между иллокутивными компонентами может отсутствовать. Это не значит, что этой границы нет, это значит, что она может не иметь однозначных средств выражения.

¹ Эта работа представляет собой расширенную и дополненную версию тезисов [Янко 2023].

Последнее обстоятельство породило проблему установления границы между темой и ремой. Этот вопрос иногда решается за счет обращения к понятиям данного и нового. Они реконструируются через обращение к контексту, но, поскольку оппозиция данное vs. новое не служит прямым коррелятом оппозиции тема vs. рема, контекст также не решает вопроса.

В качестве источника материала для анализа использован Мультимедийный подкорпус Национального корпуса русского языка [МУРКО] и малый рабочий корпус автора [МРК], собранный из открытых источников. Для верификации слуховых ощущений использована система анализа звучащей речи Praat [Boersma, Weenink 2023].

В разделе 2 ниже вводится понятие иллокутивного шва, рассматриваются средства его линейно-просодического кодирования. Понятие иллокутивного шва иллюстрируется примерами предложений различных типов. В разделе 3 предложена объяснительно-описательная модель иллокутивных опущений в русских вопросительных предложениях с различной синтаксической структурой.

2. Иллокутивный шов

Обратимся к предложениям с однозначным противопоставлением иллокутивных компонентов. Как правило такие предложения представляют собой результат линейно-акцентных преобразований, примененных к исходным предложениям, где граница между иллокутивными компонентами неочевидна. Так, в примере (1) представлено подчеркнуто расчлененное произнесение двух сочиненных предложений: *Мне — двадцать один и я — Таня*. Темы *мне* и *я* в обоих предложениях несут восходящий тон типа ИК-3, по Е. А. Брызгуновой [Брызгунова 1982а: 97–122]. Акцентоносители рем словотформы *один* и *Таня* несут нисходящий акцент ремы ИК-1. Между темами и ремами сочиненных предложений в (1) имеются паузы, см. Рис. 1.

(1) *Мне — двадцать один, и я — Таня* [МУРКО].

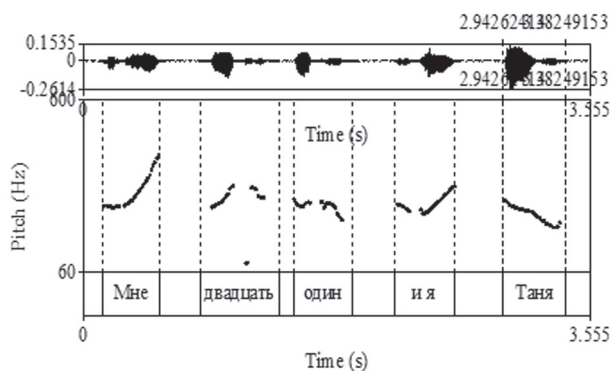


Рис. 1. Тонаграмма примера (1)

В примере же (2), которое представляет собой лабораторный аналог примера (2), произнесенный менее театральным — более обыденным — способом, пауз между темами и ремами нет, второе из сочиненных предложения имеет типичную просодию «шляпы» с подъемом в небольшом диапазоне частоты и с падением в исходе, первое предложение дополнительно несет преобразующий «шляпу» справа подъем как маркер дискурсивной незавершенности.

(2) *Мне двадцать один, и я Таня.*

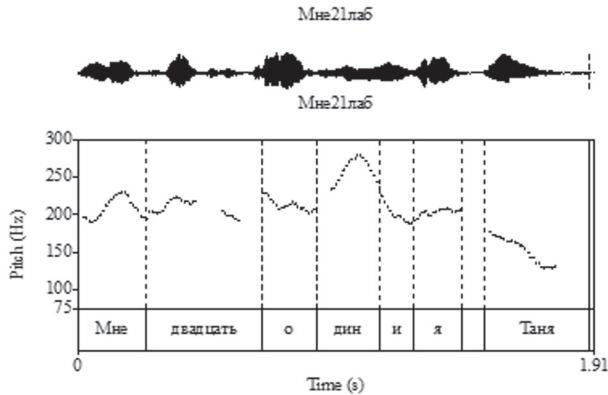


Рис. 2. Тонограмма примера (2)

Заметим, что несмотря на то, что в сложносочиненном предложении между предложениями чаще всего имеется показатель незавершенности, что мы и наблюдаем в лабораторном примере (2), где словоформа *один* несет указывающий на незавершенность восходящий акцент, в примере (1) указания на незавершенность нет. Говорящий — Таня — стоит на сцене, она представляет себя большой аудитории. Таня педальрует расчленение предложений на коммуникативные компоненты и отчленение одного компонента сочиненного предложения от другого. Произнесение (2) звучит почти в два раза короче, чем произнесение того же сегментного материала в примере (1). Пример (1) демонстрирует способ просодического расчленения сообщения на тему и рему.

Аналогично, в вопросе с вопросительным словом (3) мы наблюдаем два релъфных падения типа ИК-2 ([Брызгунова 1982а: 97–122]) на словоформах *фамилия* и *как*. Первое падение служит маркером вынесенного в препозицию известного вопроса *фамилия твоя*, а второе — постпозитивного неизвестного вопроса вопросительного слова *как*.

(3) *А фамилия твоя как?* [МУРКО]

В этом вопросе за счет вынесения вопросительного слова в постпозицию и приобретения им отдельного акцента происходит четкое разделение коммуникативно релевантных компонентов вопроса: известного (несобственно иллокутивного

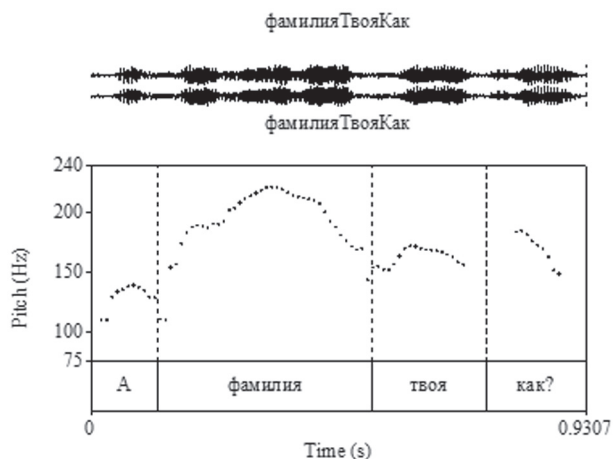


Рис. 3. Тонограмма примера (3)

компонента вопроса) и неизвестного вопроса (собственно иллокутивного компонента). Каждый компонент имеет свой акцент и свою отдельную словоформу-акцентоноситель: это словоформы *фамилия* и *как*. В вопросе (4) с той же лексико-синтаксической, что и в (3), но с иной линейно-акцентной структурой, такого разделения нет. В примере (4) вопросительное слово находится на своем дефолтном месте в начале предложения, оно несет высокий ровный тон начала, который предвещает падение в исходе вопроса на словоформе *фамилия*.

(4) *Как твоя фамилия?* [МУРКО].

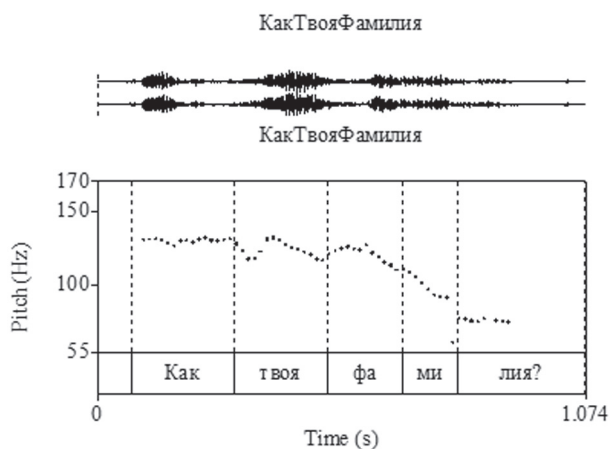


Рис. 4. Тонограмма примера (4)

В примере (4) только один релевантный акцент: это падение на словоформе *фамилия*.

Обратимся к примеру иллокутивного расчленения в *да-нет*-вопросе. Предложение (5) имеет падение на акцентоносителе препозитивного известного вопроса *Ватикана* и подъем, который маркирует иллокутивную силу *да-нет*-вопроса, на акцентоносителе вопроса *подскажете*.

(5) *А телефончик Ватикана не подскажите?* [МУРКО]

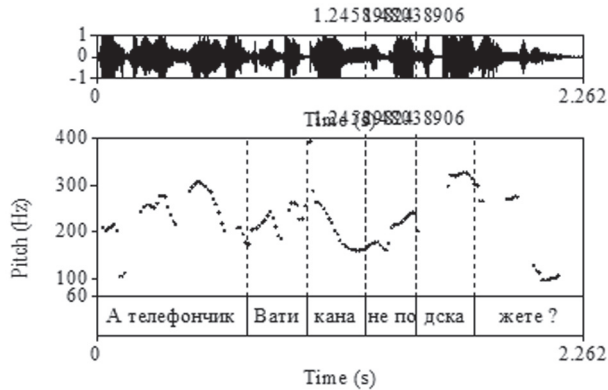


Рис. 5. Тонограмма примера (5)

В вопросе же (6) имеется только маркер неизвестного *да-нет*-вопроса. Это подъем на акцентоносителе вопроса *точно*. Вопрос (6) иллокутивного расчленения не имеет.

(6) *Он точно профессор?* [МУРКО].

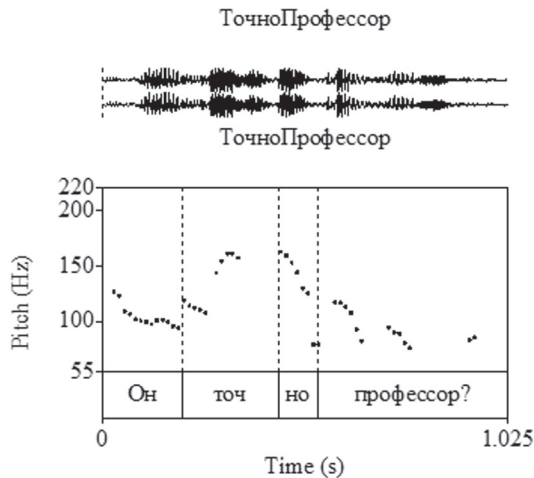


Рис. 6. Тонограмма примера (6)

И наконец пример (7) иллюстрирует расчленение на два иллокутивно релевантных компонента речевого акта императива:

(7) *Счет попросите* [МУРКО].

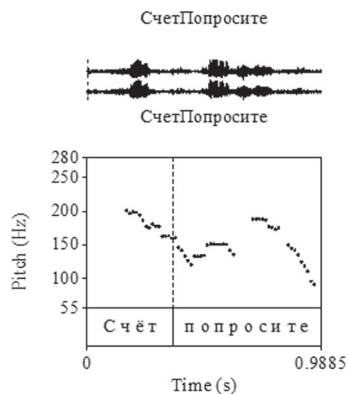


Рис. 7. Тонаграмма примера (7)

Тонаграмма фиксирует два падения типа ИК-2: на препозитивной словоформе *счет* и на императиве *попросите*. Исходный порядок слов повелительного речевого акта с имеющейся лексико-синтаксической структурой: *Попросите счет*. В примере (7) речевой акт расчленен на два компонента: про счет говорится, что его следует попросить. Речевой акт имеет два акцента и два акцентоносителя. Между тем в примере (8) с близкой лексико-синтаксической структурой имеется только один релевантный акцент — падение на словоформе *счета*. В (8) подъем на *неси* имеет чисто фонетическую компенсаторную функцию: он предвещает конечное падение.

(8) *И неси счета* [МУРКО].

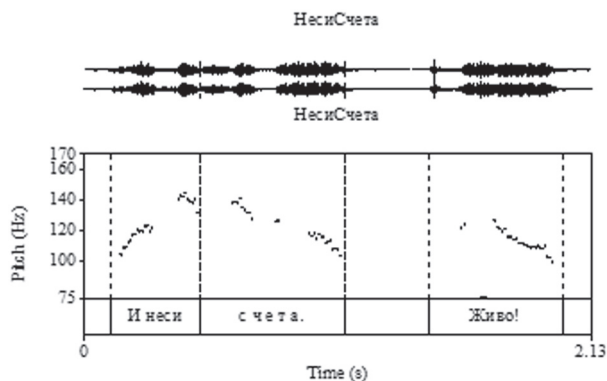


Рис. 8. Тонаграмма примера (8)

Таким образом сообщение (1), вопрос с вопросительным словом (3), *да-нет*-вопрос (5) и императив (7) имеют линейно-просодические средства расчленения речевого акта на коммуникативно релевантные компоненты. В соответствующих им предложениях (2), (4), (6) и (8) таких средств нет.

Идеи о различного рода функциональных барьерах, наблюдаемых в предложениях, высказывались во многих работах, ср. понятие барьера, сдвигающего цепочку клитик вправо в древнерусском предложении [Зализняк 1993: 286], коммуникативного барьера в работе [Циммерлинг 2013: 390 и далее], а также понятие перцептивно значимого просодического шва в работе [Кривнова 2015] и в цитированной там литературе. Вслед за этими понятиями мы предлагаем трактовать такие предложения, как (1), (3), (5) и (7), как имеющие иллокутивный шов между двумя иллокутивно релевантными компонентами: несобственно иллокутивным и собственно иллокутивным. В качестве плана выражения при формировании иллокутивного шва используются линейно-просодические средства: присутствие в предложении двух релевантных акцентов (включая их словоформы-носители) и наличной или потенциальной паузы между этими компонентами.

Вопросы с *не так ли*, *не правда ли*, *так*, *да*, *ведь так* и *а* (далее tag-questions в соответствии с англоязычной традицией) заслуживают отдельного обсуждения. Для выражения иллокутивной силы вопроса в них служит отдельный компонент типа *не правда ли*. Эти вопросы исходно иллокутивно расчлененные. В примере (9) собственно иллокутивный компонент — это слово *так*, несущее подъем типа ИК-3.

(9) Ты по отцовской тропинке решил пойти, так? [МУРКО].

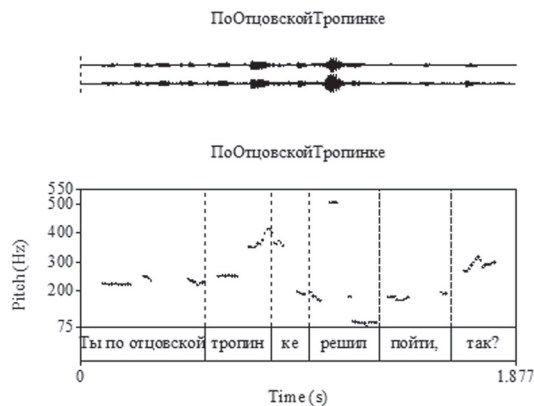


Рис. 9. Тонаграмма примера (9)

В русских tag-questions без частицы *ли* собственно вопросительный компонент несет подъем *да-нет*-вопроса, как в примере (9). Просодия же несобственно вопросительного компонента вариативна: это может быть ИК-3, как в (9), а также ИК-2 и ИК-4. Иллюстрации других возможностей мы здесь опускаем. В русских

tag-questions с частицей *ли* вариативна просодия не только несобственно вопросительного компонента, но также и собственно вопросительного. Вариативность собственно вопросительного компонента объясняется тем, что в случае с *ли* иллокутивная сила выражается сегментно — частицей *ли*. Между тем при сегментном выражении иллокутивных значений сегментные средства мажорируют суперсегментные. Соответственно, просодия тогда вариативна: она только делает иллокутивные компоненты речевого акта отдельными друг от друга, а весь речевой акт — отдельным от других речевых актов в потоке речи, но иллокутивную силу не обозначает, как это происходит, например в русском *да-нет*-вопросе.

Итак, мы показали, что предложения (1), (3), (5), (7) и (9) иллокутивно расчлененные. В дальнейшем мы будем исходить из гипотезы о том, что, если сегментная форма собственно иллокутивного компонента может быть реконструирована из контекста, языкового или неязыкового, собственно иллокутивный компонент может опускаться.

3. Иллокутивное опущение

Ниже рассмотрены основные типы опущения собственно вопросительного компонента в зависимости от его синтаксической структуры.

3.1. Вопросы с *А?*, имеющие синтаксическую форму группы

В [Русская грамматика 1982: 390] вопросы с *А?* (*а дети-то?*; *А пространство?*; *А время?*) квалифицируются как вопросы с именительным падежом. Однако интроспекция и корпусный анализ расширяют круг синтаксических контекстов вопросов с *А?* до синтаксической группы и целого предложения. В этом разделе рассматриваются вопросы с *А?*, имеющие строение синтаксической группы (*а тебя?*; *А поест?*; *А на работу?* [МУРКО]). Следующий подраздел посвящен вопросам с опущением, имеющим пропозициональную форму.

Обратимся к иллокутивному опущению, где опущенный фрагмент восстанавливается непосредственно из предшествующего контекста. В примере (10) представлен фрагмент диалога с двумя вопросами. Первый вопрос *Как тебя зовут?* — это вопрос с вопросительным словом, второй вопрос *А тебя?* — это вопрос с опущенным собственно вопросительным компонентом *как зовут*. Этот компонент известен из предтекста. Второй вопрос в (10) имеет подъем типа ИК-6 [Брызгунова 1982: 97–122] на фонетической группе *а тебя*. В вопросе, где собственно вопросительный компонент либо опущен в силу известности, либо сохранен, но также имеет статус известного, ИК-6 — это дефолтный способ интонирования.

(10) — *Как тебя зовут?* — *Илья*. — *Илья! А тебя?* — *Вова*. — *Вова! А...* [МУРКО].

ИК-6 кодирует не только вопрос с вопросительным словом, но и *да-нет*-вопрос с опущенным или известным собственно вопросительным компонентом: — *А ДК действует?* — *Да, ДК работает...* — *А парк возле него?* [МУРКО] (акцент ИК-6 фиксируется на словоформе *парк*).

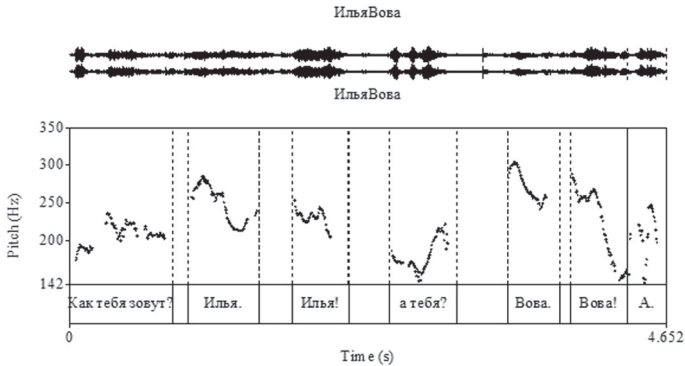


Рис. 10. Тонограмма примера (10)

Между тем пример (11) демонстрирует несколько иной взгляд на множество, относительно членов которого говорящий задает вопросы. Это пример опущения в *да-нет*-вопросе. Второй вопрос говорящего в примере (11) иллюстрирует опущение собственно вопросительного компонента *ваша*?

(11) — *Позвольте поинтересоваться. Эта клеточка ваша? — Клетка моя. — А подушечка?* [МУРКО].

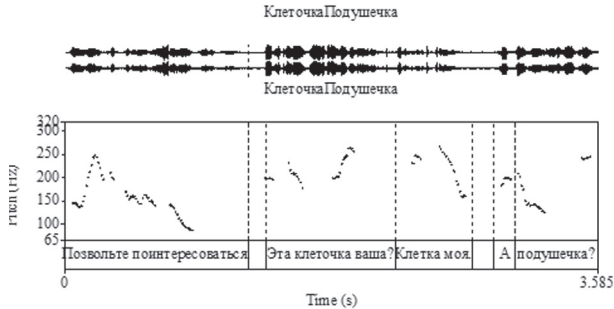


Рис. 11. Тонограмма примера (11)

Просодия второго вопроса в примере (11) — восходяще-нисходящий акцент типа ИК-4 [Брызгунова 1982а: 97–122]. ИК-4 — это хорошо изученный просодический тип вопросов с *А?* [Брызгунова 1982а: 114], [Кобозева 2005]. В [Русская грамматика 1982б: 390] такие вопросы названы напоминаниями, а в работе [Голубева-Монаткина 2004] — сопоставляющими. Возникает вопрос, чем семантика ИК-4 в вопросах с *А?* отличается от семантики ИК-6? ИК-6 — это дефолтный способ интонирования вопроса с *А?*, т. е. такой, при котором просодия вносит минимальный вклад в семантическую структуру предложения. Говорящий спросил, как зовут одного мальчика, а затем, обратив свое внимание на второго, спросил, как зовут и его. Использование же ИК-4 демонстрирует несколько иной взгляд

на вещи. В примере (11) говорящий перечисляет все предметы, которые он видит у путешествующей владелицы собаки. Это клеточка, как говорит он с подчеркнутой эмпатией, подушечка и ошейничек. Использование ИК-4 при подобном перечислении указывает на следование по виртуальному списку предметов, где каждый, не первый, предмет, получает указание на вхождение в список (упорядоченное множество). ИК-4 на словоформе *подушечка* имплицитно указывает на то, что подушечке предшествовало еще что-то. В данном случае — это «клеточка». За «подушечкой» следует «ошейничек», который также содержит отсылку назад в виде ИК-4. Итак, ИК-4 указывает не на простое множество, а на упорядоченный список, где каждый текущий элемент рассматривается с отсылкой на предыдущий в своего рода «переключке».

В [Брызгунова 1980б: 402] также указывается, что вопросы с *А?*, кроме ИК-6 и ИК-4, могут также нести акцент типа ИК-2. Действительно, обратимся к примеру (12):

(12) *А ключ-то?* [МУРКО].

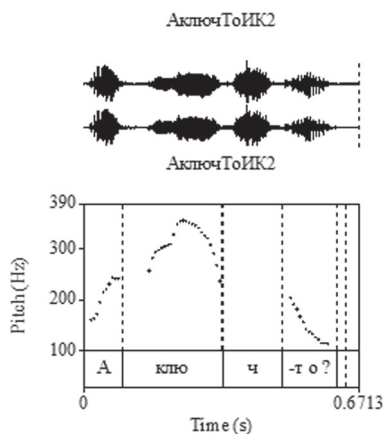


Рис. 12. Тонограмма примера (12)

Акцент ИК-2 здесь указывает на ситуацию контраста между элементами ограниченного и известного говорящему и слушающему множестве. Один из элементов множества обладает некоторым свойством (или участвует в ситуации), когда встает тот же вопрос об оставшихся членах этого множества. Так, в ситуации, разворачивающейся в контексте примера (12), человека, который задолжал своим кредиторам деньги и ключ, после возврата денег спрашивают: *А ключ?*. Акцент ИК-2 здесь отражает то, что в ситуации было только два релевантных объекта — деньги и ключ — и после того, как деньги перешли в руки их владельцев, встал вопрос о возврате и ключа. ИК-2 кодирует ситуацию контраста между элементами конечного и легко обозримого множества, известного заранее говорящему и слушающему.

3.2. Отсутствие иллокутивного показателя в вопросах, имеющих пропозициональную форму

В этом разделе мы обратимся к вопросительным предложениям, имеющим пропозициональную форму вопроса, но не имеющим стандартных средств выражения иллокутивной силы вопроса: ни вопросительного слова, ни вопросительной частицы *ли*, ни показателя вопросительности в виде восходящего акцента ИК-3, кодирующего *да-нет*-вопрос. Между тем предложения, которые анализируются ниже, могут рассматриваться как вопросы, и мы выделим ряд оснований, который позволяют считать их таковыми. Обратимся к примеру (13) из кинофильма «Криминальный талант»².

- (13) <— *На это не клюем. Дешево очень.*> — *А ты дорогая?* <— *Купить хочешь? Денег не хватит*> [МРК].

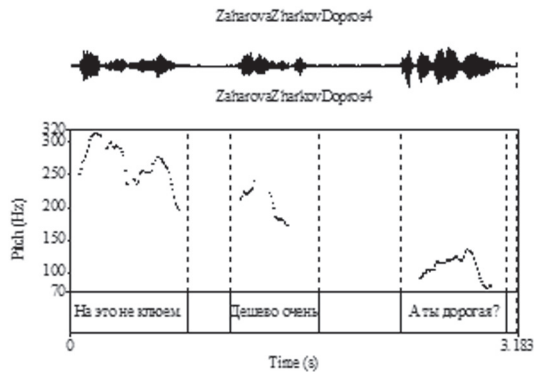


Рис. 13. Тоннограмма примера (13)

В предложении *А ты дорогая?* мы наблюдаем нисходящий акцент ИК-2 на словоформе *дорогая*. Ни вопросительного слова, ни вопросительной частицы в этом предложении нет. Перечислим аргументы за то, чтобы считать предложение *А ты дорогая?* вопросом. Основным аргументом служит то, что в Мультимедийном подкорпусе НКРЯ МУРКО графическая запись примеров с падением ИК-2, аналогичных примеру (13), — их у нас набралось около десятка, ср. например, предложение (14), — содержит вопросительный знак. Графическая запись выполнена экспертами — составителями корпуса — и мы склонны доверять их интуиции, так же, как и собственной интроспекции. Далее. Предложения типа *А ты дорогая?* входят в диалоги, содержащие ответы на гипотетический вопрос. Так, диалог (13) содержит ответ на вопрос о цене, а именно таким вопросом мы и считаем предложение *А ты дорогая?*. И, наконец, считать предложение *А ты дорогая?* не вопросом,

² Этот пример уже был рассмотрен в материалах конференции [Янко 2023]. Материалы послужили предварительной публикацией, основные положения которой мы детализируем в настоящей работе.

а иллюкуцией иной модальности, например сообщением, нам мешает акцент ИК-2, для которого требуется присутствие в сообщении других значений — контраста или указания на причину — а таких значений там нет. В простом сообщении акцентоноситель должен был бы нести акцент ИК-1.

Аналогично примеру (13) устроен и монолог (14) с вопросом *А вы, значит, руководитель?*, который несет акцент ИК-2 на словоформе *руководитель*.

- (14) *<А эт что же у вас тут происходит? Занятия кружка/ да?> А вы, значит, руководитель? <...А на чью же ставку вас взяли? У нас/ по-моему/ по смете денег на технический кружок нет> [МУРКО].*

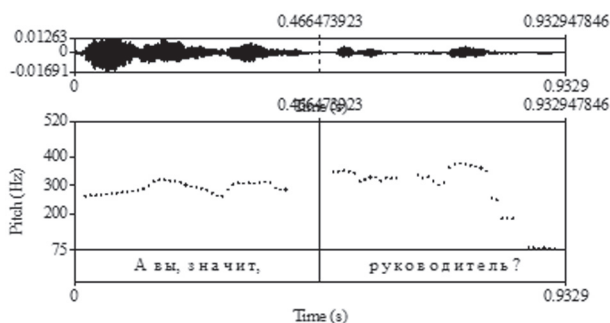


Рис. 14. График изменения частоты и осциллограмма примера (14)

Мы предполагаем, что в вопросах (13)–(14) и подобных подразумевается продолжение со значением ‘не правда ли?’. Иначе говоря, наша гипотеза состоит в том, что вопросы типа (13)–(14) — это результат опущения собственно вопросительного компонента в вопросе, имеющем структуру tag-question.

Другой способ интонирования вопроса с опущенным фрагментом типа ‘не правда ли?’ — это ИК-4. В примере (15) носителем ИК-4 служит словоформа *салуна*:

- (15) *Вы владелец этого салуна? [МУРКО]*

Таким образом, предложения (13)–(15) служат, в соответствии с нашей гипотезой результатом опущения собственно иллюкутивного компонента и наследуют просодию, которую они гипотетически могли бы иметь в соответствующем tag-question. Теоретически вопросы, служащие результатом опущения собственно вопросительного компонента, могли бы нести и акцент ИК-3, который тоже возможен в несобственно иллюкутивном компоненте в tag-question. Однако результат такого опущения был бы неотличим от *да-нет*-вопроса, поэтому от рассмотрения этой возможности приходится отказаться.

Итак, гипотеза об опущении собственно вопросительного компонента позволяет объяснить кодирование вопросов типа *А ты дорогая?* акцентом ИК-2, который в иных условиях маркером вопроса не служит. То же касается и акцента ИК-4.

* * *

Предложена модель анализа вопросительного предложения на основе вводимых понятий иллокутивного шва и иллокутивного опущения. Рассмотрены линейно-просодические трансформации, результатом которых становится эксплицитное расчленение вопроса на два компонента: собственно иллокутивный и несобственно иллокутивный. Дано описание условий опущения собственно иллокутивного компонента вопроса и способов просодического кодирования результатов такого опущения.

Литература

- Брызгунова Е. А.* Интонация // Русская грамматика. Том I. М.: Наука, 1982а. С. 98–122.
- Брызгунова Е. А.* Средства выражения неизвестного в вопросе (взаимодействие лексики, контекста и интонации) // Русская грамматика. Том II. М.: Наука, 1982б. С. 397–402.
- Голубева-Монаткина Н. И.* Вопросы и ответы диалогической речи. Классификационное исследование. М.: УРСС, 2004. 199 с.
- Зализняк А. А.* К изучению языка берестяных грамот // Янин В. Л., Зализняк А. А. Новгородские грамоты на бересте. Из раскопок 1984–1989 гг. М.: Наука, 1993. С. 191–319.
- Кобозева И. М.* Опыт разработки признаковой базы для характеристики лексико-синтаксических, семантических и прагматических свойств вопросительных реплик // Труды международной конференции «Диалог 2005». М.: РГГУ, 2005. С. 238–244.
- Кривнова О. Ф.* Глубина просодических швов в звучащем тексте (экспериментальные данные) // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции Диалог 2015. Вып. 14 (21): В 2 т. Т. 1. М.: Изд-во РГГУ. С. 338–351.
- Падучева Е. В.* Коммуникативная структура предложения (rusgram.ru) // Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru>). На правах рукописи. М., 2015.
- Русская грамматика. Том II. М.: Наука, 1980. 709 с.
- Циммерлинг А. В.* Системы порядка слов славянских языков в типологическом аспекте. М.; Языки славянской культуры. 2013. 544 с.
- Янко Т. Е.* Иллокутивный эллипсис и его просодия // «Слово и жест». Научная конференция, посвященная памяти Е. А. Гришиной («Гришинские чтения»). Москва, 8 февраля 2023 г. Материалы конференции / Отв. ред. С. О. Савчук. М.: Институт русского языка им. В. В. Виноградова РАН, 2023. С. 71–79.
- Boersma P., Weenink D.* Praat: Doing phonetics by computer. Version 6.3.20. Online: Praat: doing Phonetics by Computer (uva.nl), 2023 (accessed date: 31.10.2023).

T. E. Yanko

Institute of Linguistics RAS

(Moscow, Russia)

tanya_yanko@list.ru

PROSODY OF A QUESTION WHEN OMITTING THE ILLOCUTINARY PROPER COMPONENT

The speech act in the theories of the topic-focus structure is understood as divided into two illocutionary components: topic and focus, known and unknown of the question. Meanwhile, there may be no clearly defined boundary between the two illocutionary components. However, sentences with unambiguous opposition of illocutionary components also exist. They are the result of linear-prosodic transformations applied to original sentences where the boundary between the illocutionary components is not obvious. To analyze the mechanisms of boundary establishment, the concept of an illocutionary break between the components of a speech act is introduced. Based on the material of the Russian interrogative sentence, a description is given of the linear-prosodic coding of the illocutionary break and the results of omission of the illocutionary proper component, when it is activated in the context.

Key words: prosody, speech act, illocutionary proper component, illocutionary improper component, illocutionary break, illocutionary omission, prosodic marking

References

- Bryzgunova E. A. [Intonation]. *Russkaya grammatika. Tom I* [Russian grammar. Vol. I.] Moscow, Nauka, (1982a), pp. 98–122.
- Bryzgunova E. A. [Means of expressing the unknown in a question (interaction of vocabulary, context and intonation)]. *Russkaya grammatika. Tom II* [Russian grammar. Vol. II.]. Moscow, Nauka Publ., 1982b, pp. 397–402.
- Golubeva-Monatkina N. I. *Voprosy i otvety dialogicheskoi rechi. Klassifikatsionnoe issledovanie* [Questions and answers of dialogical speech. Classification study]. Moscow, URSS Publ., 2004. 199 p.
- Zaliznyak A. A. [Towards the study of the language of birch bark letters]. Yarin V. L., Zaliznyak A. A. *Novgorodskie gramoty na bereste. Iz raskopok 1984–1989 g.g.* [Novgorod letters on birch bark. From excavations 1984–1989] Moscow, Nauka Publ., 1993, pp. 191–319.
- Kobozeva I. M. [Experience in developing a parameters base for characterizing the lexical-syntactic, semantic and pragmatic properties of interrogative remarks]. *Trudy mezhdunarodnoi konferentsii “Dialog” (2005)*. [Papers from the Annual International Conference “Dialogue” (2005)]. Moscow, RSUH Publ., 2005, pp. 238–244.
- Krivenova O. F. [Depth of prosodic breaks in spoken text (experimental data)]. *Komp'yuternaya lingvistika i intellektual'nyeologii. Po materialam ezhegodnoi Mezhdunarodnoi konferentsii “Dialog” (2015)*. [Computational Linguistics and Intellectual

Technologies. Papers from the Annual International Conference “Dialogue”]. Issue 14 (21). Moscow, RSUH Publ., 2015, pp. 338–351.

Paducheva E. V. *Kommunikativnaya struktura predlozheniya (rusgram.ru). Materialy dl'a proekta korpusnogo opisaniya russkoi grammatiki* (<http://rusgram.ru>). Na pravah rukopisi. [Communicative structure of a sentence (rusgram.ru). Materials for the project of corpus description of Russian grammar (<http://rusgram.ru>). As a manuscript.]. Moscow, 2015.

Russkaya grammatika. Tom II [Russian grammar. Vol. II]. Moscow, Nauka Publ., 1980. 709 p.

Zimmerling A. V. *Sistemy por'adka slov slav'anskih yazykov v tipologicheskom aspekte* [Word order systems of Slavic languages in a typological aspect]. Moscow, Yazyki slav'anskoj kul'tury Publ., 2013. 544 p.

Yanko T. E. [Illocutionary ellipsis and its prosody]. “Slovo i zhest”. *Nauchnaya konferentsiya posv'ashchennaya pam'ati E. A. Grishinoi* (“Grishinskie chteniya”). [“Word and Gesture”. Scientific conference dedicated to the memory of E. A. Grishina (“Grishina’s Readings”). Moscow, February 8, 2023. Conference materials. / ed. by S. O. Savchuk]. Moscow, V. V. Vinogradov Russian Language Institute, 2023, pp. 71–79.

Boersma P., Weenink D. Praat: Doing phonetics by computer. Version 6.3.20. Online: Praat: doing Phonetics by Computer (uva.nl), 2023 (accessed date: 31.10.2023).

О. В. Федорова
 МГУ имени М. В. Ломоносова
 (Россия, Москва)
olga.fedorova@msu.ru

«РЕПОРТАЖИ О ГРУШАХ»: ИТОГИ ПЕРВЫХ ИССЛЕДОВАНИЙ И ДАЛЬНЕЙШИЕ ПЕРСПЕКТИВЫ

В статье приводится обзор исследований последних двух лет, объединенных общей тематикой изучения известного «Фильма о грушах» У. Чейфа в жанре репортажа. В отличие от более традиционного жанра пересказа, данный жанр еще плохо исследован и описан в современной лингвистической литературе. В статье кратко представлены шесть небольших корпусов, собранных автором, в том числе корпус записей, проведенных с использованием метода регистрации движений глаз (айтрекинга). Далее приводятся результаты нескольких конкретных экспериментов: о взаимосвязи между объемом вербальной рабочей памяти испытуемых и их особенностями порождения речи в ходе репортажа; исследование индивидуальных стилей речи комментаторов с привлечением данных регистрации движений глаз испытуемых; изучение известности референта как фактора активации; описание модели интродукции в «Репортажах о грушах»; исследование «Репортажей о грушах» в норме и патологии, в том числе описание языковых особенностей пациентов с заболеваниями шизофренического спектра; сравнение «Репортажей о грушах» с «Рассказами о грушах», записанных в традиционном жанре пересказа, у одних и тех же испытуемых. В качестве испытуемых в экспериментах участвовали студенты МГУ имени М. В. Ломоносова.

Ключевые слова: дискурс; репортаж; порождение речи; регистрация движений глаз; «Фильм о грушах».

Памяти Е. А. Гришиной

Введение. «Фильм о грушах» Уолласа Чейфа

В середине 1970-х годов известный американский лингвист Уоллас Чейф (1927–2019) и его ученики приступили к реализации проекта, который в 1980 году воплотился в коллективной монографии «Рассказы о грушах: Когнитивные, культурные и языковые аспекты порождения повествования» [Chafe (ed.) 1980]. В этом проекте Чейф с коллегами демонстрировали участникам исследования специально

снятый шестиминутный фильм, а затем собирали письменные пересказы этого фильма. Эта процедура варьировалась с испытуемыми разных возрастов, с носителями разных языков и с различными временными интервалами между просмотром фильма и записью пересказа. Хотя с момента выхода книги прошло почти 50 лет, «Фильм о грушах» по-прежнему активно используется в качестве стимульного материала в самых разных научных исследованиях, о чем говорят, в частности, материалы одноименного симпозиума, проведенного в 2015 году в рамках Объединенной Евро-Азиатско-Тихоокеанской конференции по когнитивной науке в Турине [Airenti et al. 2015], а также большой европейский проект на материале 10 языков по созданию телекомментария для незрячих людей [Mazur, Chmiel 2012].

В «Фильме о грушах» участвуют следующие пять персонажей, упорядоченные по важности: (1) главный герой мальчик, который действует на протяжении всего фильма; (2) герой садовник, который участвует во многих ключевых эпизодах фильма; (3) второстепенный персонаж три мальчика, который появляется в нескольких эпизодах; а также (4) и (5) эпизодические персонажи мужчина с козой и девочка, которые принимают участие только в одном эпизоде фильма. Краткий сюжет фильма таков:

- (1) Садовник на дереве собирает груши. Мимо проходит мужчина с козой. Приезжает мальчик на велосипеде, берет одну корзину с грушами и уезжает. По дороге навстречу мальчику едет девочка на велосипеде; у мальчика слетает шляпа; велосипед наезжает на камень; мальчик падает; груши рассыпаются. Три подошедших мальчика помогают собрать груши и отдают потерянную шляпу. Мальчик дает им три груши; они уходят в разные стороны. Садовник на дереве продолжает собирать груши; он обнаруживает пропажу одной корзины с грушами. Мимо проходят три мальчика, жуя подаренные груши. Садовник смотрит им вслед.

«Фильм о грушах» многократно использовался и в исследованиях с русскоязычными испытуемыми, в том числе в мультимодальном проекте (сайт <https://multidiscourse.ru>; [Kibrik, Fedorova 2018]). В 2022 году мы впервые использовали его для записи не традиционных пересказов (будем называть их «Рассказами о грушах»), а репортажей, то есть комментирования фильма непосредственно в процессе просмотра. Идея сбора репортажей возникла в 2013 году в разговоре с Е. А. Гришиной при обсуждении европейского проекта для незрячих людей, упомянутого выше. Настоящая статья посвящена ее памяти. В дальнейшем изложении мы сначала опишем результаты проведенных за последние два года исследований (раздел 1), а затем остановимся на описании текущих проектов (раздел 2).

1. «Репортажи о грушах»: методика и первые исследования

Жанр репортажа пока мало изучен в (психо)лингвистике. Р. Томлин применил похожий метод для изучения фокусного внимания, залога и порядка слов в «Фильме о рыбах» [Tomlin 1995], однако испытуемые порождали не дискурсы, а только отдельные высказывания. В работе [Eaton et al. 1999] при помощи похожего метода

были получены детские пересказы, однако трехминутный видеоролик семь раз останавливался и детей просили рассказать, что произошло в текущем фрагменте.

Нам известна только одна русскоязычная монография [Малышева 2011], посвященная когнитивному описанию спортивного комментирования. Как мы покажем далее, однако, «репортажи о грушах» во многом отличаются от спортивных репортажей.

За два прошедших года мы собрали следующие шесть корпусов записей, которые перечислены ниже в хронологическом порядке, по состоянию на 1 марта 2024 года:

- **Корпус Zoom.** Корпус из 16 репортажей, записанных в начале 2021 года на платформе Zoom со студентами отделения ТиПЛ филфака МГУ имени М. В. Ломоносова. Инструкция при записи этого и всех описанных далее корпусов была одинакова: «Вам надо будет комментировать фильм по ходу развития действия как можно более подробно. Представьте себе, что рядом с вами сидит незрячий человек и вам нужно детально описать ему все, что происходит на экране. Старайтесь описывать не только происходящие события, но и окружающую обстановку». Кроме основного теста каждый испытуемый проходил также тест по определению объема его вербальной рабочей памяти при порождении речи. Результаты описаны ниже в разделе 1.1.
- **Корпус Reportage.** Корпус из 52 репортажей, записанных в 2022–2023 годах в специально оборудованной кабине звукозаписи со студентами разных факультетов МГУ имени М. В. Ломоносова. Помимо аудиозаписи велась регистрация движений глаз при помощи айтрекера Tobii Spectrum с частотой 600 Гц. Первые результаты описаны в разделе 1.2.
- **Корпус Duple Reportage.** Корпус из 12 «двойных» репортажей (записанных в 2023 году там же на том же оборудовании), когда каждый испытуемый — студент МГУ проходил этот тест дважды с интервалом в полгода. Результаты описаны в разделе 1.3.
- **Подкорпус Introduction.** Подкорпус из 25 репортажей корпуса Reportage, записанных весной 2022 года со студентами отделения ТиПЛ филфака МГУ имени М. В. Ломоносова. Результаты описаны в разделе 1.4.
- **Корпус 16–25 male.** Корпус из 90 репортажей, записанных с юношами 16–25 лет. Первые результаты описаны в разделе 2.1.
- **Корпус Reportage+Retelling.** Корпус из 50 записей «Фильма о грушах» (записанных в 2023 году там же на том же оборудовании), когда каждый испытуемый — студент МГУ сначала проходил тест по комментированию фильма, а сразу затем пересказывал этот же фильм. Возможные пути исследований описаны в разделе 2.2.

1.1. Когнитивные особенности репортажа

По интуитивным представлениям, жанр репортажа требует от человека довольно больших когнитивных ресурсов (сравнимых, возможно, с когнитивными затратами синхронных переводчиков), поэтому мы начали исследование этого

жанра с изучения особенностей вербальной рабочей памяти (далее ВРП) испытуемых. Для первого исследования мы выбрали три критерия «успешности» репортажа и предположили, что «успешный» репортаж — это такой репортаж, в котором испытуемый говорит (1) непрерывно, (2) быстро и (3) лексически разнообразно. Все записи были аннотированы в программе Praat, тексты были разбиты на дискурсивные фрагменты (которые не всегда соответствовали «классическим» элементарным дискурсивным единицам (ЭДЕ) по работе [Кибрик, Подлесская (ред.) 2009]); при разбиении на дискурсивные фрагменты аннотаторы прежде всего руководствовались длительностью абсолютных пауз.

Тест по определению ВРП также проводился онлайн на платформе Zoom и был аналогичен тесту из работы [Daneman, Green 1986]. В ходе теста испытуемому предлагалось читать на экране отдельные слова русского языка, сгруппированные в блоки по 2, 3, 4, 5 или 6 слов, всего 100 слов, по 5 блоков каждого уровня. Между блоками слов были вставлены пустые слайды. Испытуемый должен был запоминать появляющиеся на экране слова, а при появлении пустого слайда придумывать с каждым из них по предложению, причем слова должны были стоять точно в той же форме, в которой они появлялись. Кроме того, целевое слово не могло стоять на первом месте придуманного предложения.

Статистический анализ репортажей 16 испытуемых показал, что, действительно, мы наблюдаем положительную корреляцию между объемом ВРП испытуемых и темпом речи и лексическим разнообразием. Однако **корреляция между объемом ВРП и непрерывностью репортажа вопреки ожиданиям оказалась отрицательной**; мы предположили, что навык избегать длинных незаполненных пауз хезитации относится к коммуникативной сфере и поэтому не требует больших затрат когнитивных ресурсов. Подробнее об этом исследовании см. статью [Fedorova 2022].

1.2. Индивидуальные стили комментаторов

Данное исследование дополняет проект, описанный в разделе 1.1, использованием метода регистрации движений глаз (айтрекинга), позволяющим с высокой точностью фиксировать, куда направлен взгляд испытуемого в каждый момент просмотра фильма.

Регистрация движений глаз велась при помощи айтрекера Tobii Spectrum с частотой 600 Гц. В данной работе нас интересовал вопрос, чем в принципе могут различаться комментарии разных испытуемых с точки зрения глазодвигательной активности и как это можно связать с различиями в объеме их ВРП. В настоящем разделе мы сравним два репортажа испытуемых с разным объемом ВРП: у первого испытуемого, назовем его X, объем ВРП был небольшим (56%), а у второго испытуемого Y — большим (79%).

В вышеупомянутой книге Е. Г. Малышевой «Русский спортивный дискурс» (2011) автор выделяет семь коммуникативных типов языковых личностей спортивных комментаторов, фигурально наделяя каждого своими характерными чертами: (1) комментатор-«репортер», например, Виктор Гусев; (2) комментатор-«аналитик»,

Владимир Маслаченко; (3) комментатор-«знаток», Владимир Маслаченко; (4) комментатор-«ироник», Василий Уткин; (5) комментатор-«балагур», Георгий Черданцев; (6) комментатор-«болельщик», Георгий Черданцев; (7) комментатор-«актер», Н. Н. Озеров, Коте Махарадзе. Почти все похожие черты можно найти и у комментаторов «Фильма о грушах», однако мы остановимся на противопоставлении лишь двух из них: комментатора-«репортера» и комментатора-«аналитика».

Для комментатора-«репортера» Х в целом характерна стратегия «что вижу, о том и говорю»: Х очень детально описывает всё, что видит в данный момент на экране, легко переключаясь с одного кадра на другой, иногда даже не договаривая предыдущее высказывание. Важной особенностью данного репортажа, которую можно выявить при помощи айтрекинга, является отсутствие больших задержек между взглядом испытуемого на тот или иной объект и названием этого объекта: в среднем эта задержка составляет 0,5–1 с, максимум 2 с. Отметим также непрерывность данного репортажа, то есть практически отсутствие абсолютных пауз между отдельными высказываниями испытуемого.

Совсем иначе строит свой комментарий испытуемый Y, которого можно отнести к категории комментатора-«аналитика». В его речи заметно больше абсолютных пауз, он дорассказывает эпизоды до конца, поневоле откладывая повествование о последующих событиях; он не только комментирует, но и интерпретирует происходящие события. Иногда подобные задержки составляют около 10 с, например, когда испытуемый Y видит на заднем плане мужчину с козой, он сначала заканчивает рассказ о садовнике: *повязал платок обратно на шею, вздохнул, поднялся, отряхнул колени* и только потом переключается на козу: *на заднем плане видно подходит мужчина с каким-то животным, видимо козленок*. В среднем у испытуемого Y задержка между взглядом и названием составляет 3–5 с. Кроме того, у него значимо чаще встречаются абсолютные паузы хезитации. Подробнее об этой работе см. [Федорова 2023a].

1.3. Известность референта как фактор активации¹

В ходе общения в голове у собеседников актуализировано некоторое количество референтов — людей, объектов и под. Каждый раз, упоминая референт, говорящий осуществляет референциальный выбор — выбирает одно из референциальных выражений: полную ИГ (*садовник*), местоимение (*он*) или нулевую референцию ($\emptyset$). Ключевым является вопрос, на основании каких факторов говорящий осуществляет выбор между полным и редуцированным средством. Согласно многофакторной модели А. А. Кибрика [Kibrik 2011], этот выбор обусловлен степенью активации референта в рабочей памяти говорящего, а процедура подсчета коэффициента активации учитывает факторы активации двух типов — дискурсивный контекст и внутренние свойства референта. В данном исследовании мы предлагаем ввести еще один фактор активации — который, как нам представляется, образует третий

¹ Настоящее пилотное исследование выполнено совместно с Т. Е. Тыщишиной (МГУ имени М. В. Ломоносова).

тип, отличный от двух других — степень известности референта. Мы предполагаем, что известность референта говорящему повышает коэффициент активации.

В отличие от более традиционного жанра «Рассказов о грушах», когда испытуемый порождает дискурс после окончания просмотра фильма, при комментировании он описывает то, что видит в данный момент, не имея возможности оценить значимость происходящего в масштабах всего сюжета. Однако ситуация меняется, когда при повторном комментировании примерно через полгода испытуемый забывает некоторые детали, однако помнит сюжет, так что строит свой репортаж, держа в уме знание этого сюжета. В данном исследовании мы предварительно подтвердили гипотезу о том, что известность сюжета повышает активацию основных персонажей фильма — мальчика и садовника. Подробнее см. статью [Тыщишина, Федорова 2023].²

1.4. Модель интродукции в «Репортажах о грушах»

Интродукцией мы будем называть введение нового референта в долговременную память адресата. В прототипическом случае интродуктивное предложение состоит из трех основных элементов: (1) общей позиции (англ. common ground), (2) бытийного оператора и (3) номинация референта. Проанализировав интродуктивное поведение испытуемых в «Репортажах о грушах» и сравнив их с «Рассказами о грушах», мы показали, что оно заметно отличается от классического спортивного репортажа, в котором основной акцент делается на ключевых событиях спортивного соревнования, определенных заранее (пример из работы [Малышева 2011]):

- (2) Итак ↑ внимание! ↓ / Стрельба Гараничева в эстафетной гонке! ↑ // Гараничев вторым выстрелом мажет ↑ // У словенца много промахов ↓ // Уходит немец ↑ / это Лессер! ↑ // Гараничев один промах! ↑

Типичная усредненная репортажная интродукция выглядит таким образом:

- (3) [садовник] (0.3) (ээ0.2) вижу дерево, к нему приставлена лестница, (мм0.6) на которой стоит мужчина.
[мужчина с козой] (0.4) мы видим человека, (ээ0.3) в футболке и шляпе, (0.4) он ведет козу.
[мальчик] (ээ0.4) на заднем плане едет мальчик на красном велосипеде, (0.5) он тоже в шляпе.
[девочка] (0.3) ему навстречу едет (0.3) девочка на велосипеде, (0.4) с двумя косичками.
[три мальчика] (0.6) мальчик слышит стук, (0.5) поднимает голову, (0.3) и видит трех других мальчиков.

² Еще одно исследование на материале «двойных» репортажей, описанное в статье [Федорова 2023б], посвящено изучению (не)упоминания в репортажах одного своеобразного объекта — игрушки с прикрепленным к теннисной ракетке шариком (англ. paddleball).

Как нам представляется, репортажи о грушах, имея некоторые свои особенности, все же остаются намного больше похожи на классические «Рассказы о грушах», чем на истинные (спортивные) репортажи.

2. «Репортажи о грушах»: текущие исследования

В этом разделе мы опишем два текущих проекта.

2.1. «Репортажи о грушах» в норме и патологии: языковые особенности пациентов с заболеваниями шизофренического спектра³

В проекте, который мы проводим совместно с «Научным центром психического здоровья», мы исследуем языковые особенности юношей возрастной группы от 16 до 25 лет с диагностированными заболеваниями шизофренического спектра. Основная цель проекта состоит в разработке методов ранней диагностики преддебютных стадий шизофрении и прогнозирования дальнейшего течения психического расстройства.

В рамках пилотного исследования были изучены языковые особенности восьми пациентов с заболеваниями шизофренического спектра. Пациенты первой группы из четырех человек с диагнозом по МКБ-10 F21 «Шизотипическое расстройство» имели преддебютные расстройства, пациенты второй группы из четырех человек с диагнозом по МКБ-10 F20 «Шизофрения» имели манифестный психотический приступ. Группа контроля составила шесть добровольцев без психических заболеваний.

Для всех 14 испытуемых мы сосчитали (1) длину каждой ЭДЕ в словах, (2) количество ЭДЕ, (3) длительность пауз; для каждой паузы отмечалось, находится она внутри или между ЭДЕ. Оказалось, что среднее количество слов в ЭДЕ составило в первой группе $5,85 \pm 1,45$, во второй группе $5,42 \pm 2,07$, в группе контроля $7,57 \pm 1,61$, что оказалось достоверно больше, чем в клинических группах ($p < 0,05$). Среднее количество ЭДЕ составило в первой группе $61,5 \pm 11,9$, во второй группе $45,2 \pm 23,4$, в группе контроля $71,2 \pm 14,2$. Выявлена тенденция к различию ($p = 0,07$) между больными с шизотипическим расстройством и здоровыми добровольцами, также важно отметить меньшие значения этого показателя у больных с манифестным приступом шизофрении.

Длительность пауз в секундах составила в первой группе $1,36 \pm 1,93$, во второй группе — $1,87 \pm 3,67$, в группе контроля — $1,03 \pm 1,16$, то есть наибольшая длительность пауз характерна для больных с шизофренией, наименьшая — для здоровых, больные из первой группы занимают промежуточное положение. Изучение распределения пауз внутри и между ЭДЕ показало, что доля пауз между ЭДЕ составляет у группы контроля около 35%, для первой группы этот показатель ниже — около 31%, а для второй группы он составляет около 23%. Таким образом, прослеживается

³ Данное пилотное исследование выполнено совместно с М. О. Стрыгиной и М. А. Омельченко (ФГБНУ «Научный центр психического здоровья»), а также Т. Е. Тыщишиной, М. С. Кравчук и Ф. А. Садковским (МГУ имени М. В. Ломоносова).

тенденция к увеличению количества пауз внутри ЭДЕ, что может считаться показателем сбивчивости речи. Данное исследование будет продолжено на полном корпусе пациентов с заболеваниями шизофренического спектра, который предполагается увеличить до 30 пациентов в каждой из двух клинических групп.

2.2. Сравнение «Репортажей о грушах» с «Рассказами о грушах»

Одно из перспективных, на наш взгляд, направлений исследований «Репортажей о грушах» — это сравнение их с записями «Рассказов о грушах», полученных от тех же испытуемых. В этом разделе мы кратко наметим некоторые пути дальнейшего развития этой проблематики на примере записи одного из испытуемых из группы нормы (полные расшифровки приведены в приложении ниже⁴).

- **Просодические отличия** «Репортажей о грушах» от «Рассказов о грушах». По предварительным наблюдениям, в репортажах намного чаще встречаются высказывания с интонацией незавершенности (так как испытуемый не знает, что будет дальше). Интересно сравнить как общее количество высказываний с интонацией незавершенности и с интонацией завершенности, так и разное просодическое оформление одних и тех же содержательных фрагментов повествования в «Репортажах о грушах» и «Рассказах о грушах».
- **Количество и качество пауз хезитации.** Так как запись «Репортажей о грушах» ведется в реальном времени, это требует от испытуемых более серьезных когнитивных затрат, чем при записи «Рассказов о грушах». Предварительный анализ собранных материалов говорит о том, что «Репортажи о грушах» ведутся испытуемыми в более высоком темпе, с большим количеством пауз хезитации (точнее, с большим количеством абсолютных (незаполненных) пауз хезитации и меньшим количеством заполненных пауз хезитации). Данные результаты нуждаются в дальнейшей проверке.
- **Количество и качество обрывов слов и самоисправлений.** Как показывает предварительный анализ собранного материала, из-за более высоких когнитивных затрат в «Репортажах о грушах» по сравнению с «Рассказами о грушах» в них встречается большее количество разнообразных речевых сбоев, в первую очередь обрывов слов, самоисправлений или отказа от начатого высказывания. Данные результаты также нуждаются в дальнейшей проверке.
- **Лексическое разнообразие**, которое можно определять двумя способами: (1) как отношение количества уникальных полнозначных слов к общему количеству полнозначных слов и (2) просто как количество уникальных полнозначных слов, обычно отдельно для каждой части речи. Самая простая гипотеза состоит в том, что в «Репортажах о грушах» лексическое разнообразие будет меньше, чем в «Рассказах о грушах», так как они более когнитивно затратные. Возможно, однако, у испытуемых с большим

⁴ Необходимо отметить, что предварительные выводы, сделанные ниже на материале всех проанализированных записей, не всегда подтверждаются на материале приводимой расшифровки.

объемом ВРП это различие будет менее ярко выражено, чем у испытуемых с небольшим объемом ВРП.

- **Степень лексической похожести** «Репортажей о грушах» и «Рассказов о грушах». Согласно принципу экономии усилий человеку легче использовать одни и те же ментальные схемы и употреблять одни и те же лексические средства для выражения этих схем. Интересно проследить, насколько «Рассказы о грушах», порождаемые испытуемыми сразу после «Репортажей о грушах», будут повторять их на лексическом уровне.
- **Использование «кинематографического взгляда».** Согласно работе [Tannen 1980: 54], американские испытуемые чаще пересказывали «Фильм о грушах» именно как просмотренный фильм, в то время как греческие испытуемые просто рассказывали историю, не упоминая о том, что действие происходит в фильме; согласно работе [Mazur and Chmiel 2012], большинство пересказов на европейских языках не включали лексему со значением 'фильм'; в частности, только 20% каталанских, 16,7% немецких, 20% французских и 20% британских английских пересказов включали хотя бы один раз эту лексему. Кроме того, согласно [Tannen 1980], при пересказе могут быть использованы другие характерные слова и выражения, отсылающие к фильму, например: *В этот момент камера возвращается к садовнику, который собирает груши; А теперь мы видим мужчину, ведущего козу; В это время мы слышим звуки пинг-понгового мячика и видим...* Для европейских языков, рассмотренных в работе [Mazur and Chmiel 2012], в большинстве случаев, в отличие от использования лексики 'фильм', такая отсылка встречается в пересказе хотя бы один раз; в частности, для голландских пересказов эта цифра составляет 79,2%, для греческого 80%, для польского 85%. Интуитивно представляется, что в «Репортажах о грушах» кинематографический взгляд будет встречаться чаще, чем в «Рассказах о грушах».

Заключение

В данной работе мы провели обзор первых исследований в жанре репортажа — нового дискурсивного жанра, имеющего свои специфические особенности. Взяв в качестве стимульного материала известный «Фильм о грушах» У. Чейфа, мы кратко описали собранные на его основе корпуса так называемых «Репортажей о грушах», сделали обзор проведенных исследований, а также наметили пути дальнейшего развития, работа над которыми идет в настоящий момент.

За рамками рассмотрения в настоящей работе осталось много других возможных направлений исследований. Остановимся чуть более подробно на еще одном перспективном, на наш взгляд, направлении, связанном с использованием в подобных исследованиях метода регистрации движений глаз. Данный метод позволяет с высокой точностью фиксировать, куда направлен взгляд испытуемого в каждый момент просмотра и комментирования фильма. В этом обзоре мы кратко описали один относительно частный вопрос, чем могут различаться комментарии двух испытуемых с точки зрения глазодвигательной активности и как это можно связать

с различиями в объеме их ВРП. Более общая задача — это всестороннее сравнение глаголдвигательного поведения испытуемых в двух дискурсивных режимах: в режиме репортажа, с одной стороны, и режиме пересказа после просмотра фильма, с другой стороны. Данные эксперименты также уже проводятся в рамках текущего проекта.

Литература

Кибрик А. А., Подлеская В. И. (ред.). Рассказы о сновидениях: корпусное исследование устного русского дискурса. М.: ЯСК, 2009. 736 с.

Малышева Е. Г. Русский спортивный дискурс. М.: Флинта, 2011. 369 с.

Тыщишина Т. Е., Федорова О. В. Еще раз о грушах, или Известность референта как фактор активации // Язык как он есть: Сборник статей к 60-летию Андрея Александровича Кибрика / ред. Т. И. Давидюк, И. И. Исаев, Ю. В. Мазурова и др. Москва: ООО Буки-Веди, 2023. С. 135–140.

Федорова О. В. Дело о стучалке с красным шариком открывается вновь // Состав науки: Сборник статей к юбилею Веры Исааковны Подлеской / ред. Н. А. Коротаев, Н. Р. Сумбатова. М.: Буки Веди, 2024. 594 с.

Airenti G., Bara B. G., Sandini G. (eds.). Pear Stories: 40 years later // Workshop Proceedings, v. 1419, Torino, Italy, September 25–27, 2015.

Chafe W. (ed.). The pear stories: Cognitive, cultural, and linguistic aspects of narrative production. Norwood, 1980. 327 p.

Daneman M., Green I. Individual differences in comprehending and producing words in context // Journal of Memory and Language. 1986. № 25 (1). P. 1–18.

Eaton J. H., Collis G. M., Lewis V. A. Evaluative explanations in children's narratives of a video sequence without dialogue // Journal of Child Language. 1999. № 26 (3). P. 699–720.

Fedorova O. V. "Pears Film" live: Cognitive peculiarities of the reportage // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог». М.: РГТУ, 2022. № 21. С. 203–210.

Kibrik A. A. Reference in discourse. Oxford: Oxford University Press. 2011. 651 p.

Kibrik A. A., Fedorova O. V. An empirical study of multichannel communication: Russian Pear Chats and Stories // Психология. Журнал Высшей Школы экономики. 2018. № 15 (2). С. 191–200.

Mazur I., Chmiel A. Towards common European audio description guidelines: Results of the Pear Tree Project // Perspectives: Studies in Translatology. 2012. № 20 (1). P. 5–23.

Tannen D. A comparative analysis of oral narrative strategies: Athenian Greek and American English // The Pear Stories / ed. by W. Chafe. New Jersey: Ablex, 1980. P. 51–87.

Tomlin R. S. Focal attention, voice, and word order: an experimental, cross-linguistic study // Word order in discourse / ed. by P. Downing and M. Noonan. Amsterdam: Benjamins, 1995. P. 517–554.

Olga V. Fedorova
Lomonosov Moscow State University
(Russia, Moscow)
olga.fedorova@msu.ru

“PEARS REPORTAGES”: THE RESULTS OF THE FIRST RESEARCHES AND FURTHER PERSPECTIVES

This article provides an overview of the research of the last two years, united by a common theme of studying the famous “Pears Film” by W. Chafe in the genre of reportage. Unlike the more traditional genre of retelling, this genre is still poorly researched and described in modern linguistic literature. The article briefly presents six small corpora collected by the author, including a corpus of recordings conducted using the method of recording eye movements (eyetracking). The following are the results of several experiments: on the relationship between the verbal working memory span of the participants and their peculiarities of speech production during the reportage; the study of individual speech styles of commentators using data from eye movements of the participants; the study of the prominence of the referent as an activation factor; the description of the model of introduction in the “Pears Reportages”; the study of “Pears Reportages” in norm and pathology, including a description of the linguistic characteristics of patients with schizophrenic spectrum diseases; comparison of “Pears Reportages” with “Pears Stories” recorded in the traditional genre of retelling, in the same participants. Students of Lomonosov Moscow State University participated in the experiments as test subjects.

Keywords: discourse; reportage; speech production; eye tracking; Pears Film.

References

Airenti G., Bara B. G., Sandini G. (eds.). Pear Stories: 40 years later. *Workshop Proceedings*, v. 1419, Torino, Italy, September 25–27, 2015.

Chafe W. (ed.). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, 1980. 327 p.

Daneman M., Green I. Individual differences in comprehending and producing words in context. *Journal of Memory and Language*, 1986, no. 25(1), pp. 1–18.

Eaton J. H., Collis G. M., Lewis V. A. Evaluative explanations in children’s narratives of a video sequence without dialogue. *Journal of Child Language*, 1999, no. 26 (3), pp. 699–720.

Fedorova O. V. The case of the paddleball is reopened. *Sostav nauki: Sbornik statei k yubileyu Very Isaakovny Podlesskoi* [The composition of science: A collection of articles for the anniversary of Vera Isaakovna Podleskaya] / ed. by N. A. Korotaev, N. R. Sumbatova. Moscow, Buki Vedi, 2024. 594 p. (In Russ.)

Fedorova O. V. “Pears Film” live: Cognitive peculiarities of the reportage. *Komp’yuternaya lingvistika i intellektual’nyeologii. Po materialam ezhegodnoi mezhdunarodnoi konferentsii “Dialog”* [Computational Linguistics and Intellectual Technologies.

Papers from the Annual International Conference “Dialogue”]. Moscow, RSUH Publ., 2022, no. 21, pp. 203–210.

Kibrik A. A. *Reference in discourse*. Oxford: Oxford University Press, 2011. 651 p.

Kibrik A. A., Fedorova O. V. An empirical study of multichannel communication: Russian Pear Chats and Stories. *Psikhologiya. Zhurnal Vysshei Shkoly ekonomiki* [Psychology. Journal of the Higher School of Economics]. 2018, no. 15 (2), pp. 191–200.

Kibrik A. A., Podlesskaya V. I. (eds.). *Rasskazy o snovideniyah: korpusnoe issledovanie usntogo russkogo diskursa* [Night Dream Stories: A corpus study of spoken Russian discourse]. Moscow : Jazyki Slavyanskikh Kul'tur Publ., 2009. 736 p.

Malysheva E. G. *Russkiy sportivnyy diskurs* [Russian sports discourse]. Moscow : Flinta Publ., 2011. 369 p.

Mazur I., Chmiel A. Towards common European audio description guidelines: Results of the Pear Tree Project. *Perspectives: Studies in Translatology*, 2012, no. 20 (1), pp. 5–23.

Tannen D. A comparative analysis of oral narrative strategies: Athenian Greek and American English. *The Pear Stories* / ed. W. Chafe. New Jersey : Ablex, 1980, pp. 51–87.

Tomlin R. S. Focal attention, voice, and word order: an experimental, cross-linguistic study. *Word order in discourse* / ed. by P. Downing and M. Noonan. Amsterdam: Benjamins, 1995, pp. 517–554.

Tyshhishina T. E., Fedorova O. V. [Once again about pears, or the prominence of the referent as an activation factor]. *Jazyk kak on est': Sbornik statej k 60 letiju Andreja Aleksandrovicha Kibrika* [The language as it is: A collection of articles dedicated to the 60th anniversary of Andrei Alexandrovich Kibrik] / ed. by T. I. Davidjuk, I. I. Isaev, Ju. V. Mazurova et al. Moscow, Buki-Vedi Publ., 2023, pp. 135–140.

ПРИЛОЖЕНИЕ (расшифровка выполнена Т. Е. Тыщишиной)

Репортаж

Висит груша, её кто-то сорвал, у него очень морщинистые руки, и, видимо, он одет в рубашку. (1.26) Да, это действительно мужчина, (0.38) причём в усах, (0.69) достаточно пухлый, вот он уронил грушу и спускается за ней. (1.09) По виду, (0.79) я не могу сейчас точно определить, какой он национальности. (0.62) Он груши собирал все в свой фартук, (0.71) сейчас заполняет ими (0.40) (aa0.72) корзины. (1.75) Достаточно (0.39) трепетно это делает. (2.49) (mm0.84) Также (0.29) на заднем плане стоит дерево с (0.25) стремянкой, на которую он взбирался. (1.25) Продолжает всё ещё заполнять груши и осматривать их. (0.89) Также он (0.54) берёт одну из них и очищает своим (1.87) своей банданой, которая у него до этого висела на шее, (0.81) завязывает её. (1.12) На заднем плане появляется, (0.53) видимо, его (0.32) товарищ, рыжий работник, он ведёт за собой не то козла, не то осла. (0.44) Осла. (2.00) А сам герой обратно полез на дерево по стремянке. (2.48) Второй человек... (0.33) Нет, это всё-таки козёл. (0.31) И он привёл (0.26) его (0.30) на передний план, (0.49) провёл мимо кадра, (1.36) (aa0.67) камера фокусируется на нём, (0.91) спереди стоят

корзины с грушами, и он продолжает вести козла. (0.62) И вот опять кадр на дерево, и мы видим, как (0.27) первый человек вновь срывает груши. (2.03) Он достаточно красно выглядит, (0.42) у него лицо всё красное. (1.07) Сзади проехал мальчик на велосипеде. (1.75) Сегодня достаточно жарко. Они все в шляпах и панاماх. (0.78) Видимо, где-то на югах и достаточно (сс0.50) летний период, в середине где-то всё происходит. (1.57) Мальчик подъезжает к дереву, (0.45) у него такой же красный (0.18) такая же красная бандана, (0.86) он останавливается прямо под стремянкой, (0.59) (мм1.11) и, видимо, (0.92) хочет что-то сделать с грушами, взять одну или помочь человеку. (1.49) Но (0.86) он ничего из этого не делает (0.22) пока что. (0.54) Вот он (0.4) берёт, не берёт грушу, (3.26) он перетаскивает корзину (0.23) ближе к своему велосипеду. (0.66) Мальчик поднимает велосипед, (0.86) садится, видимо, на него обратно, (0.51) и увозит грушу. (0.63) Корзину с грушами. (0.94) Наверное, его мама попросила это сделать, захватить за ними. (0.69) В принципе, (0.46) садовник, который собирал, и не против. (1.42) Видимо, (2.03) это для и мальчика они и предназначались. (1.63) Он виляет из стороны в сторону на велосипеде, и немного (2.53) (мм0.39) (3.03) разбрасывает груши, (0.30) навстречу ему едет девочка (1.79) в синей блузке, (0.98) (аа0.54) кадр отдаляется, они проезжают мимо друг друга, (0.60) она смахивает с него шляпу, он спотыкается о камень, и роняет все груши, и падает сам со своим велосипедом. (1.45) Сейчас он оправляется от падения, (1.03) и (0.68) он (0.13) повредил (0.31) голень (0.13) колено. (0.79) Не совсем понятно. (0.97) Рядом с ним (0.67) ещё (0.25) три мальчика. (0.86) Один совсем маленький в зелёном помогает ему подняться, отряхивает, другой чуть постарше (0.60) в синем собирать помогает груши. (0.41) Другой самый старший (0.44) брюнет (0.51) в красной (0.31) (ээ0.30) рубашке (0.49) подле (0.54) корзин собирает груши. (1.47) У мальчика в синем (0.82) (аа0.32) ракетка с закреплённой на ней шш= мячиком. (0.74) Он играет, пока д= остальные двое (0.63) ставят корзину на велосипед (0.56) и помогают мальчику уехать. (0.65) Тот уже не садится на (0.24) сам велосипед (0.57) и везёт (0.58) его (2.36) в руках ахах идёт с ним побок. (1.21) Те трое мальчиков уходят (0.36) (мм0.60) (0.72) по дороге, находят как раз-таки его шляпу, окриками первого самого мальчика, который вёз груши. (0.86) Один из них в синем с ракеткой подходит и отдаёт свою шля= ну, отдаёт найденную шляпу мальчику. (3.03) Тот в ответ ему за это даёт (0.17) не то две, не то три, не могу понять, груши. (1.79) Все счастливы, расходятся, прощаются друг с другом и уходят (0.85) (аа0.48) в разные стороны. (4.23) Троица, (0.76) мальчик в синей рубашке делится со своими двумя друзьями по груше (0.74) и они уходят по дороге в противоположную сторону. (0.77) Как раз-таки откуда и ехал первый самый мальчик. (0.79) На велосипеде. (2.45) Кадр возвращается к садовнику, который в очередной раз снимает с дерева груши и несёт их в своём фартуке вниз, чтобы наполнить корзину ими. (1.42) Он присаживается на колени, (0.18) и нерасторопно (1.05) хотел было заполнять (0.46) корзины. (1.00) Но он увидел, что одна из них пропала. (1.67) И (0.30) на пути к нему идут три (0.58) мальчика, у которых как раз-таки (0.39) его груши. (1.33) Он заподозривает что-то неладное. (1.05) Но так и не предпринимает ничего, провожая их взглядом. (3.94) Они уходят. Затемнение экрана.

Пересказ

(0.51) В общем-то, (1.69) первый кадр был на садовнике, у которого был белый фартук. На нём х= (эз0.51) был надет (0.14) ц= была надета красная бандана, а также (0.70) не то белая панاما, не то шляпа. И он на лестнице (0.49) собирал с грушевого дерева (0.20) (мм1.20) плоды, (0.66) и наполнял фв= (0.48) вторую корзину, которая была самая левая, ими. (0.98) (0.57) Всего было корзины три, (0.44) одна посередине была полностью заполнена, фт= с= вторую слева он (0.56) он практически заполнил, третья с= самая правая была пустая. (1.12) (аа0.48) (мм0.50) Потом он опять забрался на дерево по стремянке, собирал груши. (0.51) (аа0.91) В этот момент (0.45) на заднем плане мимо проходил (0.40) мужчина, который вёл за собой козу. (0.51) Они, как я понял, знакомы (0.84) (0.52) с садовником. (0.99) Насколько я также помню, (0.34) они поздоровались (0.20) и (0.61) разошлись (0.23) в разные, ну. (0.93) Мужчина, который вёл козу, дальше пошёл по своим делам. (0.77) Потом, (0.74) если я (0.29) правильно припоминаю, (0.91) (аа0.42) оттуда, (0.94) ж= оттуда же, отку= куда ушёл мужчина с козой, (0.70) (мм0.27) ехал мальчик тоже с красной банданой (0.54) (аа0.71) (0.48) в розовой рубашке (0.60) на красном велосипеде и с белой панамой. (0.88) (аа0.65) Он (0.10) подъехал к (0.10) грушевому дереву, (0.85) (аа0.68) (0.36) остановился, (0.78) (аа0.63) (0.15) глянул на (0.24) садовника, который (0.32) насколько как я понял не заметил его собирая груши, (0.94) (аа0.29) (мм0.32) далее мальчик (0.22) (мм0.78) (2.42) посмотрел на самую левую корзину, (0.57) и решил её забрать, поставив на передний бампер своего велосипеда, и уехал. (0.56) Садовник этого тоже и не заметил, и продолжал (0.72) (0.35) собирать груши. (1.75) (аа0.31) Следующим кадром (мм0.98) отдалённая дорога, по которой справа (0.80) ехал (0.16) наш мальчик с(эз0.77) к= а нет, (0.44) до этого ещё был кадр, (0.46) (аа0.58) как мальчик ехал виляя из стороны в сторону передним колесом по дороге, разбра=(эз0.51) ну, (0.59) не разбрасывая, (0.70) (аа0.43) выронив несколько груш. (0.53) Затем ещё отдалённый кадр, на котором была (0.58) видна дорога вдалеке, справа ехал (0.77) мальчик, слева (0.49) (мм0.54) ехала девочка в синей блузке (0.55) на велосипеде. (0.41) Они засмотрелись друг на друга, когда проезжали мимо, (0.72) девочка также сбила шляпу с мальчика, случайно, кажется (0.40), тот оглянулся на шляпу, и из-за этого потерял (0.18) управление велосипедом, и в результате (1.11) упал сам и уронил груши, (эз1.01) (1.75) по всей земле. (1.01) (аа0.40) Следующем, (1.05) в следующей сцене (0.67) (аа0.58) (0.25) мальчик (0.66) растирал (0.40) место ушиба на (0.57) голени, (0.33) и оглядывал (0.19) (мм0.66) происш= место происшествия, (0.90) и (0.10) пытался (аа0.38) что-то найти, не понял так и, что он дальше ещё делал. (0.77) Потом (1.40) (аа1.27) слева как раз-таки (1.13) после дороги, куда он собирался ехать, (0.86) шл= шла тройца из (0.76) парней (1.3) различного возраста, (0.38) самый младший был (0.29) (аа0.70) по-моему блондин в зелёной (0.23) (аа0.86) (0.57) рубашке, (0.45) чуть постарше был в (0.60) светло-синей рубашке, и самый старший был (0.48) (аа0.34) в ну (0.23) тоже не то розовой, не то ярко-красной рубашке. (0.84) Он был брюнетом, насколько я помню, (0.49) они заметили (эз1.01) (0.66) (мм0.18) как раз-таки

мальчика на велосипеде, что он упал, (0.85) (aa0.49) (мм0.80) и (0.12) подошли к нему помогать. Самый младший в зелёной рубашке (0.86) (aa1.63) насколько я помню (0.69) помог мальчику (0.60) с велосипедом. (0.70) Растереть место ушиба и отряхнуться ему, (0.50) (aa0.70) самый старший в красной рубашке побыл (0.61) нашёл корзину (0.35) п= (0.43) поднял её и начал собирать (0.46) груши. (0.33) Тот, что с синей, он был с ракеткой, (0.45) ну, не ракеткой, а (0.69) игрушкой. (0.76) (aa0.67) (мм0.47) Ракеткой с мячиком закреплённым (0.43) вот, он до этого игрался с ней, он её положил в задний карман своих джинс и тоже стал собирать (0.17) груши. (0.69) Затем мальчик в зелёном отряхнул (0.74) (aa0.68) мальчика на велосипеде (0.43), помог ему поднять велосипед и тоже продолжил собирать (0.80) груши. (0.62) В итоге они поставили корзину обратно ему на бампер (0.60) и (1.32) решили распрощаться с ним и пошли дальше (ээ0.31) в= направо, откуда ехал мальчик (0.80) на велосипеде. (0.36) Мальчик запрыгнул на велосипед и было поехал, (0.42) но его окрикнули, потому что троица нашла его шляпу, которую он ранее (0.49) обронил, (0.10) ну, или панаму, не суть. (0.82) (oo0.37) Вот, как раз-таки мальчик в синем (мм0.56) опять игрался с ракеткой, (0.44) потом он её убрал, сунул (0.21) (мм0.47) в задний карман джинс, (1.02) поднял шляпу и подбежал к мальчику на велосипеде, отдал ему ее. (0.42) Взамен (0.61) мальчик на велосипеде (0.37) дал ему три груши, как раз на всех троих, (0.35) он вернулся к своей компании, разделил со всеми грушами (0.25) (ээ0.28) поделилился со всеми грушами, (1.01) (aa0.38) (0.14) и пошёл, и они все пошли далее по своему маршруту, (0.59) в результате чего они (0.23) дошли (0.47) (0.55) до (2.01) садовника, который, вновь собрав груши в фартук, спускался вниз с ними. (0.41) В этот он обма= обнаружил пропажу (0.56) одной из своих корзин, (0.27) которую забрал до этого мальчик на велосипеде. (1.18) И в этот момент мимо него проходила троица, все они ели груши, которые (0.58) были точно такие же, какие собирал садовник. (1.36) (oo0.45) В ответ на это он ничего не сделал, лишь проводив их взглядом, а те пошли дальше направо по своим делам. В ту же сторону, куда ушел мужчина с козой изначально. (0.33) Далее экран затемняется. (0.34) Конец.

М. И. Киосе¹, А. В. Леонтьева², О. В. Агафонова³

Московский государственный лингвистический университет,

Институт языкознания РАН

(Россия, Москва)

maria_kiose@mail.ru¹, lentevanja27@gmail.com²,

olga.agafonova92@gmail.com³

ПОЛИМОДАЛЬНОЕ КОНСТРУИРОВАНИЕ РЕФЕРЕНТОВ В ЭКСПОЗИТОРНОМ РУССКОЯЗЫЧНОМ ДИСКУРСЕ: СЛОВО VS. ОБЪЕКТ¹

В статье на материале записей речевого и жестового поведения участников, поясняющих различия между близкими синонимами, устанавливаются особенности полимодального конструирования ими референтов двух типов — слов и объектов. Для проведения анализа применяется метод образ-схем; разрабатывается модель трехуровневого анализа образ-схем (миметические / телесные схемы, схемы базового уровня и схемы высокого уровня абстракции — зд. СЛОВО и ОБЪЕКТ). С опорой на теории уровней образ-схем Дж. Грэди, квалиа-ролей Дж. Пустейовского и миметических схем Й. Златева описывается состав образ-схем, представляющих четыре схемы базового уровня: ЧАСТЬ–ЦЕЛОЕ, ТИП, ФУНКЦИОНИРОВАНИЕ, ТРАНСФОРМАЦИЯ, соотносимых с дейктическими, репрезентирующими, прагматическими жестами и жестами-адаптерами. Распределение частотности базовых образ-схем, а также их синхронизация с жестами разных типов указывают на сходства и различия в конструировании референта как слова или как объекта. Значимые различия обнаружены в отношении активности схем ЧАСТЬ–ЦЕЛОЕ (превалирует при конструировании референта-объекта) и ТРАНСФОРМАЦИЯ (превалирует при конструировании референта-слова); также наблюдаются различия в использовании дейктических жестов при конструировании референта-объекта. Высказывается гипотеза о том, что в основе такого варьирования лежит выбор миметических образ-схем при конструировании референтов разного типа.

Ключевые слова: полимодальное поведение, экспозиторный дискурс, образ-схема, слово, объект, миметическая образ-схема.

¹ Исследование выполнено в рамках Государственного задания «Кинетические и вокальные аспекты коммуникации: параметры варьирования» (проект № FMNE-2022-0015) в Институте языкознания РАН.

Введение

При мультимодальном анализе специфических особенностей экспозиторного дискурса, или дискурса, реализующего объяснение / пояснение [Longacre 1983], исследователи принимают во внимание свойственную ему нечеткость в представлении референции. В языковой модальности эта нечеткость проявляется в выборе выражений, заместителей и аппроксиматоров, при введении референтов в дискурс [Podlesskaya 2013], в использовании слов с обобщающим значением, демонстрирующих «не конкретно-референтный статус описываемых объектов» [Бергельсон 2007], иных маркеров обобщения — «кванторов общности, родовых ИГ, стативных глаголов» [Кибрик 2009: 10]. К анализу, соответственно, могут привлекаться либо собственно языковые маркеры замещения и аппроксимации (если весь изучаемый пассаж рассматривается как представляющий образец экспозиторного дискурса) или типы риторических отношений и жанровых схем (дополнительно, если изучаемый пассаж представляет конфигурацию дискурсов как «типов изложения»).

Однако в тех ситуациях, когда человек выстраивает пояснение в отношении вторичной картины мира, репрезентируемой средствами какой-либо семиотической системы (например, языка, изображения), обращает на себя внимание еще одна особенность такой экспозиции, а именно, выстраивание пояснения в отношении либо самого объекта первичной картины мира (конкретного или абстрактного характера), либо семантики поясняемой единицы вторичной семиотической системы. Например, поясняя различия между синонимами, человек может выстроить речь «с опорой» на свойства объекта (пример 1) или семантику называющего его слова (пример 2):

- (1) *идеал и совершенство / ну-у.. / идеал.. / это-а.. / ой-х.. / ну для меня допустим есть в какой-то сфере идеал / человек который допустим.. / что-то идеально делает...*
- (2) *чепуха и ерунда / для меня вообще они как будто бы одинаковы / ну вот.. / по с.. по значению / вот это какая-то.. / чепуха.. / ну я бы просто по-разному их сказала / что... / ну вот... / по-разному употребила бы...*

В примере (1) говорящий представляет ситуацию в *допустим есть в какой-то сфере идеал* и фрагмент ситуации первичной картины мира, репрезентированной средствами языковой картины мира в *человек который допустим.. / что-то идеально делает...* В примере (2) говорящий при пояснении различий между синонимами *чепуха* и *ерунда* уже указывает на то, что сами слова имеют разную семантику в *ну вот.. / по с.. по значению* и сфера их употребления как единиц вторичной семиотической системы будет отличаться в *ну я бы просто по-разному их сказала / что... / ну вот... / по-разному употребила бы...*

Привлечение к анализу мультимодального дискурса, эксплуатирующего, например, речевое и жестовое поведение говорящего, способно установить дополнительные особенности представления (иначе — конструирования) референтов первичной и вторичной картин мира. Исследование экспозиторного дискурса как

мультимодального с опорой на языковые маркеры замещения и аппроксимации, реализующие отношения нечеткой референции, было ранее проведено в [Iriskh-pova et al. 2023]. В работе обнаружено, что замещение и аппроксимация как две стратегии нечеткой референции значительно (и статистически значимо чаще) сопровождаются выбором говорящим разных функциональных типов жеста — адаптеров с аппроксиматорами, а репрезентирующих и прагматических жестов с заместителями. При этом были установлены различия в индивидуальном использовании жестов только одного типа, адаптеров, выбор которых в действительности может быть обусловлен иными факторами, кроме стратегий нечеткой референции. Полученные результаты свидетельствуют о том, что говорящие довольно сходным образом выстраивали мультимодальный экспозиторный дискурс; однако при анализе речевого поведения мы также обратили внимание на то, что доля пояснений «с опорой» на семантику референта-объекта или семантику называющего его слова (референта-слова) демонстрирует большую вариативность. На основе сделанных наблюдений была далее сформулирована гипотеза о том, что мультимодальное поведение определяется типом референта и ракурсом его пояснения.

В настоящей работе для верификации данной гипотезы мы воспользуемся инструментарием квалиа-ролей Дж. Пустейовского [Pustejovsky 1995] (и развитием данной концепции в работах Р. Джекендоффа [Jackendoff 2002]), которые описывают четыре роли референта любого типа, формальную — PART-WHOLE (ЧАСТЬ–ЦЕЛОЕ), конститутивную — KIND (ТИП / ХАРАКТЕРИСТИКА (далее — ТИП)), агентивную — FUNCTION (ФУНКЦИОНИРОВАНИЕ), телическую — LIFE HISTORY (ТРАНСФОРМАЦИЯ). Предположительно, не только тип референта, но и выбор говорящим квалиа-ролей, представленных в виде образ-схем, при конструировании референта в экспозиторном дискурсе задает распределение речевой и жестовой нагрузки, а значит, является дискурс-формирующим фактором.

Методология исследования

Методологической основой исследования, как мы указали выше, стала возможность изучения конструирования референтов разного типа в аспекте степени их воплощенности и «выраженности» их знаковой природы, в нашем случае, объектов и слов, с опорой на их дискурсивные роли, определение которых проводится в контексте их использования [Pustejovsky 1995; Jackendoff 2002]. При том, что множественные исследования апеллируют к определенным ролям, часто исследуемым в формате образ-схем (типа ОБЪЕКТ, ПУТЬ, КОНТЕЙНЕР и др.), например, в работе [Johnson 1998; Peña 1999], они редко рассматриваются системно. Преимущество системного подхода, развиваемого Дж. Пустейовским, заключается в возможности его использования для дискурсивных исследований, в которых анализируются способы конструирования референтов разного типа. Так, квалиа-роль ЧАСТЬ–ЦЕЛОЕ обнаруживается в тех случаях, когда референт представлен либо через его компоненты (*а боязнь там... / ну боязнь паучков..*), либо, напротив, включен в категорию референтов более генерализованного типа (*у кого вой... м-м у волка / а-а рев... у-у-у медведя*). Квалиа-роль ТИП фиксируется, если конструирование

референта происходит через фокусирование его характеристик (*схватка ну это не обязательно что-то / что-то негативное что ли*). Квалиа-роль ФУНКЦИОНИРОВАНИЕ описывает случаи агентивной манифестации самого референта (*мертвец вос... восстал из мертвых*) или случаи использования референта (*ну линия она больше в математике как бы употребляема...*), в то время как роль ТРАНСФОРМАЦИЯ представляет случаи изменения состояния референта (*а мертвец когда скорее не до конца еще умер*). Отметим, что в типологии Дж. Пустейовского случаи ФУНКЦИОНИРОВАНИЯ второго типа, иллюстрирующие использование референта, рассматриваются в рамках роли ТРАНСФОРМАЦИЯ (т. е. роль ФУНКЦИОНИРОВАНИЕ описывает только случаи конструирования агентивного референта), однако при этом для русскоязычного материала, особенно в устной речи, затруднено определение роли, например, при использовании 1) эллиптических предложений в ЭДЕ, например, в *и конечно же черта тоже в смысле математическом*, 2) возвратных глаголов, которые характеризуют референт одновременно как агенс, и как объект действия, например, в *также там я не знаю комната погрузилась во тьму*, 3) глаголов действия в метафорическом значении, как в *Знать бы откуда эти слова пошли*, 4) неодушевленных референтов как агентивных, как в *Ну вот горит огонь*, 5) объяснений через примеры соответствующих ситуаций, например, объяснение семантики слова *страх* в *медик там.. ты ставишь какой-то диагноз*. В связи со сложностью разграничения агентивного и неагентивного референта в таких случаях мы рассматривали их в рамках роли ФУНКЦИОНИРОВАНИЕ. В действительности и сам Дж. Пустейовский, и развивающий далее концепцию квалиа-структуры Р. Джекендофф указывают на то, что она в большей степени применима для анализа номинативного компонента в составе предложений с акциональными предикатами или (с определенными ограничениями) изолированных номинаций с сохранившейся мотивировкой, т. е. указывают на некоторые формальные сдерживающие факторы ее применения. Однако, на наш взгляд, это формальное ограничение может быть преодолено с учетом особенностей анализируемого языкового и дискурсивного материала. При этом само разграничение четырех ролей (представленных в виде образ-схем) применительно к изучению полимодального поведения говорящих, нацеленного на экспозицию номинативных единиц через их введение в некоторый контекст, в том числе атрибутивный и предикативный, представляется продуктивным.

Это подтверждается и многочисленными работами, исследующими полимодальное поведение, в которых типы образ-схем рассматриваются как определяющие выбор типов жеста. Так, связь жестов с образ-схемами изучается в работе И. Миттельберг [Mittelberg 2018], в которой автор указывает, что жесты могут содержать базовые характеристики и семантику таких схем. Также выявляются следующие наиболее распространенные схемы, которые часто репрезентированы в жестах рук за счет их движения и/или формы ладони: КОНТЕЙНЕР / СОДЕРЖАНИЕ (CONTAINMENT) — рука словно держит горсть чего-л.; ОБЪЕКТ (OBJECT) — руки как бы удерживают некий объект или указывают на него; ПУТЬ (PATH) — рука очерчивает определенное направление движения объекта

в пространстве; БЛИЖЕ–ДАЛЬШЕ (NEAR–FAR) — указывается расположение объекта в пространстве с помощью локализации жеста ближе/дальше от тела говорящего и т. д. В первых исследованиях, посвященных образ-схемам и их связи с жестами, А. Ченки рассматривает особенности полимодальной реализации образ-схем ПУТЬ и ПРЯМАЯ [Cienki 1998a; 1998b]. В другой работе автор описывает ряд «графических» схем (КОНТЕЙНЕР, ЦИКЛ, СИЛА, ОБЪЕКТ, ТРАЕКТОРИЯ), которые используются для классификации жестов, наблюдаемых в ходе естественной коммуникации [Cienki 2005]. Так, например, образ-схема ЦИКЛ лежит в основе анализа жестов, включающих повторяющееся круговое движение руки, вращающейся «наружу» в запястье (см. тж. в [Ladewig 2011]). Образ-схема ПУТЬ лежит в основе формирования формы и смыслообразования при конструировании визуально-пространственных объектов в жестах [Forceville, Jeulik 2011]. Приводя конкретный пример проявления образ-схемы ПУТЬ в жестовом поведении, Р. Уильямс указывает, что она может проявляться в круговых движениях руки учителя вокруг циферблата часов в процессе обучения детей тому, как определять время [Williams 2008]. Й. Златев в работе (2013) рассматривает иконические жесты и их связь с особым типом образ-схем, миметическими (иконическими и имеющими телесный характер); объектом его исследования становится процесс совместного развития миметических схем и жестов в онтогенезе. Выдвигается гипотеза о происхождении детских иконических жестов: автор предполагает, что они являются довербальными миметическими схемами, и формулирует в том числе критерии телесного мимезиса (см. [Златев 2013]). Также в его работах приводится состав миметических образ-схем, например, ДАТЬ (GIVE), ПОКАЗАТЬ (SHOW), ДОТЯГИВАТЬСЯ (REACH-TO-THING), ОТ СЕБЯ (AWAY-FROM-EGO) [Zlatev 2005, 2007]. В отечественной науке также отмечается влияние образ-схем (чаще без использования данного термина) на невербальное поведение. Так, например, чашеобразная форма ладони с немного разведенными напряженными пальцами часто сопровождает глаголы, которые содержат в себе идею удерживания через образ-схему КОНТЕЙНЕР; перебирание пальцами отражает идею передвижения или некоторых небольших движений/действий, то есть в их основе лежит образ-схема ПУТЬ. Образ-схемы могут проявляться в дейктических жестах в форме указательного или большого пальца в момент указания, например, когда такие жесты сопровождают императивные высказывания, в значении «иди туда и делай так», в основе которых лежит образ-схема БЛИЖЕ–ДАЛЬШЕ (см. [Гришина 2014]). Обращение к таким свойствам объекта, как его форма и конфигурация, определяющим выбор изобразительных жестов, находим в работе С. О. Савчук [Савчук 2021]; данные проявления могут быть рассмотрены как реализующие образ-схемы КОНТЕЙНЕР, ТИП, ОБЪЕКТ и др.

В целом, обзор приведенных работ показывает, что исследователи обращаются к образ-схемам разного уровня абстракции, более «телесным», «воплощенным» («миметическим» в терминологии Дж. Златева) и более абстрактным, вторичным (типа ЦИКЛ, СИЛА и др. [Cienki 2005]), часто не разграничивая их, что затрудняет процедуру их системного исследования в дискурсе и может приводить

к рассогласующимся результатам. Например, в работе Д. МакНила показано, что использование метафорических жестов возрастает при конструировании абстрактных референтов [McNeill 2017]. В то же время в иной работе [Zdravilova et al 2018], напротив, установлено, что конструирование именно конкретных референтов чаще сопровождается введением иконических жестов, например, репрезентирующих. На наш взгляд, такое различие в полученных результатах может, в том числе, быть вызвано тем, что абстрактный характер, «вторичность», снижение степени воплощенности референта может быть также обусловлено тем, что референт репрезентирует не первичную, а вторичную картину мира, например, языковую.

Для решения проблемы ранжирования образ-схем в аспекте степени их воплощенности и для разработки корректной процедуры анализа полимодального поведения с опорой на инструментарий образ-схем представляется возможным привлечь уровневую теорию образ-схем, предложенную Дж. Грэди [Grady 2005], согласно которой можно разграничить 1) телесные, иначе — миметические образ-схемы, или образ-схемы самого низкого уровня абстракции, например, ДАВАТЬ, ПОКАЗЫВАТЬ, КОНТЕЙНЕР, 2) образ-схемы базового уровня (абстракции), в нашем случае соотносимые с четырьмя типами образ-схем Дж. Пустейовского, ЧАСТЬ–ЦЕЛОЕ, ТИП, ФУНКЦИОНИРОВАНИЕ, ТРАНСФОРМАЦИЯ, 3) образ-схемы самого высокого (но различающегося) уровня абстракции (и самого низкого уровня воплощенности), в нашем случае (для целей данного исследования) это СЛОВО и ОБЪЕКТ, при этом ОБЪЕКТ может быть конкретный и абстрактный, но СЛОВО в любом случае отличается более низкой воплощенностью как соотносимое с референтом-объектом опосредованно. В этом случае конструирование СЛОВА и ОБЪЕКТА можно рассмотреть в полимодальном ракурсе с помощью анализа базовых образ-схем и телесных (миметических) образ-схем, синхронизированных с жестами.

Дизайн эксперимента и процедура анализа

Для установления особенностей полимодального конструирования референта как СЛОВА и как ОБЪЕКТА были использованы видеозаписи эксперимента, в котором приняли участие 19 студентов лингвистического университета (средний возраст — 20 лет); их мультимодальное поведение фиксировалось фронтальной камерой. Для стимулирования образцов экспозиторного дискурса участникам было предложено пояснить различия между парами близких синонимов типа *линия* и *черта*, *долг* и *обязанность*, *рев* и *вой*, *мгла* и *темнота* (всего 14 пар синонимов), означающих конкретные и абстрактные референты. Для настоящего исследования к анализу привлечены 70 экспозиторных образцов, каждый из которых — это пояснение различий в одной паре синонимов. С опорой на концепцию, предложенную в работах [Кибрик, Подлесская 2009; Прокофьева 2022] и описывающую процедуру сегментирования устной речи на ЭДЕ (учитывающую фонологические и синтаксические критерии), речь участников была разбита на 1325 ЭДЕ, которые далее подвергались анализу на предмет типа образ-схем трех уровней абстракции (воплощенности). Анализ жестового поведения проведен с опорой на типологию

функциональных жестов, предложенную в работах К. Мюллер, А. Ченки, О. К. Ирисхановой, ориентированных на анализ речи и жеста в условиях дискурсивной вариативности [Müller 2014; Iriskhanova, Cienki 2018], согласно которой выделяются четыре основных типа жеста, дейктические, репрезентирующие, прагматические и адаптеры с их дальнейшей детализацией (например, дейктические жесты подразделяются на указательные и жесты-прикосновения, репрезентирующие — на жесты удержания, придания формы, разыгрывания действия, воплощения, очерчивания, см. подробнее в [Iriskhanova et al. 2023]). Для установления значимости квалиа-роли, представленной четырьмя базовыми образ-схемами, в мультимодальном экспозиторном дискурсе были определены различия в использовании жестов, синхронизированных с ЭДЕ, в которых референты конструируются с помощью этих образ-схем.

Продемонстрируем процедуру установления образ-схем трех уровней. Так, на рисунке 1 участница поясняет слово *труп*, вводя его в контекст таким образом, что обозначаемый данным словом референт является объектом действия (*ты видишь перед собой труп*).



ты патологоанатом..



... ты видишь перед собой труп...



... труп как факт

Рис. 1. Использование образ-схемы ФУНКЦИОНИРОВАНИЕ с дейктическим жестом

Данная роль позволяет определить, что конструирование референта происходит с использованием образ-схемы ФУНКЦИОНИРОВАНИЕ, в данном случае речь идет о сенсорном (ментальном) действии субъекта, производимом в отношении данного референта. Таким образом, в основании образ-схемы базового уровня ФУНКЦИОНИРОВАНИЕ лежит миметическая схема ОЩУЩЕНИЕ (восприятие зрением). При этом участница использует прагматический жест, репрезентирующий компонент дискурса, а именно, само действие в отношении референта в противовес тому состоянию, на которое было указано в предшествующей ЭДЕ (*ты патологоанатом*). Следующий за ним дейктический жест будто бы указывает на описываемый референт, как если бы он находился в поле ее зрения. Можно

заметить, что в данном примере участница конструирует референт как объект, включенный в событие, представляющее его функционирование (точнее, действие, направленное на данный объект). Такие случаи мы отнесли к примерам использования образ-схемы третьего уровня абстракции ОБЪЕКТ. Данный пример в целом описывается следующим набором образ-схем: 1) Миметическая схема: ОЩУЩЕНИЕ (восприятие зрением), 2) Схема базового уровня: ФУНКЦИОНИРОВАНИЕ, 3) Схема высокого уровня абстракции: ОБЪЕКТ.

Рассмотрим другой пример пояснения того же слова (Рис. 2), в котором используется иная базовая образ-схема.



ну, труп это



*когда человек уже ну умер
полностью*

Рис. 2. Использование образ-схемы ТРАНСФОРМАЦИЯ
с прагматическим жестом

На Рис. 2 представлен пример введения одной ЭДЕ с двумя жестами. Участница в речи конструирует референт как нечто, претерпевающее переход из одного состояния в другое, поэтому наблюдается введение образ-схемы базового уровня ТРАНСФОРМАЦИЯ. При этом, соответственно, используется миметическая схема ИЗМЕНЕНИЕ СОСТОЯНИЯ. С началом ЭДЕ синхронизировано введение жеста-самоадаптера, который далее переходит в прагматический жест с функцией эмфазы уже при конструировании референта как меняющегося. Таким образом участница мануально усиливает прагматическое значение описываемого компонента дискурса. Образ-схемой абстрактного уровня будет, как и на Рисунке 1, схема ОБЪЕКТ.

На Рисунке 3 продемонстрируем пример пояснения того же слова с реализацией образ-схемы абстрактного уровня СЛОВО.

Участница начинает пояснение с неудачной попытки сравнить один референт с другим, но не может подобрать объект сравнения; далее она эксплицитно определяет референт как референт в языке (*это термин*) и указывает на его конститутивную роль ТИП — «способность» быть ориентированным на кого-либо и необходимым



*труп.. оно как бы более... это
термин... для-я... не знаю...*

для полиции то есть что...

Рис. 3. Использование образ-схемы ТИП с репрезентирующим жестом

для кого-либо, что реализует миметическую схему ВЫДВИЖЕНИЕ. В данном случае, в отличие от ситуаций, представленных на Рисунках 1 и 2, участница конструирует референт как СЛОВО. При этом используется репрезентирующий жест удержания, где описываемый референт представлен как находящийся как бы на ладонях говорящего.

По итогам анализа речевого поведения была разработана общая модель образ-схем трех уровней применительно к конструированию референтов как ОБЪЕКТОВ и как СЛОВ, далее были установлены закономерности в их распределении с учетом синхронизации с жестами.

Результаты

На первом этапе анализа в 1325 ЭДЕ был определен состав образ-схем, представляющих уровни миметических и базовых образ-схем. Приведем его в Таблице 1.

Отметим, что миметические образ-схемы определенно могут быть подвергнуты дальнейшей детализации, например, образ-схема ИСПОЛЬЗОВАНИЕ включает схемы БРАТЬ, КЛАСТЬ, ДАВАТЬ, НЕСТИ, ОПУСКАТЬ и иные [Zlatev 2005, 2007], однако подробная их классификация не входит в задачи настоящей работы; кроме того, она определяется семантическими характеристиками стимульных слов и, соответственно, характеристиками тех референтов, которые конструируют участники, и будет варьироваться при ином их составе. В то же время установление общих их типов вслед за [Pustejovsky 1995; Jackendoff 2002] представляется более продуктивным, так как позволяет инвентаризировать возможные операции их телесного (как сенсорного, так и ментального) конструирования. Как можно заметить, уровневое соотнесение образ-схем было выполнено только в отношении миметического и базового уровня, в то время как любая из четырех образ-схем

Таблица 1. Образ-схемы конструирования референтов в экспозиторном дискурсе

Миметические (телесные) образ-схемы	Образ-схемы базового уровня	Образ-схемы высокого уровня абстракции
ФРАГМЕНТАЦИЯ КОНТЕЙНЕР	ЧАСТЬ-ЦЕЛОЕ	ОБЪЕКТ СЛОВО
ВЫДВИЖЕНИЕ СРАВНЕНИЕ ОЦЕНКА	ТИП	
ИСПОЛЬЗОВАНИЕ ОБЛАДАНИЕ НАХОЖДЕНИЕ В ПРО- СТРАНСТВЕ ПЕРЕДВИЖЕНИЕ СИЛА ПОНИМАНИЕ ОЩУЩЕНИЕ	ФУНКЦИОНИРОВАНИЕ	
ИЗМЕНЕНИЕ СОСТОЯНИЯ СОЗДАНИЕ РАЗРУШЕНИЕ	ТРАНСФОРМАЦИЯ	

базового уровня может конструировать как СЛОВО, так и ОБЪЕКТ. Так, на Рисунке 4 участница при конструировании референта как СЛОВА применяет образ-схему ФУНКЦИОНИРОВАНИЕ.



нести бремя же по-моему...

Рис. 4. Использование образ-схемы ФУНКЦИОНИРОВАНИЕ при конструировании СЛОВА с прагматическим жестом

При пояснении слова *бремя* участница приводит пример его использования в языке в составе фразеологического выражения, таким образом, конструирование референта как СЛОВА представляется очевидным. При этом используется прагма-

тический жест, который несет контактоустанавливающую функцию и функцию выражения отношения. Последняя функция находит свое отражение и в речи говорящей. Как мы показали ранее на Рисунке 1, та же самая образ-схема базового уровня ФУНКЦИОНИРОВАНИЕ была использована говорящей при конструировании референта как ОБЪЕКТА.

На втором этапе с опорой на типы базовых образ-схем было установлено распределение частотности в конструировании референтов как СЛОВ или ОБЪЕКТОВ. В 1325 ЭДЕ обнаружено 230 примеров конструирования референта как СЛОВА и 419 примеров конструирования его как ОБЪЕКТА, что, безусловно, подтверждает очевидную вторичность языкового конструирования референта. В Таблице 2 приведем полученные данные распределения частотности (абсолютной и относительной) базовых образ-схем.

Таблица 2. Частотность базовых образ-схем конструирования референтов

Образ-схемы базового уровня	Конструирование ОБЪЕКТА	Конструирование СЛОВА
ЧАСТЬ-ЦЕЛОЕ	72 / 17.2 %	62 / 27 %
ТИП	143 / 34.1 %	80 / 34.8%
ФУНКЦИОНИРОВАНИЕ	152 / 36.3 %	75 / 32.6 %
ТРАНСФОРМАЦИЯ	52 / 12.4 %	13 / 6 %

Были обнаружены значимые различия в использовании образ-схем ЧАСТЬ-ЦЕЛОЕ ($\chi^2 = 5.71$, $p = 0.017$) и ТРАНСФОРМАЦИЯ ($\chi^2 = 6.19$, $p = 0.013$). Это означает, что при конструировании референта как СЛОВА говорящий чаще, чем при конструировании его как ОБЪЕКТА, представляет его в рамках структуры языка, например, в рамках фонологического уровня в *ну страх звучит мощней чем боязнь* (о страхе и боязни), лексического уровня в *обычно ну опять же там кара небесная или кары египетские..* (о каре), *а линия ну ее ведут* (о линии), *а черта это то что.. через что можно.. и переступить..* (о черте), синтаксического уровня в *то есть ты просто по факту говоришь / ну вот горит огонь* (об огне), *когда это твои родители говорят это твоя обязанность* (об обязанности). Языковыми маркерами образ-схемы ЧАСТЬ-ЦЕЛОЕ при конструировании референта как СЛОВА служат указания на его фонетический облик (*звучит*), сочетаемость в составе словосочетания, чаще фразеологического (*кара небесная, кара египетская, вести линию, переступить черту*) или предложения, в котором компоненты имеют высокую совместную встречаемость (*горит огонь, это твоя обязанность*). При конструировании референта как ОБЪЕКТА говорящий чаще, чем при конструировании его как СЛОВА, указывает на особенности его создания и преобразования, например, в *наказание может быть даже чисто вот я у тебя вот отберу у тебя что-нибудь на какое-то условно время* (о наказании), *а темнота в комнате когда ты выключил свет* (о темноте), *а-а мертвец может это тот кого уже закопали* (о мертвце). Языковыми маркерами образ-схемы ТРАНСФОРМАЦИЯ при конструировании референта как ОБЪЕКТА служат фазовые глаголы,

описывающие либо функционирование самого референта, либо действия субъекта в отношении референта; в любом случае можно установить по меньшей мере два момента во времени, которые соотносятся с двумя состояниями референта (у тебя было что-то — у тебя нет этого, свет был включенным — свет выключен, быть не закопанным — быть закопанным).

При этом обращает на себя внимание то, что наиболее частотные образ-схемы базового уровня, ТИП и ФУНКЦИОНИРОВАНИЕ, демонстрируют сходную частотность при конструировании референтов как ОБЪЕКТОВ и как СЛОВ. Полученные результаты свидетельствуют о том, что в речи ОБЪЕКТЫ и СЛОВА, в целом, конструируются сходным образом, с опорой на их характеристики и производимые ими / над ними действия. Языковыми маркерами образ-схемы ТИП служат разные атрибуты, например, об ОБЪЕКТЕ — в *страх сильнее, а боязнь не очень сильное, ложь что-то маленькое, а обязанность это наверно.. больше что-то материя-альное.., то есть она такое мимолетное что-то чувство такое боязнь.., ну вранье это прям совсем какая-то... непонятная ложь которая..., мне кажется долг это пообширнее по сильнее по помощнее*; о СЛОВЕ — в *чепуха и ерунда ну вот тут они очень похожи вообще стилистически так-то тоже достаточно, для меня вообще они как будто бы одинаковые по с.. по значению, а.. а мертвец это.. разговорное слово*. Как можно видеть из приведенных примеров, во многих случаях в них реализуются сразу несколько миметических образ-схем: помимо ВЫДВИЖЕНИЯ с помощью синтаксического фокусирования наблюдаем СРАВНЕНИЕ в <...> *они очень похожи <...>, <...> они как будто бы одинаковые <...> и ОЦЕНКУ (например, прагматическую) в *страх сильнее, а боязнь не очень сильное, <...> по сильнее по помощнее*. Языковыми маркерами образ-схемы ФУНКЦИОНИРОВАНИЕ являются глаголы, чаще с семантикой действия, например, об ОБЪЕКТЕ — в *черта она как будто что-то огораживает, кара... прилетает человеку от-т-т ч-ч чего-то неизведанного, ноша это то на что человек подписывается сам. ., ну да вот у меня как бы в воображении пламя оно как бы движется..., бремя как-то за тобой так волокется..*; о СЛОВЕ — в *что бремя больше и м-м-м у-у используется в-в м-м в литературе там допустим девятнадцатого века, кара / так для усиления или для какого-нибудь там юмористического эффекта то можно сказать что там.., никогда не скажу слово мертвец, ну линия она больше в математике как бы употребима*. Таким образом, конструирование референтов-слов и референтов-объектов в представлении их в рамках ЧАСТИ-ЦЕЛОГО и ТРАНСФОРМАЦИИ происходит сходным образом, кроме определенных различий, обусловленных, возможно, различиями в выборе миметических образ-схем.*

На третьем этапе мы провели анализ полимодального конструирования референтов. Результаты распределения образ-схем и функциональных жестов приведены в Таблице 3.

Из Таблицы 3 следует, что конструирование референтов с использованием разных образ-схем сопровождается в целом сходным распределением жестов. Проведенный анализ главных компонент показал, что индекс уникальности жестового распределения в отношении четырех образ-схем находится в диапазоне 0.00617–0.01492

Таблица 3. Распределение жестов с базовыми образ-схемами конструирования референтов

Образ-схемы базового уровня	Функциональные типы жеста				
	Дейктические	Репрезентирующие	Прагматические	Адаптеры	Всего
ЧАСТЬ-ЦЕЛОЕ	28 (13.7 %)	38 (18.6 %)	61 (29.9 %)	77 (37.7 %)	204 (20.2 %)
ТИП	17 (5.4 %)	61 (19.2 %)	102 (32.2 %)	137 (43.2 %)	317 (31.4 %)
ФУНКЦИОНИРОВАНИЕ	27 (7.1 %)	62 (16.4 %)	132 (34.8 %)	158 (41.7 %)	379 (37.5 %)
ТРАНСФОРМАЦИЯ	6 (5.4 %)	22 (19.8 %)	38 (34.2 %)	45 (40.5 %)	111 (11 %)
Всего					1011

(т. е. уникальность их жестового оформления очень низкая); однако относительно более высокая уникальность обнаружена в использовании схем ЧАСТЬ–ЦЕЛОЕ (0.01492) и ТРАНСФОРМАЦИЯ (0.01225), что может быть вызвано тем, что в речевом поведении участников наблюдались различия в выборе этих образ-схем при конструировании СЛОВ и ОБЪЕКТОВ, где образ-схема ЧАСТЬ–ЦЕЛОЕ показала более высокую частотность при конструировании СЛОВ, а ТРАНСФОРМАЦИЯ — при конструировании ОБЪЕКТОВ. Для проверки данной гипотезы обратимся к данным по распределению жестов относительно конструирования образ-схем ОБЪЕКТ и СЛОВО. В таблице 4 приведем данные абсолютных и относительных показателей.

Таблица 4. Распределение жестов с образ-схемами конструирования ОБЪЕКТА и СЛОВА

Образ-схемы высокого уровня абстракции	Функциональные типы жеста				
	Дейктические	Репрезентирующие	Прагматические	Адаптеры	Всего
ОБЪЕКТ (419)	52 (9.7 %)	96 (17.9 %)	150 (28.2 %)	237 (44.3 %)	535 (65.2 %)
СЛОВО (230)	17 (6 %)	45 (15.8 %)	83 (29.1 %)	140 (49.1 %)	285 (34.8 %)
Всего					820

Анализ на определение значимости различий в использовании жестов относительно использований образ-схем высокого уровня абстракции не показал наличие различий ($\chi^2 = 0.074$, $p = 0.786$), что свидетельствует о сходстве жестовой репрезентации данных образ-схем. Однако обращает на себя внимание то, что при конструировании ОБЪЕКТА наблюдается возрастание количества дейктических жестов, что в целом указывает на то, что конструирование ОБЪЕКТА происходит с усилением его пространственной локализации. Данное наблюдение позволяет уточнить результаты, описанные, например, в [Zdrzilova et al. 2018] в отношении использования жестов при конструировании ОБЪЕКТОВ. Как мы указали выше, в работе отмечается снижение частотности использования репрезентирующих жестов при конструировании объектов абстрактного типа. Однако, как мы обнаружили,

конструирование референта-слова, в любом случае менее воплощенного референта, чем объект, сопровождается в первую очередь ослаблением указания на этот референт, т. е. снижением частотности индексального указания на его пространственно локализуемый «образ», а не только снижением частотности иконического конструирования референта-слова с помощью репрезентирующих жестов. При этом наблюдается возрастание жестов-адаптеров, что свидетельствует об увеличении когнитивной нагрузки, связанной с поиском (экспозиторного) решения. Полученные результаты в определенной степени подтверждают мнение, высказанное в [Mittelberg 2018], о наличии определенной корреляции между семантикой жеста и семантикой образ-схем. При этом однозначных соответствий обнаружено не было, так как, например, дейктические жесты все же используются и при конструировании СЛОВА. Рассмотрим пример (Рис. 5).



*но конкретно вот большие
я как-то конкретно ерунду
встречала именно в-в-в...*

в контексте...

*ты что-то видишь
или читаешь*

Рис. 5. Конструирование СЛОВА с жестом-адаптером и дейктическими жестами

Пример на рисунке 5 иллюстрирует случай конструирования СЛОВА (ерунда) с помощью образ-схемы базового уровня ФУНКЦИОНИРОВАНИЕ (в ЭДЕ *но конкретно вот большие я как-то конкретно ерунду встречала именно в-в-в... в контексте*) и ЧАСТЬ–ЦЕЛОЕ (в ЭДЕ *ты что-то видишь или читаешь*). При использовании образ-схемы ФУНКЦИОНИРОВАНИЕ участница демонстрирует сначала жест-адаптер, который переходит в дейктический, когда она находит нужный вариант пояснения. В этом случае происходит локализация слова в контексте, т. е. участница указывает на место слова в потоке других слов; при этом задействуется миметическая схема ОЩУЩЕНИЕ. Далее при конструировании референта с использованием образ-схемы ЧАСТЬ–ЦЕЛОЕ и миметической схемы ФРАГМЕНТАЦИЯ участница уже определяет его как компонент синтаксического уровня в составе словосочетаний (*видишь ерунду, читаешь ерунду*) и указывает на него вновь с помощью дейктического жеста.

Однако намного чаще встречаются примеры использования дейктических жестов для конструирования ОБЪЕКТОВ. Рассмотрим пример на рисунке 6.



от одного человека



к другому человеку

Рис. 6. Конструирование ОБЪЕКТА с дейктическими жестами

В данном случае участница конструирует ОБЪЕКТ, используя тот же тип базовой образ-схемы, ФУНКЦИОНИРОВАНИЕ, что и в примере на рисунке 5. Поясняя слово *наказание*, она говорит, что наказание — это *такое что-то просто то что просто человек может наказать человека то есть это от одного человека к другому человеку*, указывая жестом на участников, вовлеченных в процесс перехода действия и актуализируя в речи миметическую схему ПЕРЕДВИЖЕНИЕ. Таким образом, мы наблюдаем сходные проявления полимодального конструирования, однако повторно отметим, что в целом для конструирования СЛОВ комбинация «дейктический жест и базовые образ-схемы» более редка. Возможно, однако, что указанные различия в целом связаны с выбором миметических образ-схем (которые могут быть более согласованно синхронизированы с жестами определенного типа) при представлении образ-схем базового уровня и конструировании референтов как СЛОВ и как ОБЪЕКТОВ — приведенные примеры позволяют высказать такое предположение, однако оно нуждается в более тщательной проработке.

Обнаруженные и представленные закономерности, конечно, описывают только изученный рабочий корпус примеров и помогают сформулировать некоторые общие тенденции распределения полимодального поведения. Однако мы надеемся, что и сама методика ранжирования типов образ-схем с учетом степени их воплощенности (с опорой на концепцию Дж. Грэди [Grady 2005]) позволит значительно расширить возможности анализа полимодального конструирования разных референтов в разных типах дискурса.

Заключительные замечания

Проведенный анализ полимодального поведения говорящих в экспозиторном дискурсе позволил установить следующее. Во-первых, даже при пояснении слов как единиц системы языка человек намного чаще (почти в два раза чаще) конструирует

их референты как объекты, чем как слова. При этом самыми частотными базовыми образ-схемами при конструировании и слов, и объектов являются ТИП и ФУНКЦИОНИРОВАНИЕ. Во-вторых, в целом сходное распределение образ-схем в речи, выбор типов жеста, распределение образ-схем в речи типов жеста свидетельствует в пользу того, что человек полимодально конструирует слова почти также, как он конструирует объекты, представляя их в аспекте их структуры, характеристик, функционирования и трансформации. В исследуемом материале не было обнаружено значимых различий в совместном использовании речи и жеста относительно анализируемых образ-схем, однако это может быть частично обусловлено тем, что некоторые поясняемые слова называли референты абстрактного типа (*наказание, кара, тьма, темнота, ерунда, чепуха*), что определенно могло повлиять на выбор образ-схем, но, как мы видим, почти не повлияло на распределение сопровождающих эти образ-схемы жесты. В-третьих, однако, наблюдается некоторая специфика в частотности базовых образ-схем ЧАСТЬ–ЦЕЛОЕ и ТРАНСФОРМАЦИЯ, а также в выборе дейктических жестов при конструировании слов и объектов, что в большей степени указывает на то, что слово часто рассматривается как часть (языковой) системы, а объект — как изменяющаяся во времени сущность, которая, даже демонстрируя абстрактный характер, имеет более локализованное расположение в пространстве, чем слово. Полученные результаты указывают на необходимость изучения полимодального мышления не только в отношении конструирования референтов конкретного и абстрактного типа, но и референтов, демонстрирующих разную степень воплощенности, определяемой «отрывом» от первичной картины мира и степенью выраженности их знаковой природы, что помогает проследить их роль в формировании первичной и вторичной картины мира.

Литература

Бергельсон М. Б. Социокультурная мотивированность нарративов // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2007». М.: Изд-во РГГУ, 2007. С. 38–43.

Гришина Е. А. Жесты и прагматические характеристики высказывания // Мультимодальная коммуникация: теоретические и эмпирические исследования / Под ред. О. В. Федоровой и А. А. Кибрика. М.: Буки Веди, 2014. С. 25–47.

Златев Й. Образные схемы, миметические схемы и жесты детей // Вопросы когнитивной лингвистики. 2013. № 1 (034). С. 14–31.

Кибрик А. А. Модус, жанр и другие параметры классификации дискурсов // Вопросы языкознания. 2009. № 2. С. 3–21.

Кибрик А. А., Подлесская В. И. Рассказы о сновидениях. Корпусное исследование устного русского дискурса. М.: ЯСК, 2009. 738 с.

Подлесская В. И. Нечеткая номинация в русской разговорной речи: опыт корпусного исследования // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (2013). Вып. 12 (19). М.: Изд-во РГГУ, 2013. С. 631–643.

Прокофьева О. Н. Когнитивные и прагматические факторы совместного употребления вербальных и жестовых единиц в описательном дискурсе. Дисс. ... канд. филол. наук. М.: МГЛУ, 2022.

Савчук С. О. Что вносит жестикуляционная лингвистика в изучение дискурса // Когнитивные исследования языка. 2021. № 1 (44). С. 213–220.

Cienki A. Image schemas and gesture // From Perception to Meaning: Image Schemas in Cognitive Linguistics / ed. by Beate Hampe, in cooperation with Joseph Grady. Berlin/ New York: Mouton de Gruyter. 2005. P. 421–442.

Cienki A. Metaphoric gestures and some of their relations to verbal metaphoric expressions // Discourse and cognition: Bridging the gap / ed. by J. P. Koenig. Stanford: CSLI. 1998a. P. 189–204.

Cienki A. STRAIGHT: An image schema and its metaphorical extensions // Cognitive Linguistics. 1998b. № 9(2). P. 107–149.

Forceville C., Jeulik M. The flesh and blood of embodied understanding: The source-path-goal schema in animation film // Pragmatics and Cognition. 2011. № 19 (1). P. 37–59.

Grady J. Image schemas and perception: Refining a definition // From Perception to Meaning. Image Schemas in Cognitive Linguistics. Ed. by B. Hampe, J. Grady. Berlin, New York: Mouton de Gruyter. 2005. P. 35–56.

Iriskhanova O., Cienki A. The semiotics of gestures in cognitive linguistics: Contributions and challenges // Issues in Cognitive Linguistics. 2018. № 4. P. 25–36.

Iriskhanova O. K., Kiose M. I., Leonteva A. V., Agafonova O. V. Vague reference in expository discourse: multimodal regularities of speech and gesture // Computational Linguistics and Intellectual Technologies Papers from the Annual International Conference “Dialogue” (2023). 2023. V. 22. P. 172–180.

Jackendoff R. Foundations of language: Brain, meaning, grammar, evolution. Oxford: Oxford University Press, 2002.

Johnson M. The body in the mind: the bodily basis of meaning, reason and imagination. Chicago: University of Chicago Press, 1998.

Ladewig S. H. Putting the cyclic gesture on a cognitive basis // CogniTexts. 2011. № 6.

Longacre R. The grammar of discourse. New York, Plenum, 1983. 423 p.

McNeill D. Gesture-speech unity. What it is, where it came from // Why gesture? How the hands function in speaking, thinking and communicating / ed. by S. D. Kelly, R. B. Church, M. W. Alibali. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2017. P. 77–101.

Mittelberg I. Gestures as image schemas and force gestalts: A dynamic systems approach augmented with motion-capture data analyses // Cognitive Semiotics, 2018. P. 1–21.

Müller C. Gestural Modes of Representation as techniques of depiction. // Body — Language — Communication: An international Handbook on Multimodality in Human Interaction / ed. by C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, and J. Bressen. Berlin, Germany: De Gruyter Mouton, 2014. P. 1687–1702.

Peña S. Subsidiary relationships between image-schemas: an approach to the force schema // Journal of English studies. 1999. № 1. P. 187–207.

Pustejovsky J. The Generative Lexicon. Cambridge, Massachusetts: MIT Press, 1995.

Williams R. F. Gesture as a conceptual mapping tool // *Metaphor and gesture* / ed. by A. Cienki & C. Müller. Amsterdam & Philadelphia: John Benjamins, 2008. P. 55–92.

Zdrzilova L., Sidhu D. M., Pexman P. M. Communicating abstract meaning: concepts revealed in words and gestures // *Philosophical Transactions of the Royal Society B*. 2018, № 373 (1752): 20170138.

Zlatev J. What's in a schema? Bodily mimesis and the grounding of language // *From perception to meaning: Image schemas in cognitive linguistics* / ed. by B. Hampe. Berlin: Mouton de Gruyter, 2005. P. 313–343.

Zlatev J. Intersubjectivity, mimetic schemas and the emergence of language // *Intellectica*. 2007. № 46/47. P. 123–152.

M. I. Kiose¹, A. V. Leonteva², O. V. Agafonova³

Moscow State Linguistic University

Institute of Linguistics of the Russian Academy of Sciences

(Russia, Moscow)

maria_kiose@mail.ru¹, lentevanja27@gmail.com², olga.agafonova92@gmail.com³

MULTIMODAL CONSTRUAL OF REFERENTS IN EXPOSITORY DISCOURSE: WORD VS. OBJECT

The article explores speech and gesture behavior in explaining the differences between close synonyms to identify the peculiar features in multimodal construal of two types of referents — words and objects. To proceed, we adopt the method of image schemas which are viewed within a three-level conceptual structure comprising mimetic / embodied schemas, basic-level schemas and high-level abstract schemas — here WORD and OBJECT. The three-level theory of image schemas presented by J. Grady was further elaborated to accommodate the types of mimetic schemas identified by J. Zlatev and four qualia-roles described by J. Pustejovsky corresponding to basic-level schemas PART-WHOLE, TYPE, FUNCTIONING, TRANSFORMATION. The developed framework was used to explore the alignment of basic-level schemas in speech patterns and functional gesture types, deictic, representational, pragmatic and adapters. The results show that the distribution of basic image schemas as well as their alignment with gestures are mediated by the choice of high-level abstract schemas WORD and OBJECT. Significant differences are found in the basic-level schema PART-WHOLE which prevails in the construal of referent as an object and the schema TRANSFORMATION which prevails in the construal of referent as a word. Additionally, we observe the differences in the use of deictic gestures in the construal of referent as an object. Overall, the results suffice to claim that the multimodal construal of referents as objects and words is modulated by the differences in mimetic image schemas which underlie them.

Keywords: multimodal behaviour, expository discourse, image schema, word, object, mimetic image schema.

References

- Bergel'son M. B. [Sociocultural Motivation in Narratives]. *Komp'yuternaya lingvistika i intellektual'nyeologii. Trudy mezhdunarodnoi konferentsii "Dialog-2007"* [Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue-2007"]. Issue 6 (13). Moscow, RSUH Publ., 2007, pp. 38–43 (In Russ.)
- Cienki A. Image schemas and gesture. *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. Ed. by Beate Hampe (ed.), in cooperation with Joseph Grady. Berlin/New York: Mouton de Gruyter, 2005, pp. 421–442.
- Cienki A. Metaphoric gestures and some of their relations to verbal metaphoric expressions. *Discourse and cognition: Bridging the gap*. Ed. by J. P. Koenig. Stanford: CSLI, 1998a, pp. 189–204.
- Cienki A. STRAIGHT: An image schema and its metaphorical extensions. *Cognitive Linguistics*, 1998b, no. 9 (2), pp. 107–149.
- Forceville C., Jeulik M. The flesh and blood of embodied understanding: The source-path-goal schema in animation film. *Pragmatics and Cognition*, 2011, no. 19 (1), pp. 37–59.
- Grady J. Image schemas and perception: Refining a definition. *From Perception to Meaning. Image Schemas in Cognitive Linguistics*. Ed. by B. Hampe, J. Grady. Berlin, New York: Mouton de Gruyter, 2005, pp. 35–56.
- Grishina E. A. [Gestures and pragmatic characteristics of the utterance]. *Mul'timodal'naya kommunikatsiya: teoreticheskie i empiricheskie issledovaniya* [Multimodal communications: theoretical and empirical research]. Ed. by A. A. Kibrik, O. V. Fedorova. Moscow: Buki-Vedi Publ., 2014, pp. 25–47. (In Russ.)
- Iriskhanova O., Cienki A. The semiotics of gestures in cognitive linguistics: Contributions and challenges. *Issues in Cognitive Linguistics*, 2018, no. 4, pp. 25–36.
- Iriskhanova O. K., Kiose M. I., Leonteva A. V., Agafonova O. V. Vague reference in expository discourse: multimodal regularities of speech and gesture. *Computational Linguistics and Intellectual Technologies Papers from the Annual International Conference "Dialogue" (2023)*, 22, 2023, pp. 172–180.
- Jackendoff R. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press, 2002.
- Johnson M. *The body in the mind: the bodily basis of meaning, reason and imagination*. Chicago: University of Chicago Press, 1998.
- Kibrik A. A. [Mode, genre and other parameters of discourse classification]. *Voprosy yazykoznaniya*. 2009, no. 2, pp. 3–21. (In Russ.)
- Kibrik A. A. Podlesskaya, V. I. *Rasskazy o snovideniyakh. Korpusnoe issledovanie ustnogo russkogo diskursa* [Night Dream Stories. A corpus study of spoken Russian discourse]. Moscow: Languages of Slavic Culture, 2009. (In Russ.)
- Ladewig S. H. Putting the cyclic gesture on a cognitive basis. *CogniTexts*, 2011, no. 6.
- Longacre R. *The grammar of discourse*. New York, Plenum, 1983.
- McNeill D. Gesture-speech unity. What it is, where it came from. *Why gesture? How the hands function in speaking, thinking and communicating*. Ed. by S. D. Kelly,

R. B. Church, M. W. Alibali. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2017, pp. 77–101.

Mittelberg I. Gestures as image schemas and force gestalten: A dynamic systems approach augmented with motion-capture data analyses. *Cognitive Semiotics*, 2018, pp. 1–21.

Müller C. Gestural modes of representation as techniques of depiction. In (eds.), *Body — Language — Communication: An international Handbook on Multimodality in Human Interaction*. Ed. by C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, and J. Bressen. Berlin, Germany: De Gruyter Mouton, 2014, pp. 1687–1702.

Peña S. Subsidiary relationships between image-schemas: an approach to the force schema. *Journal of English studies*, 1999, no. 1, pp. 187–207.

Podlesskaya V. [Vague names in Russian speech: a corpus study]. *Komp'yuternaya lingvistika i intellektual'nyeologii (Mezhdunarodnaya konferenciya "Dialog 2012"* [Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue" (2012)]. Issue 12 (19). Moscow, RSUH Publ., 2013, pp. 631–643. (In Russ.)

Prokofyeva O. N. *Kognitivnye i pragmaticheskie faktory sovmestnogo upotrebleniya verbal'nykh i zhestovykh edinits v opisatel'nom diskurse* [Cognitive and pragmatic factors of speech and gesture co-occurrence in descriptive discourse]. PhD thesis (Linguistics). Moscow, MSLU, 2022. (In Russ.)

Pustejovsky J. *The Generative Lexicon*. Cambridge, Massachusetts: MIT Press, 1995.

Savchuk S. O. [The contribution of gestural linguistics in the study of discourse]. *Kognitivnye issledovaniya yazyka* [Cognitive studies of language], 2021, no. 1 (44), pp. 213–220. (In Russ.)

Williams R. F. Gesture as a conceptual mapping tool. *Metaphor and gesture*. Ed. by A. Cienki & C. Müller. Amsterdam & Philadelphia: John Benjamins, 2008, pp. 55–92.

Zdrzilova L., Sidhu D. M., Pexman P. M. Communicating abstract meaning: concepts revealed in words and gestures. *Philosophical Transactions of the Royal Society B*, 2018, no. 373 (1752): 20170138.

Zlatev J. What's in a schema? Bodily mimesis and the grounding of language. *From perception to meaning: Image schemas in cognitive linguistics*. Ed. by B. Hampe. Berlin: Mouton de Gruyter, 2005, pp. 313–343.

Zlatev J. Intersubjectivity, mimetic schemas and the emergence of language. *Intellectica*, 2007, no. 46/47, pp. 123–152.

Zlatev J. Obraznye skhemy, mimeticheskie skhemy i zhesty detei [Schemas, Mimetic Schemas and Children's Gestures]. *Voprosy kognitivnoi lingvistiki* [Issues of Cognitive Linguistics], 2013, no. 1 (034), pp. 14–31. (In Russ.).

С. И. Переверзева

Российский государственный гуманитарный университет

(Россия, Москва)

P_Sveta@hotmail.com

НЕВЕРБАЛЬНАЯ МАНИПУЛЯЦИЯ И ПРОБЛЕМЫ ЕЁ ВЫЯВЛЕНИЯ: ЖЕСТ «ПОЛОЖИТЬ (СВОЮ) РУКУ НА (ЧЬЮ-ЛИБО) РУКУ»¹

В работе на примере тактильного жеста **положить (свою) руку на (чью-либо) руку** обсуждается понятие невербальной манипуляции. На материале русских художественных произведений XIX–XXI вв. анализируются значения и употребления этого жеста. Выделяются две основные группы его употреблений: 1) исполнитель жеста показывает, что испытывает тёплые чувства к адресату, сообщает ему какую-то важную информацию или даёт понять, что говорит искренне; 2) исполнитель побуждает адресата совершить некоторые действия или измениться внутренне (соблазнение, настойчивая просьба, попытка остановить слова или действия адресата, допрос, успокоение, убеждение). В ситуациях, представленных второй группой употреблений, невербальная манипуляция более вероятна. Однако, на наш взгляд, по одному только жесту невозможно с уверенностью судить о манипулятивном характере коммуникации, поскольку рассматриваемый жест (а) допускает совмещение разных значений и функций в одном контексте; (б) может быть неверно интерпретирован адресатом; (в) может выполнять одну и ту же функцию в разных коммуникативных актах. Тем самым, о манипулятивном характере диалога позволяют судить не столько семантические и функциональные особенности того или иного невербального знака, сколько конкретная ситуация его употребления.

Ключевые слова: жест, манипуляция, многозначность, коммуникативная цель, диалог.

Введение

Термин «невербальная манипуляция» в последние десятилетия широко используется как в научной, так и в научно-популярной литературе. О невербальной манипуляции говорят применительно к таким сферам коммуникативной деятельности,

¹ Автор выражает благодарность И. Б. Иткину и анонимному рецензенту за внимательное прочтение работы и ценные замечания.

как преподавание, воспитание, общение детей со взрослыми, флирт, маркетинг, медиакommunikации, деловые переговоры. Обзор российских и зарубежных работ в области психологии, в которых данное понятие является ключевым, дан в статье [Курагин 2021]. Однако, насколько нам известно, до сих пор не предпринималось системного исследования невербальных знаков как потенциальных инструментов для манипулирования. Какие особенности жестов, поз, взглядов и мимики делают их пригодными для маскировки коммуникативных намерений говорящего, а какие — нет? Существуют ли такие признаки невербального поведения (*nonverbal cues*), по которым можно с уверенностью определить подлинные цели участников диалога? Эти и другие проблемы изучались в рамках коммуникативной ситуации обмана (см. подробный обзор в работе [Vrij 2006]); в частности, было обнаружено, что универсального невербального знака, который однозначно выдавал бы лжеца, не существует — хотя бы потому, что само понятие лжи не имеет чётких границ.

В настоящей работе мы попытаемся дать самые предварительные ответы на указанные вопросы применительно к манипуляции. Вначале мы кратко остановимся на понятии «манипуляция» и его ключевых признаках. В следующих разделах будут описаны основные значения и употребления тактильного жеста **положить (свою) руку на (чью-либо) руку** (далее для краткости — **положить руку на руку**). Наконец, будут рассмотрены проблемы интерпретации данного жеста, возникающие по причине его многозначности, и его соотносённость как с побудительными, так и с непобудительными речевыми актами. Будет показано, что о манипулятивности жеста **положить руку на руку** можно судить только с учётом этих двух факторов, а также конкретной ситуации его употребления.

Предметом нашего исследования служат главным образом языковые номинации вида *X положил (свою) [руку / ладонь] на [руку / ладонь / пальцы / кисть / кулак / запястье] Y-а*; номинация части тела может быть в ед. ч. и мн. ч.; место глагола *положить* может занимать парный к нему по виду глагол *класть*, а также глагол *накрыть*. В качестве материала мы используем тексты русских художественных произведений XIX–XXI вв. (основной подкорпус Национального корпуса русского языка) и применяем метод реконструкции жеста по его обозначениям в текстах; наша выборка содержит около 150 примеров².

Замечание.

Интересно, что в случае, когда человек *X* *накрывает* своей *рукой* или, чаще, *ладонью* *руку* (*ладонь, кулак, пальцы*) человека *Y*, *X* и *Y* обычно разного пола (*X* — мужчина, *Y* — женщина — 17 примеров; *X* — женщина, *Y* — мужчина — 9 примеров; *X* и *Y* — мужчины — 1 пример; всего 27 примеров). Для более частотного в этой конструкции глагола *положить* такого явного контраста не наблюдается (*X* — мужчина, *Y* — женщина — 31 пример; *X* — женщина, *Y* — мужчина — 36 примеров; *X* и *Y* — мужчины — 29 примеров; *X* и *Y* — женщины — 10 примеров; всего 106 примеров).

² Об этом методе см., например, доклад [Крейдли 2011].

Признаки манипуляции

Психологи, социологи, лингвисты и представители других гуманитарных наук по-разному определяют термин «манипуляция». Тем не менее многие соглашаются с тем, что при успешной манипуляции манипулируемый делает нечто не потому, что этого изначально хотел он сам, а потому, что его побудил к этому манипулятор; при этом манипулируемый не осознаёт, что его поступки подчинены воле манипулятора, а не его собственной воле³. Один из важных вопросов, на который исследователи манипуляции дают разные ответы, это вопрос о том, может ли манипуляция быть во благо манипулируемому. Одни считают, что манипуляция обязательно должна противоречить его интересам (см., например, [Van Dijk 2006]). Другие понимают данный термин более широко и выделяют такую разновидность манипулятивных действий, которые не нарушают интересов манипулируемого, — например, родители обманом заставляют ребёнка принять лекарство, добавляя его в сладкую еду [Chatti 2017: 49]. При втором подходе различаются манипуляция без цели личной выгоды (позитивное воздействие на собеседника) и манипуляция с целью личной выгоды [Непряхин 2018: 33].

Значения и употребления жеста **положить руку на руку** можно с некоторыми оговорками разделить на две группы:

- 1) исполнитель показывает, что испытывает некие чувства к адресату, или передаёт ему некоторую информацию (ситуации, в которых невербальная манипуляция менее вероятна);
- 2) исполнитель побуждает адресата совершить некоторые действия или измениться внутренне (ситуации, в которых невербальная манипуляция более вероятна).

Ниже мы рассмотрим каждую из этих двух групп.

1. Первая группа употреблений жеста «X положил свою руку на руку Y-а»: «X показывает Y-у, что...»

1а. «...X испытывает к Y-у тёплые чувства»

Такая ситуация весьма частотна: она представлена не менее чем в четверти найденных нами примеров. Тёплые чувства X-а по отношению к Y-у могут быть самыми разными: симпатия, привязанность, влечение, дружба, любовь и под. Ср.:

- (1) *Он тихонько накрыл своей рукой Наденькину руку, как будто в полузабытии, и закрыл глаза. Наденька свободной рукой вытащила градусник.*

³ Ср. описание английских конструкций *Person X verb-ed person Y into doing Z* и *Person X got person Y to do Z* в книге А. Вежбицкой [Wierzbicka 2006]. Первую конструкцию А. Вежбицкая называет манипулятивной (*manipulating someone “into doing something”*) и перечисляет те же отличительные признаки, что мы привели выше: “the causee originally didn’t want to do what he or she did”; “the causee’s action is “triggered” by the causer’s will, not by the causee’s own will”; “the causee is unaware of what is happening (namely, that his or her action is “triggered” by the causer’s will)” [Wierzbicka 2006: 179].

Градусник показывал 36 и 8. Наденька положила градусник на стол: потянулась, не меняя позы, чтоб не дернуть руки, которую накрыл Башкин. [Б. С. Житков. Виктор Вавич. Книга первая (1926–1931)]

- (2) Мой собеседник через стол на мгновение **положил** свою прыгающую **руку на мою** — теплым, дружеским жестом. — Вам в самом деле не надоело еще слушать? [И. Грекова. Хозяева жизни (1960)]
- (3) Сажу в автобусе, мчащемся по какой-то стране, которая называется Эстония. <...> Рядом Софуля, она **положила свою руку на мою**, лежащую на сидении, душе теплее. [Татьяна Окуневская. Татьянин день (1998)]

Последние два примера показывают, что авторы текстов могут эксплицитно обозначать те чувства, которые выражает исполнитель жеста по отношению к адресату или которые возникают у адресата в ответ (*тёплый, дружеский жест; душе теплее*).

Жест **положить руку на руку** часто используется в ситуации романтических отношений между X-ом и Y-ом, ср.:

- (4) Света задала какой-то вопрос, Андрей на него ответил. Света взяла его за руку, Андрей **положил свою руку на её**. <...> Впереди у них были еще два года, но общего будущего у них не было. [А. П. Краснящих. О себе // «Волга», 2015]
- (5) Они сидели за маленьким столиком далеко от оркестра. Тата, розовая от вина и танцев, **положила руку на пальцы Мстислава**, слабо двигавшиеся в такт ритмическим синкопам. — Вам хорошо, Мстислав? — Очень! С вами! [И. А. Ефремов. Лезвие бритвы (1959–1963)]
- (6) «...А я, дурак, за всеми девчонками бегал по очереди, а Тоньку не видел. <...> Дурак я, дурак!» И он уверенно **положил свою руку на трепещущую руку Тони**. [В. К. Кетлинская. Мужество (1934–1938)]

Смысловой компонент ‘X испытывает к Y-у тёплые чувства’ позволяет отнести рассматриваемый нами жест в данных контекстах к классу жестов ласки (ср. семантический компонент ‘I feel something good towards you now’, который А. Вежбицкая вводит в толкование жестов (*spontaneous*) *kiss*, а *hug* и *объятие* (*an embrace*) [Wierzbicka 1995: 221–224]).

Особое место в группе (1a) занимают невербальные выражения сочувствия или сострадания, ср.:

- (7) «Да это он мучил Веру Павловну, он отбросил ее» [, —] опять мелькнуло ей. Она встала, приблизилась к Вере Павловне, **положила свою руку на похолодевшие концы пальцев ее руки** и, поцеловав в голову, быстро отошла от плеча... [М. К. Иогель. Между вечностью и минутой // «Русская Мысль», 1880]
- (8) — Бедная сестра моя... — прошептала Магда, — я знаю... <...> Магда **положила** покрытую цыпками, шершавую **руку на худую руку Вилмы** и осторожно пожала её. Это пожатие было так нежно, что Вилма почувствовала,

как успокоение подобно тёплому току проникает в каждую клетку её тела. Она откинулась на подушку и заплакала. [Н. Н. Шпанов. Ученик чародея (1935–1950)]

- (9) Мальчик ухватился за перемычку между стеклами и пристроился снаружи. Генерал погнался за машиной по петляющей дороге. <...> Травянистые обочины и разъезженные колеи то проваливались куда-то вниз, то взлетали до самых стекол, и тогда коленка мальчика бороздила грунт на уровне его подбородка. Мне надоел ледяной комок где-то в грудной клетке <...>, и я **положил свою руку на белые запястья мальчика**, вцепившегося в перемычку. Но генерал, даже не покосившись, сказал: — Отставить. И я убрал руку. [Михаил Анчаров. Золотой дождь (1965)]

Представляется, что при проявлении сочувствия или сострадания тактильный контакт не только выражает тёплые чувства исполнителя жеста по отношению к адресату, но и указывает на их единение. О жесте **hand-holding** (≈ **держат друг друга за руки**), близком к жесту **положить руку на руку** и по смыслу, и по физической реализации, А. Вежбицкая писала, что в нём «две соединённые руки передают идею единства»; они «как бы сливаются в одно целое», в силу чего в толкование **hand-holding** следует ввести компонент ‘Я хочу, чтобы ты и я были как две части чего-то одного’ [Wierzbicka 1995: 237]⁴. На наш взгляд, сказанное справедливо и для жеста **положить руку на руку** в данном круге его употреблений. Таким образом исполнитель жеста даёт понять адресату, что он разделяет его желания, цели и ценности (ср. первую стратегию позитивной вежливости в книге [Brown, Levinson 1987: 103]).

16. ‘...Х говорит с Y-ом серьёзно и искренне’

В данном круге употреблений жест **положить руку на руку** увеличивает интенсивность иллюкутивной силы высказывания, подчеркивая искренность речевого акта. Здесь мы рассмотрим только примеры непобудительных речевых актов; о побудительных (перлокутивных) речевых актах см. ниже.

Жест, сопровождающий вопрос (≈ ‘мне действительно важно знать ответ’):

- (10) М. И. Свободина <...>, **положив свою руку на мою**, пытливо посмотрела на меня своими прекрасными темно-кариыми глазами: — Обращаю к вам вопросом первую строчку стихов: «Скажи мне, ты любил на родине своей?...» [Вл. А. Гиляровский. Люди театра (1932–1935)]
- (11) Мать смотрела на него пристально и как-то по-новому, изучающе, — по-видимому, он ее удивил... — Сын, а эта женщина... — Несмелым движением **положила руку на его руку**. — Она существует? [Ю. В. Трифонов. Время и место (1980)]

⁴ Ср. оригинальный текст: “Two joint hands convey a message of unity ... merge into a single quasi-entity...”; ‘I want you and me to be like two parts of one thing’.

Жест, сопровождающий благодарность ($\approx$ ‘мне действительно важно, что ты сделал для меня что-то хорошее’):

- (12) Леонид **положил руку на Катину руку** и крепко пожал ее сверху. — Спасибо тебе, Катюшка! Кабы не ты сегодня, кормить бы мне собою крымских ваших червей... [В. В. Вересаев. В тупике (1920–1923)]

Жест, сопровождающий одобрение ($\approx$ ‘мне действительно понравилось то, что ты сделал или сказал’):

- (13) — ...Смерть идет всегда по линии наименьшего сопротивления. Грач быстро **положил свою руку на руку доктора**: — Это вы очень хорошо сказали, очень верно. [С. Д. Мстиславский. Грач — птица весенняя (1937)]

В репертуаре русского бытового языка тела существуют и другие жесты, подчёркивающие искренность сопутствующего им речевого акта. Это, например, жест **бить себя в грудь**, исполнитель которого «показывает, что он ручается собой, своей жизнью и здоровьем за истинность отстаиваемой им точки зрения» [СЯРЖ 2001: 37], и жест **прижать руку к груди**, который (в значении благодарности) показывает, что жестикулирующий «не просто благодарит адресата, а глубоко тронут его действиями или словами» [там же: 121].

1в. ‘...то, что X говорит или сейчас скажет Y-у, — важно’

Прикосновение к руке собеседника может служить для привлечения его внимания к тому, что сейчас делает или собирается сделать исполнитель жеста (как правило, это действие речевое, и внимание адресата, таким образом, привлекается к смыслу слов). В этой ситуации существенно, что непосредственно перед совершением жеста руки X-а не касались Y-а (X не вёл Y-а за руку, не обнимал его и т. п.). Приведём несколько примеров:

- (14) — ...Хотите знать, что я больше всего люблю? — Свободу, — подсказал Санин. Марья Николаевна **положила руку на его руку**. — Да, Дмитрий Павлович, — промолвила она, и голос ее прозвучал чем-то особенным, какой-то несомненной искренностью и важностью, — свободу, больше всего и прежде всего. [И. С. Тургенев. Вешние воды (1872)]
- (15) — Тебе надо идти? Я сел рядом и поглядел на часы. — Еще есть время. Софи **положила ладонь на мою руку**, и я понял, что она хочет поговорить о чем-то важном. [Виктор Пелевин. Бэтман Аполло (2013)]
- (16) — Юлька, — сказал он, **бережно накрывая рукой ее маленькую руку**, — у меня, кажется, тоже надвигаются перемены. Выражение ее лица изменилось мгновенно, рассеянной задумчивости как не бывало, она подалась к нему: — Какие? — Предполагаю, что **важные**. [В. Ф. Панова. Времена года. Из летописей города Энска (1953)]
- (17) — Что же, ваш Воронов своим колокольным звоном хочет весь народ, что ли, сделать счастливым? — спросил, наконец, Федор, неожиданно

*и ласково **накрывая своей холодной рукой Яшину руку.** — Нет, Яша! Народ по деревням с голоду пухнет, и ему не до колокольного звона. Голодному нужен хлеб, а обиженному — правда.* [Н. Д. Телешов. Крамола (1906)]

Круг употреблений (1в) близок рассмотренному выше (1б), но если в случае (1б) жест как бы усиливает речевой акт, то в случае (1в) X, как правило, сообщает Y-у нечто и при помощи жеста побуждает Y-а обратить особое внимание на смысл его слов. Эта особенность объединяет жест **положить руку на руку** с другими тактильными жестами, инициирующими контакт исполнителя с адресатом и служащими для привлечения внимания адресата к чему-либо, ср. **постучать по плечу, толкнуть в бок** (описание этого жеста см. в [СЯРЖ 2001: 138–140]), **толкнуть ногой, подёргать за одежду** и др.

2. Вторая группа употреблений жеста

«X положил свою руку на руку Y»:

‘X побуждает Y совершить некоторые действия
или измениться внутренне’

Во многих ситуациях из группы (2) жест **положить руку на руку** служит для усиления некоего вербального побуждения — например, просьбы (2б), призыва что-либо рассказать или, наоборот, замолчать (2в, 2г). Может показаться, что такие примеры противоречат данному выше определению манипуляции, согласно которому манипулируемый (Y) не осознает, что делает нечто по воле манипулятора (X): ведь Y, соглашаясь сделать то, о чем говорит X, очевидно понимает, что действует в соответствии с его желаниями. Если это верно, то никакой побудительный речевой акт не может сопутствовать манипуляции просто в силу своей природы. Однако, на наш взгляд, здесь важно, что Y может сам решать, соглашаться ему или нет, и манипуляция со стороны X-а заключается в том, что X незаметно для Y-а уменьшает вероятность отказа. Таким образом, в подобных ситуациях Y знает, что действует по воле X-а, но не знает, что его готовность выполнить желание X-а также в какой-то мере подчинена воле X-а.

На то, что тактильные жесты могут иметь перлокутивный эффект, ранее обращали внимание исследователи прикосновений, ср. «Легкие нежные прикосновения, активирующие рецепторы С-тактильной системы, <...> могут осознанно или неосознанно использоваться говорящим для усиления коммуникативного воздействия всего высказывания или ряда предшествующих высказываний» [Варламов и др. 2020: 88]; см. также исследование, показавшее, что посетители ресторана, руки которых касался официант, подавая им меню, чаще следовали его советам по выбору блюда и оставляли более щедрые чаевые [Guéguen, Jacob, Boulbry 2007].

Ниже мы рассмотрим несколько групп таких жестов. Первые две группы, на наш взгляд, уместно соотнести с первыми двумя группами в разделе 1: граница между ними представляется нам довольно условной.

2а. X соблазняет Y-а (ср. 1а '...X испытывает к Y-у тёплые чувства')

В данном круге употреблений жест **положить руку на руку** нередко сопровождается другими жестами ухаживания (взгляд в глаза, пожатие или поглаживание руки и др.):

- (18) ...Ольга Петровна выбирала из ящика фотографические карточки, показывала их Митеньке, и они оба, склонившись головами, рассматривали. <...> — А почему вы закрыли эту карточку? — спросил Митенька, **положив свою руку на ее, и, глядя ей в глаза взглядом, показавшим, что карточка только предлог, тихо, ласково сжал ее руку.** — Нельзя... — сказала молодая женщина, на секунду оставив свою руку в его руке, потом быстро отняв ее. [П. С. Романов. Русь (1922)]
- (19) — «Ничего, милый», — она говорит, но как говорит, тем своим глубоким грудным голосом, полным ангельской нежности, которым сводила меня с ума с первого дня знакомства. И **положила свою руку на мою, и нежно стиснула.** <...> «Поцелуй меня, — говорит. — Пожалуйста». [Эдуард Володарский. Дневник самоубийцы (1997)]
- (20) — Вы не старик. — Франсуаз **положила ладонь на руку Шаллера и погладила его крепкие пальцы.** — И давайте не будем долго ухаживать друг за другом. Я вам нравлюсь, вы мне тоже. [Дмитрий Липскеров. Сорок лет Чанчжоз (1996)]
- (21) Страничка кончилась, а когда Ава потянулась за следующей, готовая читать и читать, Тарас **накрыл своей прохладной ладонью ее руку, мягко сжал и, встретившись взглядом с испуганными, покорными глазами, приобнял** Аву и, не говоря ни слова, повлек в спальню. [Ольга Новикова. Мужской роман (1999)]

Однако иногда встречаются и такие контексты, где данному жесту не сопутствуют никакие другие:

- (22) Когда они подъехали к ее дому, она, **положив свою руку на его, так просто предложила зайти к ней, что он не устоял.** [В. А. Ярмолинец. Подрезка (2007) // «Волга», 2008]

2б. X показывает Y-у, что X действительно хочет, чтобы Y нечто сделал (ср. 1б 'X говорит с Y-ом серьезно и искренне')

В качестве побудительного речевого акта, усиливаемого данным жестом, чаще всего выступает просьба или даже мольба, ср.:

- (23) — Вам надо ехать, [—] произнесла она тихо, [—] но чем скорее, тем лучше. <...> Она **положила** свою морщинистую **руку на мою руку** и преврала меня: — Без горечи, без упреков, — к чему они? Я прошу вас уехать и буду говорить с вами откровенно. [Е. В. Салиас-де-Турнемир (Евгения Тур). Старушка // «Русский Вестник», 1856]

- (24) — Я без тебя буду помогать Игорю Васильевичу. Тима **мягко положил свою ладонь на ее руку**, сжимающую руль: — Ради Бога, не предпринимай без меня ничего. [Лев Дворецкий. Шакалы (2000)]
- (25) Она тоже подкатила кресло в угол, уселась, **положив руку на руку своего Рема Степановича**, шепнула: — Милый, расхмурься. [Лазарь Карелин. Последний переулочек (1983)]
- (26) — Нельзя ли переменить разговор? Всё безумие, сумасшествие, самоубийство... Я не могу больше этого слышать!.. <...> Нарочно, как только я войду, всё об одном и том же. Я три раза уходила. Можно иметь пощаду к человеку. <...> Ну, а я прошу перестать. Это безжалостно. Посмотри мои руки (обратилась она ко Л. Н. и **положила свою руку на его**), я три раза уходила, сидела на балконе, а вы всё нарочно об одном и том же. [А. Б. Гольденвейзер. Вблизи Толстого (1910)]

Последний пример интересен тем, что поводом для данного жеста служит желание говорящей (Софии Андреевны) показать, по-видимому, в какое состояние пришли её руки, и с этой точки зрения жест является дейктическим. Именно об этом желании сообщает и она сама. Однако у её жеста есть и другая, перлокутивная функция — воздействовать на адресата (Льва Николаевича), побудить его исполнить её просьбу. Эта функция как бы замаскирована первой (дейктической), что, как кажется, позволяет рассматривать прикосновение Софии Андреевны как манипулятивный приём.

Иногда жест **положить руку на руку** употребляется при других побудительных речевых актах, но таких случаев в нашем корпусе крайне мало. Приведём два примера совета:

- (27) Скиба **положил свою руку на руку мужика**: — Мой вам добрый совет. Пока молчите. А то вас заберут в тот момент, когда мы будем нужны. [В. Лихоносов. Ненаписанные воспоминания. Наш маленький Париж. Ч. 3–4 (1983)]
- (28) — Послухай меня... — Он склонился к ней и **положил свою плотную ладонь на тонкие пальцы ее руки и сжал их**. — Ты при корпусе не оставайся, делать тебе там нечего... [А. А. Фадеев. Молодая гвардия (1943–1951)]

2в. X побуждает Y-а прекратить делать нечто

Чаще всего X побуждает Y-а перестать говорить. В этой функции жест **положить руку на руку** выступает как регулятивный (о жестах-регуляторах, т. е. жестах, регулирующих ход коммуникации, см. [Крейдлин 2002: 81]):

- (29) Но когда Плещеев вдруг заговорил повышенным тоном: — Что ж это делалось, что делалось?.. — Шалагин **положил руку ему на руку, остановил**: — Потом. Не хочу об этом говорить с кондачка. Подумавши хочу говорить. [В. Ф. Панова. Рабочий посёлок (1964)]

- (30) Я вздрогнула и хотела перебить его, сказать, чтоб он не смел говорить за меня, но он, **удерживая меня, положил свою руку на мою**. [Л. Н. Толстой. Семейное счастье (1859)]
- (31) — ...Можешь ли ты сомневаться, что только по бедности мысли // Большой частью во тьме совершается жизнь человека? И, опять усталый, он закрыл глаза. — Филет, ты **не говори**, — **положил ему руку на его холодеющие руки Язон**. — Отдохни. Мы помолчим с тобой... [И. Ф. Наживин. Иудей (1933)]

См. также пример (23): ...Она **положила** свою морщинистую **руку на мою руку и прервала меня**...

Однако в некоторых случаях Y, останавливаемый X-ом, не говорит, а действует той рукой, на которую кладёт свою руку X:

- (32) Время — это что-то вроде машинки уничтожения. «Настольная гильотинка, если хотите. <...> Каждой секунде отрезают голову. <...> Дело художника — **остановить руку того, кто крутит эту машинку. Положить свою руку на его**». [Михаил Шишкин. Венерин волос (2004) // «Знамя», 2005]

См. также пример (21), где Ава протягивает руку, чтобы взять новую страницу текста, но Тарас останавливает её.

2г. X побуждает Y-а рассказать ему нечто

Типичная ситуация, в которой встречается жест **положить руку на руку** в данном значении, — допрос. В ряде работ отмечалось, что единицы невербального кода, посредством которых следователь даёт понять своему собеседнику, что он его внимательно слушает и сочувствует ему, играют важную роль для установления доверительного контакта (см., например, [Котов 2021: 21] со ссылкой на [Чуфаровский 2010: 47]). Показательны в этом отношении и художественные тексты, в которых подробно описывается допрос свидетеля, ср.:

- (33) Жеглов **устроился на ручке кресла**, в котором сидела Волокушина, и **голос у него был такой**, будто они в парке на скамеечке про жизнь и про чувства свои высокие беседуют. <...> Жеглов **наклонился к ней вплотную** и спросил **задушевно**: — Светлана Петровна, а может быть, делишки у них есть общие? <...> Жеглов, с его **плавными движениями, мягкими жестами, вкрадчивым голосом**, приковывал к себе ее внимание, как удав... <...> — И что, Аня перезванивала вам после этого? — Жеглов стоял около нее, и я все ждал, когда он поставит свой хромовый сапожок на перекладину ее стула, но он удержался, а может, это было излишним — он уже достиг с ней контакта (Аркадий Вайнер, Георгий Вайнер. Эра Милосердия).

Приведём два примера из нашего корпуса, в которых во время допроса используется жест **положить руку на руку**:

- (34) *Тамара закусила губу, устремив сквозь меня невидящий взгляд. Я неуклюже **положил руку на ее ладонь**: — Тамара, я понимаю, что вам сейчас очень тяжело. Но вы должны мне помочь.* [Аркадий Вайнер, Георгий Вайнер. Я, следовательно... (1968)]
- (35) *И ему захотелось сказать все, признаться во всем. Петр, как будто угадывая мысль его, **положил свою руку на руку сына, с тихой ласкою**. — Расскажи-ка, Алеша, как ты бежал. <...> Он ... встал, подошел к отцу и упал перед ним на колени. <...> Отец наклонился к нему и **положил ему руки на плечи, все с той же тихой ласкою**. <...> Алексей молчал. Петр обнял, прижал к себе его голову и, тяжело вздохнув, прибавил: — Ах, Алеша, Алеша, если бы видел ты сердце мое, знал скорбь мою! <...> Пожалеешь хоть ты отца. Будь другом. Аль не хочешь, не любишь?.. — Люблю, люблю, батенька родненький!.. — прошептал царевич <...>. И рассказал все, назвал всех.* [Д. С. Мережковский. Петр и Алексей (1905)]

Заметим, что в последнем примере используется приём градации, основанной на постепенном сближении Петра и Алексея в ходе допроса, которое отражается, в частности, в телесном поведении Петра (*положил свою руку на руку сына — положил ему руки на плечи — обнял*).

Замечание.

В ряде литературных произведений в ходе диалога следователя и обвиняемого (или даже палача и жертвы) между ними наблюдается на первый взгляд парадоксальное психологическое и телесное сближение — например, палач жестами выражает тёплые чувства и участие по отношению к своей жертве, при этом оставаясь палачом и продолжая истязать её (что отличает его поведение от поведения злоумышленника, который фактически переходит на сторону жертвы и пытается ей помочь, как в случае синдрома Лимы). Приведём несколько примеров: *Он [О'Брайен] был мучитель, он был защитник, он был инквизитор, он был друг. <...> Вдруг оказалось, что он сидит и О'Брайен **обнимает его за плечи**. <...> О'Брайен успокоил его, почти ласково **взяв за руку*** (Дж. Оруэлл. 1984. Пер. В. Голышева); *...И к сынку вашему, милому отроку, обращаюсь, **придвинувшись близко**: Слабо, Ермолаюшка? Говорит мне Ермолай, отшатнувшись: нечего мне вам больше поведать.* (В. Ерофеев. Попугайчик); см. также сравнение палач как *добрый врач* в песне В. В. Высоцкого «Палач». Исследование невербального поведения палача и жертвы и причин такого поведения в художественной литературе представляет собой отдельную интересную задачу.

До сих пор в разделе (2) речь шла преимущественно о тех случаях употребления жеста **положить руку на руку**, когда исполнитель побуждает адресата сделать или перестать делать нечто (группы (2а)–(2в)). Далее мы рассмотрим те ситуации, в которых исполнитель пытается влиять на чувства и мысли адресата — более локально (успокаивая адресата здесь и сейчас) или более глобально (убеждая

адресата изменить своё отношение к кому или чему-либо). Конечно, проводимая нами граница между группами (2а–в) и (2д–е) довольно зыбка: в частности, исполнитель может побуждать адресата к определённым действиям, пытаясь изменить образ его мыслей, — таковы примеры из только что рассмотренной группы (2г), в известном смысле «промежуточной».

2д. X хочет успокоить Y-а

Данная группа представлена в нашем корпусе довольно многочисленными примерами, ср.:

- (36) *Я просил ее не беспокоиться за участь отца <...>. Она расплакалась (впервые) и я, чтобы ее утешить, сказал ей что-то вроде: «Милая ба-рышня. Вы успокойтесь», и **положил свою руку на ее**, лежащую на столе или на спинке стула, как вдруг она, нагнувшись, поцеловала мне руку!* [А. Ф. Редигер. История моей жизни (1918)]
- (37) — *Ты не беспокойся, мама, — сказал Коля, **положив свою большую руку на маленькую руку Софьи Петровны**, — ты не беспокойся, мы там с Аликом прекрасно заживем...* [Л. К. Чуковская. Софья Петровна (1939–1940)]
- (38) — *В чем же оскорбление? Ведь вы не поступили дурно? — Хуже, чем дурно, — стыдно. Варенька покачала головой и **положила свою руку на руку Кити**. — Да в чем же стыдно? — сказала она. — Ведь вы не могли сказать человеку, который равнодушен к вам, что вы его любите?* [Л. Н. Толстой. Анна Каренина (1878)]

Часто исполнитель жеста одновременно и успокаивает собеседника, и просит его перестать или даже не начинать говорить (см. выше (2в)):

- (39) — *Я просто очень встревожен пьесой. Не знаю, хотите ли вы говорить со мной об этом. Вы за последние дни дали мне понять... Фёрстер перебил меня, **положив свою руку на мою**: — Пусть вещи идут своим чередом, голубчик.* [М. С. Шагинян. Своя судьба (1916)]
- (40) — *Да как, Митя, и не растеряться... И мы бы с тобой растерялись, — ведь старик-то так вдруг умер у нее на глазах... Дмитрий Андреевич сделал нетерпеливое движение. Андрей Иванович **положил свою руку на его руку**, как бы желая **успокоить и остановить его на первом слове**.* [А. К. Шеллер-Михайлов. Голь (1882)]
- (41) — *...Он пожил довольно, старики должны умирать ранее молодых. В этом закон природы. — Молчите! — разом вспыхивая гневом, возмутилась я, — молчите, или я выпрыгну сейчас же из коляски... — Полно, княжна, полно, успокойтесь! — произнес он мягко и вкрадчиво, **положив свою пухлую руку на мои** заолодевшие от смущения **пальцы**, — я не хотел огорчить вас.* [Л. А. Чарская. Вторая Нина (1909)]

Как и в примере (26), в примере (41) жест **положить руку на руку** имеет двоякую функцию — утешение, служащее поводом для исполнения жеста, и ухаживание, подлинная цель жестикулирующего (ср. (2а))⁵.

2е. Х убеждает Y-а изменить своё отношение к кому-л. или к чему-л.

- (42) — *Он мне не нужен. И тогда старик на цыпочках приблизился к кровати роженицы, сел на краешек матраса и **положил** свою теплую сухую **руку на пальцы** кормящей матери. — Ах, девушка, — сказал Ротишльд совсем тихо. — Дети превеликое счастье! В детях весь смысл жизни человеческой!* [Дмитрий Липскеров. Последний сон разума (1999)]
- (43) *Я **накрыл ладонью ее руки**, показательно послушно лежавшие на столе, и сказал тоном, каким педагоги заканчивают лекцию <...>: — Он неврастеник и трус, как всякий красавец, понявший, что красота — временное явление.* [Сергей Бурлаченко. Сорвиголова // «Дальний Восток», 2019]

В этих двух примерах Х обращается к Y-у как бы сверху вниз, воспитывает его, будучи более старшим и опытным: в примере (42) акушер уговаривает роженицу не отказываться от ребёнка, в примере (43) отец уговаривает дочь прекратить отношения с человеком, в которого она влюблена. Возможно, в этих ситуациях жест **положить руку на руку** служит невербальной манифестацией власти X-а над Y-ом — так же, как и в примере (32), где показано, что художник имеет власть над временем.

Возможны и такие ситуации, когда X и Y общаются более или менее на равных:

- (44) — *Я должна сказать всё. Вы спрашиваете, что? То, что я любила. Она **положила свою руку на его умоляющим жестом**. Он молчал. — Вы хотите знать кого? Да, его, государя.* [Л. Н. Толстой. Отец Сергей (1890)]
- (45) *Старик ... нагнулся поближе к Пиару: — Не обижайтесь, Пиар, — он **положил свою руку на руку журналиста**, в которой тот держал вилку. Этим жестом он, наверное, хотел добиться расположения Пиара, и продолжил, — в наше нелегкое время нельзя более никому доверять.* [М. А. Богатов. [Ко]миссия. [staats]роман (2001) // «Волга», 2008]

Проблемы интерпретации жеста «положить руку на руку»

Собранные нами примеры показывают, что жест **положить руку на руку** употребляется в разных ситуациях. Этот жест многофункционален, а может быть, даже

⁵ Как справедливо указал анонимный рецензент, с попыткой успокоить адресата (2д) сопряжено сочувствие к нему (примеры группы (1а)). Разделяя примеры на эти две группы, мы ориентировались во многом на прямую речь жестикулирующего персонажа (если он говорит адресату, что не стоит беспокоиться, стыдиться и под., то такие примеры попадали в группу (2д), ср. *успокойтесь* (36, 41); *не беспокойся* (37); *Да в чем же стыдно?* (38); — *Я... очень встревожен...* — *Пусть вещи идут своим чередом* (39)) и на авторское пояснение (ср. *...как бы желая успокоить и остановить его* (40)).

многозначен, причём разные его значения и функции могут совмещаться в одном контексте. Ср.:

- (46) ('тёплые чувства' и 'успокоить') — *Какой же вы мне совет дадите, monsieur Dimitri!? — спросила она погодя немного. Санин увидел, что пальцы Джеммы дрожали на ее коленях... <...> Он тихонько **положил свою руку на эти бледные, трепетные пальцы**. — Джемма, — промолвил он, — отчего вы не смотрите на меня?* [И. С. Тургенев. Вешние воды (1872)]
- (47) ('одобрение' и 'просьба') *Я спрашиваю: как в канаве очутился? «Жердина под ногой повернулась». Сама? — спрашиваю. <...> «А кто же еще?» — говорит. Я **руку на руку ему положил**. И говорю: так и отвечай всем.* [Владимир Дудинцев. Белые одежды / Третья часть (1987)]

Благодаря контекстному совмещению разных значений и функций человек, который кладёт свою руку на руку собеседника, может выдавать повод для исполнения данного жеста за истинную, скрываемую цель без каких-либо видимых противоречий (напомним пример (26), где поводом для побуждающего прикосновения служит демонстрация рук, и пример (41), где поводом является утешение расстроенной княжны). В этом отношении жест **положить руку на руку** обнаруживает сходство с приёмами манипуляции, смысл которых в том, чтобы замаскировать подлинную цель действий манипулятора.

Иногда адресат жеста или наблюдатель (в роли наблюдателя может выступать, в частности, и сам автор текста) затрудняется (или делает вид, что затрудняется) определить основную коммуникативную цель исполнителя жеста или определяет её неправильно, ср.:

- (48) — *Отчего же мне не читать Байрона? — спросила она. — По двум причинам, — сказал Александр, помолчав. Он **положил свою руку на ее руку**, для большего ли убеждения или потому, что у ней была беленькая и мягкая ручка...* [И. А. Гончаров. Обыкновенная история (1847)]

— здесь как будто неясно, является жест Александра побудительным невербальным актом или знаком, показывающим, что герою нравится девушка.

Приведём пример ошибочной интерпретации жеста адресатом:

- (49) — *Верочка, милая... — начал он мягко, как бы продолжая прежнюю мысль, и ласково **положил свою руку на ее**. «Вот оно! — как молния мелькнуло в голове Веры. — Сейчас будет объяснение». Но ее нервы слишком напряжены. <...> — Пожалуйста! Пожалуйста! Не говорите! Я и так знаю! — вырвалось у ней сдавленным криком. <...> Ошеломленный Васильцев несколько мгновений глядел на нее молча, растерянно. <...> — Я думала... что вы влюблены в меня! <...> Васильцев, как ужасленный, выпустил ее руки. — Ах, Вера, и вы не лучше других, такая же кисейная барышня! — проговорил он укоризненно...* [С. В. Ковалевская. Нигилистка (1884)]

В последнем примере реакция адресата (Веры) оказалась для исполнителя жеста (Васильцева) неожиданной. В нашем корпусе встречаются и другие примеры такого рода:

- (50) — ...*Какое неожиданное и интересное открытие, правда? — И я **положил свою руку на руку девушки**. — Да, интересно... — отчужденно произнесла Таня и быстро поднялась. — Ну, надо идти спать, уже поздно...* [И. А. Ефремов. Обсерватория Нур-и-Дешт (1944)]

Сложности, связанные с интерпретацией данного жеста, нередко возникают при попытке разграничить его близкие функции или смыслы. Так, пример (51) *Ксения **положила ладонь на мою руку** и погладила ее. Я сжал ее пальцы. — Ой, мне больно! Ты, я вижу, тоже стосковался.* [Геннадий Алексеев. Зеленые берега (1983–1984)] — вряд ли можно с уверенностью отнести только в группу (1а) ‘Х испытывает к Y-у тёплые чувства’ или только в группу (2а) ‘Х соблазняет Y-а’ — возможно, он принадлежит обоим группам сразу. Кроме того, и для побудительных, и для непобудительных речевых актов жест **положить руку на руку** используется почти в одинаковой функции (увеличение иллюкутивной силы речевого акта, маркирование его искренности, ср. (16) ‘Х говорит с Y-ом серьёзно и искренне’ и (26) ‘Х показывает Y-у, что X действительно хочет, чтобы Y нечто сделал’); тем самым, противопоставление «неманипулятивной» группы (16) и «манипулятивной» группы (26), как мы уже говорили выше, довольно условно. Воздействующий характер коммуникации определяется в первую очередь коммуникативной целью речевого акта, а жест только подчёркивает её.

Заключение

В работе было показано, что жест **положить руку на руку** может использоваться говорящим как инструмент, усиливающий речевое воздействие и/или скрывающий его истинные цели, однако, на наш взгляд, неверно было бы считать, что его появление в коммуникации однозначно свидетельствует о её манипулятивном характере. Причин тому несколько.

Во-первых, поскольку разные значения и функции жеста **положить руку на руку** могут реализоваться в одном контексте (например, один собеседник пытается одновременно и убедить в своей правоте другого, и выразить свою симпатию к нему), не всегда можно с уверенностью установить, что является главной целью исполнителя жеста, а что — только поводом (и далеко не всегда этот вопрос имеет смысл). Во-вторых, многозначные жесты вообще легко допускают разные интерпретации (ср. ...*Королева приложила палец к губам. <...> Молодой человек рассудил, что этот очаровательный жест не призывал его хранить молчание, а имел какое-то другое значение* (А. Дюма. Королева Марго. Пер. Е. Ф. Корш)), в том числе и неожиданные для исполнителя жеста, что справедливо и для рассматриваемого нами случая. В-третьих, даже одна конкретная функция невербального знака может проявляться в совершенно разных коммуникативных ситуациях: жест **положить руку на руку** используется как средство невербального усиления

и в случае просьбы, и в случае благодарности, при этом в первой ситуации попытка психологического воздействия на адресата со стороны исполнителя жеста более несомненна, чем во второй.

Тем самым, представляется, что о манипулятивном характере диалога позволяют судить не столько семантические и функциональные особенности того или иного невербального знака, сколько конкретная ситуация его употребления. Пропитируем слова Н. С. Курагина: «Невербальные средства общения сами по себе не могут создать ситуацию, в которой один человек будет манипулировать другим. Однако они могут стать хорошим импульсом для того, чтобы манипуляция состоялась и была успешной. Прикосновения, жесты, улыбки или гримасы, громкость и тональность голоса вызывают у адресата определенную эмоцию, которая, в свою очередь, дает манипулятору возможность успешно достичь своих целей» [Курагин 2021: 445].

Литература

Варламов А. А., Кравченко А. Н., Горбачева А. В., Осадчий М. А. Язык прикосновений: биологические аспекты тактильного восприятия и особенности тактильных коммуникативных сигналов // Вопросы языкознания. 2020. № 2. С. 75–92.

Котов А. А. Механизмы речевого воздействия. М.: РГГУ, 2021. 432 с.

Крейдлин Г. Е. Невербальная семиотика. Язык тела и естественный язык. М.: Новое литературное обозрение, 2002. 592 с.

Крейдлин Г. Е. Невербальный диалог в истории кинесики // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 25–29 мая 2011 г.). Вып. 10 (17). М.: РГГУ, 2011. С. 349–358.

Курагин Н. С. Невербальная манипуляция: теории и мнения // Молодой ученый. 2021. № 22 (364). С. 443–445.

Непряхин Н. Ю. Я манипулирую тобой: методы противодействия скрытому влиянию. М.: Альпина Паблишер, 2018. 372 с.

СЯРЖ 2001 — Григорьева С. А., Григорьев Н. В., Крейдлин Г. Е. Словарь языка русских жестов. М.; Вена: Языки русской культуры; Венский славистический альманах, 2001. 254 с.

Чуфаровский Ю. В. Психология оперативно-розыскной и следственной деятельности. М.: Проспект, 2010. 207 с.

Brown P., Levinson C. S. Politeness: some universals in language use. Cambridge University Press, 1987. 345 p.

Chatti S. When causation is manipulation: the case of causative GET // British Journal of English Linguistics. 2017. Vol. 5. № 3. P. 46–56.

Guéguen N., Jacob C., Boulbry G. The effect of touch on compliance with a restaurant's employee suggestion // International Journal of Hospitality Management. 2007. № 26 (4). P. 1019–1023.

Van Dijk T. Discourse and Manipulation // Discourse & Society. 2006. № 17 (2). P. 359–383.

Vrij A. Nonverbal communication and deception // The SAGE handbook of nonverbal communication / Eds. Valerie Manusov, Miles L. Patterson. Sage Publications, Inc., 2006. P. 341–359.

Wierzbicka A. Kisses, handshakes, bows: the semiotics of nonverbal communication // Semiotica. 1995. № 103, 3/4. P. 207–252.

Wierzbicka A. English: Meaning and Culture. Oxford: Oxford University Press, 2006. 352 p.

S. I. Pereverzeva

Russian State University for the Humanities

(Russia, Moscow)

P_Sveta@hotmail.com

NONVERBAL MANIPULATION AND THE PROBLEMS OF ITS DETECTION: THE CASE OF THE GESTURE “TO LAY ONE’S HAND ON SOMEBODY’S HAND”

This article discusses the notion of nonverbal manipulation taking the tactile gesture “to lay (one’s) hand on (somebody’s) hand” as an example. We analyze the meaning and usage of this gesture on the basis of Russian literary fiction of the 19th–21st centuries and group them into two classes: 1) the performer of the gesture shows that he / she feels something good towards the addressee, he / she tells the addressee something important, or he / she implies sincerity of his / her words; 2) the performer of the gesture propels the addressee to do something or change his / her state of mind (seduction, persistent request, attempt to stop the addressee’s words or actions, examination (of a witness or of a person charged with the commission of a crime), calming, persuasion). In the situations of the second type, nonverbal manipulation is more likely. However, we assume that it is impossible to confidently regard the communication as manipulative taking into account nothing but the gesture in question itself, since it (a) allows for the combination of different meanings and functions in a single context; (b) may be misinterpreted by the addressee; (c) can perform a certain function in different communicative acts.

Keywords: gesture, manipulation, polysemy, communicative goal, dialogue.

References

Brown P., Levinson C. S. *Politeness: some universals in language use*. Cambridge University Press, 1987. 345 p.

Chatti S. When causation is manipulation: the case of causative GET. *British Journal of English Linguistics*, 2017, vol. 5, no. 3, pp. 46–56.

Chufarovskii Yu. V. *Psikhologiya operativno-rozysknoi i sledstvennoi deyatel’nosti* [Psychology of the investigative activities]. Moscow, Prospekt Publ., 2010. 207 p. (In Russ.)

СЯРЖ 2001 — Grigor'eva S. A., Grigor'ev N. V., Kreidlin G. E. *Slovar' yazyka russkikh zhestov* [Dictionary of the language of Russian gestures]. Moscow, Wien, Yazyki russkoi kul'tury, Wiener Slawistischer Almanach, 2001. 254 p. (In Russ.)

Guéguen N., Jacob C., Boulbry G. The effect of touch on compliance with a restaurant's employee suggestion. *International Journal of Hospitality Management*, 2007, no. 26 (4), pp. 1019–1023.

Kotov A. A. *Mekhanizmy rechevogo vozdeistviya* [Mechanisms of the speech influence]. Moscow, Russian State University for the Humanities Publ., 2021. 432 p. (In Russ.)

Kreidlin G. E. *Neverbal'naya semiotika. Yazyk tela i estestvennyi yazyk* [Nonverbal semiotics. The body language and the natural language]. Moscow, Novoe literaturnoe obozrenie, 2002. 592 p. (In Russ.)

Kreidlin G. E. [Nonverbal dialog in the history of kinesics]. *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* (2011). 2011, issue 10 (17), pp. 349–358. (In Russ.)

Kuragin N. S. [Nonverbal manipulation: theories and opinions]. *Molodoi uchenyi*, 2021, no. 22 (364), pp. 443–445. (In Russ.)

Nepryakhin N. Yu. *Ya manipuliruyu toboi: metody protivodeistviya skrytomu vliyaniyu* [I manipulate you]. Moscow, Alpina Publ., 2018. 372 p. (In Russ.)

Varlamov A. A., Kravchenko A. N., Gorbacheva A. V., Osadchii M. A. [Language of touch: Biological aspects of social touch perception and the system of tactile communicative signals]. *Voprosy Jazykoznanija*, 2020, no. 2, pp. 75–92. (In Russ.)

Л. Л. Шестакова¹, А. С. Кулева²

Институт русского языка им. В. В. Виноградова РАН

(Россия, Москва)

lara.shestakova@mail.ru¹, an_kuleva@mail.ru²

СЛОВО ЖЕСТ В ПОЭТИЧЕСКОМ ДИСКУРСЕ

Слово *жест* в русском поэтическом языке не относится к высокочастотным, однако примеры его употребления позволяют сделать интересные наблюдения. Случаи использования слова *жест* в поэтических текстах рассматриваются через призму авторитетных лингвистических исследований по русской жестикуляции, а также работ литературоведческой, лингвопоэтической направленности, в которых жесты понимаются как знаки, образы, часть системы символов, в том числе у отдельных авторов. В многотомном «Словаре языка русской поэзии XX века», который создается по произведениям десяти видных поэтов Серебряного века (И. Анненского, А. Ахматовой, А. Блока, С. Есенина, М. Кузмина, О. Мандельштама, В. Маяковского, Б. Пастернака, В. Хлебникова, М. Цветаевой), содержится около сотни контекстов с этим словом (наибольшее их количество отмечается у Цветаевой). В поэтическом корпусе Национального корпуса русского языка обнаруживается более 450 вхождений. В статье анализируются разные типы поэтического употребления слова *жест*. В ряде случаев оно используется без конкретизаторов (для обозначения жеста как такового, неопределенного жеста), и нередко в ряду названий других невербальных проявлений; часто слово *жест* уточняется путем указания на способ, которым производится телодвижение. В поэтическом дискурсе частотны атрибутивные уточнения, характеризующие *жест* с той или иной стороны. Особый интерес представляют случаи образного осмысления слова *жест*.

Ключевые слова: авторская лексикография, жест, поэтический дискурс, словарь поэтического языка, Национальный корпус русского языка

В статье Е. А. Гришиной «Указания рукой как система (по данным Мультимедийного русского корпуса)» отмечалось, что в России «к сожалению, изучение жестикуляции до сих пор расценивается как область, интересная скорее для психологов, а не для лингвистов <...> те или иные жестикуляционные данные все еще рассматриваются как экзотический побочный продукт изучения устной речи, притом что на самом деле сопровождающие речь жесты являются надежнейшим средством анализа семантических, синтаксических, грамматических характеристик

высказывания, — не менее, а иногда и более наглядным, доступным и доказательным, чем, например, фонетические или интонационные данные» [Гришина 2012: 3–4].

За прошедшие годы наука значительно продвинулась вперед: лингвистические исследования русской жестикуляционной системы многоаспектны, опираются на обширный языковой материал. Представляется перспективным исследовать и такой аспект, как функционирование самого слова *жест* в конкретном типе дискурса, а именно в поэтическом; посмотреть, в каком окружении выступает это слово, какие жестовые ситуации воссоздает.

Прежде чем перейти к сделанным наблюдениям, к результатам проведенного исследования, сделаем несколько уточнений.

Для целей работы на данном этапе достаточны те определения значений слова *жест*, которые предлагаются в общезыковых толковых словарях, например, в новом БАСе:

1. телодвижение (преимущественно движение рукой), сопровождающее речь для усиления ее выразительности или имеющее значение какого-л. сигнала, знака;
2. *перен.* о поступке, совершенном с каким-л. умыслом, расчетом или в знак чего-л. [БАС-3, 5: 619–620].

Нас, очевидно, интересовало значение-1, но некоторые наблюдения по значению-2 тоже были сделаны.

В отборе примеров употребления слова *жест* в поэтических текстах мы отталкивались от одноименной статьи в «Словаре языка русской поэзии XX века» [СЯРП], который создается по произведениям десяти видных поэтов Серебряного века (И. Анненского, А. Ахматовой, А. Блока, С. Есенина, М. Кузмина, О. Мандельштама, В. Маяковского, Б. Пастернака, В. Хлебникова, М. Цветаевой). Статья ЖЕСТ, включающая около 40 контекстов (наибольшее их количество отмечается у Цветаевой), задала примерные хронологические рамки материала для исследования и показала возможные направления его анализа. Но основным источником стал, конечно, поэтический корпус [НКРЯ].

Примеры рассматривались через призму авторитетных лингвистических исследований по русской жестикуляции ([Григорьева, Григорьев, Крейдлин 2001; Крейдлин 2002; 2005; Гришина 2011; 2012; 2013; 2017]), а также работ литературоведческой, лингвопоэтической направленности, в которых жесты понимаются как знаки, образы, часть системы символов, в том числе у отдельных авторов ([Слово и жест 1983; Раков 1995; Бобринская 1998; Масленникова 2000; Федунина 2017] и др.).

В поэтическом корпусе (ПК) [НКРЯ] первые фиксации слова *жест* относятся к 1820–30-м гг.: К. П. Масальский (1828), М. Ю. Лермонтов (1836); последние фиксации на данный момент: Чухонцев (1997–2016), Байтов (2009). Примеров из Пушкина ПК не фиксирует, в «Словаре языка Пушкина» находим один пример из журнальной статьи: *если жесты ее будут естественнее и не столь жеманными, --- то мы можем надеяться иметь со временем истинно хорошую актрису* [СЯП, 1: 807].

В целом поиск по этому слову дает такие результаты: 422 документа, 461 вхождение; если ограничить поиск кругом поэтов выбранного временного отрезка (пользовательский подкорпус: год рождения авторов — с 1855 по 1905 г.) — 217 документов, 246 вхождений.

Поскольку речь идет о поэтическом языке, важно сразу отметить, что по положению слова *жест* в сильной текстовой позиции — позиции рифмы ПК НКРЯ (поиск: «жест, rhyme») дает такие результаты: 171 документ, 173 вхождения, в пользовательском подкорпусе: 73 документа, 75 вхождений. В рассматриваемом массиве примеров встречаются и нерифмованные произведения, поэтому имеет смысл рассмотреть конкретные случаи.

Прежде всего бросается в глаза неравномерность употребления разных грамматических форм слова. Как показал анализ, это связано не только с частотностью падежных форм в языке, но и с их версификационным потенциалом.

Так, в массиве из 246 примеров преобладают односложная форма И. ед. *жест* и форма Т. ед. *жестом*, тяготеющая к наречному употреблению, а также И. мн. *жесты* (поиск по точной форме):

И-В. ед. *жест* 110 вхождений (45 %), мн. *жесты* 31 вхождение (13 %),
Р. ед. *жеста* — 11 (4 %), мн. *жестов* — 16 (6 %),
Д. ед. *жесту* — 0, мн. *жестам* — 1,
Т. ед. *жестом* — 52 (21 %), мн. *жестами* — 5 (2 %),
П. ед. *жесте* — 12 (5 %), мн. *жестах* — 8 (3 %).

Для сравнения: в основном корпусе НКРЯ (15 171 пример), при ожидаемо меньшей употребительности форм косвенных падежей, количественное соотношение форм отличается от поэтических данных:

И-В. ед. *жест* 4801 вхождение (32 %), мн. *жесты* 1821 вхождение (12 %),
Р. ед. *жеста* — 850 (6 %), мн. *жестов* — 961 (6 %),
Д. ед. *жесту* — 214 (1,5 %), мн. *жестам* — 175 (1 %),
Т. ед. *жестом* — 4014 (26 %), мн. *жестами* — 1287 (8 %),
П. ед. *жесте* — 470 (3 %), мн. *жестах* — 348 (2 %).

Такое отличие поэтического употребления от общеязыкового может быть соотнесено с различным функционированием отдельных форм в рифменной позиции.

Немногочисленные формы *жестам*, *жестами*, *жестах* не вступают в рифменные созвучия, даже оказываясь в тексте в конце строки (в нерифмованных текстах).

Форма *жестов* (16 примеров) рифмуется в одном тексте: Я с давних пор привык там видеть их: Все так же, не меняя поз и *жестов*, Стоят — невеста и ее жених. *И вечно так — счастливая невеста* [Д. Кнут. «Тому дней пять... дней шесть тому назад...» (1929)].

Форма *жесте* (12 примеров) — в четырех: *Он кудри седые откинул, он руку Не-вольню простер в повелительном жесте*. «Должны освятить, — он промолвил, — разлуку Мы песней все вместе!» [В. Я. Брюсов. Последние поэты (1917)]; *Ветви деревьев — в жесте Прыгающих с перил. Визг озверевшей жести, Шелест шафранных крыл*. [А. И. Несмелов. Тайфун (1924–1945)]; *Воспоминаний цепь, кольцо к кольцу, /*

нижу в чужих улыбке, взгляде, **жесте** — / так зверь домашний к новому лицу / доверчиво идет на старом месте. [В. А. Меркурьева. Второй прибор (1933)]; *Будто сбитая из жести, Вот ты — будто восковая, Вот ты — будто в каждом **жесте**, Дорогая, дорогая...* [Г. Н. Оболдуев. «Ты зажмуриваешь очи...» (1938)].

Форма **жеста** (11 примеров) — в восьми, т. е. в большинстве из них: *Невеста — жеста* (Волошин, Шервинский, Цветаева), *жеста — невеста* (Цветаева); *жеста — протеста* (Цветаева), *протеста — жеста* (Саша Черный); *места — жеста* (Асеев); *место — жеста* (Несмелов).

Частотная форма **жестом** (52 примера) менее активна в рифменном употреблении (9), но зато встречается в более индивидуальных рифмах — богатых, неточных, составных: *жестом — местом* (Блок, Оболдуев), *местом — жестом* (Саша Черный); *жестом — место* (Маяковский); *Орестом — жестом* (Цветаева); *жестом — переездам* (Филипченко); *жестом — тестом* (Саша Черный); *жестом — златошерстым* (Маяковский); *неуместны — жестом* (Баркова).

То же можно сказать о форме И. мн. **жесты** (31 пример), в рифменном употреблении — 8: *невесты — жесты* (Ходасевич), *жесты — невесты* (Бальмонт, С. Соловьев, Саша Черный); *жесты — челесты* (Аксенов); *жесты — происшествий* (Асеев); *известны — жесты* (Есенин). Ср.: *Без судорожных «где ж ты?». Жду. С тишиной в родстве, Прислуживают — жесты В Психеином дворце.* [М. И. Цветаева. Попытка комнаты (1926)].

Наиболее употребительная форма И. ед. **жест** (110 примеров) оказывается и наиболее востребованной с точки зрения рифмы (39), причем особенно заметными становятся идиостилевые особенности: *жест — ест* (Мережковский); *окрест — жест* (Саша Черный, Багрицкий); *жест — надоест* (Пяст); *звезд — жест* (Цветаева, дважды); *мест — жест* (С. Соловьев, Лившиц, Несмелов), *жест — мест* (Оцуп, Тарловский); *жест — крест* (Эллис, Баркова), *крест — жест* (Эренбург, Меркурьева); *жести — жест их* (Земенков); *жест — Рождеств* (Маяковский); *жест — палимпсест* (Меркурьева); *жесть — жест* (Сельвинский), *жест — жесь* (Несмелов); *наест — жест* (Цветаева); *жест — торжеств* (Пастернак, Оболдуев), *торжеств — жест* (Антокольский, Оболдуев); *жест — божеств* (Цветаева, Антокольский), *божеств — жест* (Антокольский); *жест — лжец* (Асеев, дважды); *жест — невест* (Юрков); *жест вам — Блаженство* (Маяковский); *жест — шест* (Эйснер), *шест — жест* (Сельвинский); *заест — жест* (Цветаева); *жест — разъест* (Оцуп); *жест — Yes* (Несмелов); *жест — Веществ* (Оболдуев); *жест — Вошебств* (Оболдуев).

Очевидно, что слово *жест* в поэтическом тексте оказывается востребованным и в силу своей фонетической организации — помимо рифменных созвучий, *жест* нередко включается в аллитерацию, сближаясь, сопоставляясь со словами *торжество*, *торжественный*, *жестко*, *жестокий* и др.:

Обойдя всех сидевших на лавках, Я отвесил поклон королям; И по старым, глубоким морицинам Пробежала усталая тень; И привычно торжественным жестом Короли мне велели остаться. [А. А. Блок. Ночная фиалка (1906)];

Ему в задумчивые глазки Взглянула мама так светло! Когда ж в пруду она исчезла И успокоилась вода, Он понял — **жестом злого жезла Ее колдун увлек туда**. [М. И. Цветаева. Самоубийство (1906–1912)];

Сначала **жестко: жесты и ужимки** Под щелкающие затворы, На конференциях, в сигарной дымке, Увертливые разговоры. [М. А. Тарловский. Звуки (1927)];

И подошла к окну и вдруг кошачьим **Жестоким хищным жестом сорвала Весь коленкор**, натянутый на раме, И отворила створки. [В. А. Луговской. Сказка о сне (1942–1956)];

А он не слышит этих голосов. Не вспомнит он Саратов или Нижний, Средь хлопающих оживленных сов Ушедший в камень. Белый. Неподвижный. И все-таки коричневые орды Не одолели стойкости его. **Как мощны плечи, поднятые Гордо! Какое в этом жесте торжество!** [И. Л. Сельвинский. Аджи-Мушкой (1943–1948)].

Если говорить о роли слова *жест* в структуре и композиции стихотворения, то вполне естественно его появление в роли ремарки, причем не только в драматических произведениях (дважды у Анненского), но и в стихотворном тексте. Ср.:

Корифей (успокаивая **жестом** Хор, мягко) Твоя печаль понятна... [И. Ф. Анненский. [Меланиппа-философ, 4] (1901)];

Царевна, здесь не место вам, я вижу, Покинутый давно вас терем ждет. (Приглашая **жестом** Меланиппу во дворец.) [И. Ф. Анненский. [Меланиппа-философ, 12] (1901)];

Навея (вся в белом, в белом покрывале, властным **жестом** прекращает спор) Оставьте все его! Мне дайте место! [Т. Л. Щепкина-Куперник. Сон зимней ночи (1922)];

Гулльд (с **жестом**): «Наша промышленность, С годами выстраивающаяся ровней, Вынуждена силой за рубеж громиться, Чтобы тромбы фабрикатов рассосать вовне. [И. Л. Сельвинский. Тряпичный король (1923)].

Один раз в рассмотренном материале слово *жест* используется в заглавии стихотворения, причем характерно, что понятие жеста в антропоморфно ориентированном тексте раскрывается через образ *рук Рока, перстов судьбы* (слово *жест* в самом тексте отсутствует):

СУДЬБЫ ЖЕСТЫ

Когда судьба занесена —

На мир презрительным **указует перстом** <...>

Презрительным перстом,

Низвергая тусклейшие ряды —

Борозды, звезды ринутся,

Раздвигая ослеплений бег и пробег. <...>

Но лопнет струной золотой меридиан,

Но, звякнув, иссякнет стран поток:

Нежно **опустит руки Рок**.

[С. П. Бобров. Судьбы жесты (1914)].

Показательно, что в ситуациях, когда в тексте так или иначе описывается сам жест, речь идет обычно о жесте, совершаемом рукой, что вполне соответствует как словарным толкованиям, так и лингвистическим наблюдениям [Гришина 2012; 2013]. Однако для поэтической речи характерно и более общее, абстрагированное понимание *жеста*.

Рассмотрим несколько типов употреблений.

Слово *жест* часто выступает в парах с названиями других невербальных проявлений (сочетаясь со словами *улыбка, взгляд, взор, поза* и т. п.):

Походка и жесты — людей, живущих созерцательной жизнью. [И. Ф. Анненский. [Меланиппа-философ, 1] (1901)];

Здесь все учтиво так и чинно, / здесь взвешен каждый шаг и жест, / здесь даже бешенство картинно, / и скрыт за каждым словом крест... [Эллис (Л. Л. Кобылинский). Эпитафия (1905–1913)];

По тебе тоскует наша зала, — Ты в тени ее видал едва — По тебе тоскуют те слова, Что в тени тебе я не сказала. Каждый вечер я скитаюсь в ней, Повторяя в мыслях жесты, взоры... [М. И. Цветаева. «По тебе тоскует наша зала...» (1906–1912)];

Молчит изваянье, но мощная страсть И в жестах, и в позе — безбрежная власть — Зовет он рукою из меди Солдата к борьбе и победе. [М. О. Цетлин (Амари). Памятник (1906)];

И в каждом жесте, в каждой маске Убитых пошлостью людей Читаю призрачные сказки О красоте грядущих дней. [С. М. Городецкий. «Торжественная пляска будней...» (1913)];

Но жених твой... Гром и пушка! Ты и он, — подумай, душка: Одуванчик и лягушка, Мотылек и вурдалак. Эти жесты и улыбки, Эти брючки, эти штрипки... Весь до дна, как клейстер липкий, — Мелкий маклер и пошляк. [Саша Черный. Виленский ребус (1919–1922)];

Но в смене Ваших декораций Следила я и взгляд, и жест — И ведал тайну муз и граций Ваш полустертый палимпсест. [В. А. Меркурьева. М. Сл. (1924)];

Я с давних пор привык там видеть их: Все так же, не меняя поз и жестов, Стоят — невеста и ее жених. И вечно так — счастливая невеста [Д. Кнут. «Тому дней пять... дней шесть тому назад...» (1929)].

Встречаются и ряды таких единиц: *Воспоминаний цепь, кольцо к кольцу, / нижу в чужих улыбке, взгляде, жесте — / так зверь домашний к новому лицу / доверчиво идет на старом месте.* [В. А. Меркурьева. Второй прибор (1933)]; *Я знал в лицо всех ее знакомых, Я знал их повадки, улыбки, жесты.* [Э. Г. Багрицкий. Февраль (1933–1934)].

Ср. также: *Разбирая с достоинством каждый жест, Каждую ноту оркестра, Ноги и плечи испанских невест — Чудаки покидают кресла.* [И. В. Юрков. Театральный разъезд (1927)].

Интересно, что такие «невербальные» пары и ряды могут противопоставляться вербальным единицам, но могут и включать их, создавая более комплексную характеристику персонажа, ср.:

*Тишина посылается роком, — Тем и вечны слова, что тихи. Говорил он о самом глубоком, Баратынского вспомнил стихи; Говорил о игре отражений, О лучах закатившихся звезд... **Я не помню его выражений, Но улыбку я помню и жест.*** [М. И. Цветаева. Последняя встреча (1906–1912)];

*Я слышу Вашими ушами, Я вижу Вашими глазами, **Звук Вашей речи на устах, Ваш робкий жест в моих руках...*** [М. А. Волошин. Письмо (1904)];

*Пред искренностью страх Торжествовал и в храме и в гостиной, **В стихах и вере, в жестах и словах.*** [В. Я. Брюсов. Замкнутые (1900–1901)];

*Умел согласовать с Судьбой ты **жест и слово!*** [В. Я. Брюсов. «Вот вновь мои мечты ведут знакомый танец...» (1912)];

*Статисты жалкие — рабы полусознания, Чей **крик и жест** был никогда не нов, В сценарии под красочным названием: История народов и веков.* [Н. С. Тихонов. Статисту (1913–1919)];

*Да, я смешна, невзрачна, Нелепы **жест и речь**, Но сердце жаждет мрачно Обжечься и зажечь.* [А. А. Баркова. «Да, я смешна, невзрачна...» (1973)].

В других случаях *жест* выступает в ряду однородных понятий:

*Мы быстры и наготове, Мы остры. **В каждом жесте, в каждом взгляде, в каждом слове** — Две сестры.* [М. И. Цветаева. «Мы быстры и наготове...» (1913)];

*Глядишь на цепкие гримасы **Улыбок, жестов, слов и дел**, И тонет в океане массы Судьбой подчеркнутый удел.* [А. И. Несмелов. На блюде (1922)];

*А в залах гуденье и шорох И воздух волнисто-рябой, **Но в жестах, в словах и во взорах** Хранит свою тайну любви.* [В. Ф. Наседкин. Вокзал (1932)];

*Он был ему спутником в дальних ночах; / бывают такие — неведомы / в людской телескоп, а небесный рычаг / их движет вровень с планетами. Он помнит **каждое слово и жест, / живого лица выражение.** Планета погасла, а спутник — не лжец — / еще повторяет движение.* [Н. Н. Асеев. Маяковский рядом (1936–1939)];

*Все же воздух В оставленной квартире сохраняет **Сухие линии движений, жестов, Голов и профилей, и тени разговоров, И тени клятв, и тень моей судьбы.*** [В. А. Луговской. Поэма о вещах (1942–1956)].

Противопоставление / соположение какого-либо невербального проявления и жеста, слова и жеста нередко подчеркивается выбором эпитетов. Эпитеты в парах, тройках сочетаний обычно не совпадают, но они могут сближаться семантически и оценочно:

*«Брат! Одну за другой хоронил я надежды, Брат! От этого больше всего устают. Были **яркие речи и смелые жесты** И неполных желаний шальной хоровод. <...>».* [Саша Черный. Интеллигент (1908)];

О, по цвету голубому стосковались мы давно, И не меньше стосковались по **ликующим словам, По свободным, смелым жестам, по несбыточным мечтам!** [Саша Черный. Театр (1908)];

Где ты теперь, богиня этих мест? Встанут, как наяву, в моих бессонных грезах Твой **взор задумчивый** и твой **пластичный жест**, Пушистое боа и шляпа в мелких розах. [С. М. Соловьев. «Я вижу вас опять, знакомые места...» (1909–1912)];

Может быть, там, за чертою Дымно-праздничных деревьев, Нет гого-чущих кретинов, **Громких слов и наглых жестов**, Изменяющих красавиц, Плоско-стертых серых Лиших, Патриотов и шпионов, Терпеливо робких стонов, Бледных дней и мелочей?... [Саша Черный. На лыжах (1911)];

О, это гордое одушевление! В нем порывания горного ветра, **И жесты быстрые, и глаза горные, И речь свободная**, вся в сменах метра. [М. О. Цетлин (Амари). «О, это гордое одушевление...» (1912)];

Ибо свиданье — местность, Роспись — подсчет — чертеж — **Слов, не всегда уместных, Жестов, погрешных сплошь**. [М. И. Цветаева. Попытка комнаты (1926)];

Но трое клеймлено неизгладимо, Но трем — преображающего грима Еще изобрести не удалось. Сквозь **гордый жест**, сквозь **благородство взора** Я узнаю их наповал, насквозь: **Шпиона, проститутку и актера!** [Г. А. Шенгели. «Доверчив я. Обманут десять раз...» (1927)];

Чтоб не было в ней / ни условий, / ни места / для липких лакеев, / ханжей / и лжецов, / для **лстивого слова, / трусливого жеста**; / чтоб люди людей — узнавали в лицо. [Н. Н. Асеев. Маяковский начинается, 18 (1936–1939)];

И он изъясняется книжно, От жестов манерных не прочь, а в окна легла неподвижно Российская черная ночь! [А. И. Несмелов. В вагоне (1939)].

Ср. и один «общий» эпитет в примере: *Воспоминаний цепь, кольцо к кольцу, / нижу в **чужих улыбке, взгляде, жесте** — / так зверь домашний к новому лицу / доверчиво идет на старом месте.* [В. А. Меркурьева. Второй прибор (1933)].

Как уже отмечалось, чаще всего в поэтическом тексте упоминаются жесты, совершаемые рукой — более того, нередко контекст включает в себя, помимо слова *жест*, лексемы *рука, пальцы* и т. п., сочетания с глаголами (*делать / сделать жест, закончить жест*), обозначения совершаемых рукой движений и другие конкретизаторы.

Примеры, включающие слово *жест*:

И, быстро поднявшись, **презрительным жестом Встряхнула одеждой** над проклятым местом, Гадает... и шепчут уста. [А. А. Блок. «Потеха! Рокочет труба...» (1905)];

Вот он встал перед тобой: Посмотри на лоб и брови И сравни его с собой! То усталость голубой, Ветхой крови. Торжествует синева Каждой благородной веной. **Жест царевича и льва Повторяют кружева Белой пеной.** [М. И. Цветаева. Сергею Эфрон-Дурново, 1 (1913)];

Над желтизной правительственных зданий Кружилась долго мутная метель, И правовед опять садится в сани, **Широким жестом** запахнув шинель. [О. Э. Мандельштам. Петербургские строфы (1913)];

Далекий стон лесного храма **Широкий жест креста на лоб**: Картина темная и рама Досок нарезанных на гроб!... [Д. Д. Бурлюк. Сухопутье (1915)];

С зарницею сероглазой Из замшевой черной сумки **Вы вынули длинным жестом И** выронили — платок. [М. И. Цветаева. «Могу ли не вспомнить я...» (1915)];

А гарсон давно уже сед, **И жестом привычной прыти Сует чаевые в жилет**. [Н. В. Крандиевская. «На станцию Кемперлей...» (1921)];

Мне бросилось в глаза, с какой фриволю, Невольный вздрог улыбкой погася, Она шутя **обдернула револьвер И в этом жесте выразилась вся**. [Б. Л. Пастернак. Спекторский, 10 (1925–1931)];

Куда ни мчишь, куда ни метишь, Но попадешь, о путник, в синь, **И, если смерть ты раньше встретишь, Широким жестом шляпу скинь**. [В. Г. Шершеневич. Встречник (1926)];

Из палатки вышла дева В васильковой нежной тоге, Подошла к воде, как кошка, Омочила томно ноги **И медлительным движеньем Тогу сбросила на гравий, — Я не видел в мире жеста Грациозней и лукавей!** [Саша Черный. Пластика (1932)];

Подает он мелочь мне Жестом очень важным... [Ю. П. Одарченко. «Что такое — денег нет? ...» (1949)].

Слово жест в сочетании со словами рука, пальцы (таких примеров, передающих целую палитру совершаемых рукой жестов, особенно много):

Руки нехоленой несоразмерный жест, Чтоб худенькую прядь / на ясном лбу поправить, — **И голос, созданный / по-женскому лукавить, Но по-мужески молчать, пока не надоест...** [В. А. Пяст. Два портрета (1910–1911)];

И Вы, подошедший С доверчивым жестом несмелых рук, — Вы со страхом услышите в ночи сумасшедшей В мои темные окна минувшего стук... [Н. Г. Львова. «Почему я со страхом жду от Вас признания...» (1913)];

Многоэтажный, с видом скуки... Считаю окна, вот подъезд. **Невольным жестом ищут руки На шее — крест**. [М. И. Цветаева. «День августовский тихо таял...» (1914)];

Не подымая головы, И тем подчеркивая скуку — О, этот жест, которым Вы Мне дали руку. [М. И. Цветаева. «Прибой курчавился у скал...» (1914)];

Да, здесь мы недотроги, **И в праве. Рук — гонцы, Рук — мысли, рук — итоги, Рук — самые концы...** Без судорожных «где ж ты?». Жду. **С тишиной в родстве, Прислуживают — жесты В Психеином дворце**. [М. И. Цветаева. Попытка комнаты (1926)];

Я был у Вас, совсем влюбленный, Ушел, сжимаясь от тоски. Ужасней шапки занесенной Жест отстраняющей руки. [Н. С. Гумилев. «Что я прочел? Вам скучно, Лери...» (1916)];

*И встал во весь рост меж изгнанников старший, Спокойно-печальный. Он кудри седые откинул, он **руку Невольно простер в повелительном жесте.** «Должны освятить, — он промолвил, — разлуку Мы песней все вместе!» [В. Я. Брюсов. Последние поэты (1917)];*

***Из кулака** родилось братство: Каин первый Нашел пристойный **жест для выраженья** Родственного чувства, предвосхитив Слова иных времен: «Враги нам близкие, И тот, кто не оставит Отца и мать, тот не пойдет за мной». [М. А. Волошин. Кулак (1922)];*

***Жест на иконе сломан и пуст, Но в этом жесте такая мука,** Что летит пушинка одна, без чувств, **На переломанные руки.** [И. В. Юрков. Неудавшееся посещение (1927)];*

*«Рассказывайте вашей тете, / что вы и тут **руки не пожмете.** / Какой там принцип! / Мы служащие... мы не принцы». / И палец затем — в ладони в обе, / забыв обо всем и о микробе. / **Знаком ли товарищеский этот жест вам?** / Блаженство! / Назавтра помылся, но было поздно. / Помглавбуха — уже гриппозный. [В. В. Маяковский. Глупая история (1927)];*

***Рука сама пришла в движенье, И, проходя, я отдал честь. И этот жест в морозе лютом, В той перламутровой тиши Моим последним был салютом,** Салютом сердца и души! [А. И. Несмелов. В Нижнеудинске (1942)];*

***В жесте тонкой руки почему королевская властность?** [А. И. Несмелов. На рассвете (1942)];*

***И властным жестом только поднял руку,** Проговорив спокойно: «Вижу, нет, Еще не пережили вы разлуку [З. Н. Гиппиус. «Вот новый Дант в последний Круг пробрался...» (1945)];*

***О руках мужчины — о руках Мужественных, бережных и честных И об их заслугах полновесных** Очень трудно говорить в стихах. <...> Я люблю на загорелой коже Блекло-синий якорь узнавать, **Но когда они берут тетрадь, Этот жест мне остальных дороже,** [М. Вега. Руки (1963–1969)].*

Ср. также в переносном значении: *Скажет с красной кафедры, **жестами всех лап,** Речь громоподобную многоумный краб;* [В. Я. Брюсов. Рыбье празднество (1916)].

Слово жест в сочетании с глаголами:

*И вот жонглер у кладезя Изысканно раскланялся. Он **сделал жест,** зеленый черт, И в воздухе, как в вечности, — **Невероятный натюрморт Цилиндра и подсвечника.** [И. Л. Сельвинский. Натюрморт (1923)];*

*Застыли докладчики всех заседаний, / не могут **закончить начатый жест.** / Как были, рот разинув, сюда они / **смотрят на Рождество из Рождеств.** [В. В. Маяковский. [Про это, 6] (1923)];*

*Он вынул пилюлю, заклапанную в **жесть,** Глотнул, подняв брови, и **сделал жест:** «А славу по мне пусть лошади **выржут.**» [И. Л. Сельвинский. [Улялаевщина, 8] (1924)];*

По скользким трапециям скачет Красавица в черном трико. И музыка в ужасе тает. И, сделав решительный жест, Красивое тело взлетает На тонкий согнувшийся шест. [А. В. Эйсер. Цирк (1928)].

Ср. также:

Широкая зелень Лежит окрест Подстилкой твоим ногам! — (Рукою он делает вольный жест От сердца — И к облакам. Я узнаю в нем Свои черты, Хотя он костляв и рыж, И я бормочу себе: «Это ты Так здорово говоришь»). [Э. Г. Багрицкий. «Итак — бумаге терпеть невмочь...» (1930)];

Хозяин смотрит на работника и делает рукою жест, изображающий отказ. [Н. М. Олейников. Притча о работниках в винограднике (1936)].

Однако можно отметить и контексты, в которых описываются и другие жесты. Причем там тоже могут подразумеваться движения, совершаемые руками, но в сочетании с другими, более сложными телодвижениями. Как справедливо указано в [Гришина 2013: 91–93], жесты, за которые отвечают другие части тела, например, голова, более ограничены физически, и неудивительно, что они в меньшей степени конкретизируются лексически. Слово *жест*, как было показано, нередко сочетается со словами *рука*, *пальцы* и т. п., но в сочетании с обозначениями других частей тела его представить трудно — если движение совершается головой, то его можно обозначить глаголами типа *кивнуть*, *покачать* и др.

См. примеры с упоминанием «сложных» жестов:

Разжал я ей алые губы Своим поцелуем палящим И жестом внезапным и грубым Обнял ее круглые плечи... [А. И. Тиняков. Сны страсти (1906)];

Она беспомощно клонилась К подушке алой, в глубь и в тень... И мне казалось, что вонзилась Стрела случайная в мишень. И, жестом медленным, безвинный Убийца, я припал к устам... И миг продлился, длинный, длинный, Врата к мучительным часам. [В. Я. Брюсов. Случайная стрела (1911)];

Насильем скованный, Отточенный для миценья, — Он вместе с кровью напитался духом Святых и праведников, Им усековенных. И стала рукоять его ковчегом Для их мощей. (Эфес поднять до губ — Доныне жест военного сапюта). И в этом меч сподобился кресту — Позорному столбу, который стал Священнейшим из символов любви. [М. А. Волошин. Меч (1922)];

Не в иглу, колдовать по канве, Крест на крест налагая жесты, Не оркестра в росной траве Увидать, как слеза челесты, Не в порхание мотылька, Не в преблагословение лилий, Не в застенчивость василька, Не в лиловый плач глициний [И. А. Аксенов. «Не в иглу, колдовать по канве...» (1919–1923)];

Связан каждый жест, Ты почти недвижим, В порхающем блеске Пьёшь холодный чай. Ровен твой пробор, а — пробора ниже — На пустом лице Усики торчат. [И. В. Юрков. Благополучие (1929)];

Длится не жест, Но жест облакает вас в латы, Золотые — от груди до колен. И так чиста была битва, Что ангел несет ее вслед. [М. И. Цветаева. «Кто нам сказал, что всё исчезает...» (1939–1941)];

Что мне до этой бедуинки, Ее пустынной красоты? Она дала мне из корзинки **Понюхать смятые цветы. О, этот жест простой и ловкий!** Я помню горечь на устах, Да синеву татуировки На темно-бронзовых ногах. [Н. Н. Тuroверов. Легион (1940–1945)];

Ладонью защитив, он нес его так нежно. «Эй!» — и монет соблазна полный звон. Как удивленно оглянулся он! **Жест отрицания был царственно небрежен!** Раздумье... **Узких глаз мгновенный блеск и жар И лотос он, бедняк, мне протянул, как дар!** [Е. Н. Рачинская. Лотос (1945)];

Побывать с собою... Но без друга И на рыбалке будет тужо; Угрюмец лишь, таясь во мгле, Плывет в единственном числе: Он эгоист, он знает место, Он даже рыбу... продает! Тут **отстраняющего жеста Движение** — такой не в счет! [А. И. Несмелов. Встреча (1945)].

Жест может быть понят и в том случае, когда он никак не выражен лексически, но отсылает к внеязыковой действительности. Так, имя *Ленин* конкретизирует *жест* как жест протянутой руки, символизирующий призыв к движению в будущее:

В камне, в железе и в стали **Повторяет свой жест Ильич:** В необозримые дали Кидает безмолвный клич. [Т. Л. Щепкина-Куперник. Петербург... Ленинград (1928)];

Помнится сизое Утро в пыли. Вражьи дивизии С моря пришли. Чистеньких, грамотных Дикарей Встретил памятник Грудью своей! **Странная статуя... Жест — как сверло, Брови крылатые Гневом свело.** — Тонко сработано! Кто ж это тут? «ЛЕНИН». Ах, вот оно? Аб! — Гут! Дико из цоколя Высится шест. Грохнулся около **Бронзовый жест**, Кони хвостатые Взяли в карьер. Нет статуи, Гол сквер. Кончено! **Свержено!** [И. Л. Сельвинский. Баллада о ленинизме (1942)].

Как можно увидеть в приведенных примерах, в качестве конкретизаторов жеста нередко выступают определения (*широким жестом запахнув шинель*). Более того, для поэтического текста такая конкретизация жеста весьма характерна — диапазон используемых поэтами эпитетов необычайно широк. Но особенно выразительны эпитеты с окказиональной лексической сочетаемостью (*длинный жест, доверчивым жестом, прошлогодним жестом*). В ряде случаев именно эпитет-конкретизатор, в том числе несогласованное определение (с обозначением того, кто совершает жест, или того, кому уподобляется совершающий жест), становится единственным указанием на то, каков именно жест, как он совершается:

Родэн навеки заковал В **полубезумный жест Кентавра Несовместимость** двух начал. [М. А. Волошин. Письмо (1904)];

Здесь жизнь мертва, как гобелен старинный, / здесь радости и здесь печали нет; / льет полусвет причудливый кинкет / на каждый **жест изысканно-картинный**. [Эллис (Л. Л. Кобылинский). Сонеты-гобелены. VII (1905–1913)];

Везде борьба со скукой, Всё тот же смех и блески тех же звезд, И там, как здесь, мне будет сладкой мукой **Твой тихий жест**. [М. И. Цветаева. Привет из вагона (1910)];

О, как лелеял я гниющие мечты: Твой аромат, твой стыд, безвольность наготы, **Развратный, властный жест**, молящие черты... [А. К. Лозина-Лозинский. «Я знаю, я умру, когда мои желанья...» (1916)];

Пока труба — тарі́рора — грассировала тра-тара, И разбухался облаком под лошадыми снег, Вылез пейзажин **с жестами оратора** С затертыми потечками от лапок пенсне. [И. Л. Сельвинский. Улялаевщина, 3 (1924)];

Знаменитость Колокольным литьем командора шагала по плитам, **Тяжелозвонкими жестами** грохая трость. [И. Л. Сельвинский. Записки поэта, 38 (1925)];

Ты огни в отчужденны колеблешь, Точно улицу вертишь в руке. // И в блуждании хлопьев кутежных Тот же **гордый, уклончивый жест**: Как собой недовольный художник, Отстраняешься ты от торжеств. [Б. Л. Пастернак. Девятьсот пятый год, 1 (1925–1926)];

Я ветви яблонь приняла, / их **жест дающий и смиренный**, / почти к земле прикосновенный / изгиб крыла. [Е. И. Дмитриева (Черубина де Габриак). «Я ветви яблонь приняла...» (1926)];

Опять у вешалки, над тем же самым местом, Увидишь в зеркале знакомый поворот; Все **та же дама прошлого**ним жестом Все так же красит **прошлого**нный рот. [Саша Черный. Размышления у подъезда «Лютеции» (1927)];

На затылок / **нежным жестом** / он / **кудрей** / **закинул шелк**, / стал / барашком **златошерстым** / и заблеял, / и пошел. [В. В. Маяковский. Птичка божия (1929)];

А из пены родившись, Венера, / **пьяным жестом щеку подперев**, / эту ночь принимает на веру, / будет слушать покорно хорей. [Э. К. Чегринцева. «Зубочистки грозят, как рапиры...» (1929–1936)];

Ему бы ворочать Устои основ Буржуазного старого мира, **Сваи вбивать Кулачищами слов, Жестом борца-командира**. [В. В. Каменский. Юность Маяковского (1931)];

В **легких жестах** — ласковость кошек, Но без вкрадчивости. [А. И. Немецлов. Полдень (1931)];

И сытый друг мой, **жестами сквалыги** Сгребая крошки со стола, мычал: «Недаром хлеб — начало всех начал, Что видно из фольклора и религий»... [М. А. Тарловский. Веселый странник, 6 (1935)];

И в этом **жесте** / **мальчишеском, гордом**, / который движенье и радость таит, / хотел бы я, / чтоб стал он / над городом, / как в памяти нынче в моей / он стоит. [Н. Н. Асеев. Маяковский начинается, 17 (1936–1939)];

И лоб горячий вытирали люди, **Одежды приподнявши легкий край. Смирненным детским жестом, полным чуда...** [Б. Б. Божнев. «С прохладною смешавшись тишиной...» (1939)];

Спортивным жестом он поднес фаянс Двум дамочкам, и ловко похватили Они, смеясь, сухие бутерброды. [В. А. Луговской. Первая свеча (1942–1956)].

Ср. также о нацистском приветствии: *Да это что такое?! Он вырос, вынес к небу подбородок. Он римским жестом руку протянул, Он зашагал. Куда? Во тьму Европы.* [В. А. Луговской. Сказка о том, как человек шел со смертью (1943–1956)].

В поэтическом тексте слово *жест* может включаться в сложный образ — задавать его, быть его основой, прежде всего, за счет введения слова в неоднородный ряд, который может включать самые разные единицы:

Всю цепь промчавшихся мгновений Я мог бы снова воссоздать: И робость медленных движений, И жест, чтоб ножик иль тетрадь Сдержатъ неловкими руками, И Вашу шляпку с васильками, Покатость Ваших детских плеч, И Вашу медленную речь, И платье цвета эвкалипта, И ту же линию в губах, Что у статуи Таиах, Царицы древнего Египта, И в глубине печальных глаз — Осенний цвет листвы — топаз. [М. А. Волошин. Письмо (1904)];

Когда в душе все скомкано и смыто Бездарной толчеей российских дел И мусором бескрасочного быта — Окрошкой глупых жестов, слов и тел, Стремись к тем, чьи взоры так беспечно Со всей земли цветные соки пьют И любят все, что временно и вечно... [Саша Черный. Профан (1911)];

Ни человека, ни людей (Живые, да, но кто и что?), а сколько жестов и вещей, Ужимок, зонтиков, пальто. [Н. А. Оцуп. «Раскачивается пакет...» (1924–1926)];

Кружилось, мерцало, мелькало, мчало, / сошел аппарат с ума; / кончалась и вновь начиналась сначала / не лента, а жизнь сама! Валились на зрителя метры и жесты / смешавшихся сцен и чувств... И думалось залу: в огне происшествий / не я ли уже верчусь? [Н. Н. Асеев. «И волосы вовсе не были рыжими...» (1924)];

От отца с закусочной кровопийцы Цвет он унаследовал и волос, а от матери — предплечий спицы, Жесты, и полет, и невеселость. [М. А. Тарловский. Случай на крыше (1926)];

Вы царица экрана и моды, Вы пушисты, светлы и нахальны, Ваши платья надменно-печальны, Ваши жесты смелы от природы. [А. Н. Вертинский. Испано-Сюзиза (1928)];

А как жуужжал, звеня, пропеллер! Пропеллер сердца. Отражали — Движенья губ, дрожанье стали. Захлебываясь, руки пели. Но выпит залпом, как аккордом, Не жест, не взмах. Последний звук Моих немых отныне рук, И мне осталось — шагом твердым. [Р. П. Спинадель. «Все изломать, все сжечь, все бросить...» (1928)];

В них все — / и неслышанность пестрой одёжи, / несдержанность жестов, / несогнутость плеч, — / за ними — / толпою поток молодежи, / а против них — «Русское слово» / и «Речь». [Н. Н. Асеев. Маяковский начинается, 4 (1936–1939)].

Показательно, что в рассмотренном корпусе поэтических текстов *жест* как обозначение телодвижения встречается много чаще, чем второе значение этого слова ‘поступок, совершаемый в знак чего-л.’ [СШ: 233] (ср. приведенное выше

толкование данного значения из БАСа), казалось бы, более книжное и тем самым более «поэтическое». Ср.:

*Но на прогулку не оденусь, Нарочно сделав томный вид И говоря: «Куда я денусь, Когда любовь меня томит?» И скажут все: «Он лицемерит, **То жесты позы, не любви**»; Лишь кто сумеет, тот измерит, Как силен яд в моей крови.* [М. А. Кузмин. «Туманный день пройдет уныло...» (1907)];

*В небо кинутый флаг — **Вызов смелого жеста**. Знак вражды и протеста, Восклицательный знак!* [М. И. Цветаева. Восклицательный знак (1913)];

*Из детских книг встают, в мечтах ожив, Герои, битвы, земли; в каждой фразе — **Отважный подвиг, смелый жест, порыв**.* [В. Я. Брюсов. «На дальней полке мирным строем стоя...» (1918)];

*Красное видение, в горячих цветах, **Жест восхваления в хороводных словах**, Полднем горящая дрожь настаёт, Солнце палящее жжёт хоровод.* [Т. Вечорка. Красные быки (1918)];

— Главное — пошибом Взять: для подобных бар **Жест — наилучший дар**. [М. И. Цветаева. Крысолов, 5 (1925)];

*То-то же, как на поверку Выйдем — стыд тебя заест: Совето-российский Вертер. **Дворяно-российский жест**.* [М. И. Цветаева. Маяковскому, 4 (1930)].

Впрочем, в некоторых произведениях можно видеть или предполагать контаминацию этих значений, ср.:

*Лгу оттого, что по кладбищам Метель метет... От скрипки — от автомобиля — Шелков, огня... От пытки, что не все любили Одну меня! От боли, что не я — невеста У жениха... **От жеста и стиха — для жеста И для стиха!*** [М. И. Цветаева. «Безумье — и благоразумье...» (1915)];

*Положи мне на сердце ладонь / и внимательно слушай: / видишь — бьется на сколько ладов, / то отчетливее, то глуше... Это вовсе / **не жалобный жест**, / не желанье / смутить и растрогать, / и к тебе — / не расчетливый лжец / прижмается / локтем об локоть* [Н. Н. Асеев. Через головы критиков (1926)];

*Под лестницею химик с адвокатом Свой прошлогодний продолжают спор, Седой шофер землячке-массажистке Диктует три рецепта пельменей, а внучка бабушке — **о лебединый жест!** — Свой карандаш передает пунцовый...* [Саша Черный. Завтра (1930)].

Особого внимания заслуживают случаи, когда слово *жест* становится ключевым, мотиво- и смыслообразующим в стихотворении, иногда пронизывая весь текст (здесь *жест* может представлять в разных значениях). В таких текстах хорошо видна индивидуальная поэтика слова *жест*, включающая и те его особенности, о которых говорилось в статье. Ср. стихотворения (приведенные в сокращении или целиком):

*Благословив его на муку, Склонившись, как идут к гробам, **Ты, как святыню, принца руку, Бледнея, поднесла к губам**. // И опустили принца веки, И понял он без слов, в тиши, **Что этим жестом вдруг навеки Соединились***

две души. <...> И вот, великой силой жеста, Вы стали до скончания лет Жених и бледная невеста, Хоть не был изречен обет. [М. И. Цветаева. Камерата (1906–1912)].

Кто-то едет — к смертной победе. У деревьев — жесты трагедий. Иудеи — жертвенный танец! У деревьев — трепеты таинств. // Это — заговор против века: Веса, счета, времени, дробь. Се — разодранная завеса: У деревьев — жесты надгробий... // Кто-то едет. Небо — как въезд. У деревьев — жесты торжеств. [М. И. Цветаева. Деревья, 8 (1923)].

*«Дай руку. Спокойно... Мы в громе и мгле Стоим на летящей куда-то земле. <...> Хорошая спортсменка, мой моральный доктор, Однажды сказала, злясь и горячась: «**Никогда не ведите движений от локтя — Давайте движение всегда от плеча!**» <...> Но будут ли газы ползти по ночам, Споят ли басы орудийного рокота, — **Давайте стремительный жест от плеча, Никогда не ведите движений от локтя!** <...> Да! Может не время об этом кричать, Не время судьбе самолетами клеать, Но будем движения вести от плеча, Широко расставя упрямые локти!» [В. А. Луговской. Кухня времени (1929)].*

*За эту улыбку, за радость при встрече (За безудержной дружбы альпийский озон...), <...> За все разговоры, за все умолчания (Осторожную нежность... подкожную дрожь...), За **радость — скупую, как жест англичанина**... (Оттого, что ты где-то грустишь и живешь...), <...> За эту улыбку... И, все ж, я согласен За детскую гордость, за тайный испуг, За взгляд твой — в упор! — что зверинно-прекрасен. За **жест некрасивых любимейших рук**. За что-то, о чем и не вспомнится вдруг... За все, за ничто, нерасчетливый друг.* [Д. Кнут. Расплата (1934)].

Рассмотренный материал показывает в целом, как русская поэзия осмысляет понятие жеста, как многогранно, с сохранением семантического ядра, эксплицируется в ней само слово *жест*, как у некоторых поэтов в раскрытии, развитии мотива жестовости оно становится ключевым словом.

Литература

БАС-3 — Большой академический словарь русского языка. Т. 5 / Гл. ред. К. С. Горбачевич. М.: Наука, 2006. 694 с.

Бобринская Е. Жест в поэтике раннего русского авангарда // Хармсиздат представляет. Авангардное поведение. Сб. материалов науч. конф. Хармс-фестиваля 4 в Санкт-Петербурге. СПб., 1998. С. 49–62.

Григорьева С. А., Григорьев Н. В., Крейдлин Г. Е. Словарь языка русских жестов. М.: Языки славянской культуры, 2001. 257 с.

Гришина Е. А. Закрытые глаза: жестикуляция, пунктуация, акцентуация // Филология. 2011. № 2. С. 39–56.

Гришина Е. А. Указания рукой как система (по данным Мультимедийного русского корпуса) // Вопросы языкознания. 2012. № 3. С. 3–50.

Гришина Е. А. Указания головой как система // Вопросы языкознания. 2013. № 3. С. 90–130.

Гришина Е. А. Русская жестикуляция с лингвистической точки зрения (корпусные исследования). М.: Издательский Дом ЯСК, 2017. 744 с.

Крейдлин Г. Е. Невербальная семиотика. М.: НЛЮ, 2002. 581 с.

Крейдлин Г. Е. Мужчины и женщины в невербальной коммуникации. М.: Языки славянской культуры, 2005. 224 с.

Масленникова О. Н. Семантика жеста в прозе Андрея Белого: «Петербург», «Котик Летаев», «Крещеный китаец»: дис. ... канд. филол. наук / Ивановский гос. ун-т. Иваново, 2000. 178 с.

НКРЯ — Национальный корпус русского языка [Электронный ресурс]. URL: <http://www.ruscorpora.ru>.

Раков В. П. Семантика жеста в поэзии Марины Цветаевой // Филологические штудии. Сборник научных трудов. Иваново, 1995. С. 5–15.

Слово и жест в литературе: сб. науч. тр. / С. Вайман (гл. ред.). Воронеж: ВГПИ, 1983. 107 с.

СШ — Шведова Н. Ю. (ред.). Толковый словарь русского языка с включением сведений о происхождении слов. М.: Азбуковник, 2007. 1175 с.

СЯП — Словарь языка Пушкина: в 4 т. 2-е изд., доп. М.: Азбуковник, 2000.

СЯРП — Словарь языка русской поэзии XX века. Т. I — / Сост.: Григорьев В. П. (отв. ред.), Шестакова Л. Л. (отв. ред.), Колодяжная Л. И. (ред.), Кулева А. С. (ред.), Бакеркина В. В., Гик А. В., Реутт Т. Е., Фатеева Н. А. М.: Языки славянской культуры, Знак, 2001—.

Федунина О. В. Поэтика жеста в литературном произведении (наблюдения и комментарии) // Известия Самарского научного центра Российской академии наук. Социальные, гуманитарные, медико-биологические науки. 2017. Т. 19. № 1. С. 114–117.

L. L. Shestakova¹, A. S. Kuleva²

V. V. Vinogradov Russian Language Institute,

Russian Academy of Sciences

(Russia, Moscow)

lara.shestakova@mail.ru¹, an_kuleva@mail.ru²

GESTURE IN POETIC DISCOURSE

The word *gesture* in the Russian poetic language does not belong to high-frequency ones, but examples of its use allow us to make interesting observations. Cases of the use of the word *gesture* in poetic texts are considered through the prism of reputable linguistic studies on Russian gesticulation, as well as works of literary, linguopoetic orientation, in which gestures are understood as signs, images, part of a system of symbols, including

for individual authors. The multivolume “Dictionary of the language of Russian poetry (20th century)” (Moscow, 2001, 2003, 2008, 2010, 2013, 2015, 2017, 2019, 2021, 2022) is based on works of ten outstanding poets of the Silver Age: I. Annensky, A. Akhmatova, A. Blok, S. Yesenin, M. Kuzmin, O. Mandelstam, V. Mayakovsky, B. Pasternak, V. Khlebnikov, M. Tsvetaeva. This dictionary contains about forty contexts with this word (the largest number is noted in Tsvetaeva). There are more than 450 occurrences in the Poetry corpus of the Russian National Corpus. The article analyzes different types of poetic use of the word *gesture*. In some cases, it is used without specifiers (to denote a gesture as such, an indefinite gesture), and often — in a number of names of other non-verbal manifestations; often the word *gesture* is clarified by indicating the way in which the body movement is performed. In poetic discourse, attributive refinements are frequent, which characterize a *gesture* from one side or another. The cases of figurative interpretation of the word *gesture* are of particular interest.

Key words: author lexicography, dictionary of poetic language, gesture, poetic discourse, Russian National Corpus

References

Bol'shoi akademicheskii slovar' russkogo yazyka [Large academic dictionary of the Russian language]. Vol. 5. St. Petersburg, Nauka Publ., 2006. 694 p.

Bobrinskaya E. Zhest v poetike rannego russkogo avangarda. *Avangardnoe povedenie. Sb. materialov nauch. konf. Harms-festivalia 4*. SPb: M. K. & Harmsizdat, 1998, pp. 49–62. (In Russ.)

Grigor'eva S. A., Grigor'ev N. V., Krejdlin G. E. *Slovar' yazyka russkih zhestov*. [Dictionary of Russian Gestures]. Moscow: Yazyki slavyanskoi kul'tury Publ., 2001. 257 p.

Grishina E. A. [Closed eyes: Gesticulation, punctuation, accentuation]. *Filologiya*. 2011, no. 2, pp. 39–56. (In Russ.)

Grishina E. A. [Hand indications as a system (according to the Multimedia Russian corpus)]. *Voprosy yazykoznaniya*. 2012, no. 3, pp. 3–50. (In Russ.)

Grishina E. A. [Head indications as a system]. *Voprosy yazykoznaniya*. 2013, no. 3, pp. 90–130. (In Russ.)

Grishina E. A. *Russkaya zhestikulyatsiya s lingvisticheskoi tochki zreniya (korpusnye issledovaniya)* [Russian gestures from a linguistic perspective: A collection of corpus studies]. Moscow, Yazyki slavyanskoi kul'tury Publ., 2017. 744 p. (In Russ.)

Kreidlin G. E. *Neverbal'naya semiotika* [Nonverbal semiotics]. Moscow, NLO Publ., 2002. 581 p. (In Russ.)

Krejdlin G. E. *Muzhchiny i zhenshchiny v neverbal'noi kommunikacii* [Men and women in non-verbal communication]. Moscow: Yazyki slavyanskoi kul'tury, 2005. 224 p. (In Russ.)

Maslennikova O. N. *Semantika zhesta v proze Andreja Belogo: “Peterburg”, “Kotik Letaev”, “Kreshhenyj kitaec”*. PhD Thesis. Ivanovo, Ivanovo State Univ., 2000. 178 p. (In Russ.)

Natsional'nyi korpus russkogo yazyka [Russian National Corpus]. Available at: <http://www.ruscorpora.ru>.

Rakov V. P. Semantika zhesta v poezii Mariny Cvetaevoy. *Filologicheskie shtudii. Sbornik nauchnykh trudov*. Ivanovo, 1995, pp. 5–15. (In Russ.)

Slovar' yazyka Pushkina. V 4 t. [Pushkin's language dictionary]. Moscow, Azbukovnik Publ., 2000.

Slovo i zhest v literature: sb. nauch. tr. Vajman S. (ed.). Voronezh, 1983. 107 p. (In Russ.)

Shvedova N. Yu. (ed.). *Tolkovy slovar' russkogo yazyka s vklyucheniem svedenii o proiskhozhdenii slov* [Dictionary of the Russian language with the inclusion of information about the origin of words]. Moscow, Azbukovnik Publ., 2007. 1175 p.

Slovar' yazyka russkoi poezii 20 veka [Dictionary of the language of Russian poetry (20th century)]. Grigoriev V. P., Shestakova L. L., Kolodyazhnaya L. I., Kuleva A. S., Bakerkina V. V., Gik A. V., Reutt T. E., Fateeva N. A. (eds.). Vol. 1–. Moscow, Yazyki Slavyanskoi Kul'tury Publ., Znak Publ., 2001–.

Fedunina O. V. [The poetics of gesture in literary works (remarks and comments)]. *Izvestiya Samarskogo nauchnogo tsentra Rossiiskoi akademii nauk. Sotsial'nye, gumanitarnye, mediko-biologicheskie nauki*. 2017, no. 1 (19), pp. 114–117. (In Russ.)

III. РЕЦЕНЗИИ

[Рец. на / Review of:] **Jesse Egbert, Tove Larsson, Douglas Biber. *Doing Linguistics with a Corpus. Methodological Considerations for the Everyday User***. Cambridge University Press, 2020. 88 p. ISBN 978-1-108-74485-0, ISSN 2632-8097 (online), ISSN 2632-8089 (print)

М. В. Хохлова

*Санкт-Петербургский
государственный университет
(Россия, Санкт-Петербург)
m.khokhlova@spbu.ru*

М. V. Khokhlova

*Saint-Petersburg State University
(Russia, St. Petersburg)
m.khokhlova@spbu.ru*

Коллективная монография Джесс Эгберт (Jess Egbert), Туве Ларссон (Tove Larsson) и Дугласа Байбера (Douglas Biber) посвящена практическим вопросам, связанным с корпусными исследованиями. Работа продолжает серию изданий «Элементы корпусной лингвистики» (“Elements in Corpus Linguistics”), которые знакомят читателей с методологической базой исследований. Несмотря на то, что вопросам корпусного анализа уделялось большое внимание в учебной и научной литературе, они не теряют своей актуальности по нескольким причинам. С одной стороны, появляются разнообразные корпуса текстов и становятся доступными новые инструменты, которые могут быть использованы при проведении исследований. С другой стороны, развивается и усложняется статистический аппарат, результаты применения которого при неправильном подходе могут привести к ошибкам. В связи с этим авторы монографии пытаются ответить на ряд вопросов, связанных с: 1) выбором корпуса для исследования; 2) формулировкой гипотез и определением единиц для наблюдения; 3) оценкой применимости количественных методов и программных инструментов для решения поставленных задач. Книга рассчитана на новичков в области, но при этом может быть полезна и сложившимся исследователям. Авторы отдают себе отчет в том, что в работе рассматриваются некоторые известные базовые понятия, однако обращение к затрагиваемым вопросам и их обсуждение может способствовать большему пониманию и у широкой аудитории.

Книга состоит из 8 глав, приложения и списка литературы. Каждая глава сопровождается примерами исследований, которые относятся к определенной теме и подробно разбираются авторами, а также перечнем основных выводов. Такая структура является несомненным достоинством книги: приведенные выводы помогают обобщить информацию, которая затрагивалась в рассматриваемой главе, а также обратить внимание специалистов на ключевые вопросы, в том числе дать рекомендации по проведению корпусного анализа.

Во введении обосновывается актуальность работы, которая связана с возрастающими техническими возможностями для обработки текстов, а также новыми методами корпусного анализа. Авторы находчиво проводят интересную параллель между работой с современными корпусами текстов и вождением автомобиля. Несмотря на кажущуюся простоту, которую в современных машинах обеспечивает электроника, водителю всё равно необходимо уметь разбираться самостоятельно в некоторых вопросах, чтобы избежать аварий, а в случае работы с корпусами — серьезных ошибок в исследовании. Поэтому при проведении лингвистического анализа и для последующей корректной интерпретации результатов важно понимать, как устроен корпус, с которым работает пользователь, а также правильно формулировать исследовательские вопросы. Между тем существующие технологии снабжают ученых необходимыми ресурсами для проведения исследования, однако было бы ошибкой ожидать, что получение результатов при помощи автоматических инструментов означают достижение конечной цели. Как раз в этом заключается риск, поэтому необходимо предпринять дополнительные шаги, чтобы убедиться в том, что на все сформулированные в начале проведения исследования вопросы были получены ответы.

Глава 2 посвящена знакомству с корпусом и затрагивает важные вопросы «корпусостроения». Например, вопрос объема данных: способствует ли увеличение количества текстов улучшению качества корпуса? Какова должна быть структура корпуса и каким должно быть его наполнение? С размером корпуса связаны такие понятия как сбалансированность и репрезентативность, которые оказывают влияние на возможность получить более корректные статистические данные и обеспечить разнообразие в представлении данных. В связи с этим можно для конкретного исследования составлять новый корпус или использовать уже существующий. Авторы склоняются к первому подходу, однако отдают себе отчет, что это является затратным по времени и ресурсам, поэтому чаще всего исследователи придерживаются второго подхода — работают с уже собранными данными. Поэтому важно, какие тексты содержатся в корпусе и в каком процентном соотношении — ведь именно на основе выводов, сделанных для этой выборки, можно прийти к обобщениям, справедливым для той генеральной совокупности, которую она репрезентирует. Таким образом, перед началом исследования необходимо ознакомиться с метаданными, документацией, а также самими текстами корпуса. Авторы приводят пример исследования номинализации и соединительных наречий в научных текстах. В качестве материала предварительно были выбраны Корпус письменных текстов британского академического английского (British Academic Written English, BAWE), академический подкорпус Британского национального корпуса (British National Corpus, BNC) и академический подкорпус Корпуса современного американского английского (Corpus of Contemporary American English, COCA). В первом представлены только неопубликованные работы студентов, что делает его непригодным для последующей работы. Тексты из BNC хорошо задокументированы, однако более подробно описан материал в корпусе COCA. Хотя более подробный анализ, проведенный авторами, показал, что последний содержит не только

научные статьи, но и записи лекций, а также научно-популярные тексты, при этом недостаточно полно представлена информация о методах, которые использовались создателями СОСА для определения того, подходит ли источник. В связи с этим полученные количественные данные для рассмотренных единиц в обоих корпусах могут статистически значимо отличаться по сравнению с тем, если бы отбор текстов был произведен согласно одним и тем же принципам. Авторы подчеркивают, что нельзя утверждать, что один корпус превосходит другой по своему качеству, однако исследователь должен отдавать себе отчет в том, с каким материалом он работает, чтобы правильно интерпретировать результаты. В качестве вывода отмечается, что состав корпуса оказывает влияние на его лингвистические характеристики. Это важно для определения соответствия корпуса целевой области исследования, а также оценки того, представляет ли корпус достаточно полный спектр из этой целевой области. Поскольку конечная цель корпусных исследований связана с получением знания о том, как используется язык, то корпуса должны дать возможность сделать такие выводы.

Третья глава посвящена тому, как связать количественный корпусный анализ и лингвистическое описание, поэтому важно грамотно разработать план исследования, а также сформулировать вопросы. Эти два пункта являются взаимосвязанными: вопросы исследования определяют его план, и наоборот — если данные собраны в соответствии с планом исследования, их можно использовать только для ответа на определенные типы вопросов. Начинаящие исследователи грешат тем, что собирают данные в соответствии с одним типом исследования, но далее их используют для ответа на другие вопросы. Авторы выделяют *единицы наблюдения* — единицы языка, или языковые объекты (слова, грамматические признаки), на которых проводятся измерения. В то время как *переменные* измеряют лингвистические характеристики этих наблюдений. В главе проводится разграничение между двумя типами плана исследования: вариационным и описательным. Первый тип связан с понятием категорий, а не количественными значениями (например, исследование прямого дополнения или субъекта). Второй тип представлен цельнокорпусным и текстово-лингвистическим видами исследования. Цельнокорпусное исследование является более простым по своей структуре: в нем вычисляется встречаемость лингвистических признаков в разных корпусах. Текстово-лингвистический подход предполагает, что каждый текст представляет собой единицу наблюдения, поэтому в рамках него можно ответить на вопрос, равномерно ли распределен признак по всем текстам. В результате такой план исследования позволяет вычислить как среднюю частоту встречаемости по всем текстам корпуса, так и меру дисперсии по текстам, которая показывает, в какой степени лингвистическое явление равномерно распределено по текстам. Одной из мер дисперсии является коэффициент Жуйана, который показывает, насколько универсально или, наоборот, специализированно используется слово в текстах различной тематики: если оно встречается только в определенных источниках, то его значение приближается к 0, если частота словоупотребления примерно одинакова по всем текстам, то значение коэффициента близко к 100.

В главе 4 обсуждается взаимозависимость между лингвистически интерпретируемыми переменными и интерпретируемостью результатов, а также подчеркивается важность грамотного определения для исследуемых единиц. Авторы считают, что корпусный анализ коллокаций, который часто основывается на сложных статистических показателях, можно с трудом интерпретировать в лингвистических терминах. Разнообразие метрик как по их количеству, так и уровню сложности также является проблемой для лингвистической интерпретации полученных результатов. Еще одним примером послужил анализ ключевых слов: при его проведении традиционные меры выделяют лексику, основываясь на частоте во всем корпусе. Однако авторы предлагают обратить внимание на те метрики для нахождения ключевых слов, которые учитывают дисперсию, что поможет найти именно те слова, которые характерны для предметной области, а не для определенного текста или не для корпуса в целом (о чем обычно свидетельствуют высокие значения частотностей).

В главе 5 авторы указывают на тот факт, что при применении автоматических инструментов анализа результаты часто интерпретируются некорректно. Поэтому важно, чтобы подбираемые статистические критерии соответствовали особенностям исследуемых данных. По мнению авторов, результаты применения некоторых методов представляют сложность для последующей интерпретации, при этом многие инструменты демонстрируют свою эффективность только на материале английского языка или для специализированных узких задач. Например, автоматический анализ сложности текстов при помощи метрик удобочитаемости хотя и может демонстрировать некоторые статистические характеристики, однако нет уверенности в том, насколько подобные результаты являются показательными. В качестве примера рассматривается статья про сложность грамматических конструкций в текстах на родном и изученном английском языке. Используемая в работе мера опирается на наличие и число подчинённых предложений, но не учитывает их структуру и уровень вложенности. Результаты применения одного автоматического инструмента часто трудно интерпретировать, однако привлечение большего числа инструментов может приводить к ложным или противоречащим друг другу выводам, а также результаты работы автоматического анализа требуют последующей сложной ручной обработки. При этом авторы оптимистичны в своих прогнозах и советуют исследователям не отказываться от автоматических инструментов, однако рассматривать их применение только как выполнение первого этапа работы.

Глава 6 связана с рассмотрением более проблемного вопроса: как более сложные статистические методы анализа могут иногда создавать ненужную дистанцию между количественным анализом и описанием языковых явлений. В ней рассматриваются методы описательной статистики (например, процентное соотношение, среднее значение и стандартное отклонение) для количественного представления данных, а также уделяется внимание методам аналитической статистики (статистики вывода) (например, хи-квадрат, регрессионный анализ), которые позволяют сделать умозаключения о тенденциях в языке. Однако опасность может заключаться в том, чтобы интерпретировать результаты статистических методов (даже дескриптивных)

в отрыве от самих данных. Поэтому важно, во-первых, использовать по возможности простые и правильно выбранные статистические методы, а во-вторых, после проведения статистического анализа стоит вернуться к лингвистическим данным для интерпретации результатов статистических исследований. Так, одинаковые значения при использовании среднего арифметического на разных данных могут привести к ложному выводу о том, что они распределены равномерно по корпусам. В главе приводится таблица с информацией о том, как статистические методы могут быть сгруппированы по уровню сложности и типам анализа. Так, например, при использовании линейной модели часто исключаются выбросы или объединяются категории с небольшим количеством данных, в то время как при логарифмировании снижается возможность интерпретации лингвистических данных. Авторы предлагают не опираться только на значение уровня значимости (*p-value*), а применять методы описательной статистики, которые могут быть прямым образом интерпретируемы, и дополнить их вычислением размера эффекта (например, коэффициента Козна).

В главе 7 подчеркивается необходимость сочетать количественный корпусный лингвистический анализ и качественную интерпретацию, а также иллюстрировать все количественные модели на материале полученных примеров. Большинство исследований не демонстрируют подробного качественного анализа для ответа на поставленные исследовательские вопросы. Так, рассмотренные авторами списки частотных коллокатов для “man” («мужчина») и “woman” («женщина») нуждались в дополнительном изучении контекстов, чтобы определить характерную сочетаемость, включающую семантические признаки. В качестве еще одного примера приводится исследование диахронических тенденций, для грамотной интерпретации которых необходимо учитывать социолингвистический контекст.

Наконец, глава 8 подводит итог и обобщает те сложные вопросы и возможности, которые связаны с квантитативной корпусной лингвистикой. По мнению авторов, не стоит слепо доверять тем результатам, которые пользователь получает в корпусе. Так, результаты автоматического морфологического анализа могут содержать ошибки, а корпус академических текстов может включать и иные источники. Авторы ведут блог под названием «Лингвистика с корпусом» (“Linguistics with a Corpus”), в котором регулярно публикуют посты по темам, которые затрагивались в этой работе.

Авторы рецензируемой коллективной монографии обосновывают свои выводы, снабжая их примерами и пояснениями. Интересно отметить тот факт, что работа нацелена не только на пользователей корпусов, но и на самих разработчиков, поскольку указывает на слабые места в существующих проектах. Весьма любопытны критические замечания, имеющие отношение к *case studies*, которые разбирают авторы. Часть из них может способствовать полемике, однако заставляет критически относиться к проводимым исследованиям.

Научный журнал

**Труды Института русского языка
им. В. В. Виноградова**

2024 г., № 4 (42)

Макет: *С. В. Родионова*

Гарнитура ZRCola. Формат 70×100/16

Бумага офсетная. Печать цифровая

Печ. л. 21,0

Тираж 300 экз. Заказ №