

# **Компьютерная лингвистика и интеллектуальные технологии**

По материалам ежегодной международной  
конференции «Диалог» (2017)

Выпуск 16

Том 2 из 2

Компьютерная лингвистика:  
лингвистические исследования

# **Computational Linguistics and Intellectual Technologies**

Papers from the Annual International  
Conference “Dialogue” (2017)

Issue 16

Volume 2 of 2

Computational Linguistics: Linguistic Research

УДК 80/81; 004  
ББК 81.1  
К63

Редакционная  
коллегия:

*В. П. Селегей (главный редактор), А. В. Байтин,  
В. И. Беликов, И. М. Богуславский, Б. В. Добров,  
Д. О. Добровольский, Л. М. Захаров, Л. Л. Йомдин,  
И. М. Кобозева, Е. Б. Козеренко, М. А. Кронгауз,  
Н. И. Лауфер, Н. В. Лукашевич, Д. Маккарти, П. Наков,  
Й. Нивре, Г. С. Осипов, А. Ч. Пиперски, В. Раскин,  
Э. Хови, С. А. Шаров, Т. Е. Янко*

Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог» (Москва, 31 мая — 3 июня 2017 г.). Вып. 16 (23): В 2 т. Т. 2 — М.: Изд-во РГГУ, 2017.

Сборник включает 71 доклад международной конференции по компьютерной лингвистике и интеллектуальным технологиям «Диалог 2017», представляющих широкий спектр теоретических и прикладных исследований в области описания естественного языка, моделирования языковых процессов, создания практически применимых компьютерных лингвистических технологий.

Для специалистов в области теоретической и прикладной лингвистики и интеллектуальных технологий.

© Редколлегия сборника «Компьютерная лингвистика и интеллектуальные технологии» (составитель), 2017

## Предисловие

16-й выпуск ежегодника «Компьютерная лингвистика и интеллектуальные технологии» содержит избранные материалы 23-й международной конференции «Диалог». На основании мнений наших рецензентов для публикации в ежегоднике Редсоветом был отобран 71 доклад из числа примерно ста работ, которые были рекомендованы по результатам рецензирования для представления на конференции в 2017 году.

Работы в сборнике отражают все основные направления исследований в области компьютерного моделирования и анализа естественного языка, представленные на конференции:

- Компьютерные лингвистические ресурсы
- Компьютерный анализ документов (классификация, поиск, анализ тональности и т.д.)
- Корпусная лингвистика (создание, разметка, методики применения и оценка корпусов)
- Лингвистические онтологии и автоматическое извлечение знаний
- Лингвистический анализ Social media
- Лингвистический анализ речи
- Машинный перевод текста и речи
- Модели и методы семантического анализа текста
- Модели общения
- Теоретическая и компьютерная лексикография
- Типология и компьютерная лингвистика
- Формальные модели языка и их применение в компьютерной лингвистике

В соответствии с традициями «Диалога», старейшей и крупнейшей конференции по компьютерной лингвистике в России, отбор работ основывается на представлении о важности соединения новых методов и технологий анализа языковых данных с полноценным лингвистическим анализом и моделированием. Одной из важнейших целей конференции была и остается поддержка создания современных компьютерных ресурсов, моделей и технологий для русского языка.

В годовом цикле проведения конференции в рамках программы Dialogue Evaluation проводится тестирования технологий решения отдельных задач компьютерного анализа языка. На конференции подводятся итоги проведенных тестов, а статьи организаторов и наиболее успешных участников представляются в настоящем сборнике.

В этом году было проведено два тестирования:

1. По идентификации внешних заимствований (External Plagiarism Detection)
2. По оценке методов морфологического анализа русского языка, с акцентом на тексты Social Media.

Как обычно, результатом проведенных тестирований стали не только объективные данные о качестве работы различных методов и алгоритмов, но также и открытые для использования эталонные размеченные корпуса, т. н. золотые стандарты, позволяющие любым исследователям проводить сравнительные оценки эффективности своих технологий.

Все направления «Диалога» важны, но каждый год какие-то темы занимают особое место в программе конференции и в составе ежегодника. В этом году можно назвать две таких темы:

1. Применение методов глубинного машинного обучения: прежде всего — нейросетей и таких результатов их применения как word embeddings, как для прикладных задач, так и в лингвистических исследованиях.
2. В программе конференции этого года особенно заметны работы по использованию параллельных корпусов для лингвистических исследований. Такие корпуса уже давно и успешно используются в NLP, например, для обучения статистических моделей машинного перевода, автоматической дизамбигуации, автоматического построения языковых моделей. Но параллельные корпуса оказываются также и важным инструментом контрастных лингвистических исследований.

Статьи в сборнике публикуются на русском и английском языках. При выборе языка публикации действует следующее правило:

- доклады по компьютерной лингвистике должны подаваться на английском языке. Это расширяет их аудиторию и позволяет привлекать к рецензированию международных экспертов.
- доклады, посвященные лингвистическому анализу русского языка, предполагающие знание этого языка у читателя, подаются на русском языке (с обязательной аннотацией на английском).

Несмотря на традиционную широту тематики представленных на конференции и отобранных в сборник докладов они не могут дать полной картины направлений «Диалога». Ее можно получить с помощью сайта конференции [www.dialog-21.ru](http://www.dialog-21.ru), на котором представлены обширные электронные архивы «Диалогов» последних лет и все результаты проведенных тестирований Dialogue Evaluation.

Мы обращаем внимание авторов и читателей сборника, что его бумажный вариант, который вы держите в руках, является вторичным по отношению к сборнику, который размещается на сайте конференции и индексируется Scopus. Мы рекомендуем при цитировании использовать именно сетевую версию.

*Программный комитет конференции «Диалог»*

*Редколлегия сборника «Компьютерная лингвистика  
и интеллектуальные технологии»*

## Организаторы

Ежегодная конференция «Диалог» проводится под патронажем Российского фонда фундаментальных исследований при организационной поддержке компании АBBYY.

Учредителями конференции являются:

- Институт лингвистики РГГУ
- Институт проблем информатики РАН
- Институт проблем передачи информации РАН
- Компания АBBYY
- Филологический факультет МГУ

Конференция проводится при поддержке Российской ассоциации искусственного интеллекта.

## Международный программный комитет

Богуславский Игорь Михайлович	Институт проблем передачи информации РАН им. А. А. Харкевича, Россия
Буате Кристиан	Университет Жозефа Фурье — Гренобль 1, Франция
Гельбух Александр Феликсович	Национальный политехнический институт, Мехико
Иомдин Леонид Лейбович	Институт проблем передачи информации РАН им. А. А. Харкевича, Россия
Кобозева Ирина Михайловна	Московский государственный университет им. М. В. Ломоносова, Россия
Козеренко Елена Борисовна	Институт проблем информатики РАН, Россия
Корбетт Гревил	Университет Суррея, Великобритания
Кронгауз Максим Анисимович	НИУ «Высшая школа экономики», Россия
Лукашевич Наталья Валентиновна	НИВЦ МГУ им. М. В. Ломоносова, Россия
Маккарти Диана	Кембриджский университет, Великобритания
Мельчук Игорь Александрович	Монреальский университет, Канада
Нивре Йоаким	Уппсальский университет, Швеция
Ниренбург Сергей	Университет Мэриленда, Балтимор, США
Осипов Геннадий Семёнович	Институт системного анализа РАН, Россия
Раскин Виктор	Университет Пердью, США
Селегей Владимир Павлович	Компания АBBYY, Россия
Хови Эдуард	Университет Карнеги — Меллон, США
Шаров Сергей Александрович	Университет Лидса, Великобритания

## Организационный комитет

Селегей Владимир Павлович, <i>председатель</i>	Компания АBBYУ
Байтин Алексей Владимирович	Компания Yandex
Беликов Владимир Иванович	Институт русского языка им. В. В. Виноградова РАН
Браславский Павел Исаакович	Уральский федеральный университет
Добров Борис Викторович	НИВЦ МГУ им. М. В. Ломоносова
Захаров Леонид Михайлович	Московский государственный университет им. М. В. Ломоносова
Иомдин Леонид Лейбович	Институт проблем передачи информации РАН им. А. А. Харкевича
Кобозева Ирина Михайловна	Московский государственный университет им. М. В. Ломоносова
Козеренко Елена Борисовна	Институт проблем информатики РАН
Лауфер Наталия Исаевна	Компания Yandex
Ляшевская Ольга Николаевна	Институт русского языка им. В. В. Виноградова РАН
Толдова Светлана Юрьевна	НИУ «Высшая школа экономики»
Федорова Ольга Викторовна	Московский государственный университет им. М.В. Ломоносова
Шаров Сергей Александрович	Университет Лидса

## Секретариат

Атясова Анастасия Леонидовна, <i>координатор оргкомитета</i>	Компания АBBYУ
Белкина Александра Андреевна, <i>секретарь оргкомитета</i>	Компания АBBYУ
Гусева Анна Александровна, <i>координатор Dialogue Evaluation</i>	Компания АBBYУ
Севергина Екатерина Александровна, <i>администратор оргкомитета</i>	Компания АBBYУ

## Рецензенты

Августинова Тая  
Антонова Александра Александровна  
Азарова Ирина Владимировна  
Андрианов Андрей Иванович  
Апресян Валентина Юрьевна  
Архангельский Тимофей Александрович  
Байтин Алексей Владимирович  
Баранов Анатолий Николаевич  
Беликов Владимир Иванович  
Бенко Владимир  
Бердичевский Александр Сергеевич  
Богданов Алексей Владимирович  
Богданова-Бегларян Наталья Викторовна  
Богуславский Игорь Михайлович  
Бочаров Виктор Владиславович  
Браславский Павел Исаакович  
Васильев Виталий Геннадьевич  
Галинская Ирина Евгеньевна  
Галицкий Борис Александрович  
Гельбух Александр Феликсович  
Гецевич Юрий Станиславович  
Гращенко Павел Валерьевич  
Губин Максим Вадимович  
Даниэль Михаил Александрович  
Диконов Вячеслав Григорьевич  
Добров Борис Викторович  
Добровольский Дмитрий Олегович  
Добрушина Нина Роландовна  
Зализняк Анна Андреевна  
Захаров Виктор Павлович  
Захаров Леонид Михайлович  
Ильвовский Дмитрий Алексеевич  
Иомдин Борис Леонидович  
Иомдин Леонид Лейбович  
Катинская Анисья Юрьевна  
Клышинский Эдуард Станиславович  
Кибрик Андрей Александрович  
Князев Сергей Владимирович  
Кобозева Ирина Михайловна  
Козеренко Елена Борисовна  
Копотев Михаил Вячеславович  
Кортаев Николай Алексеевич  
Котельников Евгений Вячеславович  
Котов Артемий Александрович  
Кронгауз Максим Анисимович  
Левонтина Ирина Борисовна  
Лобанов Борис Мефодьевич  
Лопухин Константин Александрович  
Лукашевич Наталья Валентиновна  
Лютикова Екатерина Анатольевна  
Мисюрев Алексей Владимирович  
Наков Преслав  
Недолужко Анна Юрьевна  
Падучева Елена Викторовна  
Пазельская Анна Германовна  
Паперно Денис Аронович  
Панченко Александр Иванович  
Переверзева Светлана Игоревна  
Петрова Мария Андреевна  
Пивоварова Лидия Михайловна  
Пиперски Александр Чедович  
Подлесская Вера Исааковна  
Рахилина Екатерина Владимировна  
Скулачева Татьяна Владимировна  
Смирнов Иван Валентинович  
Селегей Владимир Павлович  
Слюсарь Наталия Анатольевна  
Соколова Елена Григорьевна  
Сомин Антон Александрович  
Сорокин Алексей Андреевич  
Сорокин Виктор Николаевич  
Старостин Анатолий Сергеевич  
Степанова Мария Евгеньевна  
Тихомиров Илья Александрович  
Толдова Светлана Юрьевна  
Турдаков Денис Юрьевич  
Урысон Елена Владимировна  
Федорова Ольга Викторовна  
Хохлова Мария Владимировна  
Циммерлинг Антон Владимирович  
Шаров Сергей Александрович  
Шелманов Артём Олегович  
Янко Татьяна Евгеньевна

## Contents<sup>1</sup>

### Компьютерная лингвистика: лингвистические исследования

Апресян В. Ю.

**Отрицательная и положительная поляризация:  
семантические источники** ..... 2

Апресян В. Ю., Шмелев А. Д.

**«Ксенопоказатели» по данным параллельных корпусов  
и современных СМИ: русское якобы** ..... 17

Бадрязлова Ю. Г.

**Опыт корпусного моделирования факторов метафоричности  
на примере русских глаголов** ..... 30

Баранов А. Н., Вознесенская М. М.

**Идиоматика пьянства в корпусном измерении** ..... 45

Bergelson M. B., Khudyakova M. V.

**Interaction and Empathy as Elements  
of Narrative Strategies in the Russian Clips Corpus** ..... 55

Блинова О. В.

**Побудительная реплика как элемент устного диалога:  
случай императива 1PI** ..... 68

Богданова-Бегларян Н. В., Блинова О. В., Мартыненко Г. Я., Шерстинова Т. Ю.

**Некоторые инвариантные характеристики  
русской разговорной речи: фонетика, морфология, синтаксис** ..... 82

Богуславская О. Ю., Иомдин Б. Л.

**Развитие полисемии у русских числительных  
и порядковых прилагательных** ..... 96

Добровольский Д. О., Левонтина И. Б.

**Дискурсивные частицы и способы их перевода:  
ну в романе Владимира Сорокина «Очередь»** ..... 106

Федорова О. В.

**Вклад кинетической составляющей в мультимодальную  
коммуникацию, или Еще раз о стратегиях описания танграмм** ..... 118

---

\* Доклады упорядочены по фамилии первого автора в соответствии с английским алфавитом.  
The reports of each section are ordered by the surname of the first author in compliance with  
the English alphabet.



Гращенко П. В., Кобозева И. М. <b>Семантические классы и управление прилагательных</b> .....	134
Инькова О. Ю. <b>Принципы определения степени лингвоспецифичности коннекторов</b> .....	150
Иомдин Л. Л. <b>Как нам быть с конструкциями типа <i>как быть?</i></b> .....	161
Князев С. В., Утешева А. С. <b>Коартикуляционные изменения носовых согласных как показатель наличия просодического шва и порядка применения фонологических правил в русском языке</b> .....	175
Крейдлин Г. Е. <b>Типология русских походов</b> .....	196
Кривнова О. Ф. <b>Фонетические характеристики дыхательных пауз с разной текстовой локализацией</b> .....	207
Kunilovskaya M. A. <b>Linguistic Tendencies in English to Russian Translation: the Case of Connectives</b> .....	221
Кустова Г. И. <b>Типы конструкций с адвербиалами</b> .....	234
Letuchiy A. B. <b>'Tensed' and 'Non-tensed' Predicatives in Russian</b> .....	249
Левонтина И. Б., Денисова Г. <b>Итальянское <i> magari </i> и его русские переводные эквиваленты: разные дискурсивные стратегии</b> .....	261
Литвиненко А. О., Николаева Ю. В., Кибрик А. А. <b>Аннотирование русских мануальных жестов: теоретические и практические вопросы</b> .....	271
Лобанов Б. М., Житко В. А., Харламов А. А. <b>Компьютерная система обучения интонационным конструкциям русской речи</b> .....	287
Лютикова Е. А., Татевосов С. Г. <b>Синтаксис и семантика именного контрфактива</b> .....	302
Малафеев А. Ю., Орехова Е. А. <b>Толерантность по-русски и по-английски: корпусное исследование</b> ...	321
Падучева Е. В. <b>Эгоцентричность в семантике слова <i>никак</i></b> .....	333

Piperski A. Ch.	
<b>Semantic Halo of a Meter: a Keyword-based Approach</b> .....	342
Подлеская В. И.	
<b>«Я скажу тебе с последней прямотой»: прямая и косвенная речь по данным корпуса с просодической разметкой</b> .....	355
Пожарицкая С. К., Добрушина Е. Р.	
<b>Орфоэпический взгляд на некоторые вариантные явления русского литературного языка в эпоху корпусной лингвистики</b> .....	372
Рахилина Е. В.	
<b>Говорю я, Карл...</b> .....	384
Шмелев А. Д., Зализняк Анна А.	
<b>Реверсивный перевод как инструмент лингвистического анализа дискурсивных слов</b> .....	394
Слюсарь Н. А., Алексеева С. В.	
<b>Орфографические соседи с заменой буквы при изучении механизмов лексического доступа</b> .....	407
Соума Ю. В., Кудашев И. С., Михайлов М. Н.	
<b>Отрицание в русских и финских версиях двусторонних договоров между Россией и Финляндией: опыт корпусного исследования</b> .....	419
Толдова С. Ю., Муханова Р. В.	
<b>Порядок следования прилагательных разных семантических классов в русском языке в свете корпусных данных</b> .....	429
Вилинбахова Е. Л.	
<b>Когда статья — это не статья: «отрицательные тавтологии» в русском языке</b> .....	441
Янко Т. Е.	
<b>Интонация перечисления в русском языке</b> .....	453
Циммерлинг А. В.	
<b>Русские предикативы в зеркале эксперимента и корпусной грамматики</b> .....	466
<b>Abstracts</b> .....	482
<b>Авторский указатель</b> .....	496
<b>Author Index</b> .....	497

Компьютерная лингвистика:  
лингвистические  
исследования

# ОТРИЦАТЕЛЬНАЯ И ПОЛОЖИТЕЛЬНАЯ ПОЛЯРИЗАЦИЯ: СЕМАНТИЧЕСКИЕ ИСТОЧНИКИ<sup>1</sup>

**Апресян В. Ю.** (valentina.apresjan@gmail.com)

Национальный исследовательский университет Высшая школа экономики; Федеральное государственное бюджетное учреждение науки Институт русского языка им. В. В. Виноградова РАН; Москва, Россия

Отрицательно и положительно поляризованным языковым единицам (ОП и ПП), т. е. единицам, не употребляющимся вне контекста отрицания или, наоборот, не употребляющимся в его контексте, посвящено большое количество лингвистической литературы, в том числе формально-семантической и типологической. Однако само явление поляризации рассматривается на достаточно узком языковом материале. Так, описаны отрицательно и положительно поляризованные единицы в сериях неопределенных местоимений типа *any* vs. *some*, в системе наречий времени типа *yet* vs. *already*. Известны также отрицательно поляризованные фраземы типа *not to lift one's finger* и положительно поляризованные выражения со значением отношения типа *unfortunately* и *would rather*. При этом основные направления существующих исследований состоят в попытках определить и описать лицензирующие контексты (DE operators, pop-veridicality). В данной работе вводится обширный новый материал языковых единиц, имеющих свойства ОП и ПП, предлагается семантическое объяснение поляризации для разных семантических подтипов внутри классов NPI и PPI, демонстрируется, что отрицательная и положительная поляризация пронизывают гораздо большие участки языка, чем было принято считать ранее, и показывается, что поляризация является шкалярным признаком. Рассматривается явление возникновения отдельных поляризованных значений у многозначных слов (в первую очередь глаголов), которые не имеют свойства поляризованности в основном значении. Устанавливается семантическая связь между механизмами отрицательной и положительной поляризации. Степень поляризованности коррелирует с семантической структурой ОП и ПП единиц: чем «весомее» модальная рамка, тем больше поляризована единица.

**Ключевые слова:** отрицательная полярированность, положительная полярированность, многозначность, неверидикативность, снятая утвердительность, фразема, модальная рамка, шкалярность, лингвоспецифичность

---

<sup>1</sup> Исследование выполнено за счет гранта Российского научного фонда (проект №16-18-02054, «Исследование русского языкового сознания на основе семантического, статистического и психолингвистического анализа лексической многозначности»).

# POSITIVE AND NEGATIVE POLARITY: SEMANTIC UNDERPINNINGS

**Apresjan V. Ju.** (valentina.apresjan@gmail.com)

National Research University Higher School of Economics;  
Vinogradov Russian Language Institute of the Russian Academy  
of Sciences; Moscow, Russia

Negative and positive polarity items (NPIs and PPIs) are one of the well-explored topics in formal semantics and typology. However, the phenomenon of polarization is only addressed on a very limited linguistic material, such as indefinite pronouns (*some* vs. *any*), temporal adverbs (*yet* vs. *already*), certain idioms (*not to lift one's finger*), expressions of attitude (*would rather, unfortunately*). The main focus of polarity research are licensing contexts, or contexts that allow to use NPIs or PPIs, respectively. The current paper introduces considerable new body of NPIs and PPIs, explores semantic underpinnings of polarization, demonstrates that polarization is an all-pervasive property of language and shows that it is a scalar feature. The paper provides evidence for the fact that polarization is a regular phenomenon in polysemy, especially in verbs, where it characterizes single senses of polysemous words rather than words as a whole. It further establishes semantic connections between the phenomena of negative and positive polarization. The degree of polarization depends on semantic structure: the more "weighty" is the modal frame, the more polarized an item is.

**Key words:** negative polarity, positive polarity, polysemy, non-veridicality, phraseme, modal frame, scalarity, linguospecificity

## 1. Введение

Изначально отрицательной поляризацией в семантической традиции было принято называть свойство языковых единиц, которое заключается в невозможности их употребления без отрицания; ср. отрицательно поляризованную фразу *ни на йоту*: *Это ни на йоту не изменит нашего положения* vs. невозможность *\*Это на йоту изменило наше положение*.

Впоследствии были обнаружены и другие контексты, помимо эксплицитного отрицания, в которых возможно употребление отрицательно поляризованных единиц. На данный момент, общее сочетаемое свойство отрицательно поляризованных единиц (ОП) определяется как **невозможность** употребления в «веридикативных» (veridical) контекстах, т. е. с обозначением однократного совершившегося действия в прошлом (Giannakidou 2011). Лицензирующие ОП контексты (в их число входят вопрос, условие, будущее время, итеративность, внутрисловное отрицание, разные виды модальности и многие другие) неверидикативны (Ladusaw 1980, Zwarts 1995, Haspelmath 1997, von

Fintel 1999, Giannakidou 2011) или обладают свойством «снятой утвердительности» (Богуславский 2001, 2008, Падучева 2011, 2013 и многие другие работы).

Традиционный материал для исследования ОП — неопределенные местоимения типа английского *any*, наречия типа английских *yet*, *either*, а фраземы типа *lift one's finger* 'хотя бы пальцем шевельнуть', *budge an inch* 'сдвинуться с места хоть чуть-чуть', *give a damn* 'переживать хоть чуть-чуть', *have a hope in hell* 'иметь тень надежды' (Haspelmath 1997:38) и похожие единицы в других языках, например, греческая частица *oute* 'даже' (Giannakidou 2011:1663).

Положительной поляризации, т.е. невозможности употребления языковых единиц с отрицанием, посвящено значительно меньше лингвистической литературы; ср. например, (Szabolci 2004), где делается попытка единого анализа ОП и ПП единиц. Материал, на котором рассматривается это явление — это местоимения типа *some*, наречия *already*, *too*, выражения со значением неопределенного количества или степени типа *сколько-то*, *несколько*, а также выражения со значением отношения типа *would rather*, *unfortunately*; ср. невозможность \**I haven't already finished my breakfast*.

Целью данной работы является введение в лингвистический обиход нового языкового материала, демонстрирующего свойства ОП и ПП, а также семантическое объяснение свойств ОП и ПП.

## 2. Отрицательно поляризованные единицы

В рамках работы над онлайн-версией Активного словаря русского языка (Апресян, Mikulin 2016) было проведено обследование массового языкового материала на предмет обнаружения ОП единиц. В рамках данной работы было произведено обследование онлайн-версии Фразеологического словаря А. И. Федорова<sup>2</sup>, которое позволило выявить около восьмисот отрицательно поляризованных фразем — как лексических, так и синтаксических (см. Иомдин 2003, 2010 об этом понятии). Кроме того, в рамках настоящей работы было проведено обследование электронной версии Малого академического словаря русского языка, позволившее выявить сто шестнадцать отрицательно поляризованных лексем (т.е., в терминологии Московской семантической школы, отдельных значений слов)<sup>3</sup>.

### 2.1. Материал исследования

В первых двух выпусках Активного словаря (АС) было обнаружено около тридцати слов, которые не являются поляризованными в исходном значении, однако развивают одно или более отрицательно поляризованных переносных

<sup>2</sup> Автор выражает благодарность студентам Школы лингвистики НИУ ВШЭ Д. Демкиной и Е. Проценюк, которые произвели электронную обработку словаря Федорова.

<sup>3</sup> Автор выражает благодарность студентам Школы лингвистики НИУ ВШЭ М. Бибаевой, В. Иванову, С. Стыриной и Д. Суловой, которые произвели электронную обработку МАС.

значений. Это агитировать, афишировать, брать, волновать, ворочаться, втыкать, въезжать, вылезать, выдержать, вынести, дожидаться, допустить, жаловаться, замечать, знать, идти и некоторые другие единицы<sup>4</sup>.

В словаре Федорова были обнаружены ОП фраземы не дать в обиду, не дать гроша ломаного, слова не выжмешь, не дать спуску; не знать удержу, не знать отдыха, не знать куда глаза девать, не находить себе места, не находить слов; не сводить глаз, не смыкать глаз; не сносить головы, не светит, не сахар, не надышишься, не оберешься, не нахвалится; не в бровь, а в глаз; не в своем уме, не в расположении, не в себе, не в ту степь; не до смеха, не до шуток, не до жиру; не мое дело, не твое дело; не место; не может быть и речи, не может идти ни в какое сравнение; не с руки, не в обиду будь сказано; не жилец, не жилье; не за горами; не идет на ум; не сидится, не спится, не верится и многие другие.

В МАСе были обнаружены ОП лексемы вдруг (Он нас не вдруг узнал), взвидеть (Света не взвидел), годиться (Так поступать не годится), доглядеть, досмотреть (Не доглядели/не досмотрели за ребенком), дозваться (Вас не дозовешься), клеиться (Дело не клеится), лезть (Сапоги не лезли на ногу), миновать (Нас это не минует), наглядеться (Не могу на нее наглядеться), пара (Он ей не пара), переваривать (Я таких людей не перевариваю), переносить (Не переносу таких разговоров) и многие другие единицы.

## 2.2. Отрицательно поляризованные значения неполярных языковых единиц

Извлеченные из АС и МАС лексемы с ОП можно разделить на следующие семантические классы:

- ‘затрудненное понимание’ (Он ни во что не втыкает, Я что-то не въезжаю);
- ‘затрудненное проникновение’ (Лопата не берет мерзлую землю, Гвоздь не идет в стену, Сапоги не лезут на ногу),
- ‘затрудненное движение’ (Язык не ворочается; Язык едва ворочается),
- ‘затрудненное перемещение’ (Он целыми днями не вылезает из лаборатории),
- ‘затрудненное получение’ (Не дождешься ты от него никакой помощи),
- ‘тяжелое переживание’ (Она не вынесет этого, Он не выдержал испытаний; Он ее не выносит, Меня это совершенно не волнует, Не переносу/не перевариваю таких людей, Света не взвидел),
- ‘требующее внимания восприятие’ (Он ее упорно не замечал),
- ‘недостаточность’ (Ему не доставало опыта; Не доглядела/не досмотрела за детьми),
- речевой акт (Как дела? — Я не жалуюсь),
- ‘опыт’ (Местные жители не знают электричества),
- ‘отрицательно оцениваемая попытка убеждения’ (Я тебя не агитирую вступить в партию),

<sup>4</sup> Мы пользуемся описаниями разных авторов АС, в первую очередь, Ю. Д. Апресяна, И. В. Галактионовой, М. Я. Гловинской. Кроме того, мы опираемся на описание глагола *знать* из (Апресян 1995:405–430).

- ‘отрицательно оцениваемое распространение информации’ (*Они не афишируют своих отношений*),
- ‘непрепятствование’ (*Я этого не допущу, Как ты мог такое допустить?*),
- ‘маловероятное предположение’ (*Вот уж не думал, что он окажется таким негодяем*).

Заметно, что некоторые словообразовательные модели обладают семантикой, способствующей возникновению ОП значений; ср. глагольные модели на до- и на вы-.

Перечисленные единицы неоднородны в отношении степени ОП. Некоторые из них возможны только с отрицанием, но не в других лицензирующих контекстах:

- (1) *Жители Севера не знают этих болезней* [в смысле ‘не болеют ими’]
- (2) *\*Жители Севера едва <мало, плохо> знают такие болезни*
- (3) *??Знают ли жители Севера эти болезни?*
- (4) *??Если бы жители Севера знали эти болезни,...*

Однако в целом для фразеологически не связанных лексем характерен гораздо более широкий круг лицензирующих контекстов: например, ОП лексема *грех*<sup>5</sup> лицензируется внешним, внутренним и внутрисловным отрицанием, а также контекстом противопоставления: *Не грехи выпить; Грех не выпить; Грех пройти мимо; Грех жаловаться, Грех зарывать такой талант в землю, Грех покинуть тебя в таком несчастье, Грех сидеть дома в такой вечер* (а не идти гулять). Многие лексемы допускают использование в контексте вопроса и условия, с минимизаторами и показателями усилия (см. ниже), а некоторые из них позволяют и употребление в веридикативных контекстах:

- (5) *Посредством обуви он привлекает внимание к своим ногам. То есть афиширует свой недостаток, тычет им в глаза окружающим* [Мариам Петросян. Дом, в котором... (2009)]

При этом для большинства фразем характерна невозможность употребления без эксплицитного отрицания; ср. *\*Знает ли он, куда девать глаза?; \*Если бы он попал в бровь, а не в глаз...; \*Если бы он хватал с неба звезды...; \*Он мало хватает звезды с неба* и пр.

### 2.3. Мотивация отрицательной поляризации языковых единиц

Чем объясняется различное сочетаемое поведение ОП единиц? Естественно предположить, что ОП единицы — это семантически неоднородный класс, что и мотивирует различия в лицензирующих контекстах.

На рассмотренном материале можно выделить четыре семантических класса ОП, между которыми нет четких границ, упорядоченные по своей

---

<sup>5</sup> См. ее описание в Гловинская 2014.



семантической структуре, а именно, по «весу» неассертивных и ассертивных компонентов в значении<sup>6</sup>.

Предположение состоит в том, что единицы, значение которых в основном состоит из модальной рамки<sup>7</sup>, с тривиальной ассертивной частью, имеют более сильную поляризацию. По мере уменьшения модальной рамки и увеличения ассертивной части ослабляется степень поляризации. Кроме того, есть смыслы (усилие, предел и некоторые другие), наличие которых в значении способствует поляризации единицы.

Сочетаемость особенности ОП единиц (тяготение к отрицательным контекстам, возможность употребления в прочих неверидикативных контекстах, возможность употребления в веридикативных контекстах) были проанализированы по Основному корпусу НКРЯ. Таким образом были исследованы все ОП лексемы, извлеченные из МАС, а также сто наиболее частотных фразем из общего списка, извлеченного из словаря Федорова.

### 2.3.1. Сильно поляризованные ОП

**Первый класс** состоит в основном из фразем, а также из тех немногочисленных лексем, которые отличаются высокой степенью идиоматизации (*не преминуть, не обинуясь, не взвидеть*). Из ста частотных фразем по словарю Федорова, пятьдесят три, т. е. более половины, относятся к данному классу.

Они практически не употребляются без эксплицитного отрицания, поскольку форму без отрицания невозможно интерпретировать. Ср. фразему *не в бровь, а в глаз*: противоположный смысл выражается не формой без отрицания, а ПП единицей *попасть пальцем в небо* (\**Он не попал пальцем в небо*).

Иногда противоположный смысл может выражаться другой ОП единицей, ср.: *не знать куда глаза девать* 'быть не в состоянии посмотреть на некий объект из-за вызываемого этим объектом чувства стыда' vs. *не мочь глаз отвести* 'все время смотреть на исключительно привлекательный объект'.

Бывает, что употребление ОП фраземы без отрицания порождает не противоположный смысл, а совершенно другой; ср. фраземы с разной семантикой и синтаксическими свойствами *не Бог весть* vs. *Бог весть*. Двухвалентная фразема *X не Бог весть какой Y* (*Петя не Бог весть какой ученый*) значит 'X — не слишком выдающийся представитель множества Y', а ее валентности заполняются именными группами. Фразема *Бог весть, что (как, где,...) X* значит 'Говорящий не знает, что (как, где,...) ситуация X имеет место' и имеет одну валентность — содержания незнания, заполняемую косвенным вопросом.

Еще одна возможная причина, по которой фраземы этого класса невозможны без отрицания, состоит в том, что отсутствие отрицания слишком меняет их форму, например:

<sup>6</sup> В литературе по ОП принято разделение на два класса — strict (strong) NPIs и broad (weak) NPIs (Giannakidou 2011). Исследованный нами материал показывает, что данное бинарное противопоставление не исчерпывает разнообразия ОП единиц.

<sup>7</sup> Часть значения, в которой выражается оценка описываемой ситуации говорящим и предполагаемое говорящим отношение к этой ситуации со стороны адресата ( Апресян 1995:48).

- за счет смены генитива на аккузатив; ср., *\*хватать звезды с неба*, *\*давать спуск*, *\*знать удерж/отдых/аз*; *\*выдумает порох*, *\*берет каплю в рот*;
- за счет изменения класса неопределенного местоимения; ср. *\*идти в какое-либо сравнение*;
- за счет смены вида; ср. *\*найти себе место* [в нужном значении], *\*свести глаза*, *\*снести голову*; *\*долго подумав*;
- за счет замены других частей фраземы на антонимичные: *еще не вечер* — *\*уже вечер* [в нужном значении];
- за счет изменения иллокутивного намерения говорящего: *не взыщи* [прошу меня не критиковать] — *\*взыщи* [готов принять критику].

Иногда слово, входящее во фразему, не существует в современном языке без отрицания — нет слов *удерж*, *нахвалиться* и пр.

Однако главная причина, по которой форму без отрицания невозможно интерпретировать — семантическая: а именно, невозможность формирования противоположного смысла для большинства фразем. В работе выделяется два основных класса ОП фразем, для которых это невозможно.

Многие фраземы содержат модальные рамки, которые составляют основную часть их значения. Известно, что при наличии прямого и фразеологизованного способа выражения какого-либо смысла, фразема будет обладать дополнительными оттенками смысла — модальность, оценка, желание и пр. Ср., например, выражение *попасть не в бровь, а в глаз* и его не фразеологическую экспликацию ‘сделать очень точное критическое замечание’. Последняя допускает отрицание: ‘сделать не очень точное критическое замечание’. Однако фразема дополнительно выражает смыслы исключительной точности, удовольствия говорящего по поводу такой точности и, возможно, некоторые другие, и отрицанию не поддается.

Как отмечается в работе (Апресян 1995), многие слова, в значении которых входит модальная рамка, не присоединяют к себе отрицание (ср. *\*не ведь*, *\*не якобы*)<sup>8</sup>. Модальные рамки также объясняют невозможность употребления ОП единиц типа *пальцем не пошевелить*, *(ни) на йоту, малейший* и др. без отрицания (Апресян 1995:616–627). Возникающая аномалия объясняется противоречием между отрицательной семантикой модальной рамки ОП выражений и утвердительным смыслом предложения; ср. модальную рамку для *пальцем шевельнуть*: «говорящий считает, что субъект, который может что-то сделать, не сделает ничего, потому что не хочет» (Апресян 1995:617).

Однако представляется, что существует более общее объяснение. Отрицание слов, значение которых почти целиком состоит из модальной рамки, невозможно, потому что их ассерции, если они есть, тривиальны, а основной смысл, передаваемый модальными рамками, слишком сложен для того, чтобы иметь противоположный смысл. Например, каков мог бы быть антоним

<sup>8</sup> В работе (Падучева 2013:25–26) похожее условие формулируется для возможности построения общеотрицательных предложений, однако автор апеллирует к понятию презумпции: невозможно отрицание предложений, где семантической вершиной является слово, «значение которого целиком состоит из презумпций», поскольку отрицаться может только ассерция.

частицы *ведь*? Если антонимичный смысл и существует, он выражается не отрицанием исходной единицы, а другой единицей (ср. антонимы *-то* и *даже*, *всего* и *целых* с противоположными модальными рамками).

В случае фразем с модальными рамками, в составе которых есть частица *не*, невозможность присоединения отрицания преобразуется в невозможность употребления без отрицания, в силу того же семантического закона.

Также существует много фразем, выражающих смысл предела, крайней степени; ср. *терпеть не мочь*, *ног под собой не чуют от усталости*, *муху не обидит* и др. Здесь также действует общая семантическая тенденция, не ограниченная сферой фразеологии, которая затрудняет употребление лексических единиц со значением предела в сфере действия отрицания; ср. странность или невозможность сочетаний *\*не крайне*, *\*не чуть-чуть*, *\*не изумительно* и пр. Подобные сочетания порождают малоинформативные высказывания о наличии обычной или не выходящей за пределы нормы ситуации, невозможные в силу максимы информативности Грайса. Опять-таки, в случае ОП фразем со значением предела, невозможность присоединения отрицания преобразуется в невозможность употребления без отрицания: *\*Он чует под собой ноги*, *\*Он обидит муху*, *\*Он может ее терпеть*.

**Второй класс**, семантически близкий к первому, — это фраземы и лексемы, которые имеют «весомую», но менее сложную модальную рамку, чем единицы первого класса, и которые менее идиоматизированы, а также мало меняют форму при отсутствии *не*. Поэтому они могут появляться без эксплицитного отрицания, но не с минимизаторами, поскольку не обозначают градуируемых сущностей, а указывают на стремящиеся к нулю количество или степень или на невозможность совершения действия, достижения состояния и пр.; ср.

- (6) *Он, прежде чем пальцем пошевелит, будет два часа рассуждать*
- (7) *Вряд ли что-то изменится хоть на йоту*
- (8) *Не знаю, есть ли у нас хоть малейший шанс*<sup>10</sup>
- (9) *Если имеется хотя бы тень сомнения, надо воздержаться*
- (10) *До смеха*<sup>11</sup> *ли ему сейчас?; Ему сейчас едва ли с руки этим заниматься.*

<sup>9</sup> Ср. толкование *ведь* 1 в (Бабаева 2014), почти целиком состоящее из модальной рамки. Имеет место ситуация A1; имеет место ситуация A2; говорящий объясняет существование ситуации A1 тем, что одновременно с ней существует ситуация A2; говорящий считает, что адресату известно о существовании ситуации A2, но что он не учитывает или недостаточно учитывает A2'.

<sup>10</sup> Интересно, что английский переводной аналог единицы *малейший шанс*, выражение *remote chance*, не является полностью ОП единицей; по данным Sketch Engine, неверидативные контексты доминируют, но веридикативные возможны: *Not that either reform has the remotest chance of happening vs. He gave me a remote chance of a good grade*. Для русского *малейшего шанса* второй тип контекстов невозможен: *\*Он дал мне малейший шанс*.

<sup>11</sup> Отрицательная поляризованность синтаксической фраземы *X-у не до Y-а...* отмечается в (Июмдин 2011).

Ср. также *Привыкать ли нам?; Не знаю, в своем ли он уме; Вряд ли у него вернется язык об этом рассуждать.*

Как видно из примеров, неверидикативные контексты, которые здесь приведены, не предполагают совершения действия.

Сорок две проанализированных фраземы, т.е. почти все, не вошедшие в первый класс, относятся к данному классу; к этому классу относятся также большая часть проанализированных лексем.

Интересным примером единицы второго класса является требующий дативного субъекта предикатив *впервой*, который имеет близкий синоним, отличающийся отсутствием модальной рамки и поляризации — *впервые*. Наречие *впервые* не поляризовано; наречие *впервой* обладает свойством ОП; ср.

(11) *Ничего, нам не впервой*

(12) *Не бойся, впервой ли нам!;*

при странном (или устаревшем) употреблении *впервой* без отрицания и дативного субъекта:

(13) *Я к сознательной революционной деятельности впервой подростком приобщился* [Фридрих Горенштейн. Куча (1982) // «Октябрь», 1996].

*Не впервой* обладает ассерцией ‘не в первый раз’ и при этом содержит приблизительно следующую модальную рамку

(14) ‘Творящий хочет убедить адресата в том, что нет оснований для беспокойства о том, как субъект справится с действием Р, поскольку субъект уже производил такие действия в прошлом и имеет необходимый опыт’.

Наличие ассерции позволяет употребление без эксплицитного отрицания, наличие «весомой» модальной рамки «отталкивает» веридикативные контексты.

Строгой границы между первым и вторым классом нет: например, хотя примеров в НКРЯ на употребление слова *не выжмешь* не находится, но вполне можно сконструировать фразы типа *Не знаю, выжмешь ли ты из него хоть слово?*

Как выяснилось, самыми предпочтительными неверидикативными контекстами без эксплицитного отрицания оказываются вопрос и условие; прочие, а именно, сравнение, внутрисловное отрицание, целевое дополнение и пр. намного менее частотны.

### 2.3.2. Слабо поляризованные ОП

**Третий класс** — это лексемы и очень немногочисленные фраземы, которые содержат модальную рамку наряду с нетривиальной ассерцией. В их ОП значение обычно входит указание на усилие и трудность достижения результата или на очень малый результат (см. выше). Поэтому они возможны не только с условием и вопросом, но и с минимизаторами *едва, с трудом, мало*, иногда *редко*, с наречиями отрицательной оценки типа *плохо*, а также с некоторыми уступительными единицами; ср. *Он еле лыко вяжет, Он с трудом верил своим глазам; Я как-то с трудом/плохо въезжаю, Лопата еле/плохо брала замерзшую землю, Язык едва ворочался, Он все-таки дождался справедливости,*

*Его это мало волнует; Она плохо выносит подобное обращение; Он едва притронулся/прикоснулся к еде; Он редко вылезает из лаборатории.*

Контексты в данной группе допускают совершение действия, однако не-полноценным образом.

Слабая ОП лексем, содержащих семантические компоненты ‘трудность’ и ‘усилие’, понятна: достижение результата (и, соответственно, употребление в веридикативных контекстах) менее вероятно, чем обратное. В этот же класс входит синтаксическая фразема *X-у Y-уется*, в некоторых вариантах заполнения; ср. *Мне не верится*; ср. *с трудом верится*. С заполнением *сидеться* эта фразема более отрицательно поляризована: *Ему не сиделось на месте*, при странности *?Ему плохо сиделось на месте*. С некоторыми другими вариантами заполнения она вообще не поляризована, но включает указание на слабый смысл ‘хорошо’<sup>12</sup>, который может сниматься эксплицитным указанием на обратное (*Мне сегодня работается* = ‘хорошо’, *Мне сегодня не работается*). С обозначениями естественных состояний она приобретает ОП окраску, что неудивительно: прагматически странно было бы отмечать наличие естественного состояния, более информативно отмечать его отсутствие.

Этот класс также относительно невелик (две фраземы, пятнадцать лексем из проанализированных). Из лексем сюда относятся глаголы на *-до-* и на *вы-* (*дозваться, допроситься, вынести, вытерпеть*), а также некоторые другие (*переваривать, притронуться, клеиться*).

**Четвертый класс** — это лексем и немногочисленные фраземы, в значении которых ассертивная часть намного весомее модальной рамки. Они возможны в утвердительных контекстах, но употребляются с отрицанием чаще в силу прагматических причин. Например, ситуация, когда люди *афишируют что-л.* (*Они всячески афишировали свой роман*) является более редкой, чем обратная (*Они не афишировали свой роман*)<sup>13</sup>, поскольку предполагается, что они чрезмерно широко распространяют такие сведения о себе, которые обычно не принято распространять или даже принято скрывать. Ср. употребление глагола *афишировать* в НКРЯ (по данным на 15 июня 2016 года): всего 297 употреблений, из них 265 неверидикативных употреблений и 32 фактивных.

Похожими свойствами обладают лексем *агитировать* и *жаловаться*, в следующих употреблениях: *Я тебя не агитирую ступать в партию; Я не жалуясь*. При этом ОП эффект усиливается в псевдоперформативных употреблениях от первого лица.

Четкой границы между третьим и четвертым классами нет.

Формулировка семантических особенностей этих классов позволяет не только объяснять, но и предсказывать свойство ОП у языковых единиц. Например, слово *осилить*, принадлежащее к семантическому классу трудного достижения результата, демонстрирует тенденцию к ОП: по данным НКРЯ,

<sup>12</sup> Ср. анализ этой конструкции в (Апресян 2006:36–38).

<sup>13</sup> Ср. описание этого глагола в (Апресян 2014).

в более чем трети случаев этот глагол употребляется с эксплицитным отрицанием (*Он не осилил завтрак*), а в его отсутствие — с будущим временем (*Думаю, осилю*), модальными словами (*Надо осилить*), показателями сомнения (*Сомневаюсь, что осилит*), уступительными словами (*все-таки осилил*), минимизаторами (*кое-как осилил*) и пр. Веридикативных употреблений у него немного. На преодоление трудности указывает и глагол *лезть* в значении ‘помещаться’, что способствует отрицательной поляризации: *Пробка не лезла в бутылку, при странности* <sup>13</sup>*Пробка лезла в бутылку*. На большое внутреннее усилие указывает и сильно поляризованный глагол *терпеться*: *Мне не терпится на него посмотреть, но не \*Мне терпится на него посмотреть*.

Глагол *усмотреть*, указывающий (как и *замечать*) на трудность в восприятии визуальной информации, также обладает отрицательной поляризацией: *Не усмотрели за ребенком, но не \*За ребенком усмотрели*.

ОП свойствами обладает и глагол *разглашать*, который указывает на отрицательно оцениваемое распространение информации (ср. *афишировать*): более половины употреблений содержат эксплицитное отрицание, а остальные демонстрируют тяготение к более слабым невидикативным контекстам: *запрещено разглашать, не рекомендуется разглашать, если вы разгласите* и т. п., хотя веридикативное употребление тоже возможно: *Он разгласил секретную информацию*.

Отрицательно поляризован глагол *гадать*, обозначающий маловероятное предположение (как и *думать* в одном из значений): *Вот уж не гадал, что когда-нибудь окажусь здесь, но не \*Давно гадал, что окажусь здесь*.

Среди фразем лишь считанные единицы принадлежат к данному классу; ср. (не) *брать/взять себе в голову; по зубам: Не бери в голову!*; *Но в Москве Егорушка что-то говорил о какой-то Лене, ну, я и взяла в голову...* (Г. Щербакова); *Ему эта задача не по зубам; Зато воспользоваться другой системой, тоже помогающей ориентироваться в пространстве, нам вполне по зубам* (Д. Гронский)<sup>14</sup>.

Безусловно, не все ОП единицы можно распределить по данным четырем классам; данная классификация — лишь начало работы над описанием огромного пласта ОП лексики. Ср., например, глагол *защищать*<sup>15</sup> в значении ‘закрывать от нежелательного воздействия’ (*Крем защищает от загара; Прививка защитит вас от гриппа; Тонкий свитер не защищал от холода*). Веридикативное употребление затруднено; ср. странность <sup>15</sup>*Прививка защитила его от гриппа*. Это связано с тем, что в данном значении глагол *защищать* обозначает свойство, а для глаголов со значением свойств типично употребление в усузальном значении НЕСОВ, в то время как формы СОВ они не имеют. Таким образом, можно ожидать дальнейшего расширения класса ОП единиц за счет целых аспектуальных классов.

<sup>14</sup> Как выяснилось, отдельные фраземы, которые даются в словаре Федорова с отрицанием, вообще не являются ОП единицами, поскольку одинаково свободно допускают употребление с отрицанием и без него; ср. (не) *в радость, (не) по душе: Мне это не в радость; Ей в радость были эти заботы; Мне такое не по душе; Ей пришлось по душе ее подарки*.

<sup>15</sup> Описан в (Крылова 2017).

### 3. Положительно поляризованные единицы

ПП единицы обнаруживают такие же семантические свойства, что и ОП единицы. А именно, свойством ПП обладают единицы с модальной рамкой, например

- дискурсивные слова *ведь, мол, якобы, -то, же* и др.;
- вводные обороты *к счастью, чудесным образом, как говорится, известно* и др.;
- слова, включающие указание на оценку говорящего; ср. *С него станется туда прийти, Ты у меня еще дождешься.*

Как сказано выше, для единиц, значение которых состоит в основном из модальной рамки, употребление с отрицанием практически невозможно; ср. *\*Они не якобы приехали, \*Он не уже пообедал; \*Он не ведь отличный специалист, \*Смотрю — и не впрямь он!; \*Он, как не говорится, нашел себе другую; \*С него не станется туда прийти, \*Он у меня еще не дожидется.* Кроме того, есть слабо поляризованные ПП единицы, которые включают модальную рамку (например, со значением очень положительной или отрицательной оценки), но обладают и значительной ассертивной частью. Для них употребление с отрицанием затруднено; ср. странность *?Мы съездили не замечательно, ?Город не преобразился; ?Он не расшвырял одежду по комнате, ?Это не потрясающий фильм, ?Работа не чудовищная.*

Кроме того свойством ПП обладают многие единицы с семантикой предела — в силу прагматической неинформативности их отрицания; ср. странность *\*Он купил не чуть-чуть муки, \*Это не крайне интересно, \*Он не едва коснулся ее руки.* Аналогично, прагматически неинформативным является отрицание форм неопределенного количества типа *\*не несколько, \*не сколько-то*, отчего единицы этого семантического класса также характеризуются ПП.

Область, где интересным образом сближаются положительно и отрицательно поляризованные единицы — это малая степень. Ср. следующие выражения: *slim chance, ghost of a chance, the least chance, slightest chance* vs. *призрачная надежда, слабая надежда, незначительный шанс, малейший шанс.* Выражения *slim chance, призрачная надежда, слабая надежда, незначительный шанс* положительно поляризованы, т. е. не встречаются с отрицанием (их отрицание возможно лишь при наличии интенсификатора даже — *У нас нет даже незначительного шанса*). Выражения *ghost of a chance, the least chance, slightest chance* и *малейший шанс* отрицательно поляризованы и не встречаются вне контекстов отрицания, условия или вопроса. С чем связано такое распределение единиц? Представляется, что выражения, обозначающие исчезающе малую степень, отрицательно поляризованы, т. к. их смысл практически сводим к отрицательному смыслу 'никакой' (*У нас нет ни малейшего шанса ≈ У нас нет никакого шанса*). Выражения, обозначающие малую, но не исчезающе малую степень, положительно поляризованы, так как их отрицание прагматически неинформативно: имеет смысл отрицать наличие или отсутствие чего-либо, но не отсутствие очень малой степени чего-л.

## 4. Заключение

Из сказанного можно сделать следующие выводы:

- свойства ОП и ПП являются шкалярными, т.е. оно свойственны разным единицам в разной степени;
- фраземам в целом присуща более высокая степень ОП, чем фразеологически не связанным лексемам, обладающим этим свойством;
- отрицательная и положительная поляризация объясняются одинаковыми семантическими и прагматическими механизмами;
- свойство поляризации можно в некоторой степени предсказывать по семантическим характеристикам языковой единицы, в том числе по ее словообразовательной модели.

## Литература

1. *Активный словарь русского языка.* (2014). Отв. ред. академик Ю. Д. Апресян, М. т. 1–2. Active dictionary of Russian [Aktivnyj slovar' russkogo jazyka]. Ju. D. Apresjan, ed. Moscow, 2014. vol. 1–2.
2. *Апресян Ю. Д.* (1995). Избранные труды: В 2 т. Т. II: Интегральное описание языка и системная лексикография. М.: Школа «Языки русской культуры». [Apresjan Ju. D. *Izbrannye trudy: V 2 t. T. II: Integral'noe opisanie jazyka i sistemnaya leksikografija* [Selected works: In 2 vol. Vol. II: Integrated description of language and systematic lexicography]. Moscow: Shkola «Yazyki Russkoi Kul'tury», 1995.]
3. *Апресян Ю. Д.* (2006). Уникальное значение: конструкция *Мне (хорошо) работается* // Языковая картина мира и системная лексикография / Под ред. Ю. Д. Апресяна. М., 2006. с. 36–38. Apresjan Ju. D. Unique meaning: construction 'To me it works well' [Unikal'noe znachenie: konstrukcija *mne xorosho rabotaetsja*]. Linguistic worldview and systematic lexicography [Jazykovaja kartina mira i sistemnaja leksikografija]. Ju. D. Apresjan, ed. Moscow, 2006. pp. 36–38.
4. *Апресян Ю. Д.* (2014). Статья АФИШИРОВАТЬ // Активный словарь русского языка, вып. 1. М., 2014. С.127. Apresjan Ju. D. Entry 'to flaunt' [statja AFISHIROVAT']. Active dictionary of Russian [Aktivnyj slovar' russkogo jazyka]. Ju. D. Apresjan, ed. Moscow, 2014. issue 1. P. 127.
5. *Бабеева Е. Э.* (2014). Статья ВЕДЬ // Активный словарь русского языка, вып. 1 (буквы А-Б). М., 2014. С. 42–43. Babaeva E. E. Entry 'indeed' [statja VED']. Active dictionary of Russian [Aktivnyj slovar' russkogo jazyka]. Ju. D. Apresjan, ed. Moscow, 2014. issue 1. P. 42–43.
6. *Богуславский И. М.* (2001). Модальность, сравнительность и отрицание // Русский язык в научном освещении, № 1. С. 27–51. Boguslavsky I. M. Modality, comparison and negation [Modal'nost, sravnitel'nost' i otricanije]. Russian language in a scientific perspective [Russkij jazyk v nauchnom osveshchenii]. 2001, issue 1. P. 27–51.



7. *Богуславский И. М.* (2008). Между истиной и ложью: адвербиалы в контексте снятой утвердительности // Логический анализ языка: между ложью и фантазией. М.: Индрик. С. 67–77. Boguslavsky I. M. Between the truth and lies: adverbials in the context of suspended assertion [Mezhdu istinoj i lozh'ju: adverbialy v kontekste snjatoj utverditel'nosti]. Logical analysis of language: between lie and fantasy [Logicheskij analiz jazyka: mezhdu istinoj i lozhju]. Moscow, Indrik, p. 67–77.
8. *Гловинская М. Я.* (2014). Статья ГРЕХ // Активный словарь русского языка, вып. 1. М., 2014. С. 701–702. Glovinskaja M. Ja. Entry 'sin' [statja GREX]. Active dictionary of Russian [Aktivnyj slovar' russkogo jazyka]. Ju. D. Apresjan, ed. Moscow, 2014. issue 1. P. 701–702.
9. *Иомдин Л. Л.* (2003). Большие проблемы малого синтаксиса // Труды международной конференции по компьютерной лингвистике и интеллектуальным технологиям Диалог'2003. Протвино, 2003. С. 216–222. Iomdin L. L. Big problems of microsyntax [Bolšie problemy malogo sintaksisa]. Conference Proceedings "Computational Linguistics and Intellectual Technologies" Dialogue 2003 [Trudy mezhdunarodnoj konferencii po kompjuternoj lingvistike i intellektualnym tehnologijam Dialog'2003]. Protvino, p. 216–222.
10. *Иомдин Л. Л.* (2010). Синтаксические фраземы: между лексикой и синтаксисом // Апресян Ю. Д., Богуславский И. М., Иомдин Л. Л., Санников В. З. Теоретические проблемы русского синтаксиса: Взаимодействие грамматики и словаря / Отв. ред. Апресян Ю. Д. М.: Языки славянских культур, 2010. с. 141–208. Iomdin L. L. Syntactic phrasemes: between lexicon and grammar. [Sintaksicheskie frazemy: mezhdu grammatikoj i slovarem]. Apresjan Ju. D., Boguslavskii I. M., Iomdin L. L., Sannikov V. Z. Teoreticheskie problemy russkogo sintaksisa: Vzaimodeistvie grammatiki i slovarja. Apresjan Ju. D. (ed.). Moscow: Yazyki Slavyanskikh Kul'tur, 2010.]
11. *Иомдин Л. Л.* (2014). *Хорошо меня там не было*: синтаксис и семантика одного класса русских разговорных конструкций. Die Welt der Slaven. Sammelbände. Verlag Otto Sagner, 2014. Band 55. P. 423–436. Iomdin L. L. 'Good thing I wasn't there': semantics and syntax of one class of Russian conversational constructions [Xoroshho menja tam ne bylo: semantika i sintaksis odnogo klassa russkix razgovornyx konstrukcij].
12. *Крылова Т. В.* (2017). Статья ЗАЩИЩАТЬ // Активный словарь русского языка, вып. 2 (в печати). Entry 'to protect' [statja ZASHCHISHCHAT']. Active dictionary of Russian [Aktivnyj slovar' russkogo jazyka]. Issue 2 (in print).
13. *Падучева Е. В.* (2011). ИмPLICITное отрицание и местоимения с отрицательной поляризацией. Вопросы языкознания, 1. 2011. С. 3–18. Paducheva E. V. Implicit negation and negative polarity pronouns [Implicitnoe otricanie i mestoimenija s otricatel'noj poljarizaciej]. Problems of linguistics, 1. [Vorposy jazykoznanija]. P. 3–18.
14. *Падучева Е. В.* (2013). Русское отрицательное предложение. М.: ЯСК. Paducheva E. V. Russian negative sentence [Russkoe otricatel'noe predlozhenie]. Moscow, Languages of Russian culture.

15. *Apresjan V., Mikulin N.* (2016). Dictionary as an instrument of linguistic research // Proceedings of the XVII Euralex International Congress. pp. 224–231
16. *von Fintel, Kai.* (1999). NPI-Licensing, Strawson-Entailment, and Context-Dependency. *Journal of Semantics* 16.97–148.
17. *Giannakidou A.* (2011). Positive polarity items and negative polarity items: variation, licensing, and compositionality. *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter. P. 1660–1712.
18. *Haspelmath M.* (1997). *Indefinite Pronouns*. Oxford: Clarendon Press.
19. *Ladusaw W. A.* (1980). *Polarity Sensitivity as Inherent Scope Relations*, Ph. D. dissertation, published by New York: Garland Publishing.
20. *Szabolcsi A.* (2004). *Natural Language & Linguistic Theory*. 22: 409.
21. *Zwarts F.* (1995). Nonveridical Contexts // *Linguistic Analysis*, 25. P. 286–312.

# «КСЕНОПОКАЗАТЕЛИ» ПО ДАННЫМ ПАРАЛЛЕЛЬНЫХ КОРПУСОВ И СОВРЕМЕННЫХ СМИ: РУССКОЕ *ЯКОБЫ*<sup>1</sup>

**Апресян В. Ю.** (valentina.apresjan@gmail.com)<sup>1,2</sup>,  
**Шмелев А. Д.** (shmelev.alexei@gmail.com)<sup>3,2,4</sup>

<sup>1</sup>Национальный исследовательский университет  
Высшая школа экономики; <sup>2</sup>Институт русского языка  
им. В. В. Виноградова РАН; <sup>3</sup>Московский педагогический  
государственный университет; <sup>4</sup>Православный  
Свято-Тихоновский гуманитарный университет

## “XENO” MARKERS IN THE LIGHT OF THE DATA OF PARALLEL CORPORA AND CONTEMPORARY MASS MEDIA: THE CASE OF THE RUSSIAN WORD *JAKOBY*

**Apresjan V. Ju.** (valentina.apresjan@gmail.com)<sup>1,2</sup>,  
**Shmelev A. D.** (shmelev.alexei@gmail.com)<sup>3,2,4</sup>

<sup>1</sup>National Research University Higher School of Economics;  
<sup>2</sup>Vinogradov Russian Language Institute of the Russian  
Academy of Sciences, Moscow, Russia; <sup>3</sup>Moscow Pedagogical  
State University, Moscow, Russia; <sup>4</sup>St Tikhon's Orthodox  
University

---

<sup>1</sup> Работа над данной статьей отчасти финансировалась следующими грантами: Общие принципы унидирекционального сопоставительного корпусного анализа дискурсивных слов были разработаны в рамках научно-исследовательского проекта РФФИ № 16-06-00339 А «Контрастивное корпусное исследование дискурсивных слов русского языка»; конкретный анализ на основе параллельных корпусов проведен при поддержке РФФИ (проект № 16-48-03006, «Семантический анализ переводов в сопоставительной культурной перспективе и культурная специфичность в обучении языку»); нормативные рекомендации, касающиеся употребления частицы *якобы* в современном литературном языке, разработаны при поддержке РГНФ, проект № 15-04-00488 «Изменения узуса и кодификация норм русского литературного языка» (2015-2017). Анализ сдвигов в употреблении частицы *якобы* по данным современных СМИ был произведен при поддержке гранта Российского научного фонда (проект №16-18-02054, «Исследование русского языкового сознания на основе семантического, статистического и психолингвистического анализа лексической многозначности»)

The paper considers the less known aspects in the functioning of Russian lexical “xeno” markers, in particular, of the particle *jakoby* ‘allegedly, ostensibly’. Traditionally described as expressing the falsity of a proposition contained in somebody’s utterance, in conjunction with a negative assessment of the utterer as aware of its falsity, *jakoby* displays very different usages in the language of contemporary mass media. Namely, it is frequently used as a mere marker of evidentiality, without an obligatory assessment of the proposition as false or of its source as untruthful. In fact, it can even be used to refer to statements that are treated as true within the very same text, only to indicate that the source of this information is not the writer herself but somebody else (e.g., a different news agency), in what might be termed as “safety” strategy.

Besides, *jakoby* in its mass media usages demonstrates unusual syntactic behaviors, namely shifts in scope, where it is placed before the speech verb rather than before the challenged proposition: *jakoby utverzhdat’*, *chto P’ jakoby claim that P’* instead of *utverzhdat’*, *chto jakoby P’ claim that jakoby P’*.

However, the study of the Russian-English parallel corpus reveals that these usages are not as unusual as they may appear. In Russian translations of English texts *jakoby* sometimes functions as a translation of the English *supposedly*, *allegedly*, *ostensibly* or other (e.g., verbal) markers of uncertainty, but more frequently occurs with no apparent stimulus in the source, merely to mark indirect quotation.

It appears therefore that there is a certain need in the Russian language for a neutral evidentiality marker. It is occasionally filled with *jakoby*, which in this case displays a tendency for grammaticalization: it expresses that the source of information is other than the speaker herself (but contains no other semantic components), and takes syntactic scope over the speech verb instead of the proposition it challenges.

**Key words:** “xeno” marker, evidentiality, de re, de dicto, semantic scope, syntactic scope, proposition, “safety” discourse strategy

## 1. Вступительные замечания

При передаче содержания чужих мнений или высказываний (в частности, при несовпадении представлений говорящего о положении дел с содержащимися в них представлениями субъекта передаваемого мнения или высказывания) возможны две стратегии, которые можно обозначить как стратегию de re и стратегию de dicto (названия восходят к средневековой логике). При выборе стратегии de re говорящий ориентируется на свои представления, не заботясь о том, согласится ли с такой передачей субъект передаваемого мнения или высказывания; при выборе стратегии de dicto говорящий ориентируется на представления субъекта передаваемого мнения или высказывания, даже если сам их не разделяет. При этом обе стратегии могут иногда причудливым образом совмещаться в одном акте передачи содержания мнения или высказывания.

Вообще говоря, выбор той или иной стратегии обычно никак не маркируется. Однако в русском языке есть дискурсивные слова, специализирующиеся на передаче чужих мнений или высказываний (поэтому их иногда называют «ксенопоказателями») и употребляемые преимущественно или исключительно в рамках одной из указанных стратегий.

Так, русская частица *мол* обычно не используется при буквальной передаче чужой или даже своей собственной речи [Арутюнова 1992; 2000]; она почти всегда предполагает передачу содержания «своими словами», т. е. элементы стратегии *de re*<sup>2</sup>. При этом, отклоняясь от точности в передаче того, что непосредственно сказано, говорящий, как это и предполагается стратегией *de re*, гарантирует точность передачи содержания, как он его понимает. Отклонение от того, что буквально было сказано, может быть сделано в двух направлениях, в некотором смысле противоположных. С одной стороны, частица *мол* часто маркирует экспликацию скрытого смысла высказывания или даже невербального поведения, как в примерах из «Национального корпуса русского языка» (далее — НКРЯ), приведенных в статье [Шмелев 2016, 96–97]. С другой стороны, говорящий может счесть какие-то элементы содержания несущественными и не заслуживающими передачи и редуцировать их или даже вовсе свести к нулю, ограничившись указанием на сам факт наличия мнения или высказывания. Это особенно ясно видно при использовании таких оборотов, как *так, мол, и так*, напр.: *Я бы и агитки вам доверить мог. Раз бы показал: — вот так-то мол, и так-то...* (Владимир Маяковский).

Напротив того, частица *якобы* ориентирована на максимально точную передачу того, что непосредственно выражено в чужом высказывании или невербальном поведении, и тем самым участвует в передаче содержания в соответствии со стратегией *de dicto*. Функции частицы *якобы* в рамках данной стратегии, в соответствии с описанием, данным в статье [Rakhilina 1994], и с языковой интуицией авторов настоящей статьи, заключаются в том, чтобы указать на ложность пропозиции, содержащейся в элементе чужого высказывания, входящего в сферу действия *якобы*, и дать отрицательную оценку поведения тех, кто пытается выдать эту пропозицию за истинную (т. е. субъектов передаваемых высказываний или мнений).

Частице *мол* в последние десятилетия был посвящен целый ряд специальных исследований (в дополнение к уже названным работам [Арутюнова 1992; 2000] и [Плунгян 2008] упомянем также статьи [Баранов 1994] и [Падучева 2011]). Что касается частицы *якобы*, она также была предметом ряда работ, однако представляется, что некоторые важные особенности ее функционирования в современном языке прошли мимо внимания исследователей. Именно частице *якобы* посвящена основная часть данной работы. Отправной точкой данного исследования послужили замеченные авторами нестандартные употребления этой частицы в СМИ, которые позволили сделать предположение о том, что ее значение несколько сдвинулось по сравнению с описанным прототипом и стало близко к английским словам *alleged* и *allegedly*, которые частотны в официальных и газетных текстах

---

<sup>2</sup> В статье [Плунгян 2008] делается еще более решительное утверждение: «сочетания *мол* со словами, обозначающими точную передачу сказанного, выглядят аномально». Формально говоря, это некоторое преувеличение: легко обнаруживаются примеры, в которых *мол* вводит буквальную цитату, иногда даже заковыченную. Однако во всех случаях *мол* указывает на то, что говорящий приводит именно свою интерпретацию и вербализует ее так, как считает нужным (точно так же в реферате буквальное цитирование реферируемого текста не разрушает жанр реферата.)

и указывают на нежелание говорящего разделять ответственность за недоказанные утверждения, которые могут оказаться ложными, но не содержат прямого указания на ложность пропозиции или отрицательной оценки. В параллельном корпусе Sketch Engine встречаются примеры такого рода; ср.

- (1) *Что за информация у вас якобы имеется? — Now what is this information you **allegedly** have?*<sup>3</sup>
- (2) *Единственный свидетель этого якобы убийства — Вероника Донован. — The only witness to this **alleged** murder is Veronica Donovan.*

## 2. Частица *якобы* по данным параллельных корпусов

Как уже говорилось, общий смысл частицы *якобы* при «стандартном» прототипическом употреблении сводится к указанию на ложность пропозиции, заключенной в элементе чужого высказывания, входящего в сферу действия *якобы*, и отрицательной оценке поведения тех, кто пытается выдать эту пропозицию за истинную. В стандартном понимании *якобы* говорящий, в целом адекватно передавая содержание чужого высказывания, одновременно выражает оценку его (или его части) как неистинного, причем часто намеренно неистинного: *Он утверждает, что якобы сам придумал эту теорему* значит, что он, с точки зрения говорящего, сознательно лжет. Что же могут добавить к этому описанию данные параллельных корпусов?

Как отмечалось в статье [Шмелев 2015], переводные эквиваленты высказываний с дискурсивным словом могут рассматриваться как источник информации о его семантике и прагматике. Направление перевода с иностранного на русский язык в этом отношении особенно показательно: выявление «стимулов» для появления дискурсивного слова в переводном тексте и особенно его появление при отсутствии непосредственного источника для него в оригинале позволяет выявить признаки, создающие условия для использования этого слова.

Обратимся к англо-русскому подкорпусу как самому представительному из двуязычных параллельных подкорпусов НКРЯ. Всего слово *якобы* встретилось в этом подкорпусе 101 раз.

В некоторых случаях слово *якобы* имеет непосредственный источник, т.е. появляется как соответствие некоторого слова оригинального текста. Можно было бы ожидать, что частица *якобы* будет регулярно появляться в качестве соответствия труднопереводимому слову *alleged*. Такие примеры действительно встречаются, но они не столь частотны (всего 3 примера):

- (3) *In the abridged London edition of, there are plates of an **alleged** whale and a narwhale. В сокращенном лондонском издании 1807 года имеются иллюстрации, изображающие **якобы** «кита» и «нарвала». [Герман Мелвилл. Моби Дик (И. М. Бернштейн, 1961)]*

---

<sup>3</sup> Выделение в цитатах здесь и далее наше. — В. А., А. Ш.

- (4) ...went on to deplore my **alleged** (and perfectly abstract) "democratism." — ...с большим неодобрением отозвался о моем **якобы** «демократизме». [Владимир Набоков. Память, говори (С. Ильин, 1999)]
- (5) ...reports of an **alleged** apparition of the Beatus which had, so travelers said, come to one Francis Gerard of Utah, AOL. — сообщение о духе блаженного, который, по рассказам путников, **якобы** явился некоему Франциску Джерарду родом из Юты, AOL. [Уолтер Миллер. Страсти по Лейбовицу (С. Борисов, 1999)]

Но гораздо чаще конструкции с глаголом *to allege* (или причастием *alleged*) переводятся без помощи частицы *якобы*.

Несколько чаще (в НКРЯ 6 примеров) в качестве непосредственного источника для *якобы* выступает слово *supposed(ly)*, напр.:

- (6) *He liked to know how it really was; not how it was **supposed** to be.* — Приятно было знать все, как оно есть на самом деле, а не как оно **якобы** происходит. [Эрнест Хемингуэй. По ком звонит колокол (Н. Волжина, Е. Калашникова, 1968)]
- (7) *Lee Harvey Oswald, the ex-Marine who **supposedly** shot the President with a mail-order Italian rifle* — Ли Харви Освальд, бывший моряк, который **якобы** застрелил Президента из итальянской винтовки, заказанной по почте [Курт Воннегут. Фокус-покус (М. Ковалева, 1993)]

Кстати, легко видеть, что в примере (7) в переводе (при «стандартном» понимании *якобы*) появляется смысл, отсутствовавший в оригинале: оригинальный текст вполне совместим с возможностью, что его автор не отвергает версию, вводимую словом *supposed(ly)* (и даже не исключено, что он склонен принять ее в качестве вероятной), тогда как русский перевод, строго говоря, основан на мысли, что эта версия не имеет ничего общего с действительностью.

Впрочем, в некоторых случаях устанавливаемое на основе параллельного корпуса соответствие зависит от индивидуальных предпочтений переводимого автора и переводчика. Так, Теодор Драйзер регулярно использует слово *ostensibly* для указания на недостоверность того, о чем идет речь (эксплуатируя противопоставление *ostensibly* — *really*), а его переводчик М. Волосов столь же регулярно переводит его как *якобы* (и иногда добавляет противопоставление *якобы* — *в действительности/на самом деле*, даже когда его не было в оригинале, как в примере (9)), напр.:

- (8) *Mr. Obadiah Calvin dropped in—**ostensibly** to see his partner, but **really** to look at Eugene without committing himself.* — ...мистер Обадия Кэлвин случайно заглянул к своему компаньону, **якобы** с целью поговорить с ним по делу, а **в действительности**, чтобы, не выдавая своих намерений, посмотреть на Юджина. [Теодор Драйзер. Гений (ч. 1–2) (М. Волосов, 1930)]
- (9) *Eugene went about among the various editors looking at them, **ostensibly** making their acquaintance...* — Юджин заходил к редакторам различных журналов, **якобы** для того, чтобы познакомиться с ними, а **на самом**

*деле, чтобы приглядеться к их работе...* [Теодор Драйзер. Гений (ч. 1–2) (М. Волосов, 1930)]

- (10) *Another time—several, in fact she came to the apartment in Riverside Drive, ostensibly to see how Mrs. Witla was getting along, but really, of course, to see Eugene.* — В другой раз — вернее, несколько раз — она приходила на Риверсайд-Драйв, **якобы** с целью поведать миссис Витла, но **в действительности** ради свидания с ним. [Теодор Драйзер. Гений (ч. 3) (М. Волосов, 1930)]

В некоторых случаях *якобы* в тексте перевода возникает без непосредственного источника, усиливая смысл ложности, который и без того выражен в контексте, напр.:

- (11) *Others proceed more artfully, as they imagine; and **forge** accusations against themselves...* — Другие действуют, на их взгляд, более искусно: они **придумывают** обвинения, **якобы** направленные против них... [Лорд Честерфилд. Письма к сыну (А. Шадрин, 1971)]
- (12) *...under **pretence** of their being evil-doers, many of the noblest in the land had been dragged forth and wickedly done to death.* — ...много лучших людей страны было предано страшной смерти **под тем предлогом**, что они **якобы** являются преступниками. [Генри Райдер Хаггард. Копи царя Соломона (Н. Б. Маркович, 1958)]

Кроме того, слово *якобы* весьма часто используется в переводах для передачи содержания чужого высказывания, когда в оригинале нет прямого указания на ложность содержащейся в нем пропозиции. Чаще всего в подобных случаях *якобы* используется как показатель цитирования при речевых глаголах и вводит пропозицию, которая в оригинале не оценивается как заведомо ложная:

- (13) *...she [...] started to tell my mother in bad French a pointless and quite irrelevant story about a Polish woman who had attempted to steal her vanity-bag in the dining-car.* — ...она [...] на плохом французском стала рассказывать моей матери нелепую и неуместную историю про какую-то польку, **якобы** пытавшуюся украсть у нее в вагоне-ресторане ридикюль. [Владимир Набоков. Истинная жизнь Себастьяна Найта (А. Долинин, М. Мейлах, 1991)]
- (14) *...so many times you complain that I have concealed evidence and that have hampered your investigation...* — Ранее ведь вы неоднократно жаловались, что я **якобы** скрываю улики и это мешает вашему расследованию. [Эрл Стэнли Гарднер. Дело смелой разведки (М. Кудрявцева, 1990)]
- (15) *Lieutenant, you have stated that I told you I had the gun.* — *Лейтенант, вы заявили, что я вам **якобы** сказал о том, что у меня есть определенный револьвер.* [Эрл Стэнли Гарднер. Дело смелой разведки (М. Кудрявцева, 1990)]
- (16) *...your secretary called me and told me that the gun which had been misplaced had now been found and was available* — ...ваш секретарь позвонила мне и сказала, что револьвер, который **якобы** положили не на то место,



*нашелся, и я могу его взять.* [Эрл Стэнли Гарднер. Дело смелой разведки (М. Кудрявцева, 1990)]

В частности, стимулом для появления слова *якобы* иногда бывает глагол *to claim*, напр.:

- (17) *Knights who **claimed** to be "searching for the chalice" — Рыцари, **якобы** занятые поисками «сосуда»* [Дэн Браун. Код Да Винчи (Н. Рейн, 2004)]

В следующем примере глагол *to claim* в главной клаузе приводит к появлению слова *якобы* во второй из подчиненных клауз (но не в первой). Чем обусловлено такое решение переводчика и какая пропозиция опровергается, не вполне ясно:

- (18) *They **claimed** Langdon and Neveu had arrived earlier and were turned away from the bank because they did not have proper account identification. — Они подтвердили, что Лэнгдон и Неве приезжали сюда чуть раньше этой же ночью, но затем им **якобы** дали от ворот поворот, потому как у них не оказалось соответствующих документов.* [Дэн Браун. Код Да Винчи (Н. Рейн, 2004)]

Иногда в оригинале контекст указывает лишь на отсутствие свидетельств истинности соответствующей пропозиции, тогда как *якобы* в переводе навязывает прочтение ее как заведомо ложной, напр.:

- (19) ***Rumors** existed that the Priory had vowed someday to bring the Grail back to France to a final resting place, but **certainly no historical evidence existed to suggest that this indeed had happened.** — Ходили слухи, что Приорат **якобы** поклялся в один прекрасный день вернуть Грааль во Францию, где он и будет храниться вечно. **Но никаких исторических свидетельств в пользу того, что это действительно произошло, не существовало.*** [Дэн Браун. Код Да Винчи (Н. Рейн, 2004)]

В качестве непосредственного источника также нередко выступают слова, указывающие на обманчивое внешнее впечатление от невербального поведения, намеренно создаваемое субъектом: *to pretend, to seem, seemingly, apparently* и т. д.:

- (20) *...she was forced to **seem** to restore her attention to the orchestra and look straight forward. — ...ей пришлось, **якобы** устремив свое внимание на оркестр, смотреть прямо перед собою.* [Джейн Остин. Доводы рассудка (Е. Суриц, 1988)]
- (21) *'The ribaldry of this miserable man is despicably disgusting,' said Pott, **pretending** to address Bob Sawyer, and scowling upon Slurk. — Сквернословие этого презренного человека гнусно и отвратительно, — сказал Потт, **якобы** обращаясь к Бобу Сойеру, но устремляя грозный взгляд на Слерка.* [Чарльз Диккенс. Посмертные записки Пиквикского клуба (А. В. Кривцова, Е. Л. Ланн, 1933)]

- (22) *She would be hanging about among the stalls, **pretending** to be in search of shoelaces or sewing-thread. — Она будет бродить возле ларьков **якобы** в поисках шнурков или ниток.* [Джордж Оруэлл. 1984 (В. Гольшев, 1989)]

- (23) *Ralph looked away first, **pretending** interest in a group of littluns on the sand. — Ральф первый отвел глаза, **якобы** посмотреть на играющих в песке малышей.* [Уильям Голдинг. Повелитель мух (Е. Суриц, 1985)]
- (24) *alone and **apparently** unprotected — один, **якобы** никем не охраняемый* [Артур Конан Дойл. Собака Баскервилей (Н. Волжина, 1948)]
- (25) *his **seemingly** courteous interest — вежливый и **якобы** неподдельный интерес* [Маргарет Митчелл. Унесённые ветром, ч. 1 (Т. Озерская, 1982)]
- (26) *...a number of eyes which **seem** to look everywhere except in our direction. — ...множество глаз, глядящих **якобы** в какую угодно, только не в нашу сторону.* [Владимир Набоков. Смотри на арлекинов! (С Ильин, 1999)]

Иногда идея намеренного обманчивого поведения очевидна из более широкого контекста, напр.:

- (27) *Olaf had disguised himself as a person who did not speak English well, and **as part of his disguise**, he had spoken in an accent very similar to the one they were hearing now. — Олаф изображал человека, который плохо говорит по-английски, и для этого объяснялся с **якобы** иностранным акцентом, очень похожим на тот, который они слышали сейчас.* [Лемони Сникет. Кроважандный карнавал (Н. Рахманова, 2005)]

Отдельного упоминания заслуживают случаи, когда характеристика передаваемой пропозиции как ложной принадлежит не автору (или говорящему), а персонажу, который передает слова или мнения других людей и высказывания которого передаются в форме косвенной речи, напр.:

- (28) *[He] forbade the slightest hint being dropped of his having such an intention — запрещал даже самое малейшее намек о **якобы** имевшихся у него на этот счет видах* [Джейн Остин. Доводы рассудка (Е. Суриц, 1988)]

Итак, что же нам дает материал англо-русского корпуса? По-видимому, он позволяет сделать следующие выводы.

Во-первых, частица *якобы* может указывать не только на ложность вербально выраженной пропозиции, но и на невербальное поведение, причем примеров такого рода в англо-русском подкорпусе НКРЯ довольно много. Естественно задаться вопросом, насколько такой тип употребления характерен для русских оригинальных текстов.

Во-вторых, мы видим, что в ряде случаев *якобы* указывает не на ложность пропозиции, а лишь на сомнения в ее истинности. Иногда он используется просто как показатель «чужого слова», достоверность которого ничем не обоснована.

В-третьих, в некоторых случаях сфера действия *якобы* не вытекает прямо из его позиции в предложении: имеет место некоторый сдвиг.

Все это побуждает обратиться к современным оригинальным русским текстам, чтобы проверить, обнаруживаются ли в них эти явления.

### 3. *Якобы* по данным оригинальных русских текстов

Прежде всего, следует отметить, что фразы, в которых *якобы* маркирует не ложное высказывание (или его часть), а невербальное поведение (т. е. имеет значение 'делает вид'), не так редки и в русских оригинальных текстах и не воспринимаются как сдвинутые. Приведем примеры из НКРЯ:

(29) ... *вдруг в последнее мгновение хватал гитару* (на репетиции он хватал большой треугольник) *и, якобы бряцая по струнам, запевал...*  
[Фазиль Искандер. Мученики сцены (1989)]

(30) *Слышит ли меня на восьмом ряду тот самый, сидящий в позе якобы кучера?* [Игорь Мартынов. Это наша с тобой абстиненция (1997) // «Столица», 1997.03.18]

(31) *Кулинар решил приготовить в честь дня рождения Скиннера фруктовый салат и сделал вид, что блендер неисправен. Он засунул руку в чашу устройства, оно внезапно заработало, и из руки шеф-повара якобы потекла кровь* (lenta.ru 31/03/2017)

Поэтому формулировка, касающаяся *якобы* как маркера ложности (или хотя бы сомнения в истинности) пропозиции, содержащейся в чужом высказывании, нуждается в уточнении — речь может идти не о высказывании, а о невербальном поведении.

Кроме того, в современных СМИ часто встречаются сдвинутые употребления *якобы*, причем сдвиги могут касаться разных языковых уровней употребления — семантики, прагматики, синтаксиса.

При «стандартном» употреблении *якобы* ставится перед вершиной той пропозиции, к которой семантически относится:

(32) *Он сказал, что якобы не брал у нее деньги* [подвергается сомнению то, что он не брал у нее деньги].

(33) *Он сказал, что торгует якобы лицензионными дисками* [подвергается сомнению то, что диски, которыми он торгует, лицензионные].

Однако в текстах Интернета и СМИ встречаются употребления, где это правило нарушается: *якобы* ставится перед речевым глаголом, вводящим цитируемую речь. При этом семантически *якобы* относится не нему, а либо к содержанию цитируемой речи, либо к модальным показателям в главной клаузе, ср. примеры из корпуса RuTenTen на ресурсе Sketch Engine:

(34) *Этот мысленный эксперимент позволяет якобы утверждать, что бытие, esse, existentia, есть нечто отличное от сущности* ['якобы позволяет'].

Постановка *якобы* перед речевым глаголом, вводящим косвенную речь, может свидетельствовать о начале грамматикализации этой частицы как показателя недостоверности.

Можно привести и другие примеры сдвига семантической сферы действия *якобы*; ср. следующую фразу, где синтаксически *якобы* относится к глаголу, однако по смыслу — к оценочному слову *компромат*<sup>4</sup>:

(35) *Он также назвал якобы собранный на него компромат «чушью и какими-то бумажками», отметив, что публикации похожи на «компот».*

Другой процесс, в который вовлечена частица *якобы*, — это лексикализация, которая приводит к созданию в СМИ устойчивых клише. Их употребление отражает некую «стратегию безопасности», согласно которой говорящий не хочет брать на себя ответственность за потенциально неверную информацию и использует *якобы*, чтобы от нее отстраниться. Это происходит даже в тех случаях, когда употребление клише входит в противоречие с естественной интерпретацией:

(36) *После победы Трампа на выборах демократы отказались от услуг Стила. Однако он продолжил работать над докладом вместе с бывшим журналистом Wall Street Journal Гленном Симпсоном на безвозмездной основе в надежде, что расследование якобы имевших место кибератак со стороны России позволит сделать эти материалы достоянием общественности.*

Противоречие состоит в том, что субъект установки, выраженной в слове *надежда*, совпадает с субъектом недоверия, выраженного в слове *якобы*, что прагматически неадекватно.

Злоупотребление стратегией безопасности приводит к употреблению, которые можно назвать *self-defeating*, например, когда говорящий подвергает сомнению как всю ситуацию в целом, так и ее отдельные аспекты, предполагая, что вся ситуация имеет место:

(37) *Гражданка Узбекистана Нилуфар Мамасаидова сейчас находится со своим ребенком, якобы отобранным ранее сотрудниками московской полиции. <...> В конце декабря Мамасаидова обнаружила пропажу документов, удостоверяющих личность, и сразу же обратилась с соответствующим заявлением в ОВД МВД по району Кунцево. Вскоре к ней приехали сотрудники полиции и, якобы основываясь на том, что у женщины нет документов, отобрали у нее рожденную на территории России годовалую дочь.*

В первом предложении сомнению подвергается то, что ребенка в принципе отобрали, а во втором — причины, по которым это произошло. По-видимому, в данном примере *якобы* употребляется не для того, чтобы выразить сомнение в достоверности информации, а в качестве показателя эвиденциальности — говорящий таким образом маркирует информацию, полученную из вторых рук.

---

<sup>4</sup> Этот и следующие два примера с сайта [lenta.ru](http://lenta.ru).

#### 4. Заключительные замечания

Представляется, что результаты проведенного исследования должны быть учтены в лексикографической практике, в частности при составлении нормативных словарей. В большинстве словарей говорится, что *якобы* указывает на сомнение в достоверности сообщаемого в высказывании. Тем самым не учитывается возможность отнесения частицы *якобы* к тому, что выражено в невербальном поведении<sup>5</sup>. Толкование, предложенное в статье [Плунгян 2008]: «*якобы* X в первом приближении означает ≈ 'некоторые считают, что X истинно; я так не считаю'», — также не вполне соответствует использованию *якобы* при характеристике невербального поведения, поскольку эта частица указывает на видимость того, что пытается выразить субъект, несоответствие этого действительности и намеренность обмана, тогда как компонент 'некоторые считают, что X истинно; я так не считаю' не покрывает указанные компоненты.

Отдельного обсуждения заслуживает вопрос о сдвигах в синтаксическом поведении частицы *якобы* (изменение сферы действия, чем-то сходное с «подъемом отрицания»). Наконец, необходима нормативная оценка употреблений *якобы* с семантическим и прагматическим сдвигом, когда она используется в рамках стратегии безопасности, как показателя информации, за достоверность которой говорящий не готов поручиться.

В целом мы видим, что многое из того, что могло оцениваться как неадекватность перевода вполне отвечает современным тенденциям в употреблении *якобы*. При «стандартной» интерпретации появление в переводе слова *якобы* вносит дополнительный смысл, которого не было в оригинале — 'данная пропозиция ложна'. Однако в соответствии с употреблением, встречающимся и в оригинальных русских текстах (особенно — в современных СМИ), смысл более слабый — 'я не ручаюсь за достоверность данной пропозиции', и, как правило, подобный смысл содержался и в оригинале. Тем самым появление *якобы* в переводе можно рассматривать не как ошибку переводчика, а как употребление слова *якобы* с семантическим сдвигом.

Как кажется, разные носители языка воспринимают примеры такого рода несколько по-разному. Для некоторых *якобы* — отчетливый показатель того, что говорящий считает соответствующую пропозицию ложной, и они воспринимают рассматриваемые переводы как неточные, а примеры из оригинальных русских текстов — как небрежные. Для других *якобы* — всего лишь показатель снятия с себя ответственности за содержание чужого высказывания или поведения, и они могут воспринимать переводы как вполне адекватные, а оригинальные русские примеры — как вполне нормативные.

---

<sup>5</sup> Именно при передаче чужого невербального поведения ярче всего проявляется различие между *мол* и *якобы*, соответствующие различию стратегий *de re* и *de dicto* (для передачи невербального поведения точнее было бы говорить о стратегиях *de re* и *de visu*). Частица *мол* маркирует смысл наблюдаемого поведения как его интерпретирует говорящий (стратегия *de re*), тогда как частица *якобы* указывает на мнимый смысл наблюдаемого поведения как его пытается представить субъект (стратегия *de visu*).

Возможно также, что подтверждение кажущихся неправильными употреблений *якобы* в современных СМИ данными из переводов свидетельствует о том, что в русском языке имеется потребность в нейтральном грамматикализованном показателе эвиденциальности, функцию которого и начинает выполнять *якобы*.

## Литература

1. Арутюнова Н. Д. (1992), Речеповеденческие акты в зеркале чужой речи, Булыгина Т. В. (отв. ред.) Человеческий фактор в языке. Коммуникация. Модальность. Дейксис, Наука, Москва, с. 40–52.
2. Арутюнова Н. Д. (2000), Показатели чужой речи *де, дескать, мол*, Арутюнова Н. Д. (ред.) Язык о языке, с. 437–452.
3. Баранов А. Н. (1994), Заметки о *дескать* и *мол*, Вопросы языкознания, 4, с. 114–124.
4. Падучева Е. В. (2011), Показатели чужой речи *мол* и *дескать*, Известия РАН, серия литературы и языка, Т. 70, 3, с. 13–19.
5. Плуноян В. А. (2008), О показателях чужой речи и недостоверности в русском языке: *мол, якобы* и другие, Wiemer B. & V. A. Plungjan (Hrsg.), Lexikalische Evidenzialitäts-Marker in slavischen Sprachen [Wiener Slawistischer Almanach, Sonderband 72], Sagner, München, S. 285–311.
6. Шмелев А. Д. (2016), «Невысказанные мысли» в историко-культурной перспективе, Арутюнова Н. Д. (ред.) Логический анализ языка: Информационная структура текстов разных жанров и эпох, Гнозис, Москва, с. 92–106.
7. Шмелев А. Д. (2015), Русские лингвоспецифичные лексические единицы в параллельных корпусах: возможности исследования и «подводные камни», Компьютерная лингвистика и интеллектуальные технологии, с. 584–594.
8. Rakhilina E. V. (1996), *Jakoby* comme un moyen de médiatisation en russe, Guentchéva Z. (ed.). L'énonciation médiatisée, Peeters, Paris, p. 299–304.

## References

1. Arutjunova N. D. (1992), Speech acts as mirrored in indirect speech [Rechepovedencheskie akty v zerkale chuzhoj rechi], Bulygina T. V. (ed.). Human factor in language. Communication. Modality. Deixis [Kommunikacija, Modal'nost', Dejksis], Nauka, Moskva, pp. 40–52.
2. Arutjunova N. D. (2000), Markers of indirect speech *de*, *deskat'*, *mol* [Pokazateli chuzhoj rechi *de*, *deskat'*, *mol*], Arutjunova N. D. (ed.) Language about language [Jazyk o jazyke], pp. 437–452.
3. Baranov A. N. (1994), Notes on *deskat'* and *mol* [Zametki o *deskat'* i *mol*], Issues in linguistics [Voprosy jazykoznanija], 4, pp. 114–124.
4. Paducheva E. V. (2011), Markers of indirect speech *mol* and *deskat'* [Pokazateli chuzhoj rechi *mol* i *deskat'*], Bulletin of the Russian Academy of Sciences, series of Literature and Language (Izvestija RAN, serija literatury i jazyka), vol. 70, 3, pp. 13–19.
5. Plungian V. A. (2008), On markers of indirect speech and unreliability [O pokazateljax chuzhoj rechi i nedostovernosti v russkom jazyke: *mol*, *jakoby* I drugie], Wiemer B. & V. A. Plungian (Hrsg.), Lexikalische Evidenzialitäts-Marker in slavischen Sprachen [Wiener Slawistischer Almanach, Sonderband 72], Sagner, München, S. 285–311.
6. Rakhilina E. V. (1996), *Jakoby* as a means of mediatization in Russian [*Jakoby* comme un moyen de médiatisation en russe], Guentchéva Z. (ed.). L'énonciation médiatisée, Peeters, Paris, pp. 299–304.
7. Shmelev A. D. (2016), "Unsaid thoughts" in historical and cultural perspective ["Nevyskazannye mysli" v istoriko-kul'turnoj perspektive], Arutjunova N. D. (ed.). Logical analysis of Language: Information structure of texts of different genres and periods [Logicheskij analiz jazyka: Informacionnaja struktura tekstov raznyx zhanrov i epox], Gnosis, Moscow, pp. 92–106.
8. Shmelev A. D. (2015), Russian language-specific lexical units in parallel corpora: prospects of investigation and "pitfalls" [Russkie lingvospecificchnye leksicheskie edinicy v paralel'nyx korpusax: vozmozhnosti issledovanija i "podvodnye kamni"], Computational linguistics and intellectual technologies [Kompjuter-naja lingvistika i intellektual'nye texnologii], pp. 584–594.

# ОПЫТ КОРПУСНОГО МОДЕЛИРОВАНИЯ ФАКТОРОВ МЕТАФОРИЧНОСТИ НА ПРИМЕРЕ РУССКИХ ГЛАГОЛОВ<sup>1</sup>

**Бадрызлова Ю. Г.** (yuliya.badryzlova@gmail.com)

Национальный Исследовательский Университет «Высшая  
Школа Экономики» (НИУ ВШЭ), Москва, Россия

## A CORPUS-BASED STUDY OF FACTORS AND MODELS METAPHORICITY; EVIDENCE FROM RUSSIAN VERBS

**Badryzlova Yu. G.** (yuliya.badryzlova@gmail.com)

National Research University Higher School of Economics  
(HSE), Moscow, Russia

The paper presents a tentative scheme for the analysis and annotation of the factors that contribute to metaphoricity of verbs in real-world data. The need for a comprehensive metaphor annotation scheme was demonstrated in the comparative study of four state-of-the-art automatic metaphor identification systems (Dunn, 2013c), which showed low agreement on a gold standard. The proposed scheme aims to capture a variety of factors involved in metaphoricity, their cumulative effect, and the gradual nature of metaphoricity.

The proposed inventory of metaphoricity factors includes nonbasic meanings (the less concrete, body-related and precise meaning(s) of a word); semantic shifts (additional semantic components of the meaning), newly attested meanings; morphologically motivated metaphoricity; different types of personification (a switch in the semantic class of the argument from animate to inanimate); a switch in the semantic class of the argument between concrete and abstract; the use of a metonymic argument, participation in idiomatic expressions; and direct metaphors – the direct use of a word involved in a cross-domain mapping.

We also attempt to represent the dynamic interactions between these factors using graph-like visualization. Personifications and direct metaphors are described in terms of conceptual sources and targets. Secondary metaphoricity is attested for words that participate in extended metaphors.

---

<sup>1</sup> В данной научной работе использованы результаты проекта «Точки роста в языковой системе: корпусные исследования и моделирование», выполненного в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» в 2017 году.



Metaphorically related words can be signaled by different types of cues: graphically, or by specific grammatical constructions, or by lexical cues.

The evidence is drawn from 2,036 occurrences of 100 Russian verbs sampled from the Russian syntactically annotated corpus SynTagRus.

**Keywords:** Factors of metaphoricity, metaphor corpus, metaphor annotation scheme, Russian verbal metaphor, dynamic metaphoricity

## Введение

Разработка систем автоматической идентификации метафоры является активно развивающейся областью автоматической обработки текста, в которой используется машинное обучение на основе различных технологических подходов и инструментов — как с использованием созданных вручную баз данных и онтологий, так и с применением статистических, дистрибутивно-семантических методов. Исчерпывающий обзор систем, технологий и инструментов представлен в работе (Veale et al., 2016). Однако при ознакомлении с дизайном этих систем нельзя не заметить, что они значительно различаются не только технологически, но и лежащими в их основе теоретическими положениями, а также задачей и объектом идентификации. По сути, объекты, идентифицируемые различными системами, представляют собой явления разного порядка.

### 1. Различные трактовки метафоры в корпусных ресурсах и автоматических системах

Подобное разночтение в понимании объекта метафоры среди создателей систем автоматической идентификации было наглядно продемонстрировано в проведенном Дж. Данном (Dunn, 2013c) эксперименте по сравнению эффективности четырех систем на едином золотом стандарте. Три из принявших участие в сравнении систем (1–3) основаны на методах дистрибутивной семантики, а одна (4) — на онтологических методах. Однако обратим здесь внимание на различие в понимании их разработчиками того, что собственно является метафорой.

1) В основе системы, предложенной (Li and Sporleder, 2009; Li and Sporleder, 2010), лежит предпосылка, что метафорическое употребление является семантической аномалией на фоне неметафорического контекста. Система идентифицирует фразы, являющиеся идиоматическими выражениями, например, “*to trip the light fantastic*” (=to dance), а также отдельные слова в переносном значении, например, “*During the Iraq war, he was a sparrow; he didn’t condone the bloodshed but wasn’t bothered enough to go out and protest.*” При этом из числа идентифицируемых метафор исключаются высокочастотные лексикализованные метафоры, например, ‘*summit*’ в значении «встреча на высшем уровне».

2) Система, разработанная (Turney et al., 2011), определяет степень абстрактности/конкретности слова и сравнивает ее с аналогичными

показателями контекста. Теоретическую основу составляют положения теории концептуальной метафоры (Lakoff and Johnson, 1980), согласно которым метафора представляет собой концептуальный перенос из конкретной сферы-источника на абстрактную сферу-цель. Таким образом, объект данной системы главным образом представлен лексикализованными метафорами типа “*deep appreciation*” (в отличие от неметафорического употребления “*deep snow*”).

- 3) Основанием системы, представленной в (Shutova, Sun and Korhonen, 2010), также служит теория концептуальной метафоры. Система осуществляет в тексте поиск лексических единиц, представляющих концепты сферы-источника и сферы-цели. Предполагается, что и источник, и цель должны быть выражены на лексическом уровне, причем глагол представляет сферу-источник, а существительное, находящееся к нему в синтаксическом отношении подлежащего или дополнения, представляет сферу-цель. Таким образом, система идентифицирует пары глагол-прямое дополнение (например, “*grasp theory*”, “*suppress memory*”) и субъект-глагол (например, “*tension mounted*”, “*ideology embraces*” и т. п.).
- 4) Система, описанная в работе (Dunn, 2013a), как и предыдущая, нацелена на поиск лексических единиц, репрезентирующих в тексте сферу-источник и сферу-цель, однако наличие между ними синтаксической связи не является обязательным. Поиск осуществляется на основе онтологического подхода (Nirenburg and Raskin, 2004; Niles and Pease, 2001).

В качестве золотого стандарта для сравнения перечисленных систем Дж. Данн использовал Амстердамский корпус метафоры (сокращенное название VUAMC, по месту создания — Vrije Universiteit Amsterdam) (Steen et al., 2010a), который, в свою очередь, основан на собственных теоретических предпосылках и методологических принципах.

Англоязычный корпус VUAMC, содержащий около 200 тыс. словоупотреблений, размечен в соответствии с принципами процедуры MIPVU (сокращение от Metaphor Identification Procedure Vrije Universiteit) (Steen et al., 2010b). В основе процедуры лежит определение базового значения (БЗ) слова. БЗ — это такое из зафиксированных в словаре значений слова, которое является наиболее конкретным, телесно-ориентированным и четким (ibid.: 5–6).

Если в рассматриваемом контексте слово употреблено в значении, которое обладает менее ярко выраженной конкретностью, телесноориентированностью и четкостью, чем БЗ, то такое употребление в MIPVU обозначается термином «непрямая метафора» (indirect metaphor), например:

- (1) *Mr Franklin went there at the end of the 1970s, after the collapse of Keyser Ullman...*

Установление БЗ и определение степени контраста между ним и контекстуальным значением представляет собой наиболее уязвимый и подверженный субъективизму момент процедуры MIPVU; однако он неизбежен и снимается привлечением к анализу двух и более аннотаторов и принятием согласованных решений.

Помимо непрямой метафоры, процедура MIPVU и размеченный в соответствии с ней корпус VUAMC выделяют прямую метафору (direct metaphor) и потенциальное олицетворение (possible personification).

Прямая метафора — это употребление лексической единицы в ее прямом значении для обозначения концепта, принадлежащего к иной концептосфере, чем сама эта лексическая единица, т. е. когда имеет место междоменный перенос:

(2) *He's like a favourite old coat.*

Потенциальное олицетворение определяется как употребление неантропной (не обозначающей человека) лексики в качестве аргумента лексической единицы, которая в своем базовом значении требует антропного аргумента:

(3) *Paula's stomach turned a somersault.*

Кроме того, в корпусе MIPVU размечены случаи, когда одновременно реализуются два из вышеперечисленных типов метафоры (Double), и пограничные случаи между метафорой и неметафорой (When In Doubt Leave It In, WIDLII).

В ходе тестирования по стандарту VUAMC перечисленные выше системы продемонстрировали различные показатели эффективности (см. Dunn, 2013c: 4–9). Однако наиболее интересный для нас результат состоит в невысоком уровне их согласованности (Fleiss' Каппа), варьирующемся в диапазоне от 0,259 до 0,293. Это означает, что в большинстве случаев системы принимают правильные и ошибочные решения в различных, не совпадающих друг с другом случаях, что вполне ожидаемо, учитывая представленные выше различия во взглядах на к определению предмета и теоретические предпосылки. На основании этого результата автор эксперимента делает вывод о том, что для эффективной идентификации метафоры на реальных данных необходим синтез подходов и технологий.

Такой синтезирующий подход был реализован в системе MetaNet (Dodge, Hong and Stickle, 2015), однако информации о тестировании этой системы на каком-либо золотом стандарте мы на настоящее время не располагаем.

Для данной статьи наибольший интерес представляют теоретические выводы, к которым приходит Дж. Данн по результатам сравнения систем (Dunn, 2013c):

1. Метафора представляет собой градиентное явление. Одни метафорические высказывания обладают более высокой степенью метафоричности, тогда как другие менее метафоричны. Проблема протестированных систем состоит в том, что они пытаются проводить строгое бинарное разделение между явлениями, имеющими градиентную природу.
2. Метафоричность высказывания может происходить из нескольких различных источников.

В последующем исследовании (Dunn, 2013b) автор предлагает оценивать степень метафоричности высказывания в зависимости от его насыщенности метафорическим материалом.

Эти выводы, а также наблюдения, полученные нами в ходе предшествующих исследований (разметка русскоязычного корпуса по аналогии с VUAMC (Badryzlova, Kerimov and Shekhtman, 2013; Badryzlova et al., 2013) и идентификация лексикализованной метафоры методами машинного обучения

с применением дистрибутивной семантики (Badryzlova, 2015; Panicheva and Badryzlova, 2017)) — убеждают нас в востребованности более детализированного корпусного описания факторов и источников метафоричности, включающих в себя собственную семантику лексических единиц, их синтагматические взаимодействия на семантическом и синтаксическом уровне, изменения собственной семантики и сочетаемости, происходящих в результате этих взаимодействий, а также метафорические переходы между концептосферами. Подобное описание, опирающееся на существующий лексикографический и корпусный инструментарий, призвано способствовать более полному моделированию метафорических явлений для целей автоматической идентификации метафоры.

## 2. Методика исследования

Представленная в Разделе 3 классификация факторов метафоричности и методика их описания носят предварительный характер. На настоящем этапе нами ставилась задача обеспечить как можно более подробную детализацию разметки, чтобы максимально охватить круг явлений, способствующих возникновению метафоричности. В дальнейшем эта методика может быть упрощена и адаптирована для целей автоматической аннотации. Представленные количественные результаты носят предварительный характер и для подтверждения достоверности подлежат анализу с участием дополнительных аннотаторов и определением меры согласованности, как принято в практике разметки метафорических ресурсов.

Анализ производился на материале глаголов в синтаксически размеченном корпусе СинТагРус (Boguslavsky et al., 2009), для чего были отобраны 100 глаголов с частотностью 19–25 вхождений и общим числом 2036 вхождений.

В качестве основного лексикографического источника для определения значений лексических единиц использовалась электронная версия Словаря русского языка под ред. А. П. Евгеньевой (Yevgenyeva, 1981–1984). В случае отсутствия в нем того или иного значения применялись онлайн-версия Толкового словаря С. И. Ожегова и Н. Ю. Шведовой (Ozhegov and Shvedova) и онлайн-словарь Русский Викисловарь (Russian Wiktionary).

Разметка производилась с помощью инструмента лингвистического аннотирования brat (Stenetorp et al., 2012).

## 3. Факторы метафоричности

Поскольку анализ производился на материале, представленном глаголами с определенными характеристиками (невысокая частотность, ограниченная полисемия), то предлагаемое описание не претендует на исчерпывающую полноту. Возможно, что включение лексических единиц с другими характеристиками (частота, количество значений, часть речи) потребует расширения или дополнения предлагаемого описания.

При разработке системы факторов мы опирались на исследования механизмов языкового расширения и типов производных значений, представленных в работах Е. В. Падучевой (Paducheva, 2004) и Г. И. Кустовой (Kustova, 2004).

### 3.1. Небазовые значения

Базовые значения (БЗ) определяются в соответствии с критериями, сформулированными для процедуры MIPVU.

Все значения, не попавшие в группу БЗ, относятся к группе небазовых значений (НБЗ), которые представляют собой один из факторов метафоричности.

Каждое глагольное употребление в корпусе получает помету [BM] (basic meaning) либо [NBM] (non-basic meaning), например:

(4) «Гореть»:

- ‘Поддаваться действию огня, уничтожаться огнем.’:

[BM]

На юге Москвы горит жилой дом, сообщили РИА "Новости" ...  
в столичном управлении государственной противопожарной службы.

- ‘Давать свет, пламя.’:

[BM]

Если горит зеленый свет, это значит, что предыдущий поезд  
уехал очень далеко и можно смело отправляться в путь ...

- ‘Топиться (о печи).’:

[BM]

Печка горела бойко, светлые пятна и темные тени прыгали по стене ...

- ‘Быть охваченным каким-л. сильным чувством, со страстью отдаваться чему-л. (какому-л. чувству, делу).’:

... Директором и Генеральным конструктором назначают Валентина Глушко,

[NBM]

горящего желанием поставить победную точку в давнем споре с Королевым ...

- ‘Блестеть под влиянием какого-л. чувства, мысли (о глазах).’:

[NBM]

... (ее) глаза из-под платка горели страстью очереди.

- ‘Быть под угрозой срыва, провала из-за опоздания, упущения сроков.’:

Обком партии поручал Сафронову под личную ответственность

[GRAPH\_CUE]  
[NBM]

наиболее важные, часто "горящие" объекты ...

([GRAPH\_CUE] — graphic cue: неалфавитный знак, указывающий на наличие метафоричности, в данном случае кавычки.)

В некоторых случаях бывает проблематично однозначно классифицировать значение как БЗ или НБЗ — например, когда значение, уже являясь небазовым, продолжает быть слабо связанным со значением, обладающим

характеристиками предметности или телесности. Такие значения помечаются тэгом [DUAL]:

(5) «Дрожать»: 'Испытывать опасение, тревогу; опасаться.':

Но, как это бывает с людьми, добившимися положения исключительно благодаря

свирепому комплексу неполноценности и потому дрожащими за свой авторитет, он никогда не признавал своих промахов.

([CONSTR] — construction transformation: изменение конструкции, см. п. 3.7.)

Лексемы, все значения которых являются только базовыми или только небазовыми, не обладают собственной внутренней метафоричностью, вызываемой противопоставлением БЗ и НБЗ. Несмотря на это, в некоторых случаях такие лексемы могут становиться метафоричными под влиянием факторов, описанных ниже.

Все значения лексем, не обладающих ни одним БЗ, получают специальную помету [ALL\_NBM] (all non-basic):

(6) «Уважать»:

- 'Испытывать чувство уважения, почтения по отношению к кому-, чему-л.':

Шаманизм был истреблен в советское время, хотя раньше на острове

жили около 40 шаманов, уважаемых во всем Забайкалье.

- 'Принимать что-л. во внимание, считаться с чем-л., относиться с уважением.':

По-разному можно относиться к судам, но я в любом случае

уважаю их решение как законопослушный гражданин.

- 'Любить, ценить за что-л.':

"Финские унитазы я вообще-то не очень уважаю", - плавно жуя, сказал сантехник.

### 3.2. Семантические сдвиги

Семантические сдвиги ([SEM\_SHIFT], semantic shift) значения констатируются в тех случаях, когда помимо основной семантики лексемы, выраженной словарным определением, в рассматриваемом контексте у значения появляются дополнительные семантические компоненты, не входящие в определение, например:

(7)

В годы юности Ильи Ильича девушки бегали в домотканых сарафанах - и нравились.

(Основная семантика: «Бегать»: 'Быстро, поспешно идти куда-л., откуда-л.:'; дополнительный компонент: «Носить»: 'Надевать на себя что-л., украшать себя чем-л. по принятой моде.')

### 3.3. Инновации

При образовании нового значения ([NEW\_MEANING]), в отличие от семантического сдвига, собственная семантика исходного значения практически полностью вытесняется новым семантическим наполнением:

(8)

NEW MEANING
NBM

... на сайты, заподозренные в связях с экстремистами, сбрасываются послания на арабском, подписанные "Абдуллою Ахмедом".

«Сбрасывать» (*син.* «переслать» — ‘Отослать, отправить.’ (дополнительный компонент — ‘электронное устройство’) и «записать» — ‘Комп. Сохранить данные на устройство хранения информации.’)

Особый случай представляют значения, принадлежащие к терминологическим метаязыкам. Зачастую образование термина ([TERM]) происходит путем метафоризации по основанию, которое является непрозрачным для среднего носителя языка, не являющегося специалистом в предметной области:

(9) Тридцать лет назад в ультразвуковой диагностике начали развиваться

TERM
NEW MEANING
NBM

так называемые методы согласованной фильтрации ...

### 3.4. Метафоричность, обусловленная внутренней формой

Метафоричность, обусловленная внутренней формой слова, наблюдается у лексем, не обладающих собственным БЗ или обладающих слабо выраженным БЗ, когда их морфологические компоненты этимологически связаны с другой лексемой, демонстрирующей в современном языке ярко выраженные характеристики БЗ. Например, глагол «исчерпывать», обладающий слабо выраженным БЗ ‘Полностью использовать.’, этимологически связан с «черпать», который обладает ярко выраженным БЗ ‘Захватывая чем-л., набирать, доставать (обычно снизу, из глубины) что-л. жидкое, сыпучее.’.

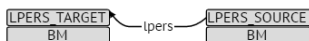
### 3.5. Олицетворение

В отличие от корпуса VUAMC, определяющего олицетворение как смену антропного аргумента на неантропный, мы предлагаем рассматривать олицетворение шире — как смену класса аргумента с одушевленного на неодушевленный. При этом мы различаем источник олицетворения [TARGET] — лексему, чей аргумент меняет класс, и цель олицетворения [SOURCE] — лексему, заполняющую позицию аргумента.

В зависимости от того, насколько одушевленность аргумента интегрирована в значение, мы предлагаем различать лексикализованное олицетворение, олицетворение, образованное контекстуальной сменой класса аргумента, и метафорическое олицетворение.

Лексикализованное олицетворение ([LPERS], lexical personification) выражается зафиксированным в словаре значением (БЗ или НБЗ) лексемы, которое, во-первых, в качестве аргумента принимает только неодушевленные объекты, и, во-вторых у этой же лексемы имеется БЗ, требующее исключительно одушевленного аргумента:

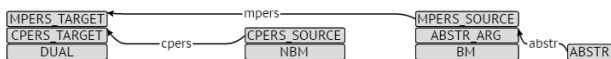
- (10) «Съедать»: ‘Причинить повреждения чему-л. (о редких, ядовитых веществах, ржавчине и т. п.).’:



Но, несмотря на все меры ухода, ржавчина понемногу съедает старые мечи.

Олицетворение, образованное контекстуальной сменой класса аргумента с одушевленного на неодушевленный ([CPERS], contextual personification) происходит, когда значение лексемы, допускающее выбор между этими двумя типами аргумента, в рассматриваемом контексте принимает неодушевленный аргумент:

- (11) «Завоевывать»: ‘Привлечь на свою сторону, расположить к себе.’:



Это произведение во все времена завоевывало сердца, покупало свободу и стоило людям жизни.

([ABSTR\_ARG и ABSTR] — смена семантического класса аргумента с конкретного на абстрактный, см. п. 3.6)

Метафорическое олицетворение ([MPERS], metaphoric personification) наблюдается, когда значение лексемы (БЗ или НБЗ) в норме требует только одушевленного аргумента, но в рассматриваемом контексте принимает неодушевленный аргумент, как «покупать» в примере (11).

### 3.6. Смена семантического класса аргумента: конкретный → абстрактный

Семантическая сочетаемость некоторых значений определена достаточно широко таким образом, что она допускает как абстрактные, так и конкретные аргументы. Когда в рассматриваемом контексте выбор происходит в пользу абстрактного аргумента [ABSTR], этот выбор может придавать глаголу [ABSTR\_ARG] некоторую степень метафоричности: см. «покупать» в (11).

### 3.7. Трансформация конструкции

Специфические модели грамматической сочетаемости ([CONSTR]) могут являться индикаторами метафоричности. Например, конструкция ДРОЖАТЬ ЗА СУЩ<sub>ВИН</sub> в примере (5) является индикатором метафорического употребления, выраженного НБЗ. А в следующем примере инструментальная конструкция ВЗГЛЯНУТЬ СУЩ<sub>ТВОР</sub> сигнализирует источник прямой метафоры (см. п. 3.10):

- (12)



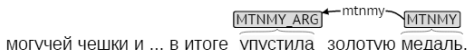
ястребом взглянул на Илью Ильича...



### 3.8. Метонимический сдвиг в аргументе

Метонимический аргумент, сам являясь результатом семантического переноса, предположительно, может оказывать влияние и на семантику глагола. (Глагол, принимающий метонимичный аргумент — [MTNMY\_ARG], метонимичный аргумент — [MTNMY]):

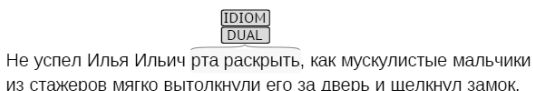
- (13) Юля Чепалова так увлеклась борьбой с полькой, что прозевала рывок



### 3.9. Вхождение в состав идиоматических единиц

Такое фундаментальное качество идиом [IDIOM] как переосмысление лексико-грамматического состава позволяет отнести их к числу факторов, способствующих метафоричности. При этом фразеологические сращения, утратившие мотивировку значения, всегда будут иметь небазовое значение; в то же время, значения фразеологических единств — единиц, в той или иной степени сохраняющих внутреннюю форму, в некоторых случаях могут быть описаны как [DUAL]:

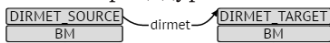
- (14)



### 3.10. Прямая метафора

Для обозначения лексических единиц в прямом значении, участвующих в схемах концептуальных переносов из сферы-источника [SOURCE] на сферу-цель [TARGET], мы используем термин «прямая метафора» ([DIRMET], direct metaphor), заимствованный из процедуры MIPVU:

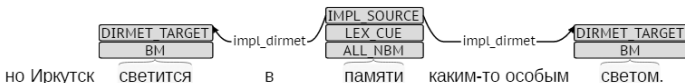
- (15) А голова у нее [птицы] красная-красная, так и горит! - пылко вскричал младший.



В этом предложении наблюдается метафорическая схема с возможной формулировкой ЯРКО-КРАСНЫЙ ЦВЕТ ЭТО ПЛАМЯ. При этом и сфера-источник, и сфера-цель выражены эксплицитно, т.е. представлены соответствующими лексическими единицами.

Однако зачастую на поверхностном уровне лингвистической структуры может быть представлена только одна из концептосфер, участвующих в переносе:

- (16) Многое, случившееся куда позже, давно позабыто,

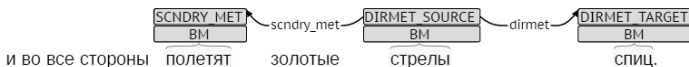


В этом примере активируется схема переноса, которую можно условно сформулировать как ХОРОШИЕ ВОСПОМИНАНИЯ ЭТО СВЕТ, при этом лексическое выражение получает только сфера-источник. Однако в предложении содержится лексический ключ [LEX\_CUE] «память» который уточняет имплицитно выраженную сферу-цель ([IMPL\_SOURCE]), указывая, к какой именно концептосфере происходит обращение.

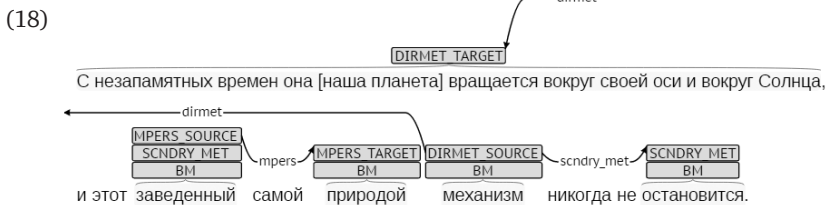
### 3.11. Сложные метафорические комплексы

Описанные выше факторы метафоричности зачастую действуют не изолированно, а совместно друг с другом, образуя сложные метафорические комплексы. Особую сложность структура таких комплексов приобретает в случаях, когда в них задействована прямая метафора. В таких сложных комплексах некоторые лексические единицы наделяются вторичной метафоричностью ([SCNDRY\_MET], secondary metaphoricity) за счет их включения в качестве сторонних элементов в схему метафорического переноса, где они участвуют в развертывании метафоры. Например, в следующем предложении глагол «полететь» задействован в развертывании метафоры, заданной схемой переноса СПИЦЫ ВЕЛОСИПЕДНОГО КОЛЕСА ЭТО СТРЕЛЫ.

- (17) [... я чувствовал щекоцущий упор верткой педали под ступней, зуд в ладонях, вспотевших на горячих резиновых ручьяках руля, миг - и деревья послушно замелькают мимо меня, и заскользят под колесом их тени,]



В следующем примере (18) присутствует схема переноса: ВРАЩЕНИЕ ЗЕМЛИ ВОКРУГ СОЛНЦА ЭТО МЕХАНИЗМ, где сфера-цель выражена целой клаузой. В эту схему переноса в качестве вторичного развертывающего элемента [SCNDRY\_MET] входят глаголы «заводить», и «остановиться». «Заводить», в свою очередь, является источником олицетворения типа [MPERS] (метафорическое олицетворение), целью которого является неодушевленный аргумент «природа».



Очевидно, что метафорические комплексы способны проявлять большое разнообразие структур, вследствие чего описание таких комплексов представляет собой обширную задачу, требующую сбора, анализа и разметки репрезентативного корпусного материала и отдельного развернутого исследования.

### Результаты анализа

Предложенная схема аннотирования направлена на выявление и описание градиентной природы метафорических употреблений. Однако, если в целях

сравнения воспользоваться общепринятым бинарным делением на метафорические и неметафорические употребления, то границу между этими двумя классами можно провести следующим образом: полностью неметафоричными можно считать употребления глаголов в базовом значении (БЗ) без каких-либо дополнительных факторов метафоричности. Употребления в небазовом значении (НБЗ), а также любые сочетания БЗ или НБЗ с дополнительными факторами можно считать метафорическими употреблениями, которые при этом различаются по степени насыщенности метафорическим материалом. С точки зрения такого бинарного разделения, 51,44% проанализированных в данном исследовании глагольных употреблений обладают метафоричностью. Это гораздо больше, чем в корпусе VUAMC, где только 18,68% глаголов были отнесены к классу метафорических употреблений (Steen et al., 2010b:194–208).

Можно предположить, что такое различие обусловлено как особенностями предложенной схемы аннотирования, так и характером выборки: она состоит из глаголов с невысокой встречаемостью в корпусе (19–25 вхождений) и небольшим количеством значений (в среднем около 5 значений на глагол), в то время как в корпусе VUAMC анализировались все глагольные вхождения, включая высокочастотные и высокополисемичные.

В Таблице 1 представлена совместная встречаемость некоторых факторов метафоричности с базовыми и небазовыми значениями (в % от всех метафорических употреблений).

**Таблица 1.** Совместная встречаемость факторов метафоричности с базовыми и небазовыми значениями (None — отсутствие дополнительных факторов метафоричности)

	NBM		
None	39,42	BM	Total
ABSTR_ARG	2,05	11,07	13,12
SEM_SHIFT	3,12	7,42	10,54
CPERS_SOURCE	5,59	2,05	7,64
DMET_SOURCE, SCNDRY_MET	1,83	4,19	6,02
CONSTR	5,7	0	5,7
MTNMY_ARG	2,15	3,12	5,27
LPERS_SOURCE	3,66	1,51	5,17
IDIOM	2,58	0	2,58
MPERS_SOURCE	0,76	1,62	2,38
NEW_MEANING	1,72	0,22	1,94
DMET_SOURCE	0,22	0,86	1,08
TERM	0,54	0,11	0,65

Как видно из таблицы, значительное число зафиксированных метафорических употреблений (39,42%) выражено небазовым значением без каких-либо дополнительных факторов (NBM). Наличие абстрактного аргумента (ABSTR\_ARG) и семантического сдвига аргумента (SEM\_SHIFT) способствуют

метафоричности в 13,12% и 10,54% случаев соответственно. Трансформация конструкции и метонимический сдвиг в аргументе встречаются в 5,7% и 5,27% метафорических употреблений соответственно.

Среди трех типов олицетворения наиболее частотным (7,64%) является олицетворение, образованное контекстуальной сменой класса аргумента с одушевленного на неодушевленный (CPERS), тогда как лексикализованное олицетворение (LPERS) встречается в 5,17% случаев. Наименее частотным из всех типов олицетворения является метафорическое олицетворение — 2,38%. Прямая метафора почти всегда реализуется в составе сложных комплексов (DMET\_S, SCNDRY\_MET), где на нее воздействует вторичная метафоричность окружающих метафорических употреблений — 6,02% (ср.: прямая метафора вне метафорических комплексов (DMET\_SOURCE) встречается только в 1,08% случаев).

## Заключение

В данной статье была предложена методика анализа и описания факторов метафоричности глагола, а также их синтагматического взаимодействия. Очевидно, что не все из предлагаемых факторов имеют одинаковый вес в формировании метафоричности; вклад различных типов факторов в метафоричность может быть исследован психолингвистическими методами. Для подтверждения предложенных положений и получения репрезентативных количественных данных в перспективе необходимо привлечение более значительного объема материала (в т. ч. представляющего другие части речи), а также проведение анализа с привлечением большего числа аннотаторов.

## References

1. *Badryzlova Y. G.* (2015), Identification of Russian verbal metaphor by methods of machine learning trained on a corpus annotated for verbal arguments [Identifikatsiya metaforicheskikh upotrebleniy russkikh glagolov metodami mashinnogo obucheniya na osnove korpusa s rametkoy glagolnykh argumentov], MA thesis, National Research University Higher School of Economics, Moscow.
2. *Badryzlova Y. G., Isaeva E. V., Kerimov R. D., Shekhtman N. G.* (2013), Using the linguistic metaphor identification procedure (MIPVU) on a Russian corpus: rules revised and extended [Pravila primeneniya protsedury lingvisticheskoy identifikatsii metaphory (MIPVU) v russkoyazychnom korpusе: lingvokognitivnyy opyt (utochneniya i dopolneniya)], Humanitarian Vector [Gumanitarnyy Vektor], Vol. 4(36), pp. 19–29.
3. *Badryzlova Y. G., Kerimov R. D., Shekhtman N. G.* (2013), The main principles and methods of the Russian-language corpus of conceptual metaphor: using BRAT annotation tool. [Osnovnyye metody i printsipy sozdaniya russkoyazychnogo korpusa kontseptualnoy metaphory: opyt razmetki s primeneniym instrumenta BRAT], The Perm University Herald. Russian and Foreign Philology [Vestnik Permskogo Universitita. Russkaya i Zarubezhnaya Filologiya], Vol. 3(23), pp. 82–92.

4. *Boguslavsky I., Iomdin L., Frolova T., Timoshenko S.* (2009), Development of a Russian Tagged Corpus with Lexical and Functional Annotation, *MONDILEX Third Open Workshop*, Bratislava, pp. 83–90.
5. *Dodge E., Hong J., Stickles E.* (2015), MetaNet: Deep semantic automatic metaphor analysis, *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Denver, pp. 40–49.
6. *Dunn J.* (2013a), Evaluating the premises and results of four metaphor identification systems, *The 2013 International Conference on Intelligent Text Processing and Computational Linguistics*, Samos, pp. 471–486.
7. *Dunn J.* (2013b), How linguistic structure influences and helps to predict metaphorical meaning, *Cognitive Linguistics*, Vol. 24(1), pp. 33–66.
8. *Dunn J.* (2013c), What metaphor identification systems can tell us about metaphor-in-language, *The First Workshop on Metaphor in NLP*, Atlanta, pp. 1–10.
9. *Kustova G. I.* (2004), Types of Derivative Meanings and Mechanisms of Language Extension [Tipy proizvodnykh znacheniy i mekhanizmy yazykovogo rasshireniya], *Languages of the Slavonic Culture [Yazyki Slavyanskoy Kultury]*, Moscow.
10. *Lakoff G., Johnson M.* (1980), *Metaphors We Live By*, The University of Chicago Press, Chicago—London.
11. *Li L., Sporleder C.* (2009), Classifier combination for contextual idiom detection without labelled data, *The 2009 Conference on Empirical Methods in Natural Language Processing*, Vol. 1, Singapore, pp. 315–323.
12. *Li L., Sporleder C.* (2010), Using gaussian mixture models to detect figurative language in context, *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Los Angeles, pp. 297–300.
13. *Niles I., Pease A.* (2001), Towards a standard upper ontology, *The 2nd international conference on Formal Ontology in Information Systems*, Ogunquit, pp. 2–9.
14. *Nirenburg S., Raskin V.* (2004), *Ontological Semantics*, MIT Press, Cambridge, MA.
15. *Ozhegov S. I., Shvedova N. Yu., editors* (1997–1999), *Explanatory Dictionary of the Russian Language [Tolkovyy Slovar Russkogo Yazyka]*, electronic resource: <http://slovari.ru/default.aspx?s=0&p=244>
16. *Paducheva E. V.* (2004), Dynamic Models in the Semantics of Vocabulary [Dinamicheskie modeli v semantike leksiki], *Languages of the Slavonic Culture [Yazyki Slavyanskoy Kultury]*, Moscow.
17. *Panicheva P., Badryzlova, Y.* (2017), Distributional semantic features in Russian verbal metaphor identification, *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2017” [Komp’yuternaya Lingvistika i Intellektual’nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2017”]*, Moscow, In press.
18. *Russian Wiktionary* (2017), electronic resource: <https://ru.wiktionary.org/wiki>
19. *Shutova E., Sun L., Korhonen A.* (2010), Metaphor identification using verb and noun clustering, *Proceedings of the 23rd International Conference on Computational Linguistics*, Beijing, pp. 1002–1010.

20. Steen G. J., Dorst A. G., Herrmann J. B., Kaal A., Krennmayr T. (2010a), VU Amsterdam Metaphor Corpus, available at: <http://ota.ahds.ac.uk/headers/2541.xml>
21. Steen G. J., Dorst A. G., Herrmann J. B., Kaal A., Krennmayr T., Pasma T. (2010b), A Method for Linguistic Metaphor Identification: From MIP to MIPVU, John Benjamins Publishing Company, Amsterdam—Philadelphia, PA.
22. Stenetorp P., Pyysalo S., Topić G., Ohta T., Ananiadou S., Tsujii J. (2012), Brat: a web-based tool for NLP-assisted text annotation, Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics, Avignon, pp. 102–107.
23. Turney P. D., Neuman Y., Assaf D., Cohen Y. (2011), Literal and metaphorical sense identification through concrete and abstract context, The 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, pp. 680–690.
24. Veale T., Shutova E., Klebanov, B. B. (2016), Metaphor: a computational perspective, Synthesis Lectures on Human Language Technologies, Vol. 9.1, pp. 1–160.
25. Yevgenyeva A. P., editor (1981–1984), Dictionary of the Russian Language, 2nd ed., Vol. 1–4, Russian Language [Russkiy Yazyk], Moscow.

# ИДИОМАТИКА ПЬЯНСТВА В КОРПУСНОМ ИЗМЕРЕНИИ<sup>1</sup>

**Баранов А. Н.** (baranov\_anatoly@hotmail.com),  
**Вознесенская М. М.** (voznesh-masha@yandex.com)

Институт русского языка им. В. В. Виноградова РАН,  
Москва, Россия

## RUSSIAN IDIOMS OF DRUNKENNESS: CORPUS-BASED ANALYSIS

**Baranov A. N.** (baranov\_anatoly@hotmail.com),  
**Voznesenskaja M. M.** (voznesh-masha@yandex.com)

Institute of Russian Language (Vinogradov's Institute), Russian  
Academy of Sciences, Moscow, Russia

The paper provides a corpus-based analysis of Russian idioms of drunkenness from the Thesaurus of Modern Russian Idioms (Moscow, 2007). As a starting point of a study three text corpora differentiated in accordance with different spheres of functioning of Russian were selected: Corpus of Russian Literature (texts from 1960 to 2000 years, 35 millions of type-tokens), Corpus of Mass-Media (texts from 1990 to 2010 years, 29 millions of type-tokens) and Corpus of Drama (plays from 1960 to 2010 years, 23 millions of type-tokens).

The main idea consists in comparing frequency of idioms of drunkenness under consideration in every corpus. The analysis shows that Russian idioms of drunkenness form four different groups in accordance their frequency characteristics. The first group is formed by low-frequency idioms from all corpora. The second group consists of idioms with the same relative high-frequency in all corpora. The third group includes idioms which demonstrate a high variations in frequencies in all corpora. The fourth group consists of idioms with a high frequency rank within every single corpus. The paper describes each of these groups providing the interpretation of the statistic data from the field of stylistics. The high frequency range of an idiom in different corpora corresponds with its stylistic features. The paper argues that there is a strict correlation between the discourse frequency of an idiom and its stylistic properties.

**Key words:** phraseology, text corpora, frequency analysis, idioms of drunkenness

---

<sup>1</sup> Работа выполнена при поддержке РФФИ, грант 17-04-00420

## Введение

Таксон пьянство-трезвость в «Тезаурусе современной идиоматики» [Тезаурус 2007] представлен достаточно большим количеством идиом — около 140 единиц. Стилистически он довольно однороден: в основном это разговорная (*принять на грудь, в доску напиться/пьян, под мухой, закладывать/заливать за воротник; что-то стало холодать* [– не пора ли нам поддать?]) и жаргонная (*залить баки, трубы горят, в [полном] отрубе, с бодуна*) идиоматика. Книжных и нейтральных идиом сравнительно мало (ср. *белая горячка, сухой закон, зелёный змей, до положения риз, огненная вода*), хотя некоторые из них занимают довольно высокий частотный ранг (см. ниже). Интересно, что данное семантическое поле не содержит обценных идиом<sup>2</sup>.

Разумеется, идиоматика трезвости существенно уступает в количественном отношении идиомам с семантикой пьянства: шесть идиом, связанных с трезвостью, против более ста, связанных с пьянством. Причина этого очевидна и прослеживается по другим семантическим полям как идиом, так и обычной лексики: отрицательное, негативное оказывается более важным в практической жизни и в силу этого язык вынужден обеспечивать для сферы отрицательного более богатый лексический репертуар.

Рассматриваемое понятийное поле с очевидностью значимо для русского языкового сознания. Оно выделяется во всех значительных лексикографических описаниях. Так, входящее в Русский идеографический словарь [РИС 2011] значительное по объёму поле хмель включает более 150 лексем и значительное количество иллюстративного материала. Собрание «Пословиц русского народа» содержит таксон пьянство с более чем 450 пословицами данной тематики [Даль 1879, с. 376–391]. Ю. С. Степанов среди «констант» словаря русской культуры выделяет словарную статью «Водка и пьянство», полушутливо указывая на «связь русской водки с законами мироздания» [Степанов 2001, с. 291–313].

Между тем, как и в других похожих случаях (например, со сферой обценной лексики), репрезентативность существующих источников относительно идиоматики пьянства-трезвости различна. Представляется интересным и важным сравнить имеющиеся корпуса с точки зрения представленности в них идиом обсуждаемой семантики.

Цель такого сопоставления в том, чтобы определить, как реализуется словарный стандарт («норма»), представленный словником таксона пьянство-трезвость «Тезауруса современной идиоматики» [Тезаурус 2007], в реальном узусе — в той мере, в которой этот узус проявляется в имеющихся корпусных источниках.

## 1. Источники — корпуса текстов

Исходной точкой отсчета здесь служат четыре корпуса, используемые в проекте описания русской фразеологии и в частности в создании «Частотного

<sup>2</sup> Самые сниженные формы: *до усёру напиться/пьян...неприл.*; *в говно напиться/пьян...неприл.*; *в жопу напиться/пьян... неприл.*



словаря русской идиоматики»: Корпус публицистики (29 млн словоупотреблений), Корпус драмы (около 23 млн словоупотреблений), Корпус детективов (около 16 млн словоупотреблений) и Корпус художественных текстов (35 млн словоупотреблений)<sup>3</sup>. Корпус детективов и Корпус драмы представляют в проекте модели разговорной речи. Тем самым, в рассматриваемом случае можно ограничиться одним из них. Таким образом, для настоящего исследования были выбраны Корпус художественных текстов, Корпус публицистики и Корпус драмы. Указанные корпуса взаимно дополняют друг друга, представляя основные сферы функционирования русского языка.

Полнота (репрезентативность) корпуса — сложная интегральная характеристика, достижение которой пока остается во многом уделом искусства составителя корпуса. Она зависит не только и не столько от объема корпуса, сколько от многих других факторов, которые можно разделить на две группы: стратификационные характеристики корпуса (множество параметров, релевантных для соответствующего дискурса и учтенных составителями) и сам тип речевых форм, которые репрезентируются корпусом. Иными словами, корпуса являются представительными относительно тех или иных явлений — речевых форм, существующих в универсуме дискурса. Для простоты будем исходить из того, что описанные выше корпуса репрезентативны относительно большинства идиом современной русской фразеологии — по крайней мере для среднечастотных идиом, какая бы конкретная величина под этим ни понималась<sup>4</sup>.

## 2. Отсутствующие и низкочастотные идиомы по всем корпусам

Анализ частоты употребления позволяет выделить следующие идиомы, отсутствующие во всех указанных корпусах: *Александр Третий*<sub>совет. жарг.;</sub> *банку держать*; *в люлю*<sub>жарг.;</sub> *в слюни напиться/пьян...*<sub>снижен.;</sub> *взять вторую производную*<sub>жарг.;</sub> *звать/кормить/пугать ихтиандра*<sub>жарг. эвф.;</sub> *как извозчик напиться/пьян...*; *Медведь пришел — медведь ушел*<sub>жарг.;</sub> *по приборам*; *пугать унитаза*<sub>эвф.;</sub> *Северный медведь*<sub>жарг.;</sub> *сухой как лист*; *сушняк долбит*<sub>жарг.;</sub> *трубы залить/прочистить*. Легко видеть, что половина из этих идиом имеют в «Тезаурусе современной идиоматики» [Тезаурус 2007] помету *жарг.* или *снижен.* — 7 из 14. Идиомы *банку держать*, *пугать унитаза*, *трубы залить/прочистить* и *в слюни напиться/пьян...* также близки к жаргонным, хотя и не получили соответствующую помету.

Отсутствие жаргонных идиом вполне понятно, поскольку жаргоны разнообразны и непостоянны по лексическому составу. Даже очень полный корпус текстов не дает гарантии вхождения (тем более репрезентативной представительности) жаргонных форм. Однако идиома *как извозчик напиться/пьян...* относится к разговорным. Впрочем, ее устаревший характер объясняет почти

<sup>3</sup> Проект реализуется в Институте русского языка РАН на основе фразеологически ориентированных корпусов и Базы данных по современной русской идиоматики. Важнейшие результаты проекта: [Тезаурус 2007; АСРФ 2015].

<sup>4</sup> Оценка частотности в рассматриваемом случае носит эвристический характер и привязана не к частоте использования идиоматики вообще, а к данному семантическому полю.

полное отсутствие употреблений в современных произведениях. В НКРЯ она встретилась три раза, причем во всех случаях либо как стилизация (ср. жанр фэнтези: *Ты пьешь, как извозчик и портишь людям праздник*. [С. Осипов. Страсти по Фоме]; *И хоть говаривали в Москве, ежели что, «пьян, как извозчик», пили они на работе мало*. Гиляровский уверял, что мог извозчик за обедом рюмку водочки пропустить, но не более. [Л. Репин. В Столешниковом Гиляровский кочергу в узел завязал]), либо как пример из соответствующей эпохи (*Он зашел в трактир и впервые в жизни, под оркестрион и в одиночку, напился как извозчик*. [Л. Леонов. Скутаревский, 1930–1932]). Интересно, что внутренняя форма здесь оказывается настолько существенной, что в явном виде влияет на употребление (хотя по Ф. де Соссюру это не должно было бы играть никакой роли, поскольку знак произволен). Отметим, что в связи с «извозчиком» в НКРЯ существенно более частотна идиома *ругаться / материться как извозчик*.

Отсутствует и идиома *по приборам*, что неудивительно, поскольку в рассматриваемом таксоне Тезауруса она попала в зону периферии, то есть ее присутствие в данном таксоне объясняется исключительно контекстами неуверенного передвижения вследствие опьянения (*Вчера гульнули не по-детски, домой возвращался по приборам*). Идиома *сухой как лист* близка к авторской. В корпусах она отсутствует (в НКРЯ ее тоже нет), а в Интернете представлена в известной миниатюре А. Петренко «Сухой как лист», относящейся к предпресторочным временам.

Идиомы данной группы, как правило, не находятся и в НКРЯ, что может рассматриваться как дополнительное подтверждение валидности полученных результатов. Исключение составляют формы *пугать ихтиандра* и *в слюни*, обнаруживаемые в единичных примерах в подкорпусе разговорной речи НКРЯ (*Нажрался / как всегда / потом всю ночь в туалете ихтиандра звал; Представь / что ты опять нахреначишься в слюни и ты точно не будешь знать как другие на это отреагируют*).

Значительное количество идиом одинаково низкочастотны по всем исследованным корпусам. Таковы, например, *под парами* (1-0,03/1-0,03/0-0,00<sup>5</sup>); *до поросычеге визга напиться/пьян...* (4-0,11/2-0,07/2-0,09); *в дупель/дупелину напиться/пьян...* снижен. (3-0,09/3-0,10/6-0,26); *третьим будешь* (5-0,14/2-0,07/5-0,22); *что-то ветер дует в спину* [– не пора ли к магазину?] (0-0,00/0-0,00/1-0,04); *спиться с круга* (4-0,11/1-0,03/3-0,13); *Северное Сияние жарг.* (2-0,06/1-0,03/1-0,04); *море разливанное* (2-0,06/2-0,07/0-0,00); *на автопилоте*<sup>6</sup> (2-0,06/1-0,03/3-0,13); *в лок напиться/пьян...* (2-0,06/2-0,07/1-0,04). Никаких особенностей у данной группы идиом не наблюдается. По большей части это разговорные идиомы, неко-

<sup>5</sup> Здесь и далее последовательно указываются абсолютные и относительные частоты по указанным корпусам в следующей последовательности: Корпус художественных текстов, Корпус публицистики и Корпус драмы. Относительная частота дается на 1 млн словоупотреблений.

<sup>6</sup> ≈ 'Будучи не в состоянии осмыслить свои действия (часто в сильном опьянении, душевном волнении, сильно устав и т. д.), делать что-л., так как это принято в соответствующей ситуации (говорить, есть, играть и т. п.), неосознанно используя интуицию или опыт осуществления такой деятельности' [АСРФ 2015].

торые из которых являются речевыми формулами (*третьим будешь; что-то ветер дует в спину* [– не пора ли к магазину?]).

Всего идиом указанного частотного типа в таксоне пьянства 45 единиц.

### 3. Идиомы с относительно высокой частотой по всем корпусам

Существенно меньше идиом, имеющих относительно высокую частоту по всем корпусам (выше 10 по абсолютной частоте). К ним относятся всего три формы: *по пьяному делу* (24-0,69/22-0,76/20-0,87), *прикладываться к бутылке / рюмке* (12-0,34/12-0,41/10-0,43), а также книжная идиома *огненная вода* (16-0,46/15-0,52/16-0,70). Понятно, что таких «универсальных» единиц и не должно быть много, поскольку специфика дискурса проявляется в использовании фиксированных дискурсивных практик, различающихся от дискурса к дискурсу. Фактически дискурс определяется уникальной совокупностью дискурсивных практик [Foucault 1971; Баранов 2004], а использование фразеологизмов является типичной дискурсивной практикой. Соответственно, частотное распределение идиом по разным дискурсам — это норма. Исключения, как мы видим, обнаруживаются, но их немного.

Интересна в этом смысле книжная по сути идиома *огненная вода*, которая удивительным образом наиболее частотна в Корпусе драмы, отражающем разговорную речь. Не исключено, что здесь играет роль фактор «литературщины»: чрезмерное уснащение речи персонажей стилистически неуместными, но привлекающими внимание формами. Легко представить себе в речи образованного городского жителя полуигровую фразу: *Не забыть затариться огненной водой!*, однако вопрос в магазине: *Огненную воду продаете?* вряд ли будет правильно понят.

В НКРЯ в основном корпусе эта идиома обнаруживает порядка 60 вхождений, а в подкорпусе публицистики — порядка 70-ти. По относительной частоте соответственно: 0,2 и 0,3.

### 4. Идиомы, обнаруживающие значительный разброс по всем корпусам

Идиом, обнаруживающим существенные расхождения в частоте употребления по корпусам, наибольшее количество — порядка шестидесяти единиц. Их распределение в явном виде связано со стилистическими характеристиками, то есть с речевыми практиками соответствующих дискурсов. Так, идиомы, относящиеся к книжным, чаще встречаются в Корпусе художественных текстов и Корпусе публицистики, а для Корпуса драмы, моделирующего устный дискурс, они нетипичны: *зелёный змий*<sub>книжн.</sub>; *напиться/упиться... до зелёного змия*<sub>книжн.</sub> (13-0,37/29-1,00/2-0,09); *до положения риз*<sub>книжн.</sub> (8-0,23/0-0,00/0-0,00). Нейтральная по стилистике идиома *сухой закон* также тяготеет к художественной литературе и публицистике: 39-1,11/64-2,21/13-0,57.

Впрочем, аналогичная по стилистике идиома *белая горячка* наиболее частотна в Корпусе драмы: 58-1,66/32-1,10/83-3,61.

Жаргонизмы и просторечия более характерны для обыденной коммуникации, соответственно они также чаще встречаются в Корпусе драмы: в [полном] *отрубe*<sub>жарг.</sub>; в [полных] *отрубях/отрубax*<sub>жарг.</sub> (1-0,03/0-0,00/6-0,26); с *бодуна*<sub>жарг.</sub> (19-0,54/13-0,45/35-1,52); на *посошок*<sub>прост.</sub> (34-0,97/10-0,34/59-2,57); для *сугрева*<sub>прост.</sub> (5-0,14/0-0,00/10-0,43); *хватить/взять/перебрать лишку/лишнего*<sub>прост.</sub> (4-0,11/8-0,28/19-0,83); *глаза/зенки/шары/фары налить/залить*<sub>жарг.</sub> (11-0,31/0-0,00/12-0,52). Близки к жаргонным формы в *завязке/завязе* (8-0,23/3-0,10/20-0,87); *трубы горят* (2-0,06/2-0,07/14-0,61); *закладывать/заливать за воротник* (7-0,20/8-0,28/17-0,74).

В некоторых случаях в использовании просторечной и жаргонной идиоматики данного типа Корпус художественных текстов, в котором особенности устной речи также находят отражение, вполне конкурирует с Корпусом драмы: *лыка/лыко не вязать*<sub>прост.</sub> (20-0,57/0-0,00/13-0,57); *со свиданьем*<sub>прост.</sub><sup>7</sup> (22-0,63/1-0,03/13-0,57).

Основная часть группы идиом с существенным частотным разбросом по корпусам относится к разговорной идиоматике и, тем самым, широко представлена в Корпусе драмы или одновременно и в Корпусе драмы, и в Корпусе художественных текстов: *пить горькую* (21-0,60/9-0,31/20-0,87); *за ваше/твое/наше здоровье* (183-5,23/35-1,21/222-9,65), *в дым, в дымину* (26-0,74/6-0,21/22-0,96)<sup>8</sup>, *не просыхать* (25-0,71/4-0,14/32-1,39); *в стельку напиться/пьян...* (24-0,69/15-0,52/39-1,70); *до чёртиков напиться/допиться...*; *до чертей напиться/допиться...* (15-0,43/5-0,17/17-0,74); *по пьяни, по пьянке* (138-3,94/47-1,62/129-5,61); *по маленькой* (27-0,77/1-0,03/63-2,74); *на брудершафт* (51-1,46/4-0,14/106-4,61); *на дорожку* (20-0,57/0-0,00/37-1,61); *трезвый...* как *стёклышко/стекло* (15-0,43/1-0,03/26-1,13).

Стилистическая распределенность идиом по корпусам текстов выдерживается достаточно последовательно, что позволяет с полной определенностью говорить о стилистических характеристиках идиоматики как дискурсивном факторе. Тем самым, связь стилистических свойств идиом с конкретным типом дискурса — это закономерное явление, а не случайность.

## 5. Идиомы с высоким рангом в каждом из корпусов

Сопоставление идиом с высоким рангом (идиом, занимающих первые десять позиций по убыванию частоты) в каждом из корпусов оказывается весьма показательным в том смысле, что отражает степень близости отраженных в них дискурсов.

<sup>7</sup> ≈ 'Ритуализованное речевое сопровождение употребления алкогольных напитков в форме поздравления участников с фактом встречи, которое обычно начинается первый акт застолья' [АСРФ 2015].

<sup>8</sup> Здесь и далее, если идиома многозначна, то частота дается только для значений, относящихся к обсуждаемому семантическому полю.

**Таблица 1.** Первые десять идиом Корпуса художественных текстов, упорядоченных по убыванию частоты

Идиома	Абсолютная частота	Относительная частота
за ваше/твое/наше здоровье	183	5,23
по пьяни, по пьянке	138	3,94
белая горячка <i>нейтр.</i>	58	1,66
на брудершафт	51	1,46
сухой закон <i>нейтр.</i>	39	1,11
под хмельком	38	1,09
на посошок <i>прост.</i>	34	0,97
в подпитии	33	0,94
по маленькой	27	0,77
в дым, в дымину	26	0,74

**Таблица 2.** Первые десять идиом Корпуса публицистики, упорядоченных по убыванию частоты

Идиома	Абсолютная частота	Относительная частота
сухой закон <i>нейтр.</i>	64	2,21
по пьяни, по пьянке	47	1,62
за ваше/твое/наше здоровье	35	1,21
белая горячка <i>нейтр.</i>	32	1,10
зелёный змий <i>книжн.</i> ; напитокся/ упиться... до зелёного змия <i>книжн.</i>	29	1,00
по пьяному делу	22	0,76
принять на грудь	16	0,55
в подпитии	16	0,55
в стельку напитокся/пьян...	15	0,52
огненная вода <i>книжн.</i>	15	0,52

**Таблица 3.** Первые десять идиом Корпуса драмы, упорядоченных по убыванию частоты

Идиома	Абсолютная частота	Относительная частота
за ваше/твое/наше здоровье	222	9,65
по пьяни, по пьянке	129	5,61
на брудершафт	106	4,61
белая горячка <i>нейтр.</i>	83	3,61
по маленькой	63	2,74

Идиома	Абсолютная частота	Относительная частота
на посошок <i>прост.</i>	59	2,57
в стельку напиток/пьян...	39	1,70
за упокой [души]	38	1,65
на дорожку	37	1,61
с бодуна <i>жарг.</i>	35	1,52

При сравнении таблиц обращает на себя внимание, во-первых, сходство частотных рангов идиом и, во-вторых, сам состав (набор) идиом с высоким рангом. Легко видеть, что первые две идиомы в списке самых частотных для Корпуса художественных текстов и Корпуса драма совпадают: *за ваше/твое/наше здоровье; по пьяни, по пьянке*. Идиома *белая горячка* оказывается третьей по частоте в Корпусе художественных текстов и четвертой в Корпусе драмы. Идиома *на брудершафт* оказывается четвертой в Корпусе художественных текстов и третьей в Корпусе драмы. Близки по рангу в этих корпусах и идиомы *по маленькой* и *на на посошок*.

Сходство частотных рангов идиом между Корпусом художественных текстов и Корпусом публицистики существенно меньше. Вторые позиции в них занимает идиома *по пьяни, по пьянке*, а первые позиции различаются (*за ваше/твое/наше здоровье* и *сухой закон* соответственно). Впрочем, ранги идиом *белая горячка* и *в подпитии* сходны.

Частотные ранги наиболее употребительных идиом Корпуса публицистики и Корпуса драмы различаются еще больше. Хотя идиома *по пьяни, по пьянке* устойчиво занимает вторую позицию (для всех корпусов!), а идиома *белая горячка* — четвертую, ранги остальных форм различны.

Таким образом, по сходству в ранговом распределении наиболее частотных идиом рассматриваемые корпуса упорядочиваются следующим образом: Корпус художественных текстов — Корпус драма > Корпус художественных текстов — Корпус публицистики > Корпус публицистики — Корпус драма.

Так же закономерность прослеживается и по параметру состава идиом. По десяти наиболее частотным формам у Корпуса художественных текстов и Корпуса драма оказывается общими 6 идиом (*за ваше/твое/наше здоровье; по пьяни, по пьянке; белая горячка; на брудершафт; на посошок; по маленькой*), у Корпуса художественных текстов и Корпуса публицистики — 5 (*за ваше/твое/наше здоровье; по пьяни, по пьянке; белая горячка; сухой закон; в подпитии*), а у Корпуса публицистики и Корпуса драма — 4 (*за ваше/твое/наше здоровье; по пьяни, по пьянке; белая горячка; в стельку напиток/пьян*). Легко видеть, что и здесь иерархия степени сходства корпусов поддержана.

Уникальными для Корпуса драма оказываются в первую очередь идиомы-поговорки, сопровождающие типовые диалогические ситуации (*за упокой [души], на дорожку*), а также жаргонная идиома *с бодуна*. Для Корпуса публицистики специфичны книжные идиомы *зелёный змий; напиток/упитие... до зелёного змия* и *огненная вода*, стилистически напоминающие типичные

журнализмы *блюститель порядка* и *каreta скорой помощи*, а также почти эфемизмы *по пьяному делу* и *принять на грудь*. В Корпусе художественных текстов уникальны идиомы *в дым*, *в дымину* и *под хмельком*.

## Заключение

Соотношение между дискурсами и корпусами текстов далеко от взаимно-однозначного соответствия. Разобранный пример показывает, что даже специально подготовленные для одной цели корпусы текстов представляют не гомогенные дискурсы, а их совокупности, в которых один дискурс проникает в другой. Так Корпус художественных текстов с очевидностью включает как нарративы, так и значительные фрагменты диалогических текстов, с разной степенью точности отражающие реальное диалогическое общение. С другой стороны, Корпус публицистики не свободен от нарративных элементов. А Корпус драмы из-за специфической стилизации речи персонажей может включать элементы нарратива. В результате речевые формы — как по составу, так и по частоте употребления — оказываются распределенными между корпусами по принципам фамильного сходства.

Несмотря на столь сложную организацию корпусов гипотезу о связи стилистических характеристик идиом с типами дискурсов, представленных в корпусах текстов [Баранов, Вознесенская, Добровольский, Киселева, Козеренко 2013], можно считать доказанной. На это указывает частотное распределение почти половины идиом, представленных в семантическом поле пьянство-трезвость: жаргонные и разговорные идиомы чаще всего встречаются в Корпусе драмы, затем в Корпусе художественных текстов. Для Корпуса публицистики типичны несколько иные идиомы, которые, кстати говоря, в рассмотренном семантическом поле почти не представлены.

## Литература

1. *АСРФ 2015* — Академический словарь русской фразеологии / Под ред. А. Н. Баранова и Д. О. Добровольского. Изд. 2. М., 2015.
2. Баранов А. Н. Метафорические модели как дискурсивные практики // Известия РАН. Сер. литературы и языка, 2004, том 63, № 1.
3. Баранов А. Н., Вознесенская М. М., Добровольский Д. О., Киселева К. Л., Козеренко А. Д. Квантитативные характеристики идиомы как показатель ее стилистических свойств // Труды международной конференции «КОРПУСНАЯ ЛИНГВИСТИКА — 2013» СПб.: С.-Петербургский гос. университет, 2013, с. 309–317.
4. Даль В. Пословицы русского народа. Изд. 2, без перемен. Т. II. СПб., М., 1879.
5. *РИС 2011* — Русский идеографический словарь: Мир человека и человек в окружающем его мире (80 концептов, относящихся к духовной, ментальной и материальной сферам жизни человека) / Отв. ред. акад. РАН Н. Ю. Шведова; РАН, Ин-т рус. яз. им. В. В. Виноградова. М., 2011.
6. Степанов Ю. С. Константы: Словарь русской культуры: Изд. 2-е, испр. и доп. М.: Академический проект, 2001

7. *Тезаурус 2007* — Словарь-тезаурус современной русской идиоматики / Под ред. А. Н. Баранова и Д. О. Добровольского. М., 2007.
8. *Foucault M. L'ordre du discours*. Paris, 1971.

## References

1. *ASRF* (2015) — Academic Dictionary of Russian Phraseology [Akademicheskij slovar' russkoj frazeologii], Baranov A. N., Dobrovol'skij D. O. (eds.), 2nd ed., Moscow.
2. *Baranov A. N.* (2004), Metaphoric models as discursive practices [Metaforicheskie modeli kak diskursivnye praktiki], Russian Academy of Sciences Bulletin. Literature and Language [Izvestija AN. Ser. literatury i jazyka], Vol. 63, № 1.
3. *Baranov A. N., Voznesenskaja M. M., Dobrovol'skij D. O., Kiseleva K. L., Kozenko A. D.* (2013), Correlation between idiom frequency and its stylistic properties [Kvantitativnye harakteristiki idiomy kak pokazatel' ee stilisticheskikh svojstv], Proceedings of the International Conference "Corpus linguistics-2015" [Trudy mezhdunarodnoj konferencii «Korpusnaja lingvistika — 2013»], St. Petersburg University, St. Petersburg, pp. 309–317.
4. *Dal' V.* (1879), Russian Proverbs [Poslovice russkogo naroda], 2nd ed., St. Petersburg, Moscow.
5. *Foucault M.* (1971), The Order of Discourse [L'ordre du discours], Paris.
6. *RIS* (2011) — Russian Ideographic Dictionary [Russkij ideograficheskij slovar'], Shvedova N. Ju. (ed.), Moscow.
7. *Stepanov Ju. S.* (2001), Constants: Dictionary of Russian Culture [Konstanty: Slovar' russkoj kul'tury] 2nd ed., Akademicheskij proekt, Moscow.
8. *Tezaurus* (2007) — Thesaurus of modern Russian idioms [Slovar'-tezaurus sovremennoj russkoj idiomatiki], Baranov A. N., Dobrovol'skij D. O. (eds.), Moscow.



# INTERACTION AND EMPATHY AS ELEMENTS OF NARRATIVE STRATEGIES IN THE RUSSIAN CLIPS CORPUS<sup>1</sup>

**Bergelson M. B.** (mirabergelson@gmail.com),  
**Khudyakova M. V.**

National research University Higher School of Economics,  
Moscow, Russia

This paper focuses on evaluation of discourse abilities of speakers with brain damage: people with dynamic aphasia (PWA(d)) and right hemisphere damage (RHD) as compared to healthy speakers of Russian language. The study is based on the material from the Russian CliPS corpus that contains retellings of the Pear Film produced by PWA and RHD, as well as neurologically healthy controls.

The nature of the narratives in the corpus allows for a comparative investigation of discourse on the level of micro-structure (grammatical and lexical phenomena) and on the macro level: narrative structure, coherence and cohesion, interactional patterns and narrative discourse strategies. In this paper we present results of the comparative analysis of some macro level discourse strategies: the way interaction and empathy are realized in the stories by PWA(d), RHD and healthy speakers.

We have found significantly higher numbers of attitude expression markers, as well as significantly lower numbers of cognitive difficulties markers, in healthy speakers as compared to PWA(d). These results support what is known about difficulties that PWA(d) demonstrate in discourse production tasks. While PWA(d) use interactive markers to get a break from keeping with the story plan, they avoid using epistemic predicates whose subjects are the story characters.

We also present qualitative analysis of the discourse strategies of healthy speakers.

**Key words:** corpus linguistics, narratives, aphasia, interaction markers, empathy, discourse strategies

---

<sup>1</sup> The paper was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2016–2017 (grant №16-05-0024) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program

# ИНТЕРАКЦИЯ И ЭМПАТИЯ КАК СОСТАВЛЯЮЩИЕ НАРРАТИВНЫХ СТРАТЕГИЙ НА МАТЕРИАЛЕ КОРПУСА RUSSIAN CLIPS

**Бергельсон М. Б.** (mirabergelson@gmail.com),  
**Худякова М. В.**

НИУ Высшая школа экономики, Москва, Россия

## 1. Introduction

The clinical discourse studies are aimed at studying language use by persons with acquired language disorders due to focal brain damage or persons with psychiatric and neurodegenerative diseases. The 'discourse approach' in clinical linguistics can be attributed to the change of perspective from assessing separate linguistic skills in various disorders to the idea that communication skills should be assessed as a whole (see discussion in Linnik et al., 2015). Also, it has been discovered that difficulties at the word and sentence level do not exclude a sensible story and a communicative success (Olness et al., 2010). The reverse may be true as well: spared grammar and lexis not necessarily result in a coherent story (Sherratt & Bryan, 2012). On the other hand, investigation of language in clinical populations provides a chance to test models of discourse comprehension and production and investigate the correlation between linguistic skills in the domain of general linguistics.

Our project is focused on evaluation of discourse abilities of speakers with brain damage: people with dynamic aphasia syndrome (PWA(d)) and right hemisphere damage (RHD) as compared to healthy speakers of Russian language. The study is based on the material from the Russian CliPS corpus that contains retellings of the Pear Film produced by people with different types of aphasia, RHD, and neurologically healthy speakers.

In this paper we present results of the comparative analysis of macro level discourse strategies. Our goal is to look at the way *interaction* and *empathy* are realized in the stories by PWA(d), RHD and healthy speakers in order to decide whether these linguistic concepts can be used as parameters for discriminating between the pathological and healthy discourse. We call the units that signal interaction and empathy *interaction markers* (IMs), and *empathy markers* (EMs), respectively.

The two clinical groups (PWA(d) and RHD) were chosen for this kind of analysis because the nature of their specific language deficit is known to affect the narration (see 3.2).

Our hypothesis predicts that certain IMs allow compensating for the linguistic deficits of the clinical populations, while some other IMs are only cognitively accessible to the healthy speakers. To prove or falsify this hypothesis we perform intergroup comparisons.

## 2. Interaction and empathy

The crucial difference between such narratives as report, chronicles, or traditional myths, and stories lies in the fact that the former don't 'interact' with the addressee. Stories, on the contrary, are not so much about providing facts, but about expressing opinions on and giving evaluations of the facts described using all discourse production types, or discourse passages (Plungyan, 2008; Polanyi, 1989).

We understand *linguistic interaction* in the specific situations of telling stories as an interpolation of the following concepts:

- narrative vs. discursive mode (see Paducheva, 2008: нарративный vs. речевой режим);
- fore- and backgrounding (Weinrich, 1964);
- the world of story vs. the world of narration (Norrick, 2000).

When telling a story the primary system of communication is the *world of story* and the collateral system corresponds to the *world of narration* (an instance of communication at certain time and place). Building on that, we look at the narrative communicative event from the perspective of the discourse strategies used by the narrator (the speaker) to *interact* with the recipient of the story (the addressee). Realization of these strategies does not interfere with the regular story components (Fleischman, 1991; Hoey, 1994; Labov, 1972; Labov & Waletzky, 1967; Longacre, 1983; Polanyi, 1989) as they belong to two different dimensions, or worlds, of narration and story, respectively.

Linguistic devices that earmark interaction (IMs) may be found inside all types of the discourse passages (Kibrik, 2009; Longacre, 1990)—descriptive, narrative, instructive, argumentative, and expository. Some of them mark the end or the beginning of the elementary discourse units (EDUs), utterances, or whole episodes. They are different in their specific functions, but all of them serve to signal switching from the world of story to the world of narration.

- (1) *мужчина интересного вида да конец лета надо полагать собирает урожай залез на видимо грушу по приставной лестнице*  
[a funny man, **yes**, it is the end of the summer, **I believe**, (he) harvests the fruit, is climbing the, **seemingly**, pear tree using the ladder]

Interaction allows the narrator to have a break from the story world with its demands for informational consistency: the necessity to balance narration and description, logical and chronological sequence of events, complications, and denouement.

The world of narration is comprised as two separate communication grounds. One is the real situation of storytelling where IMs are directed onto the addressee as part of the conversation. The other is a mental space where the narrators communicate with themselves while planning next portions of discourse. IMs then, are those linguistic elements that reveal narrators' cognitive difficulties with planning, retrieval, or physical producing discourse. For the purposes of analysis that involves clinical discourse, this latter 'communication ground' and IMs serving it are especially prominent.

- (2) одна **нет** ещё пустая почти **значит** во вторую **значит** в одну он кладёт та груша груша которую упала  
[one, **no**, (it is) almost empty, **well**, in the second, **well**, he puts that pear, pears in that one, the one that fell down]

On the other hand, *empathy* in the storytelling stands for the ability of the narrators, so to say, immerse themselves in the world of the story, to communicate what is going there from the character's viewpoint. Such understanding builds upon the concept of empathy in syntax and semantics (Kuno, 1987; Kuno & Kaburaki, 1977), though is closer to the stylistic phenomena in the paradigm of interactional linguistics (Kalliokoski & Verschueren, 1991)—see below our definition in 3.3.1. If the narrators tell us about feelings and decisions of the characters (*the boy understood that..., he decided to..., he did not like...*) it would mean they are putting themselves in the characters' shoes, taking, as it were, their roles in the story world, inventing the ways another person could think or feel.

### 3. Method and Material

#### 3.1. The Russian CliPS corpus

The Russian CliPS corpus contains retellings of the “Pear story” film (Chafe, 1980) by people with different aphasia types (efferent, dynamic, acoustic-mnemonic, and sensory), RHD and healthy speakers. The stories were audio-recorded and transcribed in ELAN (Wittenburg et al., 2006) on various linguistic levels, including lexical and grammatical information, segmentation into elementary discourse units, coherence and coreference annotation, as well as error classification (see Khudyakova et al., 2016; Toldova et al., 2016 for details).

#### 3.2. Subcorpus for the study

Texts by three groups of speakers from the corpus were selected for the study of interaction and empathy components: PWA(d), RHD and healthy speakers. These two groups are united by difficulties they experience on the macro level of discourse.

Dynamic aphasia is a language impairment, resulting from damage to the language-dominant hemisphere (usually left), characterized by non-fluent speech output. The primary deficit in dynamic aphasia is in the planning of the utterance stage, while motor components and understanding are intact (Akhutina, 2015; Luria, 1972; Luria & Hutton, 1977).

Damage to the right (language non-dominant) hemisphere does not result in aphasia, but people with RHD demonstrate speech disturbances on discourse level: lower coherence, inability to maintain the main theme of the narrative. Nevertheless, they do not have any deficits on microlinguistic level (Glosser & Deser, 1990; Marini et al., 2005).

For the study we have analyzed 8 narratives by PWA(d) (5 females; mean age—49,5; age range—41–58), 5 narratives by people with RHD (2 females; mean age—54,2; age range—41–74) and 10 narratives by healthy speakers (4 females; mean age—65; age range—42–84).

### 3.3. Annotation

#### 3.3.1. Empathy

For this specific communicative task, we define *empathy* as the narrator reporting of inner states (feelings, opinions and decisions) of the characters in the movie. The retellings of the movie are in no way personal stories, and the narrator has not experienced these feelings, nor made these decisions in his/her previous experience or even had positive information to guess about them. To report them one has to take a certain perspective, has to empathize with the subject of these inner states. Thus, in our annotation system we tag all predicates expressing inner states (we call them quasi-events) if their subject is a story character and the corresponding stance (discourse passage) belongs to the main line of the story.

- (3) и тут мальчик подумал (*Emp*), что груши можно унести  
[so the boy thought (*Emp*) that he might take the pears away]

#### 3.3.2. Interaction markers

The IMs are also annotated as a separate annotation layer. Everything that pertains to the world of storytelling where the speaker interacts with the listener—directly by appealing to him/her or by demonstrating to the listener his/her cognitive difficulties—is regarded as interaction. Fillers, feedback markers, appellations to the listener, but also: word search, false starts, repetitions.<sup>2</sup> Interaction markers can occur both within the clause (e.g., *let's call the boy Vovochka*, where *let's* is an element of interaction within a descriptive clause) or comprise a separate clause (e.g., *if you say so*).

Functionally, all the IMs fall into several groups.<sup>3</sup> Besides those that serve to reveal cognitive difficulties (see 3.3.2), there are pace regulating markers, attention drawing markers and expressing attitudes IMs.

##### The list of tags for the IMs by groups:

Pace regulating markers:

- *Fill*—(often repetitive) discourse markers (DMs) of cognitive search
- *Reg*—DMs mostly signalling the beginning or the end of EDUs, utterances or stances (passages) that serve to pace narration and mark its chunks

Cognitive difficulties markers:

- *Rp*—repetitions of words and phrases
- *Fs/Rep*—false starts followed by repairs
- *Rep*—repairs that happen without preceding false starts
- *SFs*—structural false starts

<sup>2</sup> Another approach to the macro level analysis of the stories is based on the genre schema and its components, for example: abstract, introduction, main line events, evaluation, coda. The latter two, being part of the world of narrator, also belong to interactive aspect of storytelling. We will be dealing with componential analysis of stories in CliPS at the next stage of our project.

<sup>3</sup> See Fraser 2009 for a comprehensive classification of discourse markers.

Attention drawing markers:

- ***Apl***—DM that target to draw the addressee's attention
- ***DM***—other DM<sup>4</sup>

Attitude expression markers:

- ***Est***—markers of epistemic attitude
- ***Jdg***—markers of axiological evaluation

The tag ***IMP*** refers to clausal or clause-like markers of interaction, as opposed to one- or two-word discourse markers.

Repetitions (***Rp***) may signal either the importance of the repeated item for the story, or difficulties in the process of mental search, which is especially true for the pathological discourse. Still the healthy speakers' stories often demonstrate abundance of the mental search problem as part of the specific *I-am-not-Committing-Myself* discourse strategy.

### 3.3.3. Annotation procedure

The annotation of IMs and EMs was performed by two raters independently for all texts. Each item in the Lexical transcript tier was interpreted as a start of a new marker (with identification of its type), a word belonging to the previously annotated marker, or as a neutral word.

For empathy annotation the percentage agreement between raters on 5943 items was 93.6%, Cohen's Kappa = 0.356. For interaction marker annotation the percentage agreement between raters on 5943 items was 80.9%, Cohen's Kappa = 0.501. Mismatches in annotation were resolved by the third annotator.

## 4. Results

### 4.1. Empathy markers

A marginally significant effect of the speaker group on the amount of EMs in narratives was found with ANOVA (one-way, between subject),  $F(2, 20) = 3,28$ ,  $p = 0,06$ ; as shown by posthoc analysis with Tukey HSD, healthy speakers use more EMs than people with dynamic aphasia ( $p_{adj.} = 0,09$ ), see Figure 1 for summary and of the distribution of EMs across speakers and groups.

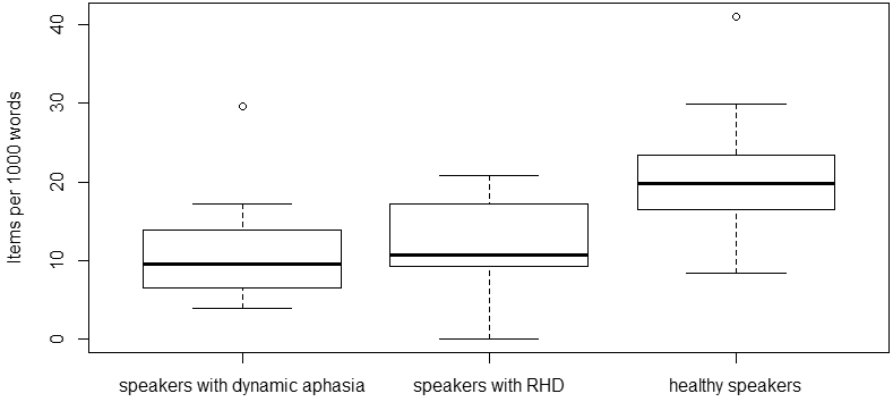
### 4.2. Interaction markers

Given the non normal distribution of IM data, we have run four separate Kruskal-Wallis tests for each group of IM, with adjusted p values using Holm method, as well as Dunn posthoc analysis. We have found no significant differences between the speakers' groups for the use of pace-regulating and attention drawing markers. PWA(d) use significantly more cognitive difficulties markers ( $p_{adj.} = 0,1$ )

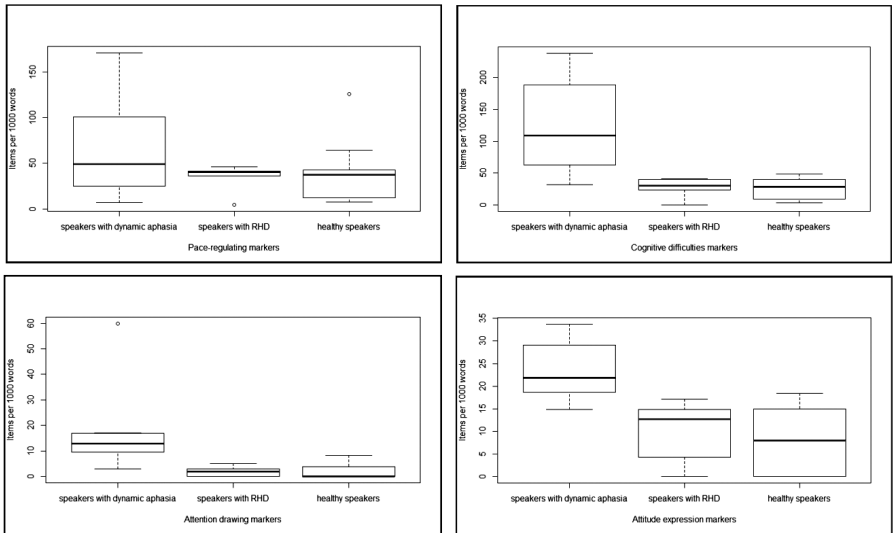
---

<sup>4</sup> Interaction markers ***DM*** and to some extent ***Apl*** are introduced here to cover those cases of interaction that can not be easily ascribed to other more specific classes.

and significantly less attitude expression markers ( $p < 0,05$ ) than healthy speakers. We have found no significant differences between the use of any type of IM by people with RHD and PWA(d) or healthy speakers. See Figure 2 for summary of the distribution of IMs across speakers and groups.



**Figure 1.** Distribution of empathy markers in groups



**Figure 2.** Distribution of IMs across groups

## 5. Discussion

Our results demonstrate that PWA(d) use significantly more markers of cognitive difficulties than healthy speakers and people with RHD. Obviously, this is a natural effect of the non-fluent speech output. Also, PWA(d) use significantly less attitude expression markers as compared with the healthy speakers. This corresponds with what is known about difficulties that PWA(d) experience in narration. Akhutina (1989/2012, p.45; cited in Akhutina 2015) points out that in dynamic aphasia, the story schema is disintegrated, so the patient is unable to produce and maintain the plan of the utterance. Complexity of planning depends on the task. As shown in Akhutina 1975/2002, p.50 (cited in Akhutina 2015), the hardest is narration on a given topic (3 points out of 3), while retellings may vary from 1 to 3. Too many cognitive resources of PWA(d) are spent on maintaining the story schema, keeping the main-line of the story, balancing narration and description, so there is not enough left for expressing attitudes towards the described events.

Our data show that discourse in RHD does not show significant differences from that of healthy speakers in the amount of IMs, or EMs, used. However, they use less EMs than healthy speakers, though this difference does not reach significance. As stated in Traxler, (2012: 531) regarding the RHD and challenges in narration, “the right hemisphere appears to play a vital role in keeping track of the main idea or theme of a story as well as using the theme to organize the various subcomponents of the story into a coherent macrostructure”. More specifically, discussing the retelling task, Traxler (ibid) notes that RHD omit more important ideas than patients with LHD. We don’t know what specific disorder (aphasia type) is meant here by the LHD (left hemisphere damage), but in our experiment PWA(d) don’t demonstrate significant differences in empathy as compared with RHD discourse. Still, speaking about RHD, we can use the same reasoning as for the PWA(d) lack of expressing attitude markers and imagine that their difficulties with keeping themes and ideas which leads to omitting some of them, will result in certain, not intended, prioritizing of what is more important. The most important thing in the story is the main line of events. This may explain the RHD lower figures for empathy.

Our results and the results we got for another group of PWA (acoustic-mnestic aphasia type—see (Bergelson et al., 2016) show that there exist correlations between the prevalence of specific types of interactional elements and type of disorder. In this and other papers (Bergelson et al., 2015, Bergelson et al., 2016) we check these data against the interactional strategies used by the healthy speakers in the same experiment, which brings into comparison additional factors, such as situational context, personal goals, social distance and others. Altogether they fall into two parameters: the parameter of the task, which is stable and in this case amounts to retelling a film in the experimental setting, and the individual discourse strategies that healthy speakers use to cope with this specific task. We call these strategies compensation for a communication setting and task that fall out of the normal regular experience of healthy speakers. We all tell personal stories about the experiences we had, but retelling a film in an experimental setting is a somewhat awkward situation. Participants feel that some of their abilities (memory? narrative skills?) are being tested and respond to this correspondingly.

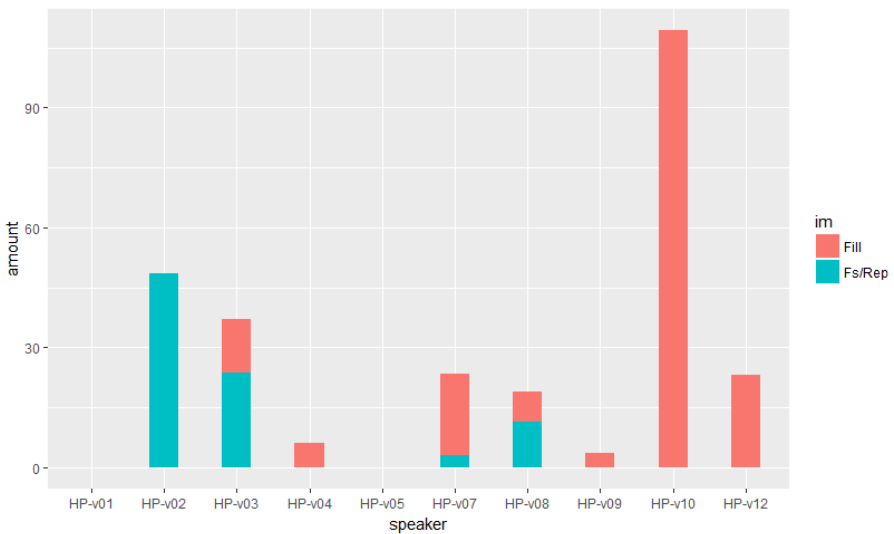


This project data provides additional evidence for the existence of at least four discourse strategies employed by healthy speakers in their retellings of The Pear Stories film in attempts to protect their social face.

- Strategy 1: Expressing Uncertainty, or *I-am-not-Committing-Myself*
- Strategy 2: Fine-grained approach, or *Staying-on-a-Safe-Side*
- Strategy 3: ‘As is’ story, or *No-Worries*
- Strategy 4: Discussing and evaluating, or *Moralist-Syndrome*

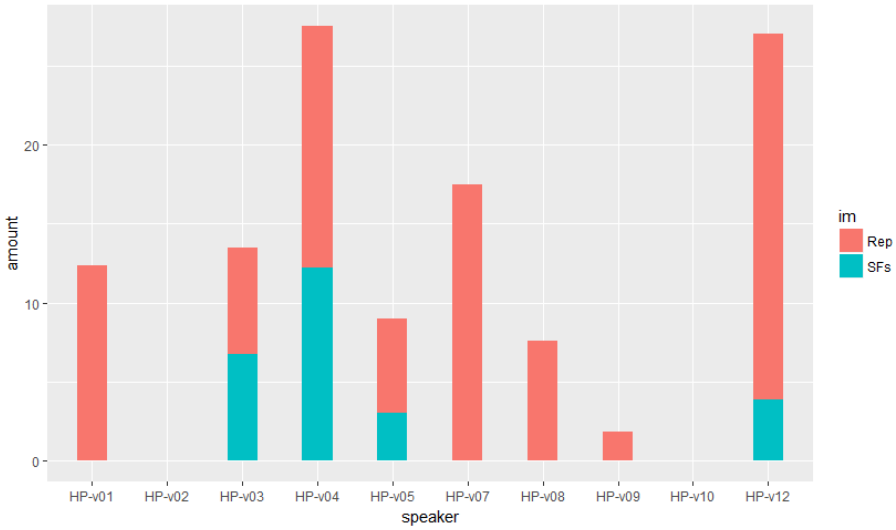
The four strategies of healthy speakers correlate with four groups of IMs, with each strategy having IMs of a certain group as its champion.

Strategy 1: Expressing Uncertainty, or *I-am-not-Committing-Myself* is manifested by abundant use of fillers (marker *Fill*), disclaimers, false starts, self-corrections, and laughter, non-standard reference and irony, some cases of epistemic markers—see Figure 3 for the use of false starts and *Fill* markers. Participants HP-v02 and HP-v10 use this strategy. While HP-v10 bases this strategy on pace regulating IMs, HP-v10 sticks to false starts and repairs.



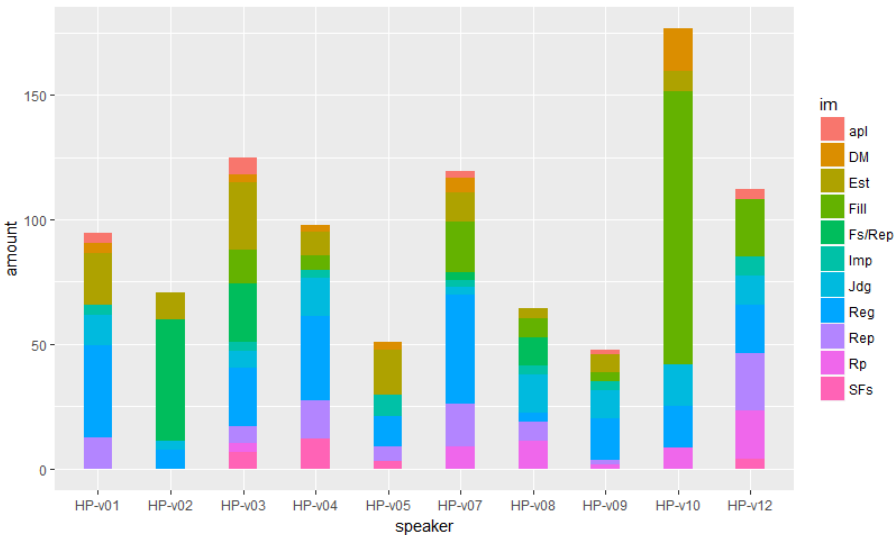
**Figure 3.** Distribution of markers of *I-am-not-Committing-Myself* strategy

Strategy 2: Fine-grained approach, or *Staying-on-a-Safe-Side* reveals itself through lots of details in descriptions of the characters. Discourse marker *ny* is associated with this strategy. Narrators try to cover everything, which leads to abundant use of structural false starts (*SFs*) and repairs (*Rep*) unmotivated by false starts. Participants HP-v04 and HP-v12 represent this strategy as shown in Figure 4. demonstrating that markers of cognitive difficulties are inherent in the healthy discourse as well.



**Figure 4.** Distribution of markers of *Staying-on-a-Safe-Side* strategy

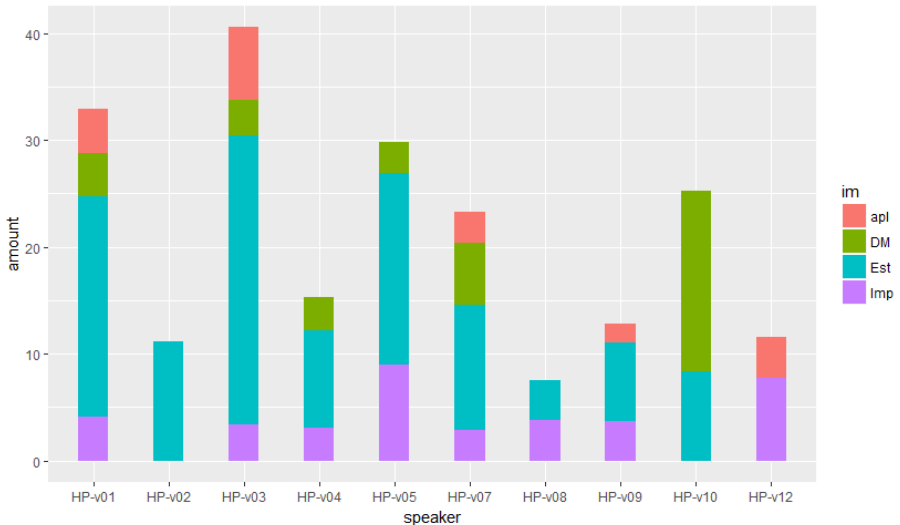
Strategy 3: ‘As is’ story, or *No-Worries* is demonstrated by minimal use of any IMs. This is best demonstrated by the participant HP-v09 and to a certain degree by HP-v05—see Figure 5.



**Figure 5.** Distribution of all types of interaction markers

Strategy 4: Discussing and evaluating, or *Moralist-Syndrome* is more difficult to capture with our data, as it involves using lots of evaluations, prolonged and detailed coda with evaluations of characters’ actions. Markers expressing attitudes

(*Est*, *Jdg*) are characteristic of this strategy. In Figure 6 we see the champion for this strategy—HP-v03.



**Figure 6.** Distribution of markers of *Moralism-Syndrome* strategy

We believe that more research is needed to be able to thoroughly describe various compensating strategies and the discourse means used by narrators for this purpose. We know that these and other compensating strategies are used by people with various types of aphasia and RHD. Yet another step will be to compare them across the groups.

## References

1. Akhutina, T. (2015). Luria's classification of aphasias and its theoretical basis. *Aphasiology*, (August), 1–20.
2. Bergelson, M. B., Akinina, Y. S., Dragoy, O. V., Iskra, E. V., & Khudyakova, M. V. (2015). Markers of word production difficulties in normal and clinical discourse production: continuity of norm in language and discourse [Zatrudnenija pri porozhdenii slov v diskurse i ix formal'nye markery: norma i patologija, ili o nediskretnosti normy v jazyke. *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference Dialogue*, 14(1), 41–51.
3. Bergelson, M. B., Khudyakova, M. V., Grabovskaya, M., Ivushok, E., & Shapiro, M. (2016). Markers of expressing empathy and interaction in healthy and pathological speech (study based on the Russian CLiPS corpus material) [Markery vyrazheniya e'mpatii i interakcii v norme i patologii (na materiale korpusa Russian CLiPS)]. In *Proceedings of the seventh international conference on cognitive science [Sed'maya mezhdunarodnaya konferenciya po kognitivnoj nauke: Tezisy dokladov]* (pp. 147–148). Moscow, Russia: Institute of Psychology RAS.
4. Chafe, W. (1980). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. (W. Chafe, Ed.). Norwood, New Jersey: Ablex.

5. *Fleischman, S.* (1991). Toward a Theory of Tense-Aspect in Narrative Discourse. In J. Gvozdanovic & T. A. J. Janssen (Eds.), *The Function of Tense in Texts* (pp. 75–97). Amsterdam.
6. *Fraser, B.* (2009). An Account of Discourse Markers. *International Review of Pragmatics*, 1 (2009), 293–320.
7. *Glosser, G., & Deser, T.* (1990). Patterns of Discourse Production among Neurological Patients with Fluent Language Disorders. *Brain and Language*, 49, 67–88.
8. *Hoey, M.* (1994). Signalling in Discourse: a functional analysis of a common discourse pattern in written and spoken English, 26–45.
9. *Kalliokoski, J., & Verschuere, J.* (1991). Emphathy as motivation for style shifting in narrative. In *Levels of linguistic adaptation. Selected papers of the International Pragmatics Conference* (Vol. 2, p. 147–161).
10. *Khudyakova, M. V., Bergelson, M. B., Akinina, Y. S., Iskra, E. V., Toldova, S., & Dragoy, O. V.* (2016). Russian ClipS: a Corpus of Narratives by Brain-Damaged Individuals. In *LREC Proceedings*. Portoroz, Slovenia.
11. *Kibrik, A. A.* (2009). Mode, genre and other parameters of discourse classification [Modus, zhanr i drugie parametry klassifikacii diskursov]. *Topics in the Study of Language [Voprosy Jazykoznanija]*, 2, 3–21.
12. *Kibrik, A. A., & Podlesskaya, V. I.* (Eds.). (2009). *Night Dream Stories: A corpus study of spoken Russian discourse [Rasskazy o snovidenijah: korpusnoe issledovanie ustnogo russkogo diskursa]*. Moscow: Languages of Slavonic Culture.
13. *Kuno, S.* (1987). *Functional syntax: Anaphora, discourse and empathy*. University of Chicago Press.
14. *Kuno, S., & Kaburaki, E.* (1977). Empathy and syntax. *Linguistic Inquiry*, 627–672.
15. *Labov, W.* (1972). The Transformation of experience in narrative syntax. In *Language In The Inner City: Studies In The Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
16. *Labov, W., & Waletzky, J.* (1967). Narrative analysis. In J. Helm (Ed.), *Essays on the Verbal and Visual Arts* (pp. 12–44).
17. *Linnik, A., Bastiaanse, R., & Höhle, B.* (2015). Discourse production in aphasia : a current review of theoretical and methodological challenges. *Aphasiology*, 30(7), 765–800.
18. *Longacre, R. E.* (1983). Spectrum, profile and constituency structure in text analysis. In S. Hattori & K. Inoue (Eds.), *Proceedings of the 13th International Congress of Linguists, August 29-September 4, 1982, Tokyo* (pp. 1024–27). Tokyo: Comité International Permanent des Linguistes.
19. *Longacre, R. E.* (1990). *Storyline Concerns and Word Order typology in East and West Africa*. Los Angeles, California: University of California.
20. *Luria, A. R.* (1972). Aphasia reconsidered. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 8(1), 34.
21. *Luria, A. R., & Hutton, J. T.* (1977). A modern assessment of the basic forms of aphasia. *Brain and Language*, 4(2), 129–151.
22. *Marini, A., Carlomagno, S., Caltagirone, C., & Nocentini, U.* (2005). The role played by the right hemisphere in the organization of complex textual structures. *Brain and Language*, 93(1), 46–54.

23. *Norrick, N.* (2000). *Conversational Narrative: Storytelling in everyday talk*. Benjamins.
24. *Olness, G. S., Matteson, S. E., & Stewart, C. T.* (2010). "Let me tell you the point": How speakers with aphasia assign prominence to information in narratives. *Aphasiology*, 24(August 2013), 697–708.
25. *Paducheva, E. V.* (2008). Discourse words and categories: interpretation regimes [Diskursivnye slova i kategorii: rezhimy interpretacii]. In V. Plungyan (Ed.), *Studies in grammar theory [Issledovaniya po teorii grammatiki]* (Vol. 4, pp. 56–86). Moscow: Gnozis.
26. *Plungyan, V.* (2008). Discourse and grammar. Preface [Diskurs i grammatika. Predislovie]. In V. Plungyan (Ed.), *Studies in grammar theory [Issledovaniya po teorii grammatiki]* (Vol. 4, pp. 7–36). Moscow: Gnozis.
27. *Polanyi, L.* (1989). *Telling the American Story. A Structural and Cultural Analysis of Conversational Storytelling*. Cambridge, Massachusetts: MIT Press.
28. *Sherratt, S., & Bryan, K.* (2012). Discourse production after right brain damage: Gaining a comprehensive picture using a multi-level processing model. *Journal of Neurolinguistics*, 25(4), 213–239.
29. *Toldova, S. Y., Bergelson, M. B., & Khudyakova, M. V.* (2016). Coreference in Russian Oral Movie Retellings (the Experience of Coreference Relations Annotation in "Russian CliPS" corpus ). In *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2016"* (Vol. 15, pp. 769–782). Moscow: RSUH.
30. *Traxler, M. J.* (2012). Introduction to psycholinguistics. *Understanding language science*.
31. *Weinrich* (1964). *Tempus: besprochene und erzählte Welt*. Stuttgart: Kohlhammer.
32. *Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H.* (2006). ELAN: a Professional Framework for Multimodality Research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.

# ПОБУДИТЕЛЬНАЯ РЕПЛИКА КАК ЭЛЕМЕНТ УСТНОГО ДИАЛОГА: СЛУЧАЙ ИМПЕРАТИВА 1PL

**Блинова О. В.** (o.blinova@spbu.ru)

Санкт-Петербургский государственный  
университет, Санкт-Петербург, Россия

В статье рассматриваются употребления императивных конструкций 1PI с *давай(те)*, *дай(те)*. Провозглашается разговорно-ориентированный подход к анализу императивных реплик. С опорой на диалогическое соседство и особенности протекания диалога выделяется четыре семантических подтипа императива 1PI: пропозитив (выражает 'предложение', предполагает вербальную реакцию слушающего), гортатив (выражает 'распоряжение', не предполагает вербальной реакции слушающего), фактитивный пермиссив (выражает запрос на устранение препятствий к действию, не предполагает вербальной реакции слушающего), пермиссив (выражает 'разрешение', завершает микродиалог). Императивные конструкции исследуются на материале двух речевых корпусов. Разница между встречаемостью подтипов императива в корпусах объясняется разными долями повседневных и институциональных разговоров в выборках. В корпусе с преобладанием институциональных разговоров между продавцами и покупателями подавляющее число употреблений императива 1PI (70%) — пропозитив. Это обусловлено особенностями речевых стратегий продавцов. Материалы исследования могут быть использованы для автоматического моделирования диалога на русском языке.

**Ключевые слова:** русский устный диалог, императивные конструкции 1PI, микродиалоги, иницирующие реплики, реактивные реплики, повседневные разговоры, институциональный дискурс

# DIRECTIVE UTTERANCE AS AN ELEMENT OF SPOKEN DIALOGUE: THE CASE OF IMPERATIVE 1PL

**Blinova O. V.** (o.blinova@spbu.ru)

Saint Petersburg State University, Saint Petersburg, Russia

The article deals with Russian imperative constructions 1PI with particles *davay(te)* / *day(te)*. It is argued that conversation-oriented approach

in analysis of these imperative utterances can be very productive. On the basis of the neighboring utterances and specifics of dialogue organization one can distinguish four semantic subtypes of 1PI imperative: propositive (which is an expression of 'suggestion' and presupposes a verbal reaction from the listener), hortative (which expresses an 'mild order' and no verbal reaction of the listener is expected), factitive permissive (expresses a request to remove obstacles for action, no verbal reaction of the listener is expected), permissive (expresses 'permission', the utterance appears as a reaction in the final part of the microdialogue). Imperative constructions are studied in two speech corpora. The difference in frequency for the four imperative subtypes in the two corpora can be explained by the difference in the amount of everyday and institutional talks in the sampling. In the corpus where institutional talks between «sellers» and «customers» are predominant most of the 1PI imperative usages (70%) are propositive. This is determined by the fact that such are specific strategies used by the «sellers». The research results can be used for the modeling of Russian dialogue.

**Key words:** Russian spoken dialogue, imperative constructions 1PI, microdialogues, initial utterance, reactive utterance, daily conversation, institutional discourse

Imperatives are a window to social structures and the ethnography of communication—the unwritten rules we abide by» (Aikhenvald 2010, 8).

## 1. Постановка проблемы

Настоящая статья посвящена изучению устного диалогического взаимодействия, точнее, рассмотрению императивной реплики как элемента диалога. Изложение строится вокруг центральной идеи: **для интерпретации употреблений императива, для приписывания императивным репликам значений необходимо использовать информацию об их диалогическом окружении.** Под «значениями императива» имеются в виду и некоторые «частные семантические интерпретации императивного значения» — 'приказ', 'просьба', 'инструкция', 'предложение', 'разрешение', 'совет' (Храковский, Володин 1986), и элементы разнообразных классификаций типов побуждения в терминах речевых либо «диалогических актов» (речевых актов в контексте диалога, см. Jurafsky 2006), и «семантические подтипы императива» согласно (Гусев 2005, 52–87).<sup>1</sup>

Приписывание употреблением русского императива частных значений затруднено тем, что инвентарь морфологических и синтаксических средств, которые используются для образования императивных форм и конструкций,

<sup>1</sup> 'Прохибитив', 'превентив', 'пермиссив', 'вежливый императив', 'некатегорический императив', 'категорический императив' и др.

невелик, а набор возможных интерпретаций — широк.<sup>2</sup> Реплики с однородным формально-грамматическим составом могут трактоваться по-разному в зависимости от речевого и социального контекста. Ср., в частности, следующие определения Л. А. Бирюлина : «Императивное высказывание понимается как ‘приказ’ <...> при следующих необходимых условиях: (а) императив выражает фактитивную прескрипцию, при которой импульс каузации исходит от Говорящего, (б) **приказ отдаётся тем лицом, чей социальный статус в социальной иерархии выше, и направлен тому лицу, чей социальный статус ниже** <...>», «Интерпретация императивного высказывания как ‘просьбы’ <...> возможно при следующих необходимых условиях: <...> (б) **просьба обычно исходит от того лица, чей социальный статус ниже, и направлена тому лицу, чей социальный статус выше, либо используется в ситуациях, где субординативные отношения отсутствуют** <...>» и т.д. (Бирюлин 1992, 26). Таким образом, «частные семантические интерпретации императивного значения» — это «**социально значимые** интерпретации базового значения императива ‘предписание’» (там же).

Частные значения императива «привязаны» к определённым интонационным контурам, ср.: «значение императива зависит от его интонации. *Приходи!*, сказанное с интонацией мольбы, означает то же самое, что *Я умоляю тебя прийти; Приходи!*, сказанное с интонацией совета, означает то же самое, что *Я советую тебе прийти*, и т.д.» (Вежбицка 1985, 256, цит. по ТФГ 1990, 203). Но каждая интонационная конструкция «соотносится более чем с одной семантической интерпретацией повелительных предложений» (там же, см., однако, Кодзасов 2009, Янко 2008, 178–180, 295), соответственно, отождествление частных значений императивных реплик на основе просодической информации проблематично.<sup>3</sup>

Вывод из сказанного (в применении к русскому императиву) может быть таким: «... выделение и классификация семантических интерпретаций императивного значения не является собственно лингвистической задачей, поскольку объектом классификации фактически выступают ситуации произнесения императивных высказываний» (Храковский, Володин 1986, 136).

Между тем, интерпретировать употребление императива можно с привлечением информации о **речевых (в частности, вербальных) реакциях собеседников** (каузируемых слушающих). Речевая реакция — это результат интерпретации императивной реплики в ситуации говорения. Эта интерпретация (в случае императивов 1 и 2 лица) производится участниками ситуации, которые в общем случае обладают достаточной информацией для согласованной трактовки волеизъявления. Более того, при несогласованной трактовке волеизъявления говорящим и слушающим могут происходить подлежащие фиксации **коммуникативные сбои** (отражающиеся, например, в **перебивах и наложениях речи** — периодах одновременного говорения), см. (Влинова 2016).

<sup>2</sup> Ср.: частных семантических интерпретаций столько, сколько в данном языке несинонимичных глаголов речевой каузации (ТФГ 1990, 204).

<sup>3</sup> Обратное (описание просодии реплик с определённым значением) выполнялось в рамках проекта «Интонация русского диалога» (Кодзасов и др. 2006).



Итак, анализу подлежит не изолированная императивная реплика, а **микродиалог**, в простейшем случае состоящий из иницирующей (стимульной) и реактивной реплики.<sup>4</sup> Такой разговорно-ориентированный подход в настоящей статье применён к побудительным репликам с одной лично-числовой формой императива (1PI) и опробован на материале двух корпусов.

Насколько можно судить, приписывание различным лично-числовым формам русского императива значений с опорой на диалогический контекст ещё не производилось (в том числе императив сколько-нибудь систематически не рассматривался в свете вербальных реакций).<sup>5</sup> В (Баранов, Крейдлин 1992а, 96–97) среди других минимальных диалогов обсуждается пример:

— *Принеси другую лампу.*

— *Какую?*

— *Ту, что висела в прошлый раз у тебя в комнате* (Запись диалога из телеспектакля).

Целью этого обсуждения является не анализ значения побудительной реплики на основании информации об особенностях протекания диалога, а уточнение свойств минимального диалога (в частности, через рассмотрение возможности отбрасывания отдельных реплик в его составе).

Однако в цитируемой работе содержится принципиально важное заключение о том, что приведённую последовательность реплик нельзя считать минимальным диалогом (или минимальной диалогической единицей, МДЕ). Для соответствия вводимому авторами определению минимального диалога следует восстановить завершающую **нулевую реплику**, так как иначе отношение иллокутивного вынуждения является невыполненным; результатом его выполнения может быть молчаливое осуществление действия (resp. восстанавливаемая нулевая реплика), вербальное выражение согласия или отказа осуществить действие.

## 2. Характеристика материала

Предметом рассмотрения являются микродиалоги с участием императивных конструкций 1PI. Рассматриваются конструкции с частицами *давай(те)*,

<sup>4</sup> В модели диалога (Баранов, Крейдлин 1992а) вводится противопоставление между иллокутивно независимыми и иллокутивно зависимыми репликами (такими, значение которых определяется репликами предшествующими). Соответственно, минимальный диалог начинается с «абсолютно независимого» речевого акта и заканчивается «абсолютно зависимым», при этом речевые акты связаны отношением иллокутивного вынуждения. Благодарю анонимных рецензентов за указание на эту влиятельную работу. Учёт диалогического соседства используется и в западной корпусной прагматике для присвоения элементам диалога меток определённых диалогических актов, см., например, (Stolke et al. 2000).

<sup>5</sup> Вербальные реакции на императив 2Р типа *Ладно / Хорошо / Сделаю* упоминаются в (Шатуновский 2009) при рассмотрении императивов НСВ (как знак того, что каузируемый «уже был успешно побуждён делать Р, но почему-либо медлит с его выполнением»). Возможные отклики на императивы 2Р *Смотри, Слушай* используются в ходе выделения некоторых значений этих форм в (Баранов, Крейдлин 1992б).

дай(те) и формами будущего времени либо инфинитива смысловых глаголов типа *давай(те) сделаем, дай(те) сделаем, давай(те) делать, дай(те) сделать*.

Первый источник языковых данных — корпус повседневной речи «Один речевой день» (ОРД), размеченный подкорпус объёмом в 239 500 слов. Подкорпус содержит преимущественно устные спонтанные диалоги (о корпусе см. Bogdanova-Beglarian et al. 2016).

Второй источник языковых данных — Акцентологический корпус НКРЯ, пользовательский подкорпус «Устная непубличная речь» (типы текста: беседа, микродиалог, разговор, спор). Объём подкорпуса — 216 119 слов.<sup>6</sup>

### 3. Микродиалоги с императивом 1РІ: схема и результаты рассмотрения

#### 3.1. Императив в реактивных репликах

При выделении частных значений императива 1РІ в первую очередь следует учитывать, является императивная реплика стимульной (иницирующей) или реактивной. Реактивные императивные реплики — это выражения с пермиссивной каузацией (т.е. импульс каузации исходит от слушающего, который задал вопрос, поинтересовался, *можно ли сделать X* и т.д.). Такие реактивные реплики выражают 'согласие'. Будем обозначать эти употребления как «**пермиссив**», ср.:

[И04, жен.] *а / это сначала выпить / извини //*

[Ж01] *ну ладно / давай(те) выпьем //* [ОРД, Празднование дня рождения]

[Илья, покупатель] *а здравствуйте. А не могли бы посчитать окна?*

[Олег, менеджер] *Да / конечно. Давай(те) посчитаем* [НКРЯ, Заказ пластиковых окон (2015)]

#### 3.2. Императив в стимульных репликах

При определении значения реплик на втором шаге следует учитывать, следует ли за репликой вербальная реакция на волеизъявление. Сначала рассмотрим случаи, когда вербальная реакция есть.

В. И. Подлеская в (Подлеская 2005) пишет о существовании двух типов конструкций с *дать/давать*: «гортатива» и «пермиссива». Употребляя гортатив, говорящий ожидает санкции слушающего — подтверждения «предположения говорящего о том, что слушающий не возражает против того, чтобы действие было осуществлено» (Подлеская, 2005, 96). В (Гусев 2008, 189) конструкциям

---

<sup>6</sup> Применение тех же параметров при формировании пользовательского подкорпуса в Мультимедийном корпусе НКРЯ (где есть не только транскрипт, но и звук) позволило получить совсем небольшую выборку из 11 538 слов. Пользовательский же подкорпус с указанными характеристиками, сформированный на основе Устного корпуса НКРЯ для текущего этапа исследования слишком велик (11 349 008 слов). Именно поэтому решено взять материалы Акцентологического корпуса.

с *давай*(*те*) и смысловым глаголом в будущем времени предложено присваивать значение «говорящий предлагает слушающему согласиться на некое действие». Такие употребления названы «пропозитивными императивами», см также (Гусев 2005, 85–87).

Итак, «гортатив» Подлесской и «пропозитив» Гусева — это типы **иницирующих** реплик, которые призваны вызывать **вербальную реакцию** каузируемого слушающего (это может быть подтверждение того, что слушающий не возражает против совершения действия или, напротив, выражение несогласия с совершением действия). Нормальное течение микродиалога в таком случае может быть схематически обозначено так:

[Каузатор говорящий] 'ПРЕДЛОЖЕНИЕ' *Давай сделаем X!* ПАУЗА

[Каузируемый слушающий] 'СОГЛАСИЕ' *Давай!* или 'ОТКАЗ'

В такой схеме после инициирующей реплики должна следовать **пауза**, которую делает говорящий для того, чтобы каузируемый слушающий мог ответить.

Таким образом, можно выделить семантический подтип императива, который вызывает вербальную реакцию слушающего. Для такого императива будем использовать обозначение «**пропозитив**». В терминах речевых актов эти реплики можно обозначать как 'предложение', ср.:

[Ж1] *давайте мы вас в друзья добавим //*

[ИЗ9] *ой добавьте //* [ОРД]

[Р1, мальчик] *давай окно откроем //*

[И4, мама] *нет Хомочка / если мы откроем окно / мы замёрзнем //* [ОРД]

Если вербальной реакции нет, и перед нами простая последовательность реплик<sup>7</sup> (стимульная императивная реплика + «нулевая» реактивная реплика, за которой следует следующая, новая стадия развития диалога, следующий речевой акт), то скорее всего стимульной императивной реплике можно приписать значение «**гортатива**». Такой гортатив в узком смысле по значению является 'распоряжением', а не 'предложением'. Каузатор-говорящий не ждёт вербальной реакции слушающего, не делает паузу после выражения волеизъявления и продолжает говорить. Важно, что значение гортатива можно уверенно приписать **лишь при наличии информации о социальном контексте**: распоряжение отдаётся лицом, находящимся в социальной иерархии выше, чем лицо, которое его получает (или лицом, обладающим в ситуации коммуникации особыми полномочиями — врачом, учителем, инспектором ГАИ). Это распоряжение, однако, носит существенно менее категоричный характер, чем 'приказ', ср.:

[И14, аспирантка] *преподавала устную () практику ...*

[Ж1, научный руководитель] [перебивает] *во во во //* **давайте / напишем //**

[И14, аспирантка]  $\emptyset$

<sup>7</sup> Сложные последовательности реплик рассматриваются ниже в пункте 3.6.

[Ж1, научный руководитель] [пишет и произносит] \**П преподавала / \*П устную / \*П устную практику* / [ОРД, Написание отзыва научного руководителя]

[И44, врач] *так / давай / послушаем тебя* //

[Р1, пациент, девочка] ∅

[Ж1, мама девочки] \**П нормально сиди* / [ОРД, На приёме у врача]

[И37, учитель рисования] *ну вот давай всё / давай убирать потихонечку* //

[Р1, ученица, девочка] ∅

[И37, учитель рисования] \**Ш всё уже ?* [ОРД, Урок рисования]

Следует учитывать просодические, синтаксические и лексические **показатели завершенности/незавершенности высказываний**. Например, в диалоге ниже стимульная реплика не завершена. По-видимому, вероятность, что слушающий захочет вклиниться в незавершенное высказывание, менее велика (отсутствие вербальной реакции может быть поддержано также и этим фактором), ср.:

[И40, врач] *то есть по всем... по всем показателям ... \*П вот / и давайте (э-э) запланируем где-нибудь ... (э-э) (...) не знаю /*

[М1, пациент] ∅

[И40, врач] *так / сегодня ... \*П у нас что получается / значит мы (э-э) клинический анализ крови сдавали вчера / правильно ?*

[М1, пациент] *угу // \*П позавчера //* [ОРД, На приёме у врача]

Наконец, несогласованная интерпретация гортатива говорящим и слушающим (когда говорящий распоряжается, а слушающий думает, что говорящий интересуется его мнением о совершении действия) может выражаться в одновременном говорении, наложениях речи.

Таким образом, при определении значения реплики учитывается и информация о **социальной дистанции между говорящими** (их статусах), **ситуации коммуникации** и других факторах, которые влияют на организацию диалогического взаимодействия. Признаки, связанные с равновесием/неравновесием статусов, соотносятся с признаком **субординации** 'говорящий ставит себя выше слушающего' (Храковский, Володин 1986, 137). Ситуация коммуникации соотносится с **типом дискурса** и выражается в различении **институционального дискурса и повседневных бытовых разговоров** (подробнее см. п. 3.5 ниже).

Рассмотрение стимульных императивных реплик и опора на работу (Подлеская 2005) позволило выделить ещё один (четвёртый) семантический подтип императива 1P1 — это императив типа *дай(те) сделаем*, который образуется с помощью частицы *дай(те)* и передаёт значение 'говорящий просит слушающего или устранить препятствия, или воздержаться от того, что могло бы препятствовать действиям говорящего' (там же). Такой императив обозначим термином «**фактитивный пермиссив**», ср.:

[И19, жен.] *ой / дай-ка(?) уберём это / () отсюда / чтобы(?) тебе не мешало* // [ОРД]

### 3.3. Подтипы императива 1PI: обобщение

Итак, мы выделили четыре частных значения конструкций императива 1PI:

1. **«пропозитив»** — употребляется в иницирующих репликах, предполагает вербальную реакцию слушающего, значит ‘говорящий предлагает слушающему согласиться или не согласиться с совершением действия’. В терминах речевых актов — ‘предложение’.
2. **«гортатив»** — употребляется в иницирующих репликах, не предполагает вербальной реакции слушающего; используется доминирующим говорящим при неравновесии статусов говорящего и слушающего, значит ‘говорящий каузирует слушающего совершить действие’.
3. **«фактитивный пермиссив»** — употребляется в иницирующих репликах, не предполагает вербальной реакции слушающего; говорящий каузирует слушающего «устранить препятствия, или воздержаться от того, что могло бы препятствовать действиям говорящего», значение см. (Подлесская, 2005, 96).
4. **«пермиссив»** — употребляется в реактивных репликах, завершает микродиалог, значит ‘говорящий выражает согласие с совершением действия, о котором ранее упоминал слушающий’.

Распределив признаки, противопоставляющие четыре типа конструкций, получаем такую таблицу:

Таблица 1. Признаки императивных конструкций 1PI

свойства конструкции	горта- тив	«фактитивный пермиссив»	пропо- зитив	пермис- сив
употребляется прежде всего с частицей	<i>давай- (те)</i>	<i>дай(те)</i>	<i>давай- (те)</i>	<i>давай- (те)</i>
употребляется в составе иницирующих реплик	да	да	да	нет
употребляется в составе реактивных реплик	нет	нет	нет	да
ожидается прежде всего вербальная реакция слушающего	нет	нет	да	нет
ожидается прежде всего выполнения каузируемого действия	да	да	нет	нереле- вантно
статус Говорящего выше статуса Слушающего	да	нет	нет	да
статус Слушающего выше статуса Говорящего	нет	да	нет	нет
Говорящий и Слушающий равноправны	нет	нет	да	нет

### 3.4. Императивные конструкции 1PI по данным корпусов

Рассмотрение материала корпусов (размеченного подкорпуса ОРД объёмом в 239 500 слов и пользовательского подкорпуса Акцентологического корпуса НКРЯ объёмом в 216 119 слов) дало результаты, представленные в Таблице 2 и Таблице 3 ниже.

**Таблица 2.** Встречаемость конструкций Imp 1PI в ОРД

конструкция	количество употреблений	ipm
гортатив	47 (28,7%)	196,24
«фактитивный пермиссив»	1 (0,6%)	4,18
пропозитив	81 (49,4%)	338,20
пермиссив	35 (21,3%)	146,14
всего	164 (100%)	(239 500 слов)

**Таблица 3.** Встречаемость конструкций Imp 1PI в АК НКРЯ<sup>8</sup>

конструкция	количество употреблений	ipm
гортатив	13 (18,6%)	60,15
«фактитивный пермиссив»	0	0,00
пропозитив	49 (70%)	226,73
пермиссив	8 (11,4%)	37,02
всего	70 (100%)	(216 119 слов)

### 3.5. Императив 1PI в институциональном дискурсе

Различия в долях употреблений разных типов конструкций в обследованных выборках довольно существенны. Так, употребления пропозитива в ОРД составляет около половины всех вхождений императивных конструкций 1PI. В АК НКРЯ значение пропозитива можно присвоить 70% примеров.

Представляется, что эта разница связана прежде всего с особенностями представленных типов дискурса. Так, в ОРД зафиксированы в основном **бытовые повседневные разговоры**, а **институциональные разговоры** встречаются спорадически. В выборке же из Акцентологического корпуса НКРЯ

<sup>8</sup> Поиск вёлся по запросу V,(imper | imper2),pl, объём пользовательского подкорпуса устной непубличной речи в Акцентологическом корпусе (типы текстов «беседа», «микродиалог», «разговор», «спор») — 180 документов, 34424 предложения. Реплики, образованные на основе форм imper2 (*пойдёмте, идёмте* и др.) при подсчётах не учитываются.

большая часть документов — это **разговоры по телефону между менеджером (продавцом-консультантом) и клиентом (покупателем)**.<sup>9</sup>

Противопоставление между повседневным и институциональным дискурсом призвано отражать различия в ситуациях коммуникации. Так, институциональное общение отличается следующими свойствами: (1) участники коммуникации ориентированы на достижение целей, которые навязаны им институционально обусловленными идентичностями (такими как «продавец» и «покупатель», целью взаимодействия которых является акт купли-продажи); (2) взаимодействие подразумевает особые ограничения на формы участия (к примеру, «продавцы» не могут ничего требовать от «покупателей»); (3) взаимодействие связано с нормами и процедурами, специфическими для конкретных институциональных контекстов, см. (Heritage & Clayman 2010: 34–35).

Продавец-консультант выстраивает диалогическое взаимодействие как совместную деятельность, направленную на удовлетворение потребностей клиента. Эта деятельность имеет цель — продать товар или услугу. Речь консультанта — по преимуществу предложение товара. Соответственно, пропозитив как специализированная конструкция, выражающая 'предложение', играет в этом типе дискурса особую роль.<sup>10</sup>

### 3.6. Сложные последовательности реплик

В обследованных материалах корпусов встречаются разнообразные микродиалоги, состоящие из более сложных последовательностей, чем «стимульная реплика + реактивная реплика». К таким последовательностям относятся, в частности, приведённые ниже примеры, в которых стимульные реплики сочетают каузацию с запросом информации. Встречаются и употребления пропозитива, являющегося запросом на подтверждение некоего порядка действий (или порядка разговора), иногда после пропозитива употребляется «встречный» пропозитив, ср.:

[Евгений, муж] *Если у вас поступлений не было с шестнадцатого числѣ. А именно на какую сумму пополняли/ дайте посмотрим.*

<sup>9</sup> Например, [Телефонный разговор с менеджером турагентства (2015)], [Заказ по телефону билетов на концерт (2015)], [Заказ цветов по телефону (2015)], [Звонок в службу поддержки сотового оператора (2015)], [Разговор с менеджером фирмы по установке пластиковых окон (2010)], [Заказ фотосессии в фотосалоне (2015)], [Звонок в фитнес-клуб (2015)], [Разговор по телефону с консультантом магазина косметики (2015)], [Разговор по телефону с администратором курсов иностранных языков (2010)], [Заказ еды на дом (2015)] и др.

<sup>10</sup> Например, в пособии «Как продать слона» (Барышева 1999) конструкции с *давайте* многократно приводятся в качестве примера стандартных фраз, которые нужно использовать в общении с покупателем. Ср.: «*Давайте посмотрим, у вашего стеллажа будут такие же полочки, высота будет на 45 сантиметров меньше <...> Конечно, трудно с помощью небольшой детали представить себе, каким будет весь стеллаж, но давайте попробуем*», «*Давайте вместе подключим это устройство и посмотрим, как оно работает*», «*Давайте мы с вами посмотрим мебель, а я подержу вашу собачку на руках*».

[Лилия, жен] *á вóсьмь/ девятьсóт пять...* [НКРЯ, Звонок в службу поддержки сотового оператора (2015)]

[Екатерина, жен] *Потóм/ без чегó опя́ть же не мо́жет обойт́сь кáждая жéнщина/ **дава́йте оставимся... останóвимся на пома́де.** Аа вы каки́е текстúры бо́льше любите/ ма́товые либо с перламúтром?*

[Вероника, жен] ***Дава́йте попрóбуем ма́товые/** потому́ что с перламúтром у меня́ аа мно́го вся́ких таки́х блéсков валя́ется и прóчего/ но вóт с ма́товой я́ ещё не прóбовала.*

[Екатерина, жен] *Та́к/ хорошó/ отл́чно. Как рáз ма́товый щáс цвёт в трéнде.* [НКРЯ, Разговор по телефону с консультантом магазина косметики (2015)]

Можно выделить ряд разновидностей сложных последовательностей реплик (или сложных микродиалогов). Этот вопрос заслуживает отдельного внимания в том числе потому, что внутри сложных микродиалогов действуют свои правила обмена репликами. Более того, по ряду примеров из НКРЯ видно, что консультанты в общении с покупателями сознательно используют набор стандартных речевых ходов (в частности, сочетание пропозитива с альтернативным вопросом типа *Давайте посмотрим, Вам больше нравится X или Y?*).

#### 4. Заключение

Использование конверсационно-ориентированного подхода к анализу значения императивных реплик (опора на диалогическое окружение реплики и особенности протекания диалога) в сочетании с информацией о соотношении статусов участников коммуникации позволило выделить четыре семантических подтипа императива IP, а именно: пропозитив, гортатив, фактитивный пермиссив, пермиссив. Между тем, обследование выборок двух речевых корпусов показало существенную разницу во встречаемости выделенных подтипов. Эта разница может быть объяснена тем, что в сравнимых по объёму выборках содержатся разные доли бытовых и институциональных разговоров. Так, корпус ОРД содержит преимущественно бытовые, повседневные диалоги. Выборка из Акцентологического корпуса НКРЯ включает прежде всего материалы институционального дискурса — диалоги по телефону между покупателями (клиентами различных служб) и консультантами (продавцами, менеджерами). В диалогах продавец — покупатель особенно активно употребляется форма пропозитива.

Результаты настоящего исследования будут использованы при формировании на материале ОРД озвученной базы данных, включающей микродиалоги с употреблением различных лично-числовых форм императива. При разметке реплик внутри базы данных будут учитываться три типа информации: во-первых, информация о внутренних свойствах реплик, входящих в микродиалоги, их функциях в диалоге; во-вторых, информация о диалогическом контексте (соседних репликах, длине пауз между репликами, перебивах, наложениях речи); в-третьих, информация о сходстве между стимульными и реактивными



репликами. Представленная статья позволила уточнить формат разметки, установить инвентарь частных семантических интерпретаций 1PI.

Рассмотрение микродиалогов с императивом на материале большого речевого корпуса (сейчас в ОРД 1 млн слов в расшифровках) позволит, в частности, произвести анализ и классификацию типов вербальных реакций на императив, установить связи между определёнными типами вербальных реакций и свойствами инициирующих реплик, то есть существенно уточнить наши представления о функционировании императива в устном диалогическом дискурсе.

## Литература

1. Баранов А. Н., Крейдлин Г. Е. (1992а), Языковое взаимодействие в диалоге и понятие иллокутивного вынуждения, в Вопросы языкознания, № 2, сс. 84–99.
2. Баранов А. Н., Крейдлин Г. Е. (1992б), Структура диалогического текста: лексические показатели минимальных диалогов, в Вопросы языкознания, № 3, сс. 84–93.
3. Барышева А. В. (1999), Как продать слона, или 51 прием заключения сделки, Изд-во УРАО, Москва.
4. Бирюлин Л. А. (1992), Теоретические аспекты семантико-прагматического описания императивных высказываний в русском языке, дисс. ... докт. филол. наук, на правах рукописи, Ленинград.
5. Блинова О. В. (2016), Императивные конструкции 1PI в русском устном диалоге: возможности интерпретации с опорой на реактивные реплики, в Компьютерная лингвистика и интеллектуальные технологии: по материалам международной конференции «Диалог 2016» (Москва, 1–4 июня 2016), <http://www.dialog-21.ru/media/3446/blinovaov.pdf>.
6. Вежибицка А. (1985), Речевые акты, в Новое в зарубежной лингвистике, Вып. 16, сс. 251–275
7. Гусев В. Ю. (2005), Типология специализированных глагольных форм императива, дисс. ... канд. филол. наук, на правах рукописи, Московский государственный университет, Москва.
8. Гусев В. Ю. (2008), Пропозитивные конструкции с давай в русском языке, в Русская филология 19: Сборник научных работ молодых филологов, Тартуский университет, Тарту, сс. 189–193.
9. Кодзасов С. В. (2009), Исследования в области русской просодии, Языки славянских культур, Москва.
10. Кодзасов С. В., Архипов А. В., Бонч-Осмоловская А. А., Захаров Л. М., Кривнова О. Ф. (2006), База данных «интонация русского диалога»: побудительные реплики, в Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2006», Бекасово, сс. 236–242.
11. Подлеская В. И. (2005), Русские глаголы дать/давать: от прямых употреблений к грамматикализованным, в Вопросы языкознания, № 2, сс. 89–103.
12. Теория функциональной грамматики (ТФГ). Темпоральность. Модальность (1990), Наука, Ленинград.

13. Храковский В. С., Володин А. П. (1986), Семантика и типология императива: Русский императив, Наука, Ленинград.
14. Шатуновский И. Б. (2009), Проблемы русского вида, Языки славянских культур, Москва.
15. Янко Т. Е. (2008), Интонационные стратегии русской речи в сопоставительном аспекте, Языки славянских культур, Москва.

## References

1. Aikhenvald, A. Yu. (2010), Imperatives and Commands. Oxford University Press, New York.
2. A theory of Functional Grammar. Temporality. Modality [Teoriya funkcional'noj grammatiki. Temporal'nost'. Modal'nost'] (1990), Science [Nauka], Leningrad.
3. Baranov A. N., Krejdlin G. E. (1992a), Language interaction in a dialogue and the notion of illocutionary constraining [Jazykovoje vzaimodejstviye v dialoge i ponyatije illokutivnogo vynuždenija], in Issues in linguistics [Voprosy jazykoznanija], № 2, pp. 84–99.
4. Baranov A. N., Krejdlin G. E. (1992b), Structure of a text of a dialogue: lexical markers of minimal dialogues [Struktura dialogičeskogo teksta: leksičeskije pokazateli minimalnyh dialogov], in Issues in linguistics [Voprosy jazykoznanija], № 3, pp. 84–93.
5. Barysheva A. V. (1999), How to sell an elephant, or 51 methods to conclude a bargain, URAE Press, Moscow.
6. Birjulin L. A. (1992), Theoretical aspects of semantic and pragmatic description of the imperative utterances in Russian, D. Sc. dissertation, Leningrad.
7. Blinova O. V. (2016), Imperative Constructions 1Pl in Russian Spoken Dialogue: Possibilities for Interpretation Based on Reactive Utterances [Imperativnye konstrukcii 1Pl v russkom ustnom dialoge: vozmožnosti interpretacii s oporoj na reaktivnye repliki], in Computational Linguistics and Intellectual Technologies: Papers from the International Conference Dialogue-2016 (Moscow, June 1–4 2016), <http://www.dialog-21.ru/media/3446/blinovaov.pdf>
8. Bogdanova-Beglarian N., Sherstinova T., Blinova O., Ermolova O., Baeva E., Martynenko G., Ryko A. (2016), Sociolinguistic Extension of the ORD Corpus of Russian Everyday Speech, in Ronzhin, A. et al. (eds.) SPECOM 2016, Lecture Notes in Computer Science, LNCS 2016. Vol. 9811, Springer, Switzerland, pp. 659–666.
9. Goussev V. Ju. (2005), A typology of specialized imperative verb forms [Tipologija specializirovannykh glagol'nykh form imperativa], unpublished PhD dissertation, Moscow.
10. Goussev V. Ju. (2008), Propositional constructions with davaj in Russian [Propozitivnye konstrukcii s davaj v russkom yazyke], in Russian Philology 19: Collection of works by young philologists [Sbornik nauchnyx rabot molodyx filologov], Tartu University Press, Tartu.
11. Heritage J., Clayman S. (2010), Talk in Action. Interactions, Identities, and Institutions. Wiley-Blackwell, Chichester.

12. *Jurafsky D.* (2006), *Pragmatics and Computational Linguistics*, in Horn L., Ward G. (eds.) *The handbook of pragmatics*, Blackwell Publishing, Oxford, pp. 578–604.
13. *Kodzasov S. V.* (2009), *Studies in Russian prosody [Issledovaniia v oblasti russkoi prosodii]*, *Yazyki slavyanskih kul'tur*, Moscow.
14. *Kodzasov S. V., Arkhipov A. V., Bonch-Osmolovskaia A. A., Zakharov L. M., Krivnova O. F.* (2006), Database “Intonation of Russian Dialogue”: imperative utterances [Baza danykh “Intonatsiya russkogo dialoga”: pobuditel'nye repliki], in *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2006” [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2006”]*, Bekasovo, pp. 236–242.
15. *Podlesskaya V. I.* (2005), Russian verbs *dat'/davat'*: from the direct to grammaticalized usages [Russkie glagoly *dat'/davat'*: ot prjamyh upotreblenij k grammatikalizovannym], in *Issues in Linguistics [Voprosy yazykoznaniya]*, № 2, pp. 89–103.
16. *Stolcke A. et al.* (2000), Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech, in *Computational Linguistics*, 26(3), pp. 339–373.
17. *Wierzbicka A.* (1985), Speech acts, in *Novelty in foreign linguistics*, Iss. 16, pp. 251–275.
18. *Xrakovskij V. S., Volodin A. P.* (1986), Semantics and typology of imperative: Russian imperative [Semantika i tipologija imperativa: Russkij imperativ], Science [Nauka], Leningrad.
19. *Shatunovskij I. B.* (2009), Problems of the Russian aspect [Problemy russkogo vida], *Yazyki slavyanskih kul'tur*, Moscow.
20. *Yanko T.* (2008), Intonational strategies of the Russian speech from a contrastive perspective [Intonatsionnye strategii russkoj rechi v sopostavitel'nom aspekte]. *Yazyki slavyanskih kul'tur*, Moscow.

# НЕКОТОРЫЕ ИНВАРИАНТНЫЕ ХАРАКТЕРИСТИКИ РУССКОЙ РАЗГОВОРНОЙ РЕЧИ: ФОНЕТИКА, МОРФОЛОГИЯ, СИНТАКСИС<sup>1</sup>

**Богданова-Бегларян Н. В.** (n.bogdanova@spbu.ru),

**Блинова О. В.** (o.blinova@spbu.ru),

**Мартыненко Г. Я.** (g.martynenko@spbu.ru),

**Шерстинова Т. Ю.** (t.sherstinova@spbu.ru)

Филологический факультет СПбГУ, Санкт-Петербург, Россия

Исследование осуществлено на материале звукового корпуса ОРД в рамках проекта, посвященного рассмотрению социолингвистической вариативности русской разговорной речи, с целью выявления диагностических признаков, характеризующих речь разных социальных групп. Результаты показали, что почти на каждом языковом уровне выявляются лингвистические параметры, в отношении которых все социолекты ведут себя одинаково: в частности, это наблюдается в дистрибуции фонем, частей речи, в частотности синтаксических структур. Распределение фонем определено на подкорпусе в 172000 аллофонов. Наиболее частотными в речи всех социальных групп являются фонемы /a/ (18,18%), /i/ (9,04%), /t/ (6,36%), /o/ (5,43%), /u/ (4,49%), /n/ (4,11%), /j/ (3,82%), /e/ (3,57%), /k/ (3,35%), /ы/ (3,01%). Распределение частей речи в повседневной бытовой речи определено на лингвистически аннотированном подкорпусе объемом в 125437 словоупотреблений и имеет следующее распределение: V (17,43%), S (15,29%), S-PRO (14,13%), PART (13,35%), CONJ (9,47%), PR (7,09%), ADV-PRO (5,30%), ADV (4,51%), A-PRO (4,30%) и др. На синтаксическом уровне наиболее частотными в речи всех социальных групп являются одноэлементные структуры: D (3,73%), S (2,26%), V (1,88%). Статистический анализ левосторонних и правосторонних зависимостей глагольных групп показал, что русский язык в его устной повседневной разновидности существенно левосторонний для всех социолектов. Представленные данные отражают некоторые постоянные, универсальные, свойства русского языка повседневного общения в целом и могут быть использованы для оптимизации систем распознавания и синтеза русской разговорной речи.

**Ключевые слова:** Русский язык, повседневная разговорная речь, речевой корпус, частотные списки, дистрибуция фонем, дистрибуция частей речей, синтаксические структуры

---

<sup>1</sup> Исследование проведено в рамках проекта «Русский язык повседневного общения: особенности функционирования в разных социальных группах», поддержанного грантом РНФ № 14-18-02070.

## SOME INVARIANT FEATURES OF RUSSIAN EVERYDAY SPEECH: PHONOLOGY, MORPHOLOGY, SYNTAX

**Bogdanova-Beglarian N. V.** (n.bogdanova@spbu.ru),  
**Blinova O. V.** (o.blinova@spbu.ru),  
**Martynenko G. Ya.** (g.martynenko@spbu.ru),  
**Sherstinova T. Yu.** (t.sherstinova@spbu.ru)

Philological Faculty of St. Petersburg State University,  
St. Petersburg, Russia

The presented research was carried out on the material of the ORD speech corpus in the framework of the project, dedicated to study sociolinguistic variation of Russian speech and aimed at identifying diagnostic features characterizing everyday speech of major social groups (age-, gender-, status-, professional-related, etc.). The obtained results showed that practically on each linguistic level one may observe the features exhibiting a very high similarity between different sociolects. In particular, the coincidence is observed in the distribution of phonemes, distribution of parts of speech, and the frequency of some syntactic structures. The distribution of phonemes was determined on the subcorpus of 172,000 allophones. The following ten phonemes are the most frequent in speech of all social groups: /a/ (18,18%), /i/ (9,04%), /t/ (6,36%), /o/ (5,43%), /u/ (4,49%), /n/ (4,11%), /j/ (3,82%), /e/ (3,57%), /k/ (3,35%), /s/ (3,01%). The distribution of parts of speech in everyday speech was obtained on the linguistically annotated subcorpus of 125,437 tokens and has the following breakdown: V (17,43%), S (15,29%), S-PRO (14,13%), PART (13,35%), CONJ (9,47%), PR (7,09%), ADV-PRO (5,30%), ADV (4,51%), A-PRO (4,30%), A (3,73%), PRAEDIC (1,84%), INTJ (1,41%), NUM (1,29%), PARENTH (0,56%), ANUM (0,27%), PRAEDIC-PRO (0,01%). At the syntactic level, one-element structures are prevailing in everyday speech of all social groups, the most frequent among them being D (particle / discursive word) (3,73%), S (2,26%), and V (1,88%). Statistical analysis of the left-branching and right-branching verb groups has showed that the first ones significantly prevail in speech of all sociolects. The revealed features reflect some constant, universal properties of everyday spoken Russian and can be used for adjustment and improvement of speech synthesis and recognition systems.

**Keywords:** Russian language, everyday spoken speech, speech corpus, frequency lists, distribution of phonemes, distribution of parts of speech, syntactic structures

### Введение

Корпус русской повседневной речи «Один речевой день» (ОРД) создан и продолжает пополняться в рамках большого проекта, направленного на решение ряда фундаментальных научных задач, таких как мониторинг и фиксация

звукового материала естественного языка, поддержка систем распознавания и синтеза речи, организация информационной среды и программного инструментария для нужд интегрального моделирования речи, создание источника новых учебных материалов как для изучения и преподавания языка в его звуковой форме, так и для реализации различных лексикографических и других прикладных проектов (см. подробнее: *Asinovsky et al.* 2009; *Богданова-Бегларян и др.* 2015).

На сегодняшний день корпус ОРД характеризуется следующими количественными показателями: около 1250 часов звучания, более 2800 коммуникативных макроэпизодов, 130 информантов (из них 69 мужчин и 61 женщина в возрасте от 18 до 83 лет), более 1000 основных коммуникантов. Объем текстовых расшифровок корпуса превышает 1 млн. словоупотреблений.

Основной целью, на решение которой направлено настоящее исследование, является описание современного русского языка повседневного общения и анализ особенностей его функционирования в разных социальных группах (см.: *Богданова-Бегларян и др.* 2016; *Bogdanova-Beglarian et al.* 2016 a, b). Одной из главных задач проекта является анализ материалов корпуса ОРД с целью выявления тех лингвистических параметров, по которым наиболее явно видны различия в речи говорящих, принадлежащих к различным социальным группам городского континуума. Решение данного вопроса видится как создание своеобразных *речевых портретов* городских социолектов.

Задача эта весьма нетривиальна, поскольку городская речь характеризуется несомненным *многоязычием*, а любому горожанину свойственен так называемый *полиглотизм* (термины Б. А. Ларина — *Ларин* 1977), т. е. в его речи соединяются не только черты, присущие разным социолектам (в широком смысле слова), но и разные языковые стили кодифицированного языка, на нее оказывает влияние множество внешних факторов, таких как коммуникативная ситуация, психологические характеристики всех участников разговора, их социальные роли по отношению друг к другу и собственно иерархия этих отношений, а также такие мелочи, как «внешняя обстановка: наличие или отсутствие постороннего шума, погодные условия (если разговор происходит не в помещении)» (*Фонетика спонтанной речи* 1988: 6) и др.

Несмотря на все эти допущения, мы все же предприняли попытку описать речевые портреты различных социальных групп современного города, сделать некий «*средний вывод* (курсив наш. — *Авт.*) из известного количества индивидуальных языков» (*Бодуэн де Куртенэ* 1917: 41) информантов и коммуникантов ОРД — насколько это удалось в ходе проведенного нами многоуровневого и многоаспектного анализа корпусного материала. В результате получились некоторые «штрихи к речевым портретам» разных социолектов.

Для проведения социолингвистического исследования был подготовлен расширенный исследовательский подкорпус: из общего объема транскриптов корпуса ОРД в объеме 1 млн. словоупотреблений было отобрано 100 макроэпизодов для 100 информантов, сбалансированно отражающих повседневную речь анализируемых социальных групп. Кроме речи основных 100 информантов в подкорпус была включена речь 154 коммуникантов.

В результате исследовательский подкорпус содержит речевые данные для представителей 20-ти анализируемых социальных групп: а) 2-х гендерных (мужчины и женщины), б) 3-х возрастных (молодежная, средняя и старшая), в) 10-ти профессиональных (рабочие, инженеры, военнослужащие, представители естественных наук, представители гуманитарных наук, работники образования, представители сферы обслуживания, IT-специалисты, офисные служащие, творческая интеллигенция) и г) 5-ти статусных (студенты и учащиеся; наемные работники и специалисты; руководящие работники; бизнесмены и частные предприниматели; неработающие, в том числе пенсионеры).

Объем лингвистически проаннотированного речевого материала составляет 125 437 словоформ на лексическом и морфологическом уровнях, 12 020 структур на синтаксическом уровне и 172 053 аллофонов на фонетическом уровне. Методика исследования была отработана ранее на материале пилотного подкорпуса меньшего объема (10 259 словоформ), результаты пилотного анализа представлены в коллективной монографии (*Русский язык повседневного общения...* 2016).

Результаты исследования показали, что почти на каждом лингвистическом уровне, помимо диагностических<sup>2</sup> и потенциально диагностических лингвистических признаков, выявляются такие, **в отношении которых все социолекты ведут себя примерно одинаково**. Эти признаки можно считать инвариантными, то есть свойственными всем говорящим по-русски и не зависящими от социальных характеристик носителя языка. В данной статье описаны такие инвариантные признаки повседневной разговорной речи, как **распределение частотных фонем, распределение частей речи**, а также некоторые **синтаксические особенности**, свойственные русской разговорной речи в целом.

Большинство примеров в работе извлекаются из речевого материала хорошо представленных социальных групп — гендерных (мужчины и женщины) и возрастных (младшие, средние, старшие говорящие), — проаннотированные массивы речевого материала которых в корпусе ОРД наиболее объемны.

## 1. Распределение частотных фонем

Статистические данные о дистрибуции фонем были получены на материале подвыборки в 172 053 аллофона, содержащей фрагменты речи всех исследуемых социальных групп.

Фонетическая транскрипция была получена автоматически посредством программного обеспечения производства ООО «Центр Речевых Технологий»<sup>3</sup>, согласно заложенным в программе транскрибирования правилам преобразования письменного текста в последовательность аллофонов. При транскрибировании используется словарь исключений, куда экспертом вручную вносятся

---

<sup>2</sup> С точки зрения различения социолектов.

<sup>3</sup> <http://www.speechpro.ru/>

типичные редуцированные формы, возможные иные варианты произнесения слов, а также список типовых исключений (в частности, для разных форм редукции). Основная часть транскрипции выполняется по письменному тексту; для принятия решения о реализации слова в том случае, когда возможны варианты произнесения, программа обращается к звукозаписи, автоматически принимает решение, какой из вариантов является в данном случае наиболее вероятным, и транскрибирует слово в соответствии с этим решением.

Для транскрибирования используется следующий список аллофонов: [a0], [a1], [a2], [a4], [o0], [o1], [o4], [e0], [e1], [ы0], [ы1], [ы4], [u0], [u1], [u4], [i0], [i1], [i4], [b], [b'], [p], [p'], [d], [d'], [t], [t'], [g], [g'], [k], [k'], [c], [ch], [v], [v'], [f], [f'], [z], [z'], [s], [s'], [zh], [sh], [sc], [h], [h'], [m], [m'], [n], [n'], [l], [l'], [r], [r'], [j]. Для гласных звуков 0 обозначает ударный слог, 1 — предударный, 2 — второй предударный (только для /a/), 4 — заударный.

Результаты транскрибирования и одновременного автосегментирования речевого потока на аллофоны визуализируются непосредственно в среде многоуровневого лингвистического аннотирования ELAN<sup>4</sup>, в которой осуществляется разметка корпуса ОРД (см. рис. 1). Это позволяет провести их экспертную оценку и коррекцию по мере необходимости.

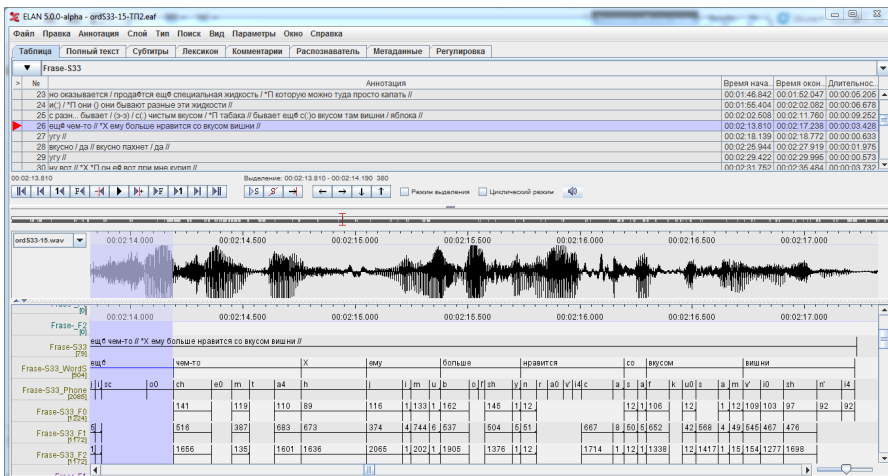


Рис. 1. Пример фонетического аннотирования корпуса ОРД (сегментация на слова и аллофоны, фонетическое транскрибирование, ЧОТ, значения первых двух формант)

Проведенное исследование показало, что наиболее частотными в речи всех социальных групп являются фонемы /a/ (18,18%), /i/ (9,04%), /t/ (6,36%), /o/ (5,43%), /u/ (4,49%), /n/ (4,11%), /j/ (3,82%), /e/ (3,57%), /k/ (3,35%), /s/ (3,01%). Более того, как минимум, до 10-го ранга не наблюдается абсолютно

<sup>4</sup> <https://tla.mpi.nl/tools/tla-tools/elan/>



никакого различия в рангах отдельных фонем между разными социальными группами (см. табл. 1).

**Таблица 1** Верхняя зона частотного списка фонем (данные по всей выборке, мужчины и женщины, три возрастных группы)

Ранг	Всего		Гендерные группы				Возрастные группы					
	Всего	%	Муж.	%	Жен.	%	Млад.	%	Сред.	%	Стар.	%
1	a	18,18	a	17,80	a	18,43	a	18,44	a	17,86	a	17,91
2	i	9,04	i	9,02	i	9,05	i	8,76	i	9,21	i	9,70
3	t	6,36	t	6,48	t	6,27	t	6,13	t	6,58	t	6,58
4	o	5,43	o	5,61	o	5,31	o	5,26	o	5,60	o	5,48
5	u	4,50	u	4,34	u	4,60	u	4,44	u	4,61	u	4,25
6	n	4,11	n	4,15	n	4,08	n	4,33	n	4,01	n	3,94
7	j	3,82	j	3,83	j	3,81	j	3,88	j	3,82	j	3,62
8	e	3,57	e	3,50	e	3,61	e	3,62	e	3,46	e	3,62
9	k	3,35	k	3,30	k	3,38	k	3,32	k	3,33	k	3,59
10	y	3,01	y	3,05	y	2,99	y	3,04	y	2,96	y	3,15

Приведем для сравнения верхнюю зону рангового *распределения аллофонов*. Здесь для всех социальных групп на первом месте ожидаемо находится ударный [a0] (7%), на втором — [t] (6%), аллофоны [o0], [a1] и [a4] занимают позицию с третьего по пятый ранг, [n] — стабильно 6-е место в ранжированном списке, замыкают десятку аллофоны [j], [i4], [e0], [i1] и [k] (см. табл. 2).

**Таблица 2.** Верхняя зона частотного списка аллофонов (данные по всей выборке, мужчины и женщины, три возрастных группы)

Ранг	Всего		Гендерные группы				Возрастные группы					
	Всего	%	Муж.	%	Жен.	%	Млад.	%	Средн.	%	Старш.	%
1	a0	6,91	a0	6,81	a0	6,97	a0	7,10	a0	6,59	a0	6,73
2	t	6,36	t	6,49	t	6,27	t	6,13	t	6,58	t	6,58
3	o0	5,36	o0	5,54	a1	5,40	a1	5,29	o0	5,53	o0	5,40
4	a1	5,28	a1	5,12	o0	5,23	o0	5,19	a1	5,30	a1	5,28
5	a4	4,55	a4	4,55	a4	4,56	a4	4,61	a4	4,56	a4	4,38
6	n	4,11	n	4,15	n	4,08	n	4,33	n	4,01	n	3,94
7	j	3,82	j	3,83	j	3,82	j	3,88	i4	3,88	i4	3,91
8	i4	3,61	i4	3,73	e0	3,61	e0	3,62	j	3,82	j	3,62
9	e0	3,57	e0	3,50	i1	3,56	i1	3,62	i1	3,47	e0	3,62
10	i1	3,52	i1	3,47	i4	3,54	k	3,32	e0	3,46	k	3,59

Таким образом, в распределении отдельных аллофонов наблюдается несколько меньшая согласованность, чем в распределении частотных фонем. Однако можно предположить, что при увеличении объема исследуемой выборки ранговые последовательности аллофонов для разных социальных групп будут согласованы в большей степени.

## 2. Распределение частей речи

Распределение частей речи получено на лингвистически аннотированном подкорпусе ОРД. Объем исследовательского материала в словах составляет 125 437 словоупотреблений, из них 47 135 принадлежат речи мужчин, 78 302 — речи женщин. Возрастные группы представлены рабочими подкорпусами соответственно в 46 328 слов (младшая возрастная группа), 51 431 слов (средняя возрастная группа), 23 475 слов (старшая возрастная группа). Кроме того, в подкорпусе представлена речь коммуникантов-детей (4203 слова).

Для частеречной разметки использовалась программа-морфоанализатор TreeTagger<sup>5</sup>. Программа использует следующий набор помет: V (глагол); S (существительное); S-PRO (местоимение-существительное); PART (частица); CONJ (союз); PR (предлог); ADV-PRO (местоимение-наречие); ADV (наречие); A-PRO (местоимение-прилагательное); A (прилагательное); PRAEDIC (предикатив); INTJ (междометие); NUM (числительное); PARENTH (вводное слово); ANUM (числительное-прилагательное); PRAEDIC-PRO (местоимение-предикатив).

Проведен автоматический морфологический анализ всего отобранного для исследования речевого материала и осуществлена его ручная коррекция. Результаты распределения частей речи в аннотированном подкорпусе представлены в табл. 3.

**Таблица 3.** Распределение частей речи (данные по всей выборке, мужчины и женщины, три возрастных группы)

Ранг	Всего		Гендерные группы				Возрастные группы					
	Всего	%	Муж.	%	Жен.	%	Млад.	%	Средн.	%	Старш.	%
1	V	17,43	V	17,44	V	17,42	V	17,21	V	17,65	V	17,33
2	S	15,29	S	15,68	S	15,06	S	14,86	S	15,56	S	15,30
3	S-PRO	14,13	PART*	13,40	S-PRO*	14,63	S-PRO	14,35	S-PRO	13,75	S-PRO	14,21
4	PART	13,35	S-PRO*	13,31	PART*	13,32	PART	13,80	PART	12,79	PART	13,83
5	CONJ	9,47	CONJ	9,19	CONJ	9,64	CONJ	9,31	CONJ	9,31	CONJ	10,24
6	PR	7,09	PR	7,25	PR	7,00	PR	6,99	PR	7,19	PR	7,17
7	ADV-PRO	5,30	ADV-PRO	5,68	ADV-PRO	5,08	ADV-PRO	5,32	ADV-PRO	5,52	ADV-PRO	4,88
8	ADV	4,51	ADV	4,41	ADV	4,57	ADV*	4,64	ADV*	4,53	A-PRO*	4,52
9	A-PRO	4,30	A-PRO	4,29	A-PRO	4,31	A-PRO*	4,14	A-PRO*	4,34	ADV*	4,42
10	A	3,73	A	3,92	A	3,61	A	3,72	A	3,93	A	3,41
11	PRAEDIC	1,84	PRAEDIC	1,89	PRAEDIC	1,82	PRAEDIC	1,77	PRAEDIC	1,95	PRAEDIC	1,73
12	INTJ	1,41	NUM*	1,44	INTJ*	1,55	NUM*	1,58	INTJ*	1,37	INTJ*	1,25
13	NUM	1,29	INTJ*	1,17	NUM*	1,20	INTJ*	1,47	NUM*	1,23	NUM*	0,88
14	PARENTH	0,56	PARENTH	0,60	PARENTH	0,54	PARENTH	0,56	PARENTH	0,57	PARENTH	0,61
15	ANUM	0,27	ANUM	0,31	ANUM	0,25	ANUM	0,26	ANUM	0,30	ANUM	0,20
16	PRAEDIC-PRO	0,01	PRAEDIC-PRO	0,01	PRAEDIC-PRO	0,02	PRAEDIC-PRO	0,02	PRAEDIC-PRO	0,01	PRAEDIC-PRO	0,01

<sup>5</sup> <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

Из табл. 3 видно, что самые распространенные части речи — это глагол (17,43%), существительное (15,29%), местоимение-существительное (14,13%), частица (13,35%) и союз (9,47%). В целом по подкорпусу части речи имеют следующие ранги: глагол (1), существительное (2), местоимение-существительное (3), частица (4), союз (5), предлог (6), местоимение-наречие (7), наречие (8), местоимение-прилагательное (9), прилагательное (10), предикатив (11), междометие (12), числительное (13), вводное слово (14), числительное-прилагательное (15), местоимение-предикатив (16). Наблюдается почти полное совпадение ранговых последовательностей для рассмотренных 5 социальных групп.

На фоне стабильного единообразия в дистрибуции частей речи обращают на себя внимание некоторые незначительные колебания. Они относятся к гендерным группам и отмечены в таблице символом (\*). Во-первых, это высокочастотная пара **S-PRO** (местоимения-существительные) и **PART** (частицы): первые чаще используют женщины, вторые — мужчины. Во-вторых, это пара менее частотных классов слов: **NUM** (числительные) и **INTJ** (междометия), имеющие 12–13 ранг (числительные более характерны для мужской речи, а междометия — для женской).

Для проверки статистической значимости различий между употреблением данных частеречных классов в речи мужчин и женщин был проведен тест с использованием стандартного критерия Стьюдента для оценки разностей выборочных долей (*Математические методы...* 1978: 181). Результаты оказались таковы:

- для частиц (**PART**) наблюдаемая разность между долями оказалась недоуверенной при всех уровнях значимости, т. е. на данном материале нельзя утверждать, что мужчины используют частицы чаще, чем женщины;
- для местоимений-существительных (**S-PRO**) разность достоверна при 5%-ом уровне значимости, но недостоверна при более высоких уровнях. Поскольку в лингвистических исследованиях 5%-ый уровень значимости считается общепринятым, можно считать, что разность достоверна, т. е. женщины в повседневной речи употребляют личные местоимения чаще, чем мужчины. Эти данные согласуются и с результатами, полученными в (*Шерстинова 2016а*);
- для двух других классов слов — числительных (**NUM**) и междометий (**INTJ**) — разность между долями оказалась существенной при всех (5-, 1- и 0,1%-ом) уровнях значимости. Это означает, что мужчины действительно чаще используют в своей речи числительные, а женщины — междометия.

Ранговое распределение частей речи у представителей трёх рассматриваемых возрастных групп также практически идентично. Различия наблюдаются в рангах наречия (**ADV**) и местоимения-прилагательного (**A-PRO**): у младших говорящих и говорящих среднего возраста наречие и местоимение-прилагательное имеют 8 и 9 ранг соответственно. У старших говорящих картина обратная. При этом относительные количества употреблений упомянутых частей речи в процентах отличаются незначительно — на десятки доли процента.

Интересно, что материалы речи попавших в выборку детей (которые были собеседниками, коммуникантами информантов) и в отношении относительных цифр, отражающих употребительность слов различных частеречных классов, и в отношении их рангового распределения не показывают существенных отличий от речи взрослых говорящих.

### 3. Синтаксические особенности

Синтаксическое аннотирование корпуса ОРД выполнено вручную для подвыборки объемом в 13 208 элементарных предложений (клауз). Выполнялась разметка структуры глагольной группы, при этом высчитывалось количество левых и правых зависимых от вершины членов, анализировалась структура именной группы, маркировались некоторые особенности разговорного синтаксиса.

Варианты разнообразных синтаксических структур оказались весьма многочисленны, что затрудняет их формальный анализ. Наиболее частотными в речи всех социальных групп оказались одноэлементные структуры: D (частица/дискурсивное слово) (3,73%) — *вот //, так //, да //, ладно //* и т. п., S (2,26%) — *понедельник //, завал //, супчик?* и др., V (1,88%) — *звони! ушел //, поехали //* и др., а также простые нераспространенные фразы структуры SV (1,07%) — *я сваливаю //, дождь кончился //, ты сядь!* или группы дискурсивных слов {D} (1,95%) — *вот так //, ну вот //, да да да //* и др.

Важный результат был получен при исследовании линейризованных глагольных групп с точки зрения их левой и правой ширины (Фитиалов 1968) относительно вершины в структуре зависимостей. Так, каждая глагольная группа была представлена в виде кода с символом-разделителем «двоеточие», слева от которого указывается число ее левых членов (расположенных перед глаголом), а справа от него — число правых членов. При этом рассматриваются лишь количественные характеристики, при полном отвлечении от качественного состава зависимых членов в разных позициях. Ниже приведено несколько примеров таких конструкций.

2:0	<p>1 2 V Я тебя спрашиваю!</p>
1:1	<p>1 V 1 А вчера прогревали батареи.</p>
1:3	<p>1 V 1 2 3 Я пойду к зубному сегодня в четыре часа.</p>
6:0	<p>1 2 3 4 5 6 V ... потому что тут уже частично часть розеток у меня моими руками уже переделана.</p>

**Рис. 2.** Примеры линейризованных глагольных групп с левой и правой шириной

В аннотированном подкорпусе объемом в 13 208 клауз число глагольных групп составило 7620 единиц (57,69%). В таблице 4 приведено ранговое распределение типов глагольных групп с точки зрения их левой и правой ширины.

**Таблица 4.** Ранговое распределение глагольных групп с точки зрения их левой и правой ширины

Ранг	Конструкция	Кол-во	%	Ранг	Конструкция	Кол-во	%
1	1:0	1919	25,18	14	0:3	22	0,29
2	1:1	1237	16,23	15	3:2	15	0,20
3	2:0	1096	14,38	16	4:1	10	0,13
4	0:1	1084	14,23	17	0:4	8	0,11
5	0:0	729	9,57	18–19	5:0	7	0,09
6	2:1	518	6,80		2:3	7	0,09
7	3:0	352	4,62	20–21	6:0	3	0,039
8	0:2	208	2,73		1:4	3	0,039
9	1:2	168	2,21	22	3:3	2	0,026
10	3:1	98	1,29	23–25	5:1	1	0,013
11	4:0	57	0,75		5:2	1	0,013
12	2:2	51	0,67		2:4	1	0,013
13	1:3	23	0,30				

На основании этих данных могут быть построены распределения левосторонних и правосторонних членов, см. табл. 5. В таблице не принимается во внимание структура вида 0:0 (строка 5 табл. 3), так как в противном случае это приведет к искажению симметричных отношений.

**Таблица 5.** Распределение левых и правых зависимых членов глагольных групп

Левая ширина	Количество	Правая ширина	Количество
0	1322	0	3434
1	3350	1	2948
2	1673	2	443
3	467	3	54
4	67	4	12
5	9	5	–
6	3	6	–
Σ	6891	Σ	6891

Оба распределения существенно разнородны с точки зрения критерия  $\chi^2$ , а средние столь различны, что не нуждаются в проверке с помощью какого-либо критерия значимости. Это означает, что повседневная устная речь с точки зрения соотношения левой и правой ширины является существенно левосторонней, что подтверждает вывод, сделанный нами ранее на более ограниченном материале (Мартыненко 2015; *Русский язык повседневного общения...* 2016: 127–129).

Далее, с помощью критерия  $\chi^2$  было показано, что распределение левых и правых зависимостей в речи мужчин и женщин, а также в речи разных возрастных групп однородно. Кроме того, с помощью критерия Стьюдента было

установлено, что между соответствующими средними величинами в указанных группах различия статистически незначительны при самых жестких критериях значимости.

В заключение раздела приведем данные о **синтаксических нерегулярностях** повседневной речи, выделяемых на материале корпуса ОРД. При выполнении синтаксического аннотирования отмечались следующие особенности:

- 1) обрыв фразы (CUT):
  - *я еле люстру ...*,
  - *а утром девятого...*;
- 2) эллипсис (EL):
  - *он там со своим животиком //*
  - *он ещё натяжные потолки //*
  - *Оксан% / я вот вам хочу //*;
- 3) парцелляция (PARC), установленная на основе разметки, в которой фиксировались, в частности, длительные и синтаксически не мотивированные паузы:
  - *дырка же там вот такая // <пауза> большая //*
  - *не помню ну ладно // <пауза> потом вспомню // <пауза> скажу*
  - *тут у нас снег лежит // <пауза> прямо вообще //*;
- 4) самокоррекция говорящего (COR):
  - *такие глу... (...) ужасы рассказывают,*
  - *они (...) им пришлось приехать сюда.*

В табл. 5 представлены данные о распределении синтаксических особенностей устной речи, обнаруженных на материале аннотированного подкорпуса в объеме 13 020 синтаксических структур.

**Таблица 5.** Распределение синтаксических нерегулярностей в русской разговорной речи (данные по всей выборке, мужчины и женщины, три возрастных группы)

Ранг	Всего		Гендерные группы				Возрастные группы					
	Всего	%	Муж.	%	Жен.	%	Млад.	%	Средн.	%	Старш.	%
1	CUT	3,66	CUT	3,38	CUT	4,06	CUT	2,98	CUT	3,72	CUT	4,45
2	EL	1,80	EL	1,82	EL	1,79	EL	1,84	EL	1,58	EL	2,13
3	COR	1,34	COR	1,36	COR	1,32	COR	1,08	COR	1,41	COR	1,59
4	PARC	0,48	PARC	0,48	PARC	0,48	PARC	0,39	PARC	0,50	PARC	0,60

Из таблицы 5 видно, что в речи всех социальных групп чаще всего встречаются обрывы (незавершенные фразы), эллипсис занимает второе место по частоте встречаемости, самокоррекция — третье, реже всего используется парцелляция. Ранговые порядки отдельных синтаксических нерегулярностей инвариантны для всех рассмотренных социолектов. Просматривается общая тенденция к увеличению доли синтаксических нерегулярностей с возрастом говорящих, однако ее статистическая состоятельность требует специальной проверки.

#### 4. Заключение

Описанные в работе данные отражают некоторые постоянные, универсальные, свойства русского языка повседневного общения в целом. Они представляют значительный теоретический интерес, а также могут быть использованы для оптимизации систем распознавания и синтеза русской разговорной речи (Шерстинова 2016 б). Принципиально важным представляется заключение о том, что разговорная речь, несмотря на ее спонтанность и диффузность, подчиняется достаточно жестким законам языка.

В данной работе не был рассмотрен большой пласт «потенциально инвариантных» характеристик русской разговорной речи, которые показали на проанализированном речевом материале некоторые количественные различия между рассматриваемыми группами, поскольку степень этих различий не позволяет на настоящий момент считать их статистически достоверными. К ним относится ряд характеристик лексического и лексико-дискурсивного уровней, а также другие фонетические и синтаксические особенности русской устной речи (см.: *Русский язык повседневного общения...* 2016). В отношении этих параметров необходимо проведение исследования на более представительном объеме проанотированного речевого материала. Не исключено, что по результатам будущих исследований число инвариантных свойств русской разговорной речи будет расширено.

#### Литература

1. Богданова-Бегларян Н. В., Асиновский А. С., Блинова О. В., Маркасова Е. В., Рыко А. И., Шерстинова Т. Ю. (2015), Звуковой корпус русского языка: новая методология анализа устной речи, *Язык и метод: Русский язык в лингвистических исследованиях XXI века. Вып. 2 / Ред. Д. Шумска, К. Озга. Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków*, сс. 357–372.
2. Богданова-Бегларян Н. В., Блинова О. В., Мартыненко Г. Я., Шерстинова Т. Ю. (2016), Русский язык повседневного общения: некоторые количественные данные в зеркале социолингвистики // *Коммуникативные исследования*, № 2 (8), сс. 81–92.
3. Бодуэн де Куртене И. А. (1917), *Введение в языковедение*. Изд. 5, Петроград.
4. Ларин Б. А. (1977), *К лингвистической характеристике города (несколько предпосылок)*, Б. А. Ларин. *История русского языка и общее языкознание*, Просвещение, Москва, сс. 189–199.
5. Мартыненко Г. Я. (2015), Синтаксис живой спонтанной речи: симметрия линейных порядков, *Корпусная лингвистика-2015. Труды международной конференции / Ред. В. П. Захаров, О. А. Митрофанова, М. В. Хохлова, Санкт-Петербург*, сс. 371–378.
6. *Математические методы в биологии. Учебное пособие (1978)*, под ред. Н. А. Плохинского. Изд-во МГУ, Москва.
7. *Русский язык повседневного общения: особенности функционирования в разных социальных группах. Коллективная монография (2016)*, отв. ред. Н. В. Богданова-Бегларян. ЛАЙКА, Санкт-Петербург.

8. Фитиалов С. Я. (1968), Об эквивалентности грамматик НС и грамматик зависимостей, Проблемы структурной лингвистики, Наука, Москва, сс. 71–102.
9. Фонетика спонтанной речи (1988), отв. ред. Н. Д. Светозарова. Изд. ЛГУ, Ленинград.
10. Шерстинова Т. Ю. (2016 а), Наиболее употребительные слова повседневной русской речи (в гендерном аспекте и в зависимости от условий коммуникации), Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 1–4 июня 2016 г.), Вып. 15 (22), Изд-во РГГУ, Москва, сс. 616–631.
11. Шерстинова Т. Ю. (2016 б), Распознавание и синтез речи, Прикладная и компьютерная лингвистика (коллективная монография), Изд. группа УРСС, Москва, сс. 94–120.

## References

1. Asinovsky, A., Bogdanova, N., Rusakova, M., Ryko, A., Stepanova, S., Sherstinova, T. (2009), The ORD Speech Corpus of Russian Everyday Communication «One Speaker's Day»: Creation Principles and Annotation / Matoušek, V., Mautner, P. (eds.) TSD 2009. LNAI, vol. 57292009. Springer, Berlin-Heidelberg, pp. 250–257.
2. Baudouin de Courtenay J. (1917), Introduction to Linguistics [Vvedenie v yazykovedenie], Petrograd.
3. Bogdanova-Beglarian, N. V., Asinovsky, A. S., Blinova, O. V., Markasova, E. V., Ryko, A. I., Sherstinova, T. Yu. (2015), Speech Corpus of Russian Language: New Methodology of Speech Analysis [Zvukovoj korpus russkogo jazyka: novaja metodologija analiza ustnoj rechi], Language and Method [Jazyk i metod: Russkij jazyk v lingvisticheskikh issledovanijah XXI veka] / Eds. Szumska, D., Ozga, K. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego, Iss. 2, pp. 357–372.
4. Bogdanova-Beglarian N. V., Blinova O. V., Martynenko G. Ya., Sherstinova T. Yu. (2016), Everyday Russian Language: Some Figures in Sociolinguistic Perspective [Russkiy yazyk povsednevnogo obshcheniya: nekotorye kolichestvennye dannye v zerkale sotsolingvistiki], Communication Studies [Kommunikativnye issledovaniya], Iss. 2 (8), pp. 81–92.
5. Bogdanova-Beglarian, N., Martynenko, G., Sherstinova T. (2015), The “One Day of Speech” Corpus: Phonetic and Syntactic Studies of Everyday Spoken Russian, Ronzhin, A. et al. (eds.) SPECOM 2015, Lecture Notes in Artificial Intelligence, LNAI, vol. 9319. Springer International Publishing Switzerland, pp. 429–437.
6. Bogdanova-Beglarian N., Sherstinova T., Blinova O., Baeva E., Martynenko G., Ryko A. (2016 а), Sociolinguistic Extension of the ORD Corpus of Russian Everyday Speech, SPECOM 2016, Lecture Notes in Artificial Intelligence, LNAI, vol. 9811, Springer, Switzerland, pp. 659–666.
7. Bogdanova-Beglarian N., Sherstinova T., Blinova O., Martynenko G. (2016 б), An Exploratory Study on Sociolinguistic Variation of Spoken Russian, SPECOM 2016, Lecture Notes in Artificial Intelligence, LNAI, vol. 9811. Springer, Switzerland, pp. 100–107.



8. Everyday Russian Language in Different Social Groups. Collective Monograph [Russkij jazyk povsednevnogo obshchenia: osobennosti funkcionirovania v raznykh social'nykh gruppakh. Kollektivnaja monografija] (2016), ed. by N. V. Bogdanova-Beglarian, LAIKA, St. Petersburg.
9. *Fitalov S. Ja.* (1968), On the Equivalence of Immediate-Constituent Grammars and Dependency Grammars [Ob ekvivalentnosti grammatik NS i grammatik zavisimostej], Problems of Structural Linguistic [Problemy strukturnoj lingvistiki], Nauka, Moscow, pp. 71–102.
10. *Larin B. A.* (1977), To the Linguistic Characteristics of the City (Several Prerequisites) [K lingvisticheskoy kharakteristike goroda (neskol'ko predposylok)], B. A. Larin, The history of the Russian language and general linguistics [Istoriya russkogo yazyka i obshchee yazykoznanie], Prosveshchenie, Moscow, pp. 189–199.
11. *Martynenko G. Ya.* (2015), Syntax of Russian Spontaneous Speech: Symmetry in Linear Order of Verb Phrase [Sintaksis zhivoi spontannoï rechi: simmetriia lineinykh poriadkov], Corpus Linguistics, Int. Conf. Proc., St. Petersburg university Publ. House, St. Petersburg, pp. 371–378.
12. Mathematical Methods in Biology [Matematicheskie metody v biologii. Uchebnoe posobie] (1978), ed. by N. A. Plokhinskiy. Moscow State University, Moscow.
13. *Sherstinova T. Yu.* (2016 a), The Most Frequent Words in Everyday Spoken Russian (in the Gender Dimension and Depending on Communication Settings) [Naibolee upotrebitel'nye slova povsednevnoy russkoy rechi (v gendernom aspekte i v zavisimosti ot usloviy kommunikatsii)], Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”. Issue 15 (22) [Komp'yuternaya lingvistika i intellektual'nye tekhnologii: Po materialam ezhegodnoy Mezhdunarodnoy konferentsii “Dialog”]. Вып. 15 (22). RSUH, Moscow, pp. 616–631.
14. *Sherstinova T. Yu.* (2016 b), Speech Synthesis and Recognition [Raspoznavanie i sintez rechi], Applied and computational linguistics [Prikladnaya i komp'yuternaya lingvistika]. URSS, Moscow, pp. 94–120.

# РАЗВИТИЕ ПОЛИСЕМИИ У РУССКИХ ЧИСЛИТЕЛЬНЫХ И ПОРЯДКОВЫХ ПРИЛАГАТЕЛЬНЫХ<sup>1</sup>

**Богуславская О. Ю.** (boguslavskaya.olga@gmail.com),  
**Иомдин Б. Л.** (iomdin@ruslang.ru)

Институт русского языка им. В. В. Виноградова РАН,  
Москва, Россия

## DEVELOPMENT OF POLYSEMY IN RUSSIAN NUMERALS AND ORDINALS

**Boguslavskaya O. Ju.** (boguslavskaya.olga@gmail.com),  
**Iomdin B. L.** (iomdin@ruslang.ru),

The Russian Language Institute of the RAS, Moscow, Russia

Words denoting numbers (cardinal and ordinal numerals, or adjectives) represent a small (although potentially infinite) lexicographic type. In this article we deal with the polysemy structure of these two lexical classes. We propose a lexicographic pattern and study standard types of semantic shifts, including regular metaphors and metonymies. The words of both classes normally develop special senses with conversion into other parts of speech. Additional senses, different for different words, can appear due to cultural conventions.

**Keywords:** polysemy, lexicography, lexicographic types, numbers, numerals

### 1. Введение

Настоящая работа посвящена ранее, насколько нам известно, не изучавшейся теме — системному описанию полисемии русских числительных. С точки зрения структуры полисемии семантический класс числительных и семантический класс порядковых прилагательных, разумеется, демонстрируют значительное единообразие. Тем не менее, предложить единые схемы для представления этих семантических классов нельзя. Обособляются группы прагматически выделенных слов, выпадающих из общей схемы. Слова, имеющие более богатую полисемию, развивающие уникальные значения, не повторяющиеся у других элементов класса, ожидаемым образом оказываются более частотными. Так,

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФ (проект 16-18-02054).

числительное *один*, имеющее максимальное число значений, имеет ранг 174 по (Ljashevskaya, Sharov 2009), а одно из имеющих сравнительно немного значений числительное *восемь* имеет ранг 1483; точно также имеющее максимальное число значений порядковое прилагательное *первый* имеет ранг 64, а имеющее минимальное в этой группе число значений *седьмой* имеет ранг 2969.

Работа преследовала две цели. Первая — предложить схемы описания количественных числительных и порядковых прилагательных, реализуемые в Активном словаре русского языка (АС)<sup>2</sup>. Эта задача имеет в значительной степени технический лексикографический характер. Предлагаемые схемы помогут единообразно представить то, что является общим и одинаковым у рассматриваемой группы слов. Но за рамками остаются многочисленные индивидуальные особенности, требующие создания «лексикографических портретов». Вторая цель — описать типы семантических переходов, представленные у слов этих семантических классов.

## 2. Количественные числительные

2.1. Приведем схему описания числительного *шесть*. Оно представляет наиболее распространенный тип многозначности среди представителей рассматриваемого семантического класса. Это числительное имеет шесть значений. Его частотный ранг по (Ljashevskaya O. N., Sharov S. A. 2009) — 984. Все числительные от *два* до *двадцать* имеют по крайней мере эти же шесть значений.

Принятое в этой работе схематическое описание многозначности предполагает заполнение следующих позиций: имя вокабулы, имя лексемы, примеры, схематическое толкование, примечания об особых употреблениях, описание управления, синонимы (факультативно), тип перехода (для производных значений).

### ШЕСТЬ

#### *шесть* 1<sup>3</sup>

**Примеры.** Число *шесть*; умножить *шесть* на *четыре*; *шестью* *шесть*; считать до *шести*; делить <умножать> на *шесть*; Три плюс два будет *шесть*.

**Схема толкования.** Число, на один больше, чем пять.

СИН: (разговорное) *шестерка*; ДЕР: *шестью* [при умножении: *шестью* *шесть*]

<sup>2</sup> Все предложенные описания публикуются впервые. Помимо этого, при анализе рассматриваемых семантических классов мы опирались на словарные статьи количественных числительных и порядковых прилагательных, опубликованные во втором томе Активного словаря русского языка, и статьи, подготовленные для публикации в третьем томе словаря. Авторы этих словарных статей — О. Ю. Богуславская, Б. Л. Иомдин, Т. В. Крылова, А. В. Санников и Е. В. Урысон.

<sup>3</sup> Нумерация значений во всех приводимых примерах носит отчасти условный характер; порядок блоков значений в словарной статье конкретного слова может быть другим. В начало ради технического удобства изложения мы помещаем значения, характерные для всех или для многих членов семантического класса, а в конец — значения, развивающиеся у отдельных представителей.

### шесть 2.1

**Примеры.** *Шесть девочек; шесть раз; шесть лет; ждать шесть часов.*

**Схема толкования.** *Шесть А1* 'Количество А1 — шесть'.

**Тип перехода.** От **шесть 1**. Сдвиг типа абстрактная сущность → значение параметра.

СИН: *шестерка [шестерка коней]; шестеро [шестеро друзей].*

### шесть 2.2

**Примеры.** *Шесть утра; шесть вечера; шесть часов утра; в шесть; в шесть часов; ровно шесть двадцать; К шести мне надо быть уже в городе; Поезд приходит в шесть с минутами.*

**Схема толкования.** 'Момент времени на шесть часов позже, чем полночь или полдень'.

**Тип перехода.** Сужение значения от **шесть 2.1**.

### шесть 2.3

**Примеры.** *В девять ноль шесть; в восемь часов шесть минут; Часы показывали двадцать один ноль шесть; Сигнал раздастся в девять ноль шесть.*

**Схема толкования.** 'Момент времени на шесть минут позже определенного часа суток'.

**Тип перехода.** Сужение значения от **шесть 2.1**.

**Примечания.** Обычно в сочетании *ноль шесть*.

### шесть 3

*Квартира шесть; дом шесть; числиться под номером шесть; глава шесть.*

**Схема толкования.** 'Номер шесть' [часто с самим словом *номер*].

**Тип перехода.** Метонимия от **шесть 1**.

### шесть 4, в функции существительного среднего рода.

**Примеры.** *Арабское шесть; старославянское шесть; Шесть — перевернутая девятка.*

**Схема толкования.** 'Цифра, обозначающая число шесть'.

**Тип перехода.** Метонимия от **шесть 1**.

СИН: *шестерка*.

2.2. Числительные *два, три, четыре* и *пять*, а также стоящее особняком *единица*, соответствующие стандартным школьным оценкам, имеют дополнительное значение: 'Оценка, указывающая на степень усвоения знаний в области А1'. Приведем схему описания этого значения на примере числительного *пять*.

### пять 5

**Примеры.** *Пять по математике; пять с плюсом; Получить пять; поставить пять.*

**Схема толкования.** ‘Оценка, указывающая на то, что знания в области А1 при испытании А2 полностью усвоены’.<sup>4</sup>

СИН: *пятерка; пять баллов; отлично.*

**Управление.**

А1: по ДАТ: *пять по английскому.*

А2: за ВИН: *пять за экзамен (по физике).*

**Тип перехода.** Метонимия от **пять 1.**

**2.3. Слово ОДИН.** Слово *один* обладает целым рядом уникальных свойств. В частности, это единственное числительное, которое подобно прилагательным согласуется с существительным по роду и числу (*одна книга, один стол, одно решение, одни ножницы*). Тем не менее ряд синтаксических свойств *один* позволяют отнести количественные употребления слова *один* к категории числительных. Подробное обоснование этого и обзор литературы по этому вопросу дается в (Mel’chuk 1985: 271–279).

Помимо количественных употреблений, у слова *один* есть ряд других значений, которые не находят параллелизма среди других числительных. *Один* в этих значениях относится к категории прилагательных. В работе (Mel’chuk 1985: 43) выделяется четыре адъективных значения *один*: 1) ‘некоторый’ (*в один прекрасный день*), 2) ‘без других’ (*Наконец, мы одни*), 3) ‘только’ (*жить одними надеждами*), 4) ‘один и тот же’ (*жить в одном доме*). В нашем описании, приводимом ниже, список неколичественных значений *один* несколько расширяется. Мы добавляем одно адъективное значение, которое, насколько нам известно, ранее никем не выделялось (*один 9*) и субстантивированное местоименное значение *одно 10*.

В Активном словаре русского языка числительное *один* является семантическим примитивом и не толкуется.<sup>5</sup> У него отсутствует значение, соответствующее типовому количественному значению 2.2: мы говорим *час дня, час ночи*, но не *\*один дня*. Нет у него и значения оценки знаний: в этом значении используется слово *единица* или (в другом речевом регистре) *кол*: *получить единицу по физике, поставили единицу, получил кол*; или сленговое *банан*. И тот, и другой синонимы возникли в результате метафорического переноса по сходству формы объекта и цифры; о развитии у слова *банан* значения ‘оценка два’ см. (Iomdin 2014). Приведем схематическое описание неколичественных значений *один*.

**один 5**, в функции прилагательного.

**Примеры.** *Один приятель; в один прекрасный день; один из приехавших.*

**Схема толкования.** *Один А1* ‘Один неопределенный объект, событие, временной отрезок или действие А1’.

<sup>4</sup> К настоящему времени в России устоялась пятибалльная система оценок, и это закрепилось в языке. Всем понятны дериваты *двоечник, пятерочник* и подобные. При переходе к другим системам, например, к десятибалльной или стобалльной, аналогичные значения разовьются и у других числительных (так, уже появился термин *стобалльник*).

<sup>5</sup> А. Вежицка относит к семантическим примитивам не только *один*, но и *два* (Wierzbicka 2001: 53).

**Тип перехода.** Суженное значение от **один 2.1**<sup>6</sup>.

**один 6**, в функции прилагательного.

**Примеры.** *Хранить все документы в одном месте; Мы с ним живем в одном доме.*

**Схема толкования.** *Один А1* 'А1 выполняет одинаковую функцию для разных лиц или объектов'.

**Тип перехода.** Метафора по коннотации исключительности от **один 2.1**.

**один 7**, в функции прилагательного.

**Примеры.** *Вышел из дома в одной рубашке; В доме один хлеб; Одна работа по-прежнему приносит радость; Один Петя меня понимает.*

**Схема толкования.** *Один А1 А2* 'Объект или явление А1 находятся в ситуации А2; другие люди, объекты или явления этого класса не находятся в ситуации А2'.

СИН: *только*.

**Тип перехода.** Метафора по коннотации исключительности от **один 2.1**.

**Примечание.** *Один* не может нести фразового ударения.

**один 8**, в функции прилагательного.

**Примеры.** *Я приду один; Трудно одной растить детей; Выхожу один я на дорогу; Детей оставили одних.*

**Схема толкования.** *Один А1 А2* 'Человек А1 находится в ситуации А2; можно было ожидать, что другие люди тоже находятся в ситуации А2, но это не так'.

**Тип перехода.** Метафора по коннотации исключительности от **один 2.1**.

**один 9**, в функции прилагательного.

**Примеры.** *Одних белых набрали две корзины; Одних золотых медалей привезли с Олимпиады пять штук.*

**Схема толкования.** *Одних А1 в А2 А3* 'Объекты А1 находятся в ситуации А2 в количестве А3; другие такие объекты находятся в ситуации А2 в большем количестве'<sup>7</sup>.

СИН: *только*.

**Тип перехода.** Метафора по коннотации исключительности от **один 2.1**.

**Примечания.** Фразовое ударение падает на существительное, подчиняющее прилагательное *один*. Именная группа с *один* преимущественно располагается в начале предложения. Предложение *Набрали две корзины одних белых* обычно интерпретируется в смысле 'набрали только белые грибы', то есть здесь представлена лексема *один 7*.

<sup>6</sup> Лексема *один 2.1*, представленная в примерах типа *Пришел только один студент; Осталось одно яблоко*, толкуется следующим образом: *Один А1* 'Количество А1 — один'.

<sup>7</sup> Аналогичное значение выделено у слова *только* в (Boguslavskij 1980).

одно 10, в функции местоимения; среднего рода.

**Примеры.** *Ей важно было одно — семья; Одного только мне не хватало — здоровья; Он одного хотел, чтобы ты могла учиться.*

**Схема толкования.** Указывает на объект, ситуацию или свойство, которые говорящий считает важными и которые описываются в тексте далее.

**Тип перехода.** Метафора по коннотации исключительности от **один 2.1**.

**Примечания.** Это местоимение употребляется катафорически — редкий случай в русском языке.

**2.4.** Слово *единица* относится к другому классу количественных слов: к существительным типа *десяток, сотня, дюжина*. Не рассматривая подробно структуру многозначности этого слова, отметим у него два значения, типичных для количественных числительных. В примерах типа *минус единица; мнимая единица; квадратный корень из единицы; прибавить единицу; равняться единице* представлено значение ‘число один’. В примерах типа *Жирная <корявая> единица; написать <вывести> единицу* представлено значение ‘цифра, обозначающая число один’.

Кроме того, как было указано выше в п. 2.2, *единица* используется для обозначения оценки усвоения знаний в области А1 «вместо» числительного *один*: *Поставить единицу по математике за списанную контрольную; написать контрольную на единицу.*

**Схема толкования.** ‘Самая низкая из возможных оценок, указывающая на то, что знания в области А1 по результатам испытания А2 полностью отсутствуют’.

**Управление**

А1 по ДАТ: *единица по химии.*

А2 за ВИН: *единица за контрольную.*

СИН: (разговорное) *кол, (сленговое) банан*; Дериват: *единичка.*

**Тип перехода.** Сдвиг типа абстрактная сущность — значение параметра от **единица 1** с сужением значения.

**2.5.** Представляет интерес специфическая конструкция, в которой могут употребляться некоторые количественные числительные в своем основном значении. Процедура умножения описывается рядом свободных сочетаний: *три умножить на пять; три взять пять раз*. Для описания умножения на два, три и четыре в устной речи употребительна конструкция с наречиями на *-жды*: *дважды пять, трижды три, четырежды восемь*. А для описания умножения на числа от пяти до десяти используется конструкция с наречием на *-ю*: *шестью три, девятью восемь, десятью семь*. Эта конструкция более употребительна для чисел первого десятка, хотя возможны и примеры с числами второго десятка.

- (1) *Он учит, и ты учишься — в том единственном случае, когда учительство заключается не в заклинании: одиннадцатью-одиннадцать-сто-двадцать-один, а в наводке на самый принцип умножения величин друг на друга (А. Найман, Г. Наринская).*

В отличие от наречий типа *дважды*, свободно употребляющихся в позиции при глаголе, наречия на *-ю* используются только для описания умножения.

- (2) *Я был там два раза = Я был там дважды.*  
(3) *Я был там пять раз, но \*Я был там пятью.*

### 3. Порядковые прилагательные

**3.1.** Порядковые прилагательные обычно имеют пять значений, сгруппированных в три блока. Приведем схему описания порядкового прилагательного *пятый*. Оно представляет наиболее распространенный тип многозначности среди представителей рассматриваемого семантического класса. Его частотный ранг по (Ljashevskaya, Sharov 2009) — 1708. Оно имеет шесть значений — дополнительное, седьмое значение, указывающее на время суток, имеется у порядковых прилагательных от *первый* до *двенадцатый*.

#### ПЯТЫЙ

##### пятый 1

**Примеры.** *Пятый по порядку; пятый знак после запятой; жить где-л. пятый месяц <год>; Мы летим уже пятый час.*

**Схема толкования.** Синтаксическая деривация от *пять*.

**Примечание.** Особое употребление в позиции определения при существительном *модель*: ‘особый вид транспортного средства или механизма, который соотносится с порядком разработки разновидностей этого транспортного средства или механизма’. *Пятая модель «Жигулей»*. Тип перехода — сужение значения.

**пятое 2.1,** в функции существительного; среднего рода.

**Примеры.** *Пятое мая; перенести встречу на пятое, улететь пятого.*

**Схема толкования.** ‘Пятый день месяца A1’.

**Тип перехода.** Сужение значение от **пятый 1** с изменением части речи.

**пятое 2.2,** в функции существительного; среднего рода.

**Примеры.** *Говорю одно, вижу другое, представляю третье, чувствую четвёртое, не могу вспомнить пятое* (В. Катаев).

**Схема толкования** ‘Пятый член перечисления’. [Может образовывать отдельное предложение].

**Тип перехода.** Сужение значение от **пятый 1.1** с изменением части речи.

##### пятый 3.1

**Примеры.** *Пятая часть; одна пятая доля наследства.*

**Схема толкования.** *Пятая A1* ‘A1, в пять раз меньшая, чем целое’ [A1 — часть, доля].

**Тип перехода.** Сужение значения от *пять* с изменением части речи.

**пятая 3.2,** в функции существительного; женского рода.

**Примеры.** *Одна пятая от ста; три пятых от ста будет шестьдесят.*



**Схема толкования.** ‘Часть объекта А1, в пять раз меньшая, чем А1’.

**Тип перехода.** Изменение части речи от **пятый 3.1**.

**пятый 4**, в функции существительного; мужского рода.

**Примеры.** *В начале пятого мы начали беспокоиться; Который час? — Половина пятого уже; В четверть пятого начали запись.*

**Схема толкования.** ‘Отрезок времени от четырех часов до пяти часов дня или ночи’.

**Тип перехода.** Сужение значения от **пятый 1** с изменением части речи.

**3.2.** Порядковые прилагательные, образованные от числительных, обозначающих десятки (*десятый, двадцатый*), в позиции определения при существительном *число* имеют особые употребления в рамках основного, первого значения: ‘Дни месяца, предпоследняя цифра в порядковом номере которых — один (два)’: *Лед тронулся в десятых числах апреля; Он должен вернуться в двадцатых числах октября.*

Все порядковые прилагательные, образованные от числительных, обозначающих десятки, в позиции определения при существительном *год* имеют особые употребления в рамках основного, первого значения: ‘Тоды, предпоследняя цифра в порядковом номере которых — один (два, три...)’ *Он окончил медицинский факультет в начале двадцатых годов; В начале семидесятых годов от оттепели ничего не осталось.* Тип перехода — метонимия. Обычно такое словоупотребление относится к десятилетиям XIX или XX веков. Сейчас часто употребляется также название *нулевые годы* применительно к первому десятилетию XXI века.

Некоторые порядковые прилагательные, образованные от числительных, обозначающих круглые числа, в позиции определения при слове *раз* имеют особые употребления в рамках основного, первого значения, указывающие не на точное число, а на то, что что-либо повторяется очень много раз: *В сотый <десятый> раз тебе повторяю.* **Тип перехода** — эмоциональное усиление, затемняющее собственно количественное значение.

Прагматически выделенные представители рассматриваемого семантического класса — *первый, второй, третий*. У них развивается целый ряд дополнительных значений. Для примера рассмотрим дополнительные значения прилагательного *первый*.

**первое 5**, в функции существительного; среднего рода.

**Примеры.** *На первое выбрали солянку; Он не стал есть первое.*

**Схема толкования.** ‘Суп в составе обеда из нескольких блюд, обычно подаваемый после закуски и перед вторым блюдом’.

**Тип перехода.** Сужение значения от **первый 1** с изменением части речи.

**первый 6.1**

**Примеры.** *Первый министр; первый заместитель; первая премия.*

**Схема толкования.** ‘Самый важный’.

**Тип перехода** — метафора от **первый 1**.

### первый 6.2

**Примеры.** *Первый сорт; первый ученик; вагон первого класса; команда первого эшелона.*

**Схема толкования.** 'Лучший по официальной классификации'.

**Тип перехода** — метафора от **первый 1**.

**Примечания.** Некоторые классификации выше первого сорта или разряда ставят *высший сорт* или *высший разряд*. Однако в языке закрепилось *первосортный, перворазрядный*, но не *\*высшесортный, \*высшеразрядный*.

### первый 6.3

**Примеры.** *Первый голос; первая скрипка.*

**Схема толкования.** 'Исполняющий основную партию в полифоническом музыкальном произведении'.

СИН: *прима*.

**Тип перехода.** Сужение значения от **первый 6.1**.

## 4. Выводы

Систематическое изучение полисемии количественных числительных и порядковых прилагательных показало, что вопреки распространенному мнению<sup>8</sup> все числительные и порядковые прилагательные обладают развитой многозначностью.

Разные значения многозначных слов, а также контекстно обусловленные модификации значения лексем возникают в результате действия одних и тех же семантических переходов. Для рассмотренных семантических классов характерны метафора, сужение значения и переход от абстрактного понятия (числа) к значению параметра количества. Сужение значения часто сопровождается изменением частеречной принадлежности<sup>9</sup>, в первую очередь — субстантивацией.

---

<sup>8</sup> Так, О. Ф. Жолобов в работе, посвященной истории числительных в русском языке, рассматривает многозначность числительного *один* и приходит к выводу, что «Хотя полисемия — рутинная, природная черта лексического строя языка, числительные ее полностью лишены, если не считать начала счетного ряда» [Zholobov 2004: 209].

<sup>9</sup> О неопределенности состава класса числительных см., например, [Sichinava 2012].

## Литература

1. *Apresjan Ju. D.* (ed.). (2014). Active Dictionary of Russian. Vol. 1 (A–B), Vol. 2 (V–G) [Aktivnyj slovar' russkogo jazyka. Tom 1 (A–B), tom 2 (V–G)]. Jazyki slavjanskih kul'tur, Moscow.
2. *Boguslavskij I. M.* (1980). Semantics of particle TOL'KO. [Semantika chasticy TOL'KO]. Semiotika i informatika [Semiotics and Information Science]. I. 14, Moscow, pp. 134–158.
3. *Iomdin B. L.* (2014). Polysemous words in and out of the context. [Mnogoznachnyje slova v kontekste i vne konteksta]. Voprosy jazykoznanija [Issues in Linguistics]. Vol. 4. Moscow, pp. 87–103.
4. *Ljashevskaya O. N., Sharov S. A.* (2009). Frequency dictionary of modern Russian (based on the National Russian Corpus). [Častotnyj slovar' sovremennogo russkogo jazyka (na materialax Nacional'nogo korpusa russkogo jazyka)]. Azbukovnik, Moscow.
5. *Mel'chuk I. A.* (1985). The surface syntax of Russian numeral expressions. [Poverxnostnyj sintaksis russkix chislovyx vyraženij]. Wiener slawistischer Almanach. Sonderband 16.
6. *Sichinava D. V.* (2012). Numeral [Čislitel'noe]. Russkaja korpusnaja grammatika [Russian Corpus Grammar]. <http://rusgram.ru>.
7. *Wierzbicka A.* (2001). Understanding cultures through their keywords. [Ponimanie kul'tur čerez posredstvo ključevyx slov]. Jazyki slavjanskoj kul'tury, Moscow.
8. *Zholobov O. F.* (2004). Notes on Old Russian numerals. II: “1”, “3”, “4”. [Zametki o drevnerusskix čislitel'nyx. II: “1”, “3”, “4”]. Russkij jazyk v naučnom osveščenii [Russian Language in Linguistic Perspective]. Moscow. No. 2 (8). Pp. 207–221.

# ДИСКУРСИВНЫЕ ЧАСТИЦЫ И СПОСОБЫ ИХ ПЕРЕВОДА: НУ В РОМАНЕ ВЛАДИМИРА СОРОКИНА «ОЧЕРЕДЬ»<sup>1</sup>

**Добровольский Д. О.** (dobrovolskij@gmail.com),  
**Левонтина И. Б.** (irina.levontina@mail.ru)

Институт русского языка РАН, Москва, Россия

Художественное представление устной речи в литературе сильно отличается от реальной устной речи. В этом смысле интересно изучить роман Владимира Сорокина «Очередь», поскольку это произведение, полностью состоящее из диалога, производит впечатление аутентичной передачи устной речи.

В работе рассматривается употребление частицы *ну*, встречающейся в этом романе около 300 раз. Цель работы — проанализировать немецкие и английские эквиваленты в разных типах употребления *ну*, что позволит более рельефно представить функционирование этой русской частицы. С другой стороны, сопоставление с переводами позволит лучше понять роль этой частицы в структуре художественного текста.

Лексикографические описания показывают, что в качестве стандартного соответствия русского *ну* в немецком языке используется частица *na*. Однако, как видно из анализа, во многих случаях выбираются другие варианты перевода. В контексте уговаривания (*Ну, скажите, не вредничайте*) чаще предлагаются эквиваленты *aber* и *schon*, а в функции преодоления смущения (*Ну что вы..*) — частицы *na*, *naja*, *ja*, *also*, *so*, *eben* или междометия *ach*, *tja*. В английском переводе эта игра на дискурсивных частицах не передается. Это свидетельствует о том, что необходимо уточнить сопоставительно-словарное описание соответствующих частиц с опорой на представительный корпус параллельных текстов.

Анализ показывает, что Сорокин использует частицу *ну* не просто как маркер разговорности, а как ключевое слово. В переводах эта функция не сохраняется или сохраняется не полностью.

**Ключевые слова:** дискурсивные частицы, семантика, устная речь, язык художественной литературы, перевод, русский язык, немецкий язык, английский язык

---

<sup>1</sup> Работа выполнена при поддержке РФФИ (грант 16-06-00339).

## DISCOURSE PARTICLES AND THEIR TRANSLATION: *HY* IN VLADIMIR SOROKIN'S NOVEL *THE QUEUE (OCHERED')*

**Dobrovol'skij D. O.** (dobrovolskij@gmail.com),  
**Levontina I. B.** (irina.levontina@mail.ru)

Russian Language Institute of the Russian Academy  
of Sciences, Moscow, Russia

The representation of oral speech in literature differs sharply from real spoken language. It is therefore interesting to study Vladimir Sorokin's novel *Ochered'* (*The Queue*), which consists entirely of dialogue and gives the impression of conveying authentic orality.

The present study examines the use of the particle *Hy*, which occurs some 300 times in the novel. The aim is to analyze various uses of its German and English equivalents, which will allow us to better describe the functioning of the Russian particle. On the other hand, comparison with translations will lead to a better understanding of this particle in the structure of literary texts.

Lexicographical descriptions indicate that the standard equivalent of Russian *Hy* in German is the particle *na*. As is clear from our analysis, however, in many cases translations select a different variant. In a context involving coaxing (*Nu, skazhite, ne vrednichayte*), the equivalents *aber* and *schon* are found more often, while the particles *na, naja, ja, also, so, eben* or the interjections *ach, tja* are used more frequently to express overcoming embarrassment (*Nu chto vy...*). In English this play on discourse particles is not conveyed. This indicates that we need to refine the contrastive and lexicographic description of these particles on the basis of a representative corpus of parallel texts.

Our analysis shows that Sorokin uses *Hy* not merely as a colloquial marker but also as a keyword, as this term is understood in linguistic poetics. In translation this function is not preserved or preserved only partially.

**Key words:** discourse particles, semantics, oral speech, the language of fiction, translation, Russian, German, English

*Взаимоотношения устной и письменной  
речи усложняются, как только  
мы переходим к сфере искусства.  
(Ю. М. Лотман, 1978)*

1. В последние десятилетия устная речь — один из любимых объектов лингвистических исследований<sup>2</sup>, и за это время сделано очень много; см. обзор в [Крысин 2016]. Лингвисты в своей работе постоянно обращаются

---

<sup>2</sup> Разумеется, эта тема затрагивалась и раньше, см. [Якубинский 1923], но всерьез лингвисты стали изучать это явление во второй половине XX в.

и к аудиозаписям, и к транскрипциям, и к буквальным расшифровкам звучащей речи, так что сейчас уже никого не надо убеждать в том, что устная речь радикально отличается от письменной.

Однако еще не так давно лингвисты отмечали: «Можно сказать, что носитель письменной культуры нового времени — от только что выучившегося читать школьника до профессионального филолога — существует в обстановке своего рода **легенды, мифа** об устной речи, сформировавшегося в условиях и традициях письменной культуры. Оказывается чрезвычайно трудным прорваться сквозь стилизованную культурную репродукцию данного феномена и реконструировать его «домифологические» черты. Парадоксальный факт — приходится говорить о реконструкции явления, повседневно наблюдаемого и сопровождающего нас на каждом шагу» [Гаспаров 1978: 65]. Это выдержка из программной статьи Б. М. Гаспарова «Устная речь как семиотический объект», опубликованной в знаменитом в свое время Тартуском сборнике «Семантика номинации и семиотика устной речи» (Тарту 1978), в который вошла также статья Ю. М. Лотмана «Устная речь в историко-культурной перспективе», из который взят эпиграф к настоящей работе, и статья И. А. Паперно «Устная речь в историко-культурной перспективе (кружковая речь и домашняя литература в пушкинскую эпоху)». В том же году в Москве вышла и книга «Русская разговорная речь. Тексты» [PPP 1978], которая шокировала многих: люди не могли поверить, что они действительно так говорят.

С тех пор многое изменилось — и в части наших представлений об устной речи, и в части ее функционирования: появился интересный феномен устно-письменной речи, особенно характерный для переписки в социальных сетях. С возникновением корпусной лингвистики, в особенности устных корпусов (впрочем, пока не очень больших) появилась возможность исследовать устную речь в ее реальном бытовании, в частности с использованием количественных методов.

И теперь уже хотелось бы вернуться к вопросу о том, как, собственно, устроена «культурная репродукция» устной речи в письменных текстах. Понятно, что в большинстве случаев имеет место редукция признаков устной речи при переводе в письменный текст (если это, скажем, расшифровка интервью<sup>3</sup>), либо, наоборот, включение каких-то дискурсивных маркеров для имитации устности, если это художественный текст, в который включается, например, сцена разговора.

Однако в тех произведениях, где устная стихия является важным структурным элементом художественного текста, впечатление аутентичности создается не копированием устного дискурса, а скорее концентрированным выражением его характерных свойств.

2. В этом смысле интересно изучить тексты такого выдающегося стилиста, как Владимир Сорокин, в особенности его небольшой ранний роман «Очердь» (1985). Весь текст состоит из нескончаемого диалога с меняющимися

---

<sup>3</sup> В этом отношении очень показательна книга Соломона Волкова «Диалоги с Бродским», где очень тонко создается впечатление устной речи Бродского, при том что в действительности текст, конечно, сильно отредактирован.

собеседниками. При чтении этот текст производит впечатление поразительной аутентичности. Трудно вспомнить еще какое-нибудь произведение русской литературы, где разговорная речь передавалась бы столь убедительно. В частности, яркой приметой русского языка, особенно в его устной реализации, является частота и раннообранние дискурсивных частиц, что, естественно, в полной мере реализуется в повести Сорокина:

*Я слышала вчера по два давали. — По два? — Ага.  
Сначала-то по четыре, а потом по два. — Мало как! Так и стоять смысла нет... — А вы займите две очереди. Тут приезжие по три занимают. — По три? — Ага. — Так это целый день стоять! — Да что вы. Тут быстро отпускают. — Чего-то не верится. Мы вон с места не сдвинулись...<sup>4</sup> [...] — На Ленинском давали молодую, вполне хорошая. — Да ну! Одни листья. — Молодая очень полезная. — Вон подходят как, совсем обнаглели. Мужчина, зачем вы пропускаете?! Что нам целый день стоять?! Подходят, подходят! — Они занимали, отошли просто... — Да ничего они не занимали! — Мы занимали, чего вы кричите. — Ничего вы не занимали! [...] — Да занимали они, успокойтесь!*

Сорокин — признанный автор. Его произведения переводятся на разные языки, и переводчики придают большое значение тому, чтобы сохранить в переводе стилистические особенности его прозы. Мы взяли перевод на немецкий (немецкий язык, как и русский, отличается большим количеством и частотой употребления частиц). Переводчик «Очереди» Петер Урбан — известный и высококвалифицированный специалист. Для сравнения мы взяли также перевод на английский. В английском языке частиц не так много, и используются они не столь интенсивно. Переводить «Очередь» на английский очень трудно. Кстати, и в предисловии переводчицы отмечается, что диалог, состоящий их простейших реплик, «[is] not always easily transferable to English» [Laird 1988: x].

В этой работе мы рассмотрим одну из самых ярких русских частиц — частицу *ну*. В небольшом произведении Сорокина (29 643 слова) содержится около 300 вхождений *ну*.

Наша цель двоякая: проанализировать немецкие и английские эквиваленты в разных типах употребления *ну*, что позволит более рельефно представить функционирование этой русской частицы; с другой же стороны, сопоставление с переводами позволит лучше понять роль этой частицы в структуре художественного текста.

3. Частица *ну* неоднократно становилась объектом лингвистического исследования. Подробное описание различных контекстов употребления этой частицы и ее значений содержится в работе [Баранов, Кобозева 1978].

Структура многозначности *ну* достаточно детально описана в словарях [Морковкин 1997; Шимчук, Шур 1999]. Функционирование *ну* в одном из самых

<sup>4</sup> См. разбор этого употребления *вон* в [Левонтина 2016].

важных для этого слова контекстов — контексте повторной просьбы — рассмотрено в работе [Левонтина 1999].

Словари выделяют у *ну*-междометия следующие значения:

- 1) призыв или побуждение к действию; ср. **Ну, давай плясать**, **Ну, выкладывай**: *что тебе нужно*.
- 2) требование не надоедать или желание избавиться от чего-то надоевшего (**Ну тебя**, **Ну ее**);
- 3) (обычно в сочетании с частицами *и*, *уж*, *жс*) удивление, восхищение или недовольство, негодование, ирония и другие чувства; **Ну бал!** **Ну Фамусов!** *умел гостей называть!*; **Ну уж!**; **Ну и гроза!**

У *ну* в качестве частицы МАС выделяет следующие разновидности:

- 1) вопросительная (в сочетании с частицей *да*: *да ну*): **Да ну?**
- 2) вопросительная; **Ну?** в качестве отдельной ответной реплики: *Ты здесь работаешь?* — **Ну!**
- 3) вопросительная (в сочетании со словом *как*) — в значении 'что, если': **Ну, как с безумных глаз Затеет драться он** (Грибоедов).
- 4) усилительная: **Ну, конечно**; **Ну как не порадеть родному человечку**; **Ну что за жизнь**; *ну а они...*
- 5) употребляется в значении союза, указывающего на вывод, заключение или вводное замечание: **Ну, и гости бывают**; **Ну, наступил вечер**, **Ну, словом...** В начале реплики в значении 'итак, наконец': — **Ну, я к себе**.
- 6) в сочетании с инфинитивом: **И ну кричать**.
- 7) употребляется в значении 'допустим, положим'. — *Вы первый курс?* — **Ну, первый**.<sup>5</sup>

Четвертое, пятое и седьмое значение скрывают за собой разные, сложным образом взаимосвязанные виды *ну*. Отчасти соответствующие контексты описаны в работе [Баранов, Кобозева 1988: 51–61]. Там, в частности, выделен интересный вид *ну*, связанного с преодолением коммуникативного затруднения. К нему относятся следующие разновидности: *ну* «поиска (гезитации)» (**Ну... пять раз**); *ну* «выбора» (**Ну про школу расскажи**); и *ну* «вызова», внутри которого различаются *ну* «признания» (**Ну плохой я человек!**) и *ну* «непонимания цели» (**Ну в среду, а что?**). Эта идея преодоления коммуникативного затруднения (объясняющая, кстати, почему *ну* так легко превращается в «слово-паразит») имеет нечто общее с побудительностью: говорящий как бы подталкивает сам себя<sup>6</sup>.

Разнообразные употребления *ну* все же образуют некое единство. Большинство значений этой частицы концентрируются вокруг побудительного. Это значение для *ну* чрезвычайно важно: именно с побудительностью *ну* в первую очередь ассоциируется у носителей языка, да и этимологически именно это значение первично (изначально *ну* возникает как «погоняющий» окрик [Фасмер 1986–1987]).

<sup>5</sup> Можно заметить, что это даже и не всегда значения: некоторые из этих разновидностей — это типы сочетаемости или указания на самую общую функцию и т.д.

<sup>6</sup> Ср., например, употребления побудительной частицы *-ка* в контекстах самопобуждения (*Пойду-ка я домой*; ср. [Левонтина 1991]).



В работе [Левонтина 1999] предложены следующие компоненты толкования для «побудительного» ну: 1) я хочу, чтобы ты сделал Р; 2) я знаю, что ты знаешь об этом моем желании; 3) я не располагаю новыми аргументами в свою пользу или не хочу прибегнуть к аргументам.

С помощью ну говорящий не просто побуждает, но действительно всегда как бы погоняет, торопит (чувства его могут быть при этом разными: от радостного нетерпения до раздражения).

Если адресат сам заинтересован в выполнении действия и предполагается, что только скромность или робость мешают ему, то фраза с побудительным ну может быть произнесена с поощрительной или подбадривающей интонацией; ср.:

— Ну, ешьте шоколад, ну, я вас прошу! (И. Ильф и Е. Петров. Двенадцать стульев).

Если же в выполнении действия криво заинтересован говорящий, но возможности требовать или приказывать у него нет, тогда то же ну сопровождается умоляющей интонацией; ср.:

*Давайте помиримся! Ну давайте помиримся, Федор Иванович! И опять начнем заниматься ботаникой!* (В. Дудинцев. Белые одежды).

Таким образом, кажется, что различия между разными видами «побудительного» ну обусловлены контекстом (соотношением интересов и позиций говорящего и адресата) и интонацией.

В сущности, то же значение представлено и еще в некоторых типах изолированного употребления ну, в частности, когда оно выражает призыв скорее сообщить интересующую собеседника информацию или успокоиться; ср.:

— Ну? Ну? — Ну и ну, — отвечал Остап раздраженно, — один шанс против одиннадцати (И. Ильф и Е. Петров. Двенадцать стульев);

— Ну, ну, пошутил (И. Ильф и Е. Петров. Двенадцать стульев).

Возможность употребления «побудительного» ну в контексте вопроса также совершенно понятна: ведь вопрос представляет собой, в сущности, просьбу об информации. В отказах или при аргументации ну также взаимодействует с неявной целью речевого акта. Когда умоляющее ну используется при аргументации, создается своеобразный эффект. Само значение ну предполагает волевой и эмоциональный нажим на собеседника, а не рациональное обоснование просьбы. На первый взгляд, оно должно лишать аргумент убедительности. Однако вместо этого ну усиливает аргумент, как бы предлагая воспринимать его не на рациональном, а на эмоциональном уровне.

Очень часто два разных ну соседствуют в диалоге: одно сопровождает просьбу, а другое — согласие (неохотное или после колебания); ср.:

— Мам, дай еще хлеба! — Нет, хватит. — Ну дай! Видишь они поклевали. Дай! — Ну на, на... (В. Сорокин. Очередь);

— Ну, немного совсем <налить вина>? — Ну, немного можно (В. Сорокин. Очередь).

Замечательно, что здесь Сорокин обыгрывает именно многозначность частицы.

Впрочем, в работе [Баранов, Кобозева 1988] предлагается иная интерпретация побудительных контекстов с *ну*. Побудительность, как можно понять, рассматривается здесь как элемент иллокутивного контекста, а не часть значения самой частицы. Значение же *ну* определяется как «отсылка к упомянутому» (с. 62–64).

Однако *ну* может и не иметь побудительного значения в контексте побуждения, а может, напротив, придавать императивность непобудительным контекстам. Его значение взаимодействует при этом с целью речевого акта: например, при обещаниях с компонентом 'говорящий хочет, чтобы слушающий ему поверил'. В примерах типа *Это же Витька Иванов! Ну в параллельном классе учился!* говорящий при помощи *ну* призывает слушающего сосредоточиться и понять, что он имеет в виду. Ср. подробнее о переходах между значениями *ну* в [Левонтина 1999].

Сравнение с немецким переводом показывает, что в значительном числе случаев русское *ну* передается при помощи немецкого *na*.

В «Новом большом немецко-русском словаре» [НБНС 2010: 1025] дискурсивная частица *na* описывается следующим образом (с сокращениями):

**na** I *prtc disc* разг. ну!; ~? ну (и что там у тебя?); ~ warte! ну, погоди!; ~, so was! кто бы подумал! ~, wird's bald? ну долго там ещё?; ~, ich danke! *упон.* ну, спасибо!, вот уж спасибо!; ~, der wird sich wundern! ну, и удивится же он!; ~, es wird sich schon zeigen ну что ж, там видно будет; ~, ~, so schlimm wird es (schon) nicht sein! ну не так уж это страшно!; ~, wo bleibst du denn? куда ты пропал?; ~, wie geht's? ну, как дела? **○** ~ **und?** ну и (что)?; ~ **ja**, ~ **gut [schön]** ну ладно; ~ ja, dann mach, was du willst! ну ладно, тогда делай, что хочешь!; ~ ja, ich weiß schon ну да, я знаю; ~ **also!**, ~ **bitte!** ну, вот видишь [что я говорил]! **◇** ~, **komm doch [schon]!** ну же!

Как видно из этой словарной статьи, наиболее стандартный эквивалент *na* в русском — это *ну*. Выбор того или иного соответствия в переводе художественного текста, естественно, зависит от особенностей описываемой ситуации, но в самых обычных, не отягощенных дополнительными смыслами контекстах (таких, как в иллюстративной части словарной статьи) этот эквивалент оказывается самым употребительным. Обращает на себя внимание, что компонент *ну* присутствует даже в большинстве идиоматических конструкций: грамматических фразем и коллокаций (за знаком **○**) и в идиоме (за знаком **◇**).

4. Как видно из приведенных описаний русского *ну* и немецкого *na*, семантически они весьма сходны, поэтому неудивительно, что в немецком переводе «Очереди» в значительном числе случаев *ну* переводится именно так. Однако в остальных случаях переводческие решения разнообразны и ситуативны.

Наиболее показательны, пожалуй, два фрагмента, где в русском *ну* оказывается структурообразующим элементом.

В первом фрагменте перед нами контекст уговаривания:

— Простите, а вас как зовут? — А зачем вам? — Очень нужно. — Ничего вам не нужно. Не скажу. — **Ну**, скажите, не вредничайте. — Неа. — **Ну**, скажите, пожалуйста. — **Ну**, а зачем вам? — **Ну**, что вам жалко что ли?

*Verzeihen Sie, und wie heißen Sie? — Wozu wollen Sie das wissen? — Ich muß es wissen. — Nichts müssen Sie. Ich sage es Ihnen nicht. — Na, sagen Sie es schon, seien Sie kein Spielverderber. — Nee. — Sagen Sie schon, bitte. — Aber wozu wollen Sie das wissen? — Fällt Ihnen doch keine Perle aus der Krone, oder?*

Заметим, что из четырех ну ни одно не переведено самым стандартным образом — как *na*. Одному соответствует *na... schon*, второму *schon*, третьему *aber* (строго говоря, *aber* соответствует не просто ну, а ну а), а в четвертом случае переводчик просто перевел всю фразу *Ну, что вам жалко что ли?* при помощи идиомы, которая буквально значит: *Что, жемчужина из короны выпадет?*

Теперь обратимся к английскому переводу этого фрагмента:

— *So tell me now, what's your name? — And why should I tell you? — Because I need to know, that's why. — You don't need to know, and I am not telling. — Go on, don't be mean, tell me. — But why should I tell you? — Why shouldn't you? Is it such a big deal?*

Как мы видим, в английском тексте переводчица использовала две формы для передачи «уговаривательного» ну: *go on* и противительный союз *but* (интересно, что это решение полностью соответствует немецкому *aber*). Что касается последней фразы *Ну, что вам жалко что ли?*, то здесь идея уговаривания «разлита» по всему контексту.

Проанализированный эпизод весьма характерен. Дело в том, что роман Сорокина — об одиночестве человека в толпе и о его несвободе. Повторяющаяся частица ну маркирует здесь усилие человека, пытающегося одиночество и несвободу преодолеть, и девушка Лена — это первая, неудачная попытка героя.

Еще более важен другой эпизод: герой, выйдя из очереди переждать дождь, попадает домой к незнакомой девушке Люде. Можно сказать, что эта сцена полностью строится на ну. Здесь перед нами контекст преодоления смущения (в скобках после каждой фразы, содержащей ну, помещены те единицы, которые ей соответствуют в немецком и английском переводах; знаком «—» показано, что частица вообще никак не переводится). Два незнакомых человека стремятся к сближению, и ну помогает им преодолеть коммуникативное затруднение на каждом шаге этого сближения — и в то же время каждый шаг оформить как неохотное согласие:

— *Да пойдёмте. Что ж вы на лестнице дрожать будете. [...] Ну ладно (Also gut / oh well). [...] — Что не успели добежать вовремя? — Ну да. (Na ja / Oh) [...] — Эпизонская живопись... — Ну, почему. (Ach / No) [...] — Ничего интересного. Тоска зеленая. — Ну, это, видимо, зависит от места работы... (Na / Well) [...] — И долго вы стоите? — Да не очень... — Очередь громадная. Я таких давно не видела. — Ну так их же раз в полгода*

выбрасывают. (*Naja / Well*) [...] — **Ну** вот... (*So / —*) рубашка ваша готова... держите... [...] — Бабушка умерла, я в ее комнате. — В общей квартире? — Да. — **Ну** и как? (*Und / —*) — Вполне. Соседи приличные люди. В центре живу. — **Ну**, это хорошо, когда соседи такие... (*Ja / —*) [...] — **Ну** вот. (*Eben / So*) [...] — **Ну**, а кто ж в такой дождь торговать будет? (*Na / Well*) Все промокнет. [...] — А у меня он любимый поэт. — **Ну**, а что. (*Ja klar / —*) [...] — **Ну**, что вы! (*Aber was glauben Sie! / But imagine!*) [...] — Вадим, а вы есть хотите? Я лично хочу. — **Ну** я немножко чего-нибудь... (*Naja / Well*) [...]. — **Ну** вот. (*Also schön / Right*) [...] За встречу. — **Ну** что ж... (*Ja / Alright then*) за встречу... — [...] — **Ну**, это наверно из ранних. (*Na / —*) [...] — Стоял в паскудной очереди, толкался, ждал чего-то. И вдруг пью вино с очаровательной женщиной... — **Ну**, не преувеличивайте. (*Na / Come on*) [...] — Вы удивительно похожи на одного человека. — На какого? — **Ну**... просто это было очень давно... (*Ach / Oh*) [...] — **Ну**, учитывая, что я вас старше, это была не я. (*Na / Well*) [...] — **Ну**, немного совсем? (*Na / Go on*) — **Ну**, немного можно... (*Gut / Well*) хватит, хватит! — За ваши глаза. Чудные карие глаза. — **Ну** почему все время за меня?! (*Und / Come on*) [...] — Нет, нет, за ваши глаза. — Господи, **ну** что они вам дались... (— / —) [...] — **Ну** хорошо. (*Aber gut / —*) [...] — Вадим, **ну** что такое? (*was ist denn / Go on*) [...] — **Ну** вот видите, вы не хотите... (*Na sehen Sie / There, see*) [...] — Потому что вы мне очень нравитесь. — Ой, Вадим. **Ну** что вы...(— / —) [...] — Если для вас смех, то для меня нет... **ну**, давайте? (*also / —*) [...] — Ты не заметил? — Неа. — **Ну**! (*Tja / —*) [...] — Да. Я так давно не видел настоящего уюта... — **Ну**, **ну**, поплачь еще... (*Na, na / Ah*) [...] — А ты любишь тащить? — Любил. Сто лет уж не танцевал. Давай? — **Ну**, пошли. (*Also gut / Okay*)

Объем небольшой статьи не позволяет подробно рассмотреть все варианты переводов и оттенки значений *ну* во всех случаях. Можно, однако, сделать ряд наблюдений.

Как уже было отмечено, лексикографические описания показывают, что немецкое *na* довольно хорошо соответствует русскому *ну*, даже не так просто сформулировать явные семантические различия между ним. Однако, как видно из этого фрагмента, во многих случаях хороший переводчик часто выбирает другие варианты перевода.

Там, где *ну* используется в контексте уговаривания (*Ну, скажите, не вредничайте*), в качестве эквивалентов чаще предлагаются частицы *aber* и *schon*, а там, где *ну* служит преодолению смущения (*Ну вот, Ну что вы...*)<sup>7</sup> — скорее частицы *na*, *naja*, *ja*, *also*, *so*, *eben* и междометия *ach*, *tja*. В английском переводе эта игра на дискурсивных частицах передается хуже. Возможно, по-английски эти нюансы вообще трудно выразить. Можно лишь отметить, что в контексте преодоления смущения эквивалентом-фаворитом, по-видимому, является *well*.

<sup>7</sup> Строго говоря, преодолению смущения служит не просто *ну*, а соответствующие клише, произносимые с определенной интонацией.

Эти факты свидетельствуют о том, что необходимо уточнить сопоставительно-словарное описание соответствующих частиц с опорой на представительный корпус параллельных текстов, что эти еще никем не описанные тонкости, должны быть четко зафиксированы и проиллюстрированы.

Говоря об использовании частицы *ну* в романе Сорокина, нельзя не остановиться на вопросе о ее частотности. Действительно, при чтении кажется, что Сорокин использует эту частицу необычайно интенсивно. При этом, как отметил анонимный рецензент «Диалога», «Проверка относительной частоты этой частицы по разным подкорпусам НКРЯ показывает, что на фоне других текстов Сорокина она, действительно, высокая, но все они, в отличие от „Очереди“ не являются полностью диалогическими текстами. Сравнение „Очереди“ с „устным подкорпусом“ показывает, что отношение *ну* к общему числу слов в них совпадает, а в кинодиалогах МУРКО *ну* встречается гораздо чаще, представляя 0,015 от всех слов против 0,010 в „Очереди“». Однако здесь существенна не просто частота частицы в произведении в целом, а интенсивность ее использования в отдельных фрагментах. Так, если частота в «Очереди» в целом действительно сопоставима с ее частотой в устной речи, то в анализируемом фрагменте она существенно выше и составляет примерно 0,025.

Экстремальная концентрация *ну* в этом фрагменте показывает, что Сорокин использует здесь частицу *ну* не просто как маркер разговорности, а как ключевое слово — в том смысле, как этот термин понимается в лингвистической поэтике. При этом в немецком переводе используются разные переводные эквиваленты, и эффект ключевого слова частично пропадает. По-английски же игру на дискурсивных частицах вообще трудно передать.

## Тексты

1. Сорокин В. Очередь. Париж: Синтаксис, 1985.
2. Sorokin V. The Queue / Novel / trans. Sally Laird. New York: Review Books, 1988.
3. Sorokin V. Die Schlange / Roman / Aus dem Russischen von Peter Urban. Zürich: Haffmanns Verlag, 1990.

## Литература

1. Баранов А. Н., Кобозева И. М. Модальные частицы в ответах на вопрос // Логический анализ языка. Прагматика и проблемы интенциональности. М., 1988.
2. Гаспаров Б. М. (1978), Устная речь как семиотический объект // Семантика номинации и семиотика устной речи. Тарту, 1978.
3. Крысин Л. П. Лексикографическое представление разговорной речи в толковом словаре // Русский язык за рубежом 4/2016 (257). С. 12–19.
4. Левонтина И. Б. Словарная статья частицы *-ка* // Семиотика и информатика, вып 32. М., 1991. С. 140–145.
5. Левонтина И. Б. Стратегии уговаривания: частицы в повторных просьбах // Язык. Культура, Гуманитарное знание. Научное наследие Г. О. Винокура и современность. М., 1999. С. 188–201.

6. Левонтина И. Б. Иду это я... О дискурсивных употреблениях дейктических слов. // Труды Института русского языка им. В. В. Виноградова. Вып. 11. М., 2016. С. 160–172.
7. Морковкин В. В. и др. Словарь структурных слов русского языка / Под ред. В. В. Морковкина. М., 1997.
8. НБНС — Новый большой немецко-русский словарь. В 3 т.: около 500 000 лексических единиц / под рук. Д. О. Добровольского. Т. II: G—Q. М., 2010.
9. PPP — Русская разговорная речь. Тексты / Отв. ред. Е. Л. Земская, Л. А. Кападзе. М., 1978.
10. Фасмер М. Этимологический словарь русского языка. В 4 т. 2-е издание. М., 1986–1987.
11. Шимчук Э., Щур М. Словарь русских частиц. Berlin, 1999.
12. Якубинский Л. П. О диалогической речи // Русская речь / под ред. Л. В. Щербы. Петроград, 1923. С. 96–194.
13. Laird S. Foreword // Sorokin V. The Queue / Novel / trans. Sally Laird. New York, 1988, pp. i–x.

## Sources

1. Sorokin V. Ochered'. Sintaksis, Paris, 1985.
2. Sorokin V. The Queue (novel), trans. Sally Laird. Review Books, New York, 1988.
3. Sorokin V. Die Schlange (Roman). Aus dem Russischen von Peter Urban. Haffmanns Verlag, Zürich, 1990.

## References

1. Baranov A. N., Kobozeva I. M. (1988), Modal particles in responses [Modal'nye chastitsy v otvetakh na vopros], Logical analysis of language. Pragmatics and issues of intensionality [Logicheskij analiz yazyka. Pragmatika i problemy intensional'nosti], Moscow.
2. Gasparov B. M. (1978), Oral speech as an object of semiotics [Ustnaya rech' kak semioticheskij ob'ekt], Semantics of nomination and semiotics of oral speech [Semantika nominatsii i semiotica ustnoy rechi], Tartu.
3. Krysin L. P. (2016), Lexicographic representation of colloquial language in a monolingual dictionary [Leksikograficheskoe predstavlenie razgovornoj rechi v tolkovom slovare], The Russian language abroad [Russkij yazyk za ru-bezhom] 4 (257), pp 12–19.
4. Laird S. (1988), Foreword, Sorokin V. The Queue (novel), trans. Sally Laird. Review Books, New York, pp. i–x.
5. Levontina I. B. (1991), Particle *-ka*: a dictionary entry [Slovarnaya stat'ya chastitsy *-ka*], Semiotics and informatics [Semiotika i informatika], 32, Moscow, pp. 140–145.
6. Levontina I. B. (1999), Strategies of coaxing: particles in repeated requests [Strategii ugovarivaniya: chastitsy v povtornykh pros'bakh], Language. Culture. Humanities. Scientific heritage of G. O. Vinokur today [Yazyk. Kul'tura. Gumanitarnoe znanie. Nauchnoe nasledie G. O. Vinokura i sovremennost'], Moscow, pp. 188–201.

7. *Levontina I. B.* (2016), Once I went... On discursive use of deictic words [Idu éto ya... O diskursivnykh upotrebleniyakh deyktichestikh slov], Proceedings of the Russian Language Institute of the Russian Academy of Sciences [Trudy Instituta russkogo yazyka im. V. V. Vinogradova], 11, Moscow, pp. 160–172.
8. *Morkovkin V. V. et al.* (1997), Dictionary of auxiliary words in Russian [Slovar' strukturnykh slov russkogo yazyka], ed. by V. V. Morkovkin, Moscow.
9. *NBNS* (2010) — The new German-Russian comprehensive dictionary. In 3 volumes [Novyy bol'shoy nemecko-russkiy slovar'. V 3 t.], ed. by D. O. Dobrovolskiy. Vol. II: G—Q, Moscow.
10. *RRR* (1978) — Russian spoken language. Texts [Russkaya razgovornaya rech'. Teksty], ed. by E. L. Zenskaya and L. A. Kapapadze, Moscow.
11. *Shimchuk È., Shchur M.* (1999), Dictionary of Russian particles [Slovar' russkikh chastits], Berlin.
12. *Vasmer M.* (1986–1987), Russian etymological dictionary. In 4 volumes [Ètimologicheskij slovar' russkogo yazyka. V 4 t.], 2nd edition, Moscow.
13. *Yakubinskiy L. P.* (1923), On dialogic speech [O dialogicheskoy rechi], The Russian language [Russkaya rech'], ed. by L. V. Shcherba, Petrograd, pp. 96–194.

# ВКЛАД КИНЕТИЧЕСКОЙ СОСТАВЛЯЮЩЕЙ В МУЛЬТИМОДАЛЬНУЮ КОММУНИКАЦИЮ, ИЛИ ЕЩЕ РАЗ О СТРАТЕГИЯХ ОПИСАНИЯ ТАНГРАММ<sup>1</sup>

**Федорова О. В.** (olga.fedorova@msu.ru)

МГУ имени М. В. Ломоносова, РАНХиГС, Москва, Россия

Мультимодальная коммуникация — слова, просодия, жестикуляция, позы, движения глаз — фундаментальная составляющая языка. Принято считать, что коммуникация более эффективна в том случае, когда она сопровождается жестикуляцией. Для более систематической проверки интерфейса «речь–жестикуляция» мы использовали метод референциальной коммуникации, первоначально разработанный Крауссом и модернизированный Кларком. В описываемом эксперименте два испытуемых сидели на татами, разделенные низкой или высокой перегородкой (фактор «видимость»), перед каждым был комплект из 12 карточек-танграмм. Инструктор, карточки которого были разложены в заданном порядке, должен был донести до Раскладчика, карточки которого были разложены в беспорядке, информацию о заданной последовательности. У каждой из 16 пар было по 4 попытки. Полученные описания (общей длительностью более 3 часов) были аннотированы в программе Praat на вокальном уровне и в программе ELAN на кинетическом уровне. Мы рассмотрели 6 конкретных гипотез: (1) любая кинетическая активность помогает коммуникации; (2) иконические жесты и позы помогают коммуникации; (3) иконические позы помогают коммуникации; (4) видимая кинетическая активность помогает коммуникации; (5) видимые иконические жесты и позы помогают коммуникации; (6) иконические жесты и позы, совершенные в 1-й попытке, помогают коммуникации. Однако ни одна из этих гипотез не подтвердилась. Наиболее правдоподобное объяснение полученных результатов связано с комплексным характером взаимодействия между речью и жестикуляцией: кинетический канал оказывает не аддитивный, а скорее интерактивный эффект.

**Ключевые слова:** мультимодальная коммуникация, жесты, позы, танграммы, речь

---

<sup>1</sup> Работа выполнена при финансовой поддержке РГНФ (грант 14-04-00390). Автор благодарен трем анонимным рецензентам, предложившим целый ряд важных улучшений данной статьи.



# THE CONTRIBUTION OF THE KINETIC COMPONENT TO THE MULTIMODAL COMMUNICATION, OR TANGRAM DESCRIPTION STRATEGIES REVISITED

**Fedorova O. V.** (olga.fedorova@msu.ru)

Lomonosov Moscow State University, RANEPa, Moscow, Russia

Multimodal communication — words, prosody, gesticulation, postures, eye gaze — is a fundamental part of language use. It is widely thought that a communication is more effective when it is accompanied by gesticulation than when it is not. In search of systematic evidence how the interface “speech–gesticulation” works we turn to the referential communication task originally devised by Krauss and specified by Clark. In our experiment, two participants were seated on tatami separated by a low or high plastic screen (the within-participants factor “visibility”), in front of each were twelve cards of so-called Tangram figures. For the Director the cards were already arranged in a target sequence, and for the Matcher the same figures lay in a random sequence. The Director’s job was to get the Matcher to rearrange her figures to match the target ordering. Thirty two participants carried out the task in four trials. All descriptions (over three hours long) were annotated, in Praat at the vocal channel and in ELAN at the kinetic channel. We consider six concrete hypotheses: (1) any kinetic activity benefits communication; (2) iconic gestures and postures benefit communication; (3) iconic postures benefit communication; (4) visible kinetic activity benefit communication; (5) visible iconic gestures and postures benefit communication; (6) iconic gestures and postures of the 1-st trial benefit communication. However, none of these hypotheses was confirmed. The most likely explanation of the obtained results is associated with the complex interface between speech and gesticulation: kinetic channel demonstrates not the additive effect, but the interaction one.

**Key words:** multimodal communication, gesture, posture, tangram, speech

## 1. Метод референциальной коммуникации «Танграммы»

В самых общих чертах суть метода референциальной коммуникации<sup>2</sup> состоит в том, что собеседники совместными усилиями решают некоторую когнитивную задачу, успешность решения которой во многом зависит от слаженности их вербального поведения. Инструктор (И.) видит нечто, что он должен вербально донести до второго собеседника, Раскладчика (Р.), который этого не видит.

---

<sup>2</sup> Мы будем использовать дословный перевод английского термина «referential communication», другой принятый перевод — «референтное общение» (Самойленко 2010).

Существует два способа проведения исследований — через специальный непрозрачный экран и по телефону, а также два типа задания — пройти определенным путем по лабиринту (методика «лабиринт») или разложить в заданном порядке предметы, первоначально лежащие в беспорядке (методика «беспорядок»).

Данный метод был разработан Р. Крауссом (Krauss, Weinheimer 1966), после чего многократно воспроизводился (на русском материале см. Беляева, Самойленко 1983; Самойленко 2010; Ананьева, Харитонов 2011). Большой вклад в усовершенствование методики «беспорядок» внес Г. Кларк в серии экспериментов с танграммами (Clark, Wilkes-Gibbs 1986) — фигурками из китайской игры-головоломки, которые с трудом поддаются вербальному описанию<sup>3</sup> (рис. 1).

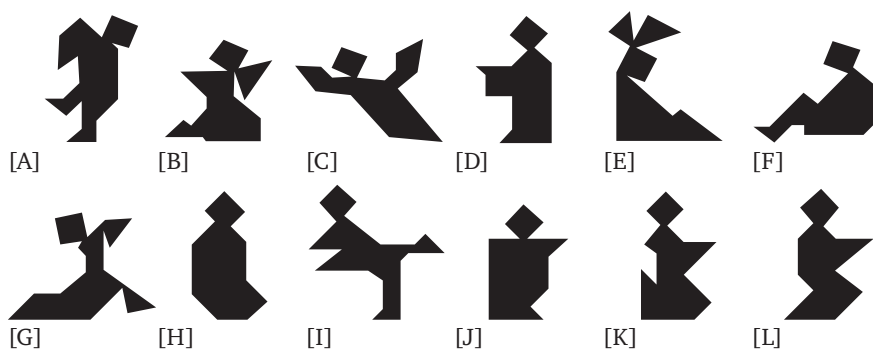


Рис. 1. Стимульный материал из Clark, Wilkes-Gibbs 1986

На русском материале метод «Танграммы» был использован в исследованиях с взрослыми испытуемыми (Федорова 2014), детьми 10–12 лет (Федорова и др. 2013) и пациентами с афазией (Зинова и др. 2011). В частности, в Федорова и др. 2013 было показано, что холистическая стратегия описания танграмм оказывается более успешной по сравнению с фрагментированной стратегией.

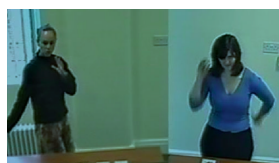
## 2. Мультиmodalность и метод «Танграммы».

### Новый корпус TanTatami<sub>rus</sub>

Как понятно из определения, традиционно метод референциальной коммуникации исключает визуальный контакт, поэтому большинство подобных исследований (обзор см. в Федорова 2014) было ограничено вокальной modalностью. Однако в работе Holler, Wilkin 2011 авторы модифицировали метод, предоставив испытуемым возможность в половине случаев видеть друг друга. Дизайн, который был использован при сборе корпуса TanTatami<sub>rus</sub>, максимально близок к исследованию Karpiński et al. 2015, при проведении

<sup>3</sup> Как пишет Кларк, идею использовать китайские танграммы ему подсказала А. В. Беляева во время его визита в Россию. Однако Кларк был не первым — согласно В. Лефельту, впервые танграммы были использованы в научных целях Э. Эспером еще в 1933 г. (Levelt 2013).

которого испытуемые сидели на полу на татами; эта модификация была выбрана из-за удобства фиксации кинетической активности испытуемых, см. рис. 2. Таким образом, испытуемые задействуют все каналы коммуникации, что дает возможность проводить полноценный анализ мультимодального (мультиканального по Кибрик 2017) общения.



Holler, Wilkin 2011



Karpiński et al. 2015



данное исследование

**Рис. 2.** Сопоставление дизайнов исследований

Корпус был собран летом 2016 г. и записан на видеокамеру Panasonic 4K HC-WX970, которая имеет два объектива и позволяет снимать сразу два видеоплана; результирующий видеофайл устроен по принципу «два в одном» (основной участник — И. — изображен на крупном плане, Р. — во встроенном окне, см. рис. 3), что решает проблему синхронизации видеофайлов. У каждой из 16 пар (29 женщин и 3 мужчины, средний возраст 22 года) было 4 попытки, в каждой из которых они раскладывали по 12 карточек-танграмм; в двух попытках испытуемый выступал в роли И., в двух других — Р.; в нечетных попытках собеседники видели друг друга (низкая перегородка, закрывавшая только карточки) в четных — не видели (высокая перегородка), рис. 3. Общая длительность корпуса составила 3 часа, средняя длительность записи 12 мин.



Попытка №1



Попытка №2



Попытка №3



Попытка №4

**Рис. 3.** Общий дизайн исследования

Собранный материал, включающий 16 записей, 64 попытки и 768 описаний танграмм, был расшифрован и аннотирован на вокальном уровне в программе Praat (<http://fon.hum.uva.nl/praat/>) и на кинетическом уровне в программе ELAN (<https://tla.mpi.nl/tools/tla-tools/elan/>).

### 3. Кинетическая составляющая мультимодальной коммуникации

Кинетическая составляющая мультимодальной коммуникации, или язык тела, имеет сложную организацию, которая включает как динамический компонент — различного рода движения и жесты: мимические жесты, жесты головы, рук и тела, так и статический: выражение лица, позы, проксемику. Несмотря на то, что в реальной коммуникации важны все составляющие, большинство современных исследований посвящены жестикуляции и в первую очередь иллюстративным мануальным жестам (Kendon 2004, McNeill 2005). Так, наиболее известная классификация мануальных жестов МакНилла (2005) включает указательные, иконические (изображающие конкретные объекты или действия), метафорические (представляющие абстрактные понятия) и ритмические (выделяющие важные фрагменты речи) жесты.

Взаимодействие речи и сопровождающей ее жестикуляции — один из самых обсуждаемых топиков в мультимодальных исследованиях; в частности, еще не решен вопрос о том, насколько речь и жесты связаны между собой в когнитивной системе человека. МакНилл утверждает, что речь и жесты одновременно активируются в общем источнике (McNeill 2005), а авторы модели Interface (Kita, Ozuyrek 2003), наоборот, полагают, что они планируются в разных модулях (Федорова и др. 2016). К «вечным» мультимодальным вопросам относятся и два следующих:

- упрощает ли жестикуляция коммуникацию?
- мы жестикулируем ради облегчения собственного речепроизводства или в целях лучшего понимания нашего сообщения адресатом?

Включая в традиционную проблематику танграмм кинетическую составляющую, данное исследование вносит определенный вклад в решение этих «вечных» вопросов.

#### 3.1. Кинетическая аннотационная схема

В данной работе мы выделяем 3 типа кинетических явлений:

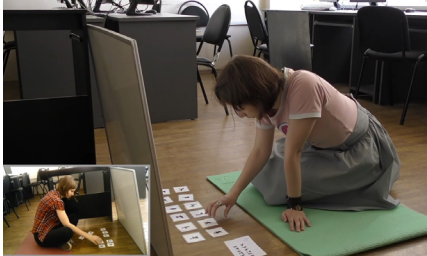
**1. Вспомогательная кинетическая активность (auxiliary).** К этому типу относятся ритмические жесты по МакНиллу, указательные жесты рук (рис. 4а), головы и тела, действия с карточками (рис. 4б и 4в) и наклоны туловища (рис. 4г). Все обозначенные движения объединяет то, что они ничего не сообщают о семантике танграмм, в отличие от других типов. Важно отметить также, что мы выделяем не только жесты рук, но и жесты всего тела. Кроме того, практически все подобные движения носят повторяющийся характер.



а. показывает



б. берет в руки



в. поправляет



г. наклоняется

**Рис. 4.** Вспомогательная кинетическая активность

**2. Иконические жесты (iconic gesture).** К данному типу мы относим только мануальные жесты, изображающие танграммы, рис. 5.



№1



№6



№7



№11

**Рис. 5.** Танграмма J (см. рис. 1): иконические жесты

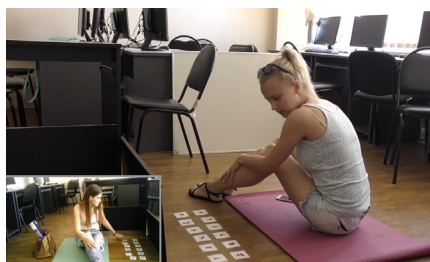
**3. Иконические позы (iconic posture).** Насколько нам известно, данный тип кинетической активности раньше не выделялся (ср., напр., деление на указательные и иконические жесты в de Ruiter et al. 2012). В отличие от иконических жестов, позы изображаются всем телом и являются статичными — человек замирает в выбранной позе на некоторое время (рис. 6).



№5



№1



№12



№14

**Рис. 6.** Танграмма F (см. рис. 1): иконические позы

В ходе кинетической аннотации поведения И.<sup>4</sup> каждому описанию был поставлен в соответствие один из пяти зафиксированных вариантов: ‘0’ (нет активности); ‘Aux’ (только вспомогательные жесты); ‘Aux & Icon Gesture’ (вспомогательные и иконические жесты), ‘Aux & Icon Posture’ (вспомогательные жесты и иконические позы) и ‘Aux & Icon Gesture & Icon Posture’ (все три типа)<sup>5</sup>. Затем мы суммировали цифры для каждой записи и для всех записей в целом (табл. 1).

<sup>4</sup> В этой работе мы анализируем только кинетическую активность И.; совместная кинетическая активность И. и Р. — тема отдельного исследования в будущем.

<sup>5</sup> В собранном материале иконические жесты и позы всегда сопровождалось вспомогательными движениями; насколько это действительно универсальное явление, подтвердят будущие исследования.

**Таблица 1.** Распределение жестикуляции И. по записям (числами обозначено количество описаний танграмм, сопровождаемых жестами данного типа, из 48 возможных)

№	0	Aux	Aux & Icon Gesture	Aux & Icon Posture	Aux & Icon Gesture & Icon Posture
1	13	25	6	0	4
2	1	37	1	9	0
3	0	27	13	5	3
4	44	4	0	0	0
5	29	6	1	10	2
6	5	21	16	5	1
7	31	7	5	2	3
8	12	27	5	2	2
9	5	17	11	5	10
10	10	30	7	1	0
11	15	11	9	7	6
12	0	15	19	6	8
13	5	29	7	2	5
14	32	8	3	5	0
15	3	31	12	2	0
16	3	25	9	8	3
sum	208	320	124	69	47

Переформулируем теперь 1-й из «вечных» вопросов в виде трех гипотез, от более общей к более частным:

H1. Любая кинетическая активность помогает коммуникации

H2. Иконические жесты и позы помогают коммуникации

H3. Иконические позы помогают коммуникации.

Что касается 2-го «вечного» вопроса, мы конкретизируем его в виде двух гипотез:

H4. Видимая P. любая кинетическая активность помогает коммуникации

H5. Видимые P. иконические жесты и позы помогают коммуникации<sup>6</sup>.

Наконец, в заключительном разделе мы проверим гипотезу о роли жестикуляции в 1-й попытке:

H6. Иконические жесты и позы, совершенные в 1-й попытке, помогают коммуникации.

Чтобы ответить на эти вопросы, нам осталось операционализировать понятие «успешности».

<sup>6</sup> Гипотеза об упрощающем эффекте видимых иконических поз отдельно не проверялась вследствие недостаточности собранного материала.

### 3.2. Понятие «успешности» и способы ее измерения

Инструкция испытуемым содержала задание как можно быстрее, но при этом правильно, разложить карточки Р. в том же порядке, в котором они лежат у И. В данной работе для определения успешности мы независимо измерили две эти переменные: длительность записи и правильность выполнения задачи, см. табл. 2.

**Таблица 2.** Длительность, правильность, успешность и гипотезы (в колонках Н1–Н6 обозначено количество описаний танграмм, сопровождаемых жестами данного типа, из 48 возможных)

№	длительность	правильность	успешность	Н1	Н2	Н3	Н4	Н5	Н6
1	6:43	48	1	35	10	4	20	9	9
2	14:52	46	4	47	10	9	23	7	3
3	12:36	45	3	48	21	8	24	13	11
4	6:59	45	3	4	0	0	2	0	0
5	7:47	48	1	19	13	12	12	10	8
6	10:21	45	3	43	22	6	23	16	10
7	10:30	41	4	17	10	5	10	10	9
8	10:41	47	2	36	9	4	22	7	6
9	17:17	43	4	43	26	15	20	18	12
10	8:56	46	2	38	8	1	21	7	4
11	13:06	45	3	43	22	13	20	16	11
12	19:14	40	4	48	33	14	24	19	8
13	12:46	46	3	43	14	7	22	13	9
14	8:12	48	1	16	8	5	15	8	7
15	13:53	46	3	45	14	2	24	12	9
16	12:37	47	3	45	20	11	21	14	10

**1. Длительность.** По этому параметру 16 записей логичным образом образовали 4 группы:

- 1) меньше 9 мин.: 1, 4, 5, 10 и 14
- 2) от 9 до 11 мин.: 6, 7 и 8
- 3) от 12 до 13 мин.: 3, 11, 13, 15 и 16
- 4) больше 14 мин.: 2, 9 и 12.

**2. Правильность.** По этому параметру мы разделили 16 записей на такие 4 группы:

- 1) нет ошибок: 1, 5 и 14
- 2) 1–2 ошибки: 2, 8, 10, 13, 15 и 16
- 3) 3–4 ошибки: 3, 4, 6 и 11
- 4) 5 и больше ошибок: 7, 9 и 12.

Наконец, мы выделили 4 группы по параметру «успешность»:

- 1) значения обеих переменных находятся в 1й группе: 1, 5 и 14
- 2) значения обеих переменных находятся в 1й или 2й группах: 8 и 10



- 3) значения находятся в 1й, 2й или 3й группах: 3, 4, 6, 11, 13, 15 и 16  
4) хотя бы одно значение находится в 4й группе: 2, 7, 9 и 12.

Теперь мы можем проверить сформулированные выше гипотезы.

#### **4. «Вечный» вопрос 1: Жестикуляция упрощает коммуникацию**

Хотя некоторые исследователи дают отрицательный ответ на этот вопрос (Krauss et al. 1996), большинство полагает, что жестикуляция помогает коммуникации; в частности, об этом свидетельствует мета-анализ 63 исследований в Hostetter 2011. Некоторые исследования в данной выборке проводились с заранее записанными клипами (Kelly, Church 1997), в некоторых исследовалось «живое» общение (Holler et al. 2009); в некоторых испытуемые отвечали на вопросы на понимание (Kelly, Church 1997), в других их просили воспроизвести услышанное (Alibali et al. 1997). В целом, авторы наблюдали упрощающий эффект жестикуляции: используя принципиально другой — недискретный (Kibrik 2015) — формат, жестикуляция (1) дублирует информацию, которая передается по вокальному каналу, или (2) передает дополнительную, а иногда и противоречащую, информацию (во втором случае сила эффекта обычно больше).

Обратимся теперь к трем сформулированным гипотезам. Первая из них (H1) утверждает, что любая жестикуляция И. — как несущая семантическую составляющую, так и нет — помогает коммуникантам успешнее справиться с заданием. Проверив эту гипотезу, используя коэффициент ранговой корреляции Спирмена, мы получили незначимую отрицательную корреляцию  $\text{cor} = -0.4199$ ,  $p\text{-value} > 0.05$  (при  $N=16$  уровень значимости достигается при значении 0.5 для  $p\text{-value} < 0.05$  и при значении 0.64 для  $p\text{-value} < 0.01$ ). Таким образом, данный результат можно рассматривать на уровне отрицательной тенденции (чем больше жестикуляции, тем хуже успешность), но ни о каком положительном эффекте не может быть и речи.

Вторая гипотеза (H2) исходит из предложения, что только иконические (семантически нагруженные) жесты и позы способствуют успешной коммуникации. Эта гипотеза также не подтвердилась, имея почти такой же по величине отрицательный эффект  $\text{cor} = -0.4126$ ,  $p\text{-value} > 0.05$ .

Третья гипотеза (H3) априори представлялась нам наиболее перспективной — когда И. изображает Р. танграмму всем телом, используя иконическую позу, это может заметно облегчать дальнейшее взаимодействие, как кажется на первый взгляд. Но и эта гипотеза не подтвердилась, лишь немного уменьшилось значение корреляции  $\text{cor} = -0.3831$ ,  $p\text{-value} > 0.05$ .

Подобные тенденции можно заметить и невооруженным статистикой глазом, посмотрев распределение в табл. 2. Так, с одной стороны, в трех наиболее успешных записях (№1, 5 и 14) распределение жестов очень разное; с другой стороны, в самой неуспешной по обоим параметрам записи №12 мы наблюдаем максимальное количество жестов.

Таким образом, полученные данные дают отрицательный ответ на поставленный вопрос.

## 5. «Вечный» вопрос 2: Жестикуляция «для себя» или «для других»

Существует несколько теорий, объясняющих, зачем говорящий человек может жестикулировать «для себя». Согласно гипотезе информационной упаковки (Information packaging, Kita 2000), жестикуляция помогает человеку упаковывать зрительно воспринимаемую информацию в формате, совместимом с линейным и дискретным форматом вокального канала. С другой стороны, модель лексического упрощения (Lexical facilitation, Krauss et al. 2000) утверждает, что первичная функция иконических жестов состоит в упрощении для говорящего поиска в ментальном лексиконе подходящей лексемы. В частности, в работе Morsella, Krauss 2004 было показано, что когда лексический доступ к слову затруднен, говорящий начинает больше жестикулировать; в то же время, при запрете жестикуляции речь становится менее плавной (Rauscher et al. 1996).

Таким образом, в общем случае предполагается, что «для себя» человек жестикулирует. Но встает вопрос, жестикулирует ли он «для других»? Точнее, является ли жестикуляция «для других» побочным продуктом жестикуляции «для себя», или человек намеренно жестикулирует для собеседника? В работе Mol et al. 2011 авторы подсчитали количество жестов в тех случаях, когда коммуниканты могли и не могли видеть друг друга; в первом случае жестов было значимо больше.

Обратимся к рис. 7. На всех скриншотах испытуемые изображают танграммы при помощи иконических поз в ситуации, когда адресат не может их видеть через высокую перегородку, то есть И. или (1) изображает танграммы для себя, или (2) изображает их для Р., но забывает про перегородку.



№2



№9



№11



№12

Рис. 7. Иконические позы: невидимость

Рассмотрим теперь только те случаи, когда Р. может видеть жестикуляцию И. (1я и 3я попытки). Сначала проанализируем всю видимую кинетическую активность И. (Н4). Используя коэффициент ранговой корреляции Спирмена, получаем совсем незначимый отрицательный эффект  $\text{cor} = -0.1949$ ,  $p\text{-value} > 0.05$ ; другими словами, видимая кинетическая активность Инструктора никак не влияет на успешность коммуникации. Подсчитав отдельно только иконические жесты и позы (Н5) получаем незначимую отрицательную тенденцию  $\text{cor} = -0.3537$ ,  $p\text{-value} > 0.05$ . Таким образом, и на второй «вечный» вопрос, переформулированный в виде гипотез Н4 и Н5, мы получаем отрицательный ответ — если И. и старается помочь Р. своими жестами, к более успешной коммуникации это не приводит<sup>7</sup>.

## 6. Обсуждение результатов и заключение

Итак, подсчитав корреляции между успешностью выполнения совместной когнитивной задачи и кинетической активностью И. в разных условиях, мы не получили значимой взаимозависимости. Попробуем, однако, провести более тонкий анализ и предположить причины таких результатов. Во-первых, можно заметить, что вся жестикуляция в целом, а в первую очередь семантически нагруженная жестикуляция, и особенно иконические позы, отнимают у И. дополнительное время, тем самым ухудшая параметр успешности. Оптимальная стратегия, по-видимому, состоит в использовании жестикуляции только в наиболее важные (сложные) моменты коммуникации. Можно заметить такую закономерность, что все пять лучших записей объединяет то, что в них иконические жесты и позы используются почти только в 1-й попытке; однако подсчитав корреляцию (Н6), мы опять получаем слабый отрицательный эффект  $\text{cor} = -0.1662$ ,  $p\text{-value} > 0.05$ . Кроме того, из табл. 2 видно, что, напр., в записи №7 жестикуляция тоже наблюдается только в 1-й попытке, но эта запись входит в самую неуспешную группу. Таким образом, по-видимому, использование иконических жестов и поз только в 1-й попытке, возможно, является необходимым, но точно не является достаточным условием успешной коммуникации.

Во-вторых, сопоставим данные кинетической активности с вокальной. Как представляется, наиболее важная задача для успешной коммуникации при выполнении этого задания в целом — совместными усилиями подобрать удачные ярлыки для каждой танграммы. В некоторых случаях этот ярлык удастся подобрать без помощи жестикуляции, а в некоторых бывает достаточно один раз изобразить танграмму иконическими жестами / позой и закрепить (назвать на вокальном уровне) на остальные три попытки. В большинстве неуспешных записей с большим количеством неэффективной жестикуляции это вербальное закрепление или отсутствует — испытуемые каждый раз пытаются обходиться только жестами, или в 3-й попытке сменившийся И. предлагает новый ярлык. Рассмотрим это на примере одной из двух самых сложных танграмм В (см. рис. 1).

<sup>7</sup> Этот вывод сделан по результатам анализа статистических корреляций; тем не менее, при внимательном изучении данных в табл. 1 и 2 можно увидеть интересные закономерности, которые не подтверждаются строгой статистикой.



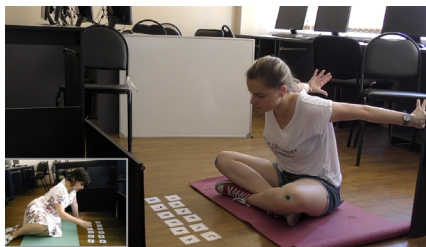
а. №1, В, п. 1



б. №1, В, п. 2



в. №11, В, п. 1



г. №11, В, п. 3

**Рис. 8.** (Не)успешные стратегии описания: танграмм В

В самой успешной записи №1 в 1-й попытке И. изображает танграмму В при помощи иконической позы (рис. 8а), закрепляя это на вокальном уровне (1). Во 2-й попытке И. обходится уже простым указательным жестом (рис. 8б), а далее вообще без жестикуляции.

(1) №1, В

п. 1

И: Это аа(0.2) похоже на= я не знаю= женщину, которая раскинула руки в две стороны(0.5), а ноги в виде ромба(0.4). И она сидит как будто бы на коленях. И у неё белый галстучек на шее.

Р: Угу.

п. 2

И: Так(0.5), вторая — (0.2) этоо(0.2) женщина с ручками и галстучком.

Р: Угу.

п. 3

И: Потом (0.6)женщина с галстучком.

(1.3)

Р: Угу.

п. 4

И: Потом женщина с галстучком.

(1.1)

Р: Угу.

В одной из неуспешных записей №11 И. в 1-й попытке также показывает танграмму В иконической позой и закрепляет вербально (рис. 8в), однако делает это неудачно, так что Р. ошибочно выбирает вместо танграммы В танграмму G (см. рис. 1) — вторую из двух самых сложных танграмм. В ходе 1-й попытки

Р. самостоятельно исправляется, поэтому во 2-й попытке сразу выбирает правильно. Однако в 3й попытке новый И. предлагает свое новое описание, изображая его новой иконической позой (рис. 8г). Р. тоже путает эту танграмму с танграммой G, дальше в ходе долгой дискуссии они раскладывают карточки правильно, однако в 4-й попытке каждая снова предлагает свое описание.

(2) №11, В

п. 1

И: Мм(0.8) (1.6)вторая — наа(0.6) (0.7)сидящего на коленях человека с квадратной головой и чем-то треугольным в зубах.

И: Вот силуэт такой (1.0)ну (1,3)вот такой(0.7) и ещё какой-то треугольничек спереди.

(1.2)

Р: Да.

п. 2

И: Аа(0.5) вторая — стоящий на коленях человек с треугольником в зубах. (0.5)

Р: Мм(2.7) (1.0)да.

п. 3

И: Ну, похоже на птицу. У неё такой треугольный клюв и крылья у неё за спиной.

Р: Ээ(0.5) стоп. (0.4)Крылья за спиной.

И: Да, такие крылья треугольные. (1.0)И хвост ещё сзади. (0.7)

Р: Ээ(0.5) (0.4)это то, про что я говорила, что оно похоже на петуха? (0.3)

И: Нет, [не на петуха].

Р: [Аа(0.7)], (0.4)всё, поняла.

И: [Нашла, да]?

Р: [Загогулина] с треугольниками и квадратом.

И: Угу.

п. 4

И: Так, (0.3)птица с= с треугольником во рту. (1.4)

Р: Которая сидящая на коленях человек?

И: Мм(0.3) может быть и похоже. (0.3)

Р: Да. (0.6)Есть.

Главный методологический итог данного исследования, таким образом, состоит в признании того факта, что только при выполнении самых простых когнитивных задач можно анализировать и сравнивать воздействие информационных каналов по отдельности, аддитивно, т. е. независимо оценивая вклад вокальной и кинетической составляющих. В более сложных задачах, к числу которых принадлежит эксперимент с танграммами, такое прямолинейное разделение оказывается искусственным и не приводит к значимым результатам.

## Литература

1. *Alibali M. W., Flevares L., Goldin-Meadow S.* (1997), Assessing knowledge conveyed in gesture: Do teachers have the upper hand?, *Journal of Educational Psychology*, Vol. 89, pp. 183–193.
2. *Ananieva K. I., Kharitonov A. N.* (2011), Methodological problems of oculomotor activity research in communicative situations [Metodologicheskie problemy issledovaniya okulomotornoy aktivnosti v kommunikativnykh situatsiyakh], N. L. Nagibina (ed.) *Psychological and psychoanalytic studies 2010–2011* [Psikhologicheskie i psikhoanaliticheskie issledovaniya 2010–2011], Moscow Institute of Psychoanalysis, Moscow, pp. 36–49.
3. *Belyaeva A. V., Samoilenko E. S.* (1983), The monologue and dialogue in the problems of verbalization image [Monolog i dialog v zadachakh verbalizatsii obraza], *Psychological studies of cognitive processes and personality* [Psikhologicheskie issledovaniya poznavatelnykh processov i lichnosti], Moscow, pp. 106–121.
4. *Clark H. H., Wilkes-Gibbs D.* (1986), Referring as a collaborative process, *Cognition*, Vol. 22(1), pp. 1–39.
5. *de Ruiter J. P., Bangerter A., Dings P.* (2012), The Interplay between gesture and speech in the production of referring expressions: Investigating the tradeoff hypothesis, *Topics in Cognitive Science*, Vol. 4, pp. 232–248.
6. *Fedorova O. V.* (2014), *Experimental discourse analysis* [Eksperimentalnyy analiz diskursa]. *Yazyki Slavyanskoy Kultury*, Moscow.
7. *Fedorova O. V., Delikishkina E. A., Slabodkina T. A., Tsipenko A. A.* (2013), Dialogue modelling in psycholinguistics: literal and analogical perspectives as bases for adults' and children's references [Modelirovanie dialoga v psikholingvistike vzroslye i detskie strategii opisaniya obektov deystvitelnosti], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2013”* [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2013”], RGGU, Moscow, pp. 230–238.
8. *Fedorova O. V., Kibrik A. A., Korotaev N. A., Litvinenko A. O., Nikolaeva Yu. V.* (2016), Temporal coordination between gestural and speech units in multimodal communication [Vremennaya koordinatsiya mezhdu zhestovymi i rechevymi edinicami v multimodalnoy kommunikatsii], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2016”* [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2016”], RGGU, Moscow, pp. 159–170.
9. *Holler J., Shovelton H., Beattie G.* (2009), Do Iconic Hand Gestures Really Contribute to the Communication of Semantic Information in a Face-to-Face Context? *Journal of Nonverbal Behavior*, Vol. 33(2), pp. 73–88.
10. *Holler J., Wilkin K.* (2011), Co-Speech Gesture Mimicry in the Process of Collaborative Referring During Face-to-Face Dialogue, *Journal of Nonverbal Behavior*, Vol. 35, pp. 133–153.
11. *Hostetter A. B.* (2011), When do gestures communicate? A meta-analysis, *Psychological Bulletin*, Vol. 137(2), pp. 297–315.
12. *Karpiński M., Jarmołowicz-Nowikow E., Czoska A.* (2015), Gesture annotation scheme development and application for entrainment analysis in task-oriented dialogues in diverse cultures, *GESPIN 4*, Nantes, 2–4 September 2015, pp. 161–166.

13. *Kelly S. D., Church R. B. (1997), Can children detect conceptual information conveyed through other children's nonverbal behaviors?, Cognition and Instruction, Vol. 15, pp.107–134.*
14. *Kendon A. (2004), Gesture. Visible action as utterance, Cambridge.*
15. *Kibrik A. A. (2015), The problem of non-discreteness and spoken discourse structure, Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2015” [Komp'yuternaya Lingvistika i Intellekturnye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2015”], RGGU, Moscow, pp. 231–240.*
16. *Kibrik A. A. (2017), Russian multichannel discourse as a promising object of research [Russkiy mul'tikanal'nyy diskurs kak perspektivnyy objekt issledovaniya], Psychological journal [Psikhologicheskii zhurnal], in press.*
17. *Kita S. (2000), How representational gestures help speaking, D. McNeill (ed.) Language and gesture, Cambridge University Press; Cambridge, UK, pp. 162–185.*
18. *Kita S., Ozyurek A. (2003), What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: evidence for an interface representation of spatial thinking and speaking, Journal of Memory and Language, 48 (1), pp. 16–32.*
19. *Krauss R. M., Chen Y., Chawla P. (1996), Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? Advances in Experimental Social Psychology, Vol. 28, pp. 389–450.*
20. *Krauss R. M., Chen Y., Gottesman R. (2000), Lexical gestures and lexical access: A process model, D. McNeill D (ed.) Language and gesture, Cambridge University Press, Cambridge, UK, pp. 261–283.*
21. *Krauss R. M., Weinheimer S. (1966), Concurrent feedback, confirmation, and the encoding of referents in verbal communication, Journal of Personality and Social Psychology, Vol. 4.*
22. *Levelt W. J. M. (2013), A History of Psycholinguistics: The pre-Chomskyan era, Oxford University Press, Oxford.*
23. *McNeill D. (2005), Gesture and thought, Chicago.*
24. *Mol L., Krahmer E., Maes A., Swerts M. (2011), Seeing and being seen: The effects on gesture production, Journal of Computer-Mediated Communication, Vol. 17, pp. 77–100.*
25. *Morsella E., Krauss R. M. (2004), The role of gestures in spatial working memory and speech, American Journal of Psychology, Vol. 117, pp. 411–424.*
26. *Rauscher F. H., Krauss R. M., Chen Y. (1996), Gesture, speech, and lexical access: The role of lexical movements in speech production, Psychological Science, Vol. 7, pp. 226–231.*
27. *Samoilenko E. S. (2010), Problems of comparison in psychological research [Problemy sravneniya v psikhologicheskom issledovanii]. Publishing House of the IP, RAS, Moscow.*
28. *Zinova Ju. A., Dragoy O. V., Fedorova O. V. (2011), Experimental study of verbal interaction: language pathology data [Eksperimentalnoe issledovanie rechevogo vzaimodeystviya: dannye yazykovoy patologii], Vestnik MGU, Ser. 9. Philology, 4, pp. 167–175.*

# СЕМАНТИЧЕСКИЕ КЛАССЫ И УПРАВЛЕНИЕ ПРИЛАГАТЕЛЬНЫХ<sup>1</sup>

**Гращенко П. В.** (pavel.gra@gmail.com)

МГУ, ИВ РАН, МПГУ, Москва, Россия

**Кобозева И. М.** (kobozeva@list.ru)

МГУ, Москва, Россия

## SEMANTIC CLASSES AND GOVERNMENT OF ADJECTIVES

**Grashchenkov P. V.** (pavel.gra@gmail.com)

MSU, IOS RAS, MSUE, Moscow, Russia;

**Kobozeva I. M.** (kobozeva@list.ru)

MSU, Moscow, Russia

The paper aims at establishing the list of semantic classes of Russian adjectives. Another goal is to enumerate semantic roles relevant for the participants of Russian adjectives and the surface tools for coding such adjectival participants. The discussion of adjectival classes starts from the classifications proposed in Western and Russian linguistics. Such classifications are often based on pure semantic grounds and sometimes have rather speculative character. To overcome such shortcomings, the paper takes into consideration participants introduced by situations described by adjectives. Such participants can be of two types. Participants of the first type are relevant for specific semantic classes of adjectives. The second type participants can be introduced by adjectives from different semantic classes (and under certain circumstances—by the adjectives that lack government at the first sight). The overall conclusion is that adjectival government is more idiosyncratic and “more semantic” in nature than the verbal one.

**Keywords:** adjectives, semantic classes, semantic roles, Russian, government

---

<sup>1</sup> Исследование выполнено в рамках проекта РНФ № 16-18-02003



## 1. Семантические классы прилагательных

Различные подходы к семантической классификации прилагательных неоднократно предлагались как в отечественной, так и в зарубежной лингвистике, см. (Bolinger 1967), (Givon 1970), (Dixon 1977), (Шрамм 1979), (Харитончик 1990), (Богуславская 2005), (Кустова 2009), (Карпова и др. 2009), (Кюсева и др. 2013).

Одной из наиболее влиятельных работ, так или иначе затрагивающих вопросы семантической классификации прилагательных, является статья (Dixon 1977). Выделение семантических классов прилагательных проводилось как на основе свойств английских прилагательных, так и с учетом типологического материала. Диксоном было установлено 7 классов прилагательных с близкими семантическими и грамматическими свойствами:

(1) Семантические классы прилагательных по (Dixon 1977)

**Размер** (*большой, маленький, длинный, короткий, широкий, узкий,...*)

**Физические свойства** (*жесткий, мягкий, тяжелый, легкий,...*)

**Цвет** (*черный, белый, красный,...*)

**Состояния человека** (*ревнивый, счастливый, добрый, умный, веселый, жестокий,...*)

**Время** (*новый, молодой, старый,...*)

**Оценка** (*хороший, плохой, правильный, совершенный,...*)

**Скорость** (*быстрый, медленный,...*)

Другой важной работой, посвященной данной проблеме, является монография (Шрамм 1979). Эта работа основана на анализе 10000 прилагательных, значения которых близки к так называемым «качественным прилагательным». Качественные прилагательные разбиваются на эмпирические и рациональные. В дальнейшем эмпирические прилагательные делятся на типы в зависимости от органа восприятия. Рациональные прилагательные подразделяются на классы, связанные с типом определяемого объекта.

Шрамм также выделяет типы значений прилагательных, «отображающих несобственные признаки предметов», или иначе класс оценочных прилагательных. Выделяются такие значения как (НЕ)СООТВЕТСТВИЕ ОПРЕДЕЛЕННОМУ СТАНДАРТУ (*хороший, отличный,...*), ОСНОВАННЫЙ НА ЧЕМ-ТО (*рационалистический, рассудочный,...*), ПРИЧИНЯЮЩИЙ (*приятный, полезный,...*), ВЫРАЖАЮЩИЙ НЕЧТО (*умный, унылый,...*), МОДАЛЬНЫЕ значения (*возможный, допустимый, нужный,...*) и т. д.

В статье (Харитончик 1990) рассматриваются примеры различных классификаций и делается вывод о множественности вариантов ad hoc классификаций и сложности построения универсальной классификации на каком-либо одном основании. Действительно, если мы хотим, чтобы семантическая классификация каким-либо образом отражала не только логическую структуру адъективного класса, но и представляла собой инструмент для описания

их грамматических свойств, выделение классов должно быть более чувствительно к грамматическим свойствам лексем. Такой подход предлагается, например, в (Bolinger 1967), где прилагательные разделяются на предикативные и атрибутивные в зависимости от их поведения и интерпретации в том или ином контексте. Похожее разделение предлагается также в (Givón 1970), где выделяются глагольные (verb-based) и именные (noun-based) прилагательные, ср. также (Кустова 2006).

Часто прилагательным — в меньшей степени, чем глаголам — свойственно задавать ситуации, имеющие более одного участника и обладающие свойством изменения во времени. Рассмотрим следующий пример из интернета: *Фердинанд сначала был недоволен такой ролью, но скоро убедился, что она могла доставить ему и почет, и славу. У прилагательного недоволен в данном примере есть не только участник-носитель выражаемого качества (Фердинанд), но и второй участник, причина недовольства (такой ролью). Наречие сначала предполагает дальнейшее развитие ситуации, вводимой прилагательным.*

Мы попытаемся сгруппировать русские адъективные лексемы, учитывая не только их семантику, но и грамматику, а именно — то, какое поверхностно-синтаксическое выражение получают участники задаваемых ими ситуаций. Здесь мы приведем основные классы управляющих прилагательных, а затем подробно опишем семантико-синтаксические свойства членов разных классов, связанные с управлением.

Введение семантических классов, ориентированных не только на значение, но и на некоторое его грамматическое отражение, удобно и для решения теоретической проблем, и в прикладной лингвистике, ср. подход в духе Московской семантической школы, где требование идентичности валентностей есть условие синонимичности лексем и т.д., см (Апресян 1974: 223).

Сделаем еще два важных уточнения относительно принципов выделения классов прилагательных. Во-первых, предлагаемая здесь классификация не затрагивает адъективных классов МЕСТО (*новгородский, речной,...*), НАЦИОНАЛЬНОСТЬ (*немецкий, японский,...*), МАТЕРИАЛ (*каменный, золотой,...*) и некоторых других, традиционно называемых относительными прилагательными. Эти значения в языках мира также зачастую связаны с группами лексем, выделяемых на основании как грамматических, так и семантических критериев. Мы, однако, оставим относительные прилагательные в стороне по причине их слабой валентной способности (последнее не относится к рассматриваемым ниже относительным прилагательным СИММЕТРИЧНЫХ СВОЙСТВ).

Во-вторых, как это бывает с любой классификацией объектов реального мира, мы будем активно пользоваться понятием прототипа, см. (Rosch 1978). Выделяемые классы прилагательных будут определяться относительно некоторых ядерных, прототипических членов класса, т.е. таких, которые обладают наибольшим количеством релевантных для всего класса свойств. Наряду с ними имеются и периферийные члены класса, у которых таких базовых свойств меньше.

Итак, вот список основных семантических классов прилагательных, которые мы считаем релевантными для описания адъективного управления (в списке включены наиболее типичные члены класса):

- (2) Семантические классы русских поливалентных прилагательных, т. е. обладающих какими-либо иными семантическими валентностями помимо валентности **носителя свойства** или **экспериенцера** состояния:
- а. СОБСТВЕННО ОЦЕНОЧНЫЕ  
*важный, второстепенный, гадкий, жуткий, забавный, загадочный, заманчивый, занятный, заразительный, интересный, омерзительный, отвратительный, скучный, удачный, ...*
  - б. ПРОСТРАНСТВЕННЫЕ  
*ближний, близкий, далекий, дальний*
  - с. СИММЕТРИЧНЫХ СВОЙСТВ  
*знакомый (с кем-л.), идентичный, одинаковый, параллельный, похожий, равный, различный, разный, схожий, тождественный, эквивалентный*
  - д. НЕСИММЕТРИЧНЫХ ОТНОШЕНИЙ  
*полушный, покорный, (не)зависимый, главный, старший ...*
  - е. ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ  
*агрессивный, безразличный, беспристрастный, благожелательный, вежливый, враждебный, гостеприимный, грубый, доброжелательный, добрый, дружелюбный, заботливый, милосердный, ...*
  - ф. ЭМОЦИОНАЛЬНОГО И ФИЗИЧЕСКОГО СОСТОЯНИЯ ЧЕЛОВЕКА  
*гордый (чем-л.), довольный, рад, сердитый, счастливый, усталый, ...*
  - г. БЕНЕФАКТИВНОЙ ОЦЕНКИ<sup>2</sup>  
*вредный, выгодный, опасный, полезный, удобный, ценный, чреватый*
  - h. ИЗВЕСТНОСТИ  
*знакомый (кому-л.), знаменитый, известный, популярный, прославленный*
  - i. ТИПИЧНОСТИ  
*обыкновенный, обычный, привычный, свойственный, специфический, типичный, характерный, ...*
  - j. НАЛИЧИЯ / ОТСУТСТВИЯ  
*богатый, бедный, полный, пустой*
- Одни классы включают небольшое количество лексем, другие (прежде всего — СОБСТВЕННО ОЦЕНОЧНЫЕ прилагательные, прилагательные ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ и ТИПИЧНОСТИ) легко могут быть дополнены. Если выходить за рамки описания адъективного управления, можно добавить следующие важные (под)классы значений прилагательных:
- (2') Семантические классы русских прототипических моновалентных прилагательных, т. е. имеющих единственную валентность на носителя свойства):

<sup>2</sup> См. [Арутюнова 1988, 68]

к. ВРЕМЕННЫЕ

*долгий, краткий, новый, старый*

l. РАЗМЕРА

*длинный, короткий, высокий, низкий, широкий, узкий, глубокий, протяженный, ...*

м. ФИЗИЧЕСКИХ СВОЙСТВ

ЦВЕТА (*черный, красный, зеленый, ...*)

ОСЯЗАНИЯ (*горячий, сухой, мягкий, ...*)

СЛУХА (*тихий, громкий, беззвучный, ...*)

ОБОНЯНИЯ (*медвяный, чадный, затхлый, ...*)

ВКУСА (*сырой, сладкий, соленый, горький, ...*)

ВЕСА (*твердый, легкий, крепкий, мягкий, ...*)

п. ЧЕРТ ХАРАКТЕРА И СПОСОБНОСТЕЙ ЧЕЛОВЕКА

*добродушный, злой, умный, веселый, ревнивый, скупой, смелый, храбрый, ...*

Прилагательные, лишённые валентности, могут переходить в другой семантический класс за счёт введения нового участника. Так происходит, например, с прилагательными ЧЕРТ ХАРАКТЕРА (*он злой*), которые, получая обязательную роль причины / стимула, переходят в класс ЭМОЦИОНАЛЬНОГО И ФИЗИЧЕСКОГО СОСТОЯНИЯ ЧЕЛОВЕКА (*Он зол на меня*) или ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ (*...она знала, что тот редко был зол к своим подданным*), в целом примеры такой «межклассовой конверсии» достаточно часты. Подобно этому поливалентные прилагательные могут изменять тип участника, переходя в другой класс (см. примеры ниже).

Ниже мы обсудим семантическую организацию отдельных классов и характерные для них актанты, а также примеры семантических участников, сочетающихся с разными классами прилагательных. Под семантическими участниками мы будем понимать как актантные, так и сирконстантные зависимые (тем более, что отделить одни от других в случае прилагательных не всегда бывает возможно).

## 2. Управление отдельных семантических классов

Количество работ, посвященных моделям управления (аргументной структуре) прилагательных невелико по сравнению с работами по глагольному управлению, см., однако: (Вольф 1978), (Stowell 1991), (Булыгина, Шмелев 1997: 58–73), (Bennis 2004), (Landau 2006), (Кустова 2006, 2016), (Meltzer-Asscher 2011), (Lyashevskaya & Kashkin 2016) и приведенные там ссылки. Ниже мы представим попытку систематизации семантических ролей прилагательного.

Примерами прилагательных, чьи участники в той или иной степени однозначно определяются семантическим классом, могут служить прилагательные СИММЕТРИЧНЫХ СВОЙСТВ и ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ, а также ПРОСТРАНСТВЕННЫЕ прилагательные.

## 2.1. Участники у прилагательных класса симметричных свойств

Прилагательные данного класса являются симметричными предикатами (см. [Июмдин 1981]). Это обуславливает синонимию типа  $\text{Syn}(A \text{ равно } B) = B \text{ равно } A^3$ . Адъективные лексемы данного класса относятся к разным классам тезауруса и могут требовать как одушевленных (*Петя и Маша знакомы*), так и неодушевленных участников (*Прямая и отрезок параллельны*). В (Вольф 1978:38) такие лексемы также выделяются в одну группу: «Имеются, однако, квалификативные структуры, не ориентированные на общую оценку ни по качеству («хорошо/плохо»), ни по количеству («много/мало»). Так, в частности, большие серии составляют П (прилагательные — П. Г., И. К.), которые соотносят предмет или событие с другими предметами или событиями по признаку сходства или несходства, ср. исп. *igual* ‘равный’, *semejante* ‘похожий’, *identico* ‘идентичный’, *diferente* ‘другой, не похожий’, *distinto* ‘другой’ и т. п.»<sup>4</sup>

Общее свойство лексем класса СИММЕТРИЧНЫХ СВОЙСТВ — их способность оформлять обоих своих участников сочиненной ИГ, как в (3), или множественным числом, как в (4):

- (3) а. *Зачастую можно увидеть пару, в которой мужчина и женщина похожи друг на друга как брат и сестра.*<sup>5</sup>  
 б. *...если нейтрин и антинейтрин тождественны, рождаются 2 электрона...*
- (4) а. *Это родственники с разницей в возрасте более 200 лет, но они похожи как никто другой!*  
 б. *Все клетки организма тождественны по природе.*

Наряду с этим второй участник может кодироваться комитативной группой с + NP<sub>ins</sub>, при том, что в позиции подлежащего находится NP<sub>nom,sg</sub><sup>2</sup>, обозначающая первого участника (ср. *Саша с Петей похожи*), или личное местоимение во мн. ч., обозначающее обоих участников (ср. *Они с Петей похожи / знакомы*) и т. д. Применительно к данным участникам мы будем говорить о семантической роли **контрагента** (подробнее см. в 3.1.).

Альтернативный способ оформления — сохранять (атрибутивное или предикативное) согласование с одним из участников отношения тождества и задавать второго при помощи предложно-падежных или косвеннопадежных конструкций:

<sup>3</sup> Прилагательное *женатый* отличается от его переводных эквивалентов в других языках тем, что имеет предикативный лексический конверсив *замужем*, исключающий синонимию данного типа (что, возможно, отражает архаические представления об отношениях между мужем и женой как несимметричных), однако во всем прочем это прилагательное ведет себя как относящееся к данному классу.

<sup>4</sup> Прилагательные *другой, разный, иной* и им подобные находятся на периферии класса СИММЕТРИЧНЫХ СВОЙСТВ — семантически они имеют два участника с одинаковыми ролями (‘X иной, чем Y’), но синтаксически такие роли выражаются редко — в отличие от многих других прилагательных данного класса.

<sup>5</sup> Приводимые примеры взяты либо из Интернета, либо из НКРЯ, В первом случае конкретный источник не приводится, а во втором случае указывается имя автора (авторов).

- (5) *Чехертма похожа на пресный куриный бульон.*
- (6) *Этот процесс тождественен многократно ускоренной природной очистке воды под воздействием солнечного света и кислорода.*

Предложно-падежное маркирование второго контрагента различно для разных прилагательных данного класса, однако демонстрирует некоторые регулярные паттерны — например, большая часть прилагательных данного класса присоединяет второго участника в дативе: *тождественный / эквивалентный / равный / параллельный кому или чему-л.* и т. д.

## 2.2. Участники у прилагательных несимметричных отношений

Прилагательные данного класса обозначают отношение между участниками, один из которых подчиняется другому, причем носителем свойства «быть членом, отношения, обозначаемого предикатом», может быть как подчиненный участник, так и «контролер». Поскольку статус члена таких несимметричных отношений определяется лексическим значением прилагательного, мы не видим смысла во введении особых семантических ролей для участников таких отношений и называем первого из них (кодируемого лексемой, с которой согласуется прилагательное) носителем свойства, а второго — контрагентом.

Естественно, что для таких прилагательных невозможны характерные для симметричных предикатов конструкции с сочинением или комитативом. Главенствующий контрагент при прилагательных с семантикой подчиненности кодируется дативом (ср. *послушный им народ*, *покорный Перуну старик одному*) или группой *от* + NP<sub>gen</sub> (*зависимые от помещика крестьяне*), а подчиненный контрагент прилагательных с семантикой главенства — группой *над* + NP<sub>ins</sub>:

- (7) *Там штатский был один, главный над ними.* (Г. Я. Бакланов)

Адъективные лексемы данного класса (также как и класса СИММЕТРИЧНЫХ СВОЙСТВ) часто допускают как одушевленных (см. (7)), так и неодушевленных участников:

- (8) *Что вода? Субстанция, послушная ветрам и Луне...* (Т. Соломатина)
- (9) *Человек превратился в самый совершенный и независимый от условий, климата, природных катаклизмов вид.*

## 2.3. Участники у прилагательных эмоционального отношения

Прилагательные ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ имеют обязательный **адресат**, который чаще всего оформляется предлогом *к* и дативом, как в (10) (о других вариантах управления см. в п. 3.1). При отсутствии такого участника прилагательное относится к классу ЧЕРТ ХАРАКТЕРА.

- (10) а. *Она вспоминала, как внимательна и заботлива к ней и отцу была мать.*  
б. *Будьте добры к родителям, и к вам будут добры ваши дети...*  
в. *Комиссар был очень вежлив к нам и разместил нас в бывших бараках иностранцев на фабрике.*

## 2.4. Участники у прилагательных эмоционального и физического состояния человека

В этом семантическом классе присутствует актант с семантической ролью **стимула** (внешняя причина или объект, вызвавший состояние экспериенцера).

Кодирование данного участника в основном лексически обусловлено, ср.: *рад кому/чему-л., гордый кем/чем-л., усталый от кого/чего-л.* и т.д.

## 2.5. Участники у пространственных прилагательных

В семантическом классе ПРОСТРАНСТВЕННЫХ прилагательных, которые образуют особый подкласс прилагательных СИММЕТРИЧНЫХ СВОЙСТВ, присутствует актант с семантической ролью **ориентира**, см. (Кустова 2006). У прилагательных *близкий, ближний* этот актант кодируется  $k+NP_{dat}$ ; у *соседний* —  $k+NP_{dat}$ , но чаще  $s+N_{тв}$ ; у *дальний, далекий* —  $от+NP_{gen}$ . То же управление сохраняется у непространственных, метафорических лексико-семантических вариантов данных прилагательных, ср. (116).

(11) а. *Вы тут далеки от источников информации, — он оглядел планетологов.* (А. и Б. Стругацкие)

б. *Эдик был близок к полной деморализации — появление Выбегаллы его доконало.* (А. и Б. Стругацкие)

Возможно также и обратное оформление ориентира: *ближний от чего-либо, дальний к чему-либо*. Группы с предлогами *от* и *к*, опять-таки, задают один и тот же семантический актант и являются неподчинимыми. В ряде случаев ориентир может быть не выражен и восстанавливается прагматически:

(12) а. *Очень уж далекая планета.* (А. и Б. Стругацкие)

б. *Но директор «Джей-станции» не видел ни звезд, ни желто-зеленого сияния над близкими скалами.* (А. и Б. Стругацкие)

Прилагательные данного класса могут также употребляться с обстоятельствами, вводящими темпоральные сирконстанты, ср. *далекая сегодня планета, близкие в тот момент скалы* и т.д. Впрочем, вопрос сочетаемости прилагательных разных классов с темпоральными сирконстантами заслуживает отдельного рассмотрения.

## 3. Семантические (гипер)роли, употребляющиеся более чем с одним классом прилагательных

Здесь мы опишем случаи, когда один и тот же участник употребляется более чем с одним классом прилагательных.

### 3.1. Контрагент

Используя применительно к прилагательным роль контрагента, изначально предложенную Ч. Филмором, мы неизбежно отходим от той ее трактовки, которая

подразумевает агентивность (активность) участника (см. Апресян 1974:127). Дело в том, что актанты прилагательных по своей природе никогда не являются агентивными. Агенсом не могут считаться даже участники прилагательных, имеющих валентность на одушевленного участника, предполагающего некоторое действие в рамках описываемых прилагательными ситуаций, таких как *агрессивный* или *грубый*. Об этом может говорить, например, невозможность сочетаться с агентивно-ориентированными наречиями, ср.: *Петя построил дом / обидел друга намеренно vs. ?намеренно агрессивный / грубый человек*.

Термин «контрагент» применительно к адъективным валентностям будет употребляться нами тогда, когда актантами прилагательного являются два принципиальных для данной ситуации объекта, независимо от степени их активности, как это имеет место для классов СИММЕТРИЧНЫХ СВОЙСТВ и ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ.

Роль контрагента у прилагательных соответствует неагентивному участнику ситуации, в которой присутствует некоторый «основной» актант с ролью носителя свойства.

У прилагательных ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ участник-адресат кодируется двумя способами: к NP<sub>dat</sub> (см. 2.3) и с+NP<sub>ins</sub>:

- (13) а. *Все сотрудники МКЖД будут одинаково вежливы с пассажирами.*  
б. *Крайне внимательна и заботлива с детьми, склонна их опекать и защищать.*  
в. *Цезарь серьезный и в тоже время ласковый и добрый со своими хозяевами...*

Второй способ допустим едва ли не для всех прилагательных данного класса. Такое управление возможно и для прилагательных СИММЕТРИЧНЫХ СВОЙСТВ:

- (14) а. *Лабео не переносит любую другую рыбу, которая схожа с ним по окраске*  
б. *По природе своей философия как рациональное познание должна иметь одинаковый с наукой тип движения в истории...*

Как представляется, и в случае адресата прилагательных ЭМОЦИОНАЛЬНОГО ОТНОШЕНИЯ, и при кодировании второго участника у прилагательных СИММЕТРИЧНЫХ СВОЙСТВ описываемое поверхностное маркирование соответствует гиперроли контрагента.

Понятие гиперроли может быть полезно при описании прилагательных, т. к. подобно его применению при построении описательной типологии падежных стратегий позволяет объединить различные семантические роли, получающие одинаковое кодирование, см. (Кибрик 1992: 191–195). Кроме гиперроли контрагента применительно к прилагательным можно говорить еще об одной гиперроли, которая будет введена в 3.2.

### 3.2. Субъект квалификации, целевой стандарт, класс сравнения, бенефактив

Рассмотрим различные семантические роли, задаваемые предложной группой для+NP<sub>gen</sub>. Но прежде введем некоторые общие понятия.



Ядро класса прилагательных состоит из градуальных лексем, которые отличаются тем, что могут принимать интенсификаторы<sup>6</sup>, т.е. модификаторы со значением степени проявления признака (*очень, слишком, совсем, немного, несколько* и т.д.). Иногда такими модификаторами выступают наречия от самих градуальных прилагательных (*жутко, страшно, удивительно,...*).

В семантике вслед за Э. Сепиром (Sapir 1949) принято связывать градуальные (скалярные, качественные) прилагательные с некоторым уровнем D на шкале проявления признака, определяемым по его положению относительно уровня, соответствующего норме или эталону (Апресян 1974: 295), (Morzycki 2015). Порождая сочетания типа *высокая девушка*, говорящий подразумевает, что рост девушки достигает значения D, превышающего норму для класса девушек. При помощи градуальных прилагательных осуществляется дескриптивная (*высокий, синий* и т.п.) или оценочная (*хороший, бесполезный* и т.п.) квалификация объекта. Вот как описывает ситуацию квалификации Е. М. Вольф: «Типовая квалификативная структура включает четыре элемента: собственно квалификацию — само П (прилагательное — П.Г., И.К.), субъект квалификации, объект квалификации и основание квалификации» (Вольф 1978:35).

В ситуации квалификации, таким образом, всегда имплицитно присутствует **субъект**, производящий сравнение уровня D на шкале, задаваемой прилагательным, с нормой / эталоном. Эксплицитно в русском языке данный участник бывает поверхностно выражен у СОБСТВЕННО ОЦЕНОЧНЫХ прилагательных, и одним из средств его выражения является группа для+NP<sub>gen</sub>:

(15) *Для имперцев Иван Грозный хорош, для националистов нет.*

(16) а. *Говорун ... испускал сильный, неприятный для непьющего Федора запах дорогого коньяка «Курвуазье». (А. и Б. Стругацкие)*

б. *Как будто каждый раз, когда он уходит в рейс, на Земле остается какое-то необычайно важное дело, самое важное для людей...*

(А. и Б. Стругацкие)

Последний пример с прилагательным *важный* сближается с бенефактивным употреблением (см. ниже). Наряду с самим субъектом оценки такое оформление может получать ИГ, с которой он либо на поверхностном, либо на семантическом уровне связан посессивным отношением:

(17) *Она испустила протяжный, неприятный для слуха визг и полоснула Харона ногтями по физиономии. (А. и Б. Стругацкие)*

Еще одна характерная роль, обслуживаемая группой для+NP<sub>gen</sub>, — роль **целевого стандарта** (см. понятие ситуативной нормы, или нормы пригодности в (Leisi 1953)). Этот участник соответствует целевой ситуации, точнее — потенциальной ситуации, реализация которой возможна только при достижении носителем свойства определенного уровня проявления этого свойства:

<sup>6</sup> Интенсификаторы обладают некоторой избирательной сочетаемостью, ср.: *совершенно новый* vs. *\*совершенно умный* и т.д. Мы, однако, воздержимся от изложения правил такой сочетаемости.

- (18) а. Экологический район удобен для проживания и работы.  
б. Объектом договора найма жилого помещения может быть изолированное жилое помещене, пригодное для постоянного проживания.

Подобная роль может при определенных условиях (о них см. ниже) присутствовать и у прилагательных, моновалентных по умолчанию, ср. лексему **высокий** класса РАЗМЕР:

- (19) Осадков на Месете выпадает мало, а горы недостаточно высокие для образования ледников... (Денис Креленко)

Еще один участник, кодируемый для+NP<sub>gen</sub> — **класс сравнения** (см. также ниже):

- (20) Сложная для ребенка задача — построение высказывания из нескольких предложений.

Подобная роль может присутствовать и у прилагательных, моновалентных по умолчанию, ср. лексему **низкий** (РАЗМЕР), употребленную в следующем примере в переносном значении и также задающую класс сравнения группой для + NP<sub>gen</sub>:

- (21) Из плюсов знаю только очевидные: хорошая проходимость и низкая для джипа цена.

Класс сравнения (ребенок, джип) ограничивает тот интервал на шкале параметра (сложности, цены и т. п.), в пределах которого определяется норма, служащая основанием для установления значения данного параметра: норма сложности задания для ребенка иная, чем для взрослого, норма цены для джипов иная, чем для машин в целом, а тем более товаров вообще.

Наконец, участник с ролью **бенефактива**, кодируемый при помощи для NP<sub>gen</sub> имеется у всех прилагательных БЕНЕФАКТИВНОЙ ОЦЕНКИ:

- (22) Через давление, опасное для психики и даже для самой жизни.  
Указанный N\_819360 есть скрытый коммунист и безусловно, вредный для порядка человек.  
... в его временном жилище я не найду ничего для себя полезного.

Если попытаться обобщить четыре рассмотренных семантических роли в одной гиперроли, то ее можно было бы назвать гиперролью **репера квалификации**, поскольку в каждом из четырех случаев участник с данной гиперролью выступает как та или иная точка отсчета для вынесения общей или дескриптивной оценки.

Заметим, что выражение ролей целевого стандарта и класса сравнения зависимой группой для+NP<sub>gen</sub> у дескриптивных прилагательных РАЗМЕРА, ЧЕРТ ХАРАКТЕРА И СПОСОБНОСТЕЙ ЧЕЛОВЕКА, ФИЗИЧЕСКИХ СВОЙСТВ затруднено, как показывают примеры ниже:

- (23) а. ??вход, узкий для (проезда) инвалидной коляски  
б. ??вода в ванной холодная для (купания) малыша  
с. ??умный для старшеклассника юноша

Однако подобное управление становится вполне приемлемым в двух случаях: 1) при предикативном употреблении в краткой форме (Арутюнова 1988: 240); 2) при употреблении с ними интенсификатора (*слишком, очень, весьма* и т. д.):

(24) а. Вход был узок для коляски

б. Вода в ванной холодна для малыша

с. Этот юноша умен для старшеклассника

(25) а. Очень умный для старшеклассника Кэндзи Койсо пытается устранить вызванную его просвещенной глупостью проблему в глобальной сети.

б. Слышала (от врачей), что если водичка в ванной *чересчур горячая для малыша*...

с. Вход в заведение *слишком узкий для коляски*.

Мы считаем, что кваликативная структура, присутствующая в моно-валентных прилагательных на семантическом уровне, при введении интенсификаторов получает «лицензию» на синтаксическое выражение. Что касается кратких форм, то сошлемся на мнение Н. Д. Арутюновой: «В значение предиката (выраженного краткой формой прилагательного — П. Г. и И. К.) имплицитно входит понятие ‘слишком’, относящееся к ситуации «перебора» с той или другой стороны нормы» (Арутюнова 1988: 240).

### 3.3. Аспект

Такой участник характерен для прилагательных СОБСТВЕННО ОЦЕНОЧНЫХ, БЕНЕФАКТИВНОЙ ОЦЕНКИ, а также для части прилагательных НАЛИЧИЯ / ОТСУТСТВИЯ, СИММЕТРИЧНЫХ СВОЙСТВ, ЭМОЦИОНАЛЬНОГО И ФИЗИЧЕСКОГО СОСТОЯНИЯ ЧЕЛОВЕКА и, спорадически, для некоторых других.

Этот участник возникает в том случае, если объект квалификации многоаспектен или сложен по составу, а квалификация, выражаемая прилагательным, может либо основываться на разных дескриптивных свойствах объекта (как у прилагательных оценочных классов), либо распространяться на разные его аспекты или составные части.

Один из способов выражения этой роли — именная группа в инструменталисе:

(26) а. По сравнению с другими ягодами черника *бедна витаминами*, их даже в 20 раз больше в ее листьях, чем в плодах.

б. Книга, о которой пойдет речь, прежде всего *интересна взглядом на современную русскую провинцию*.

с. Хек *полезен богатым набором витаминов, йода, ...*

Альтернативно аспект (основание, критерий<sup>7</sup>) оценки может выражаться предложной группой с *по / в* и (факультативно) посессивным рефлексивом:

<sup>7</sup> В [Золотова 1988, 242] такие зависимые представлены синтаксемой «критерий оценочной квалификации».

- (27) а. *Что бы и кто бы там не говорил, а Роналду гениален в своей игре!*  
б. *Все клетки организма тождественны по природе.*
- (28) а. *Во всяком случае, если это и не ад, то нечто совершенно неотличимое по своим проявлениям.* (А. и Б. Стругацкие)  
б. *Он изложил Бромбергу в высшей степени загадочную и чертовски соблазнительную в своей загадочности легенду...* (А. и Б. Стругацкие)

В последнем примере роль критерия имеет причинный оттенок.

### 3.4. Сложные предлоги

Достаточно часто семантический участник прилагательных является синтаксически факультативным («слабым» в терминологии Апресяна) и не имеет фиксированного поверхностного средства выражения. В таких случаях часто употребляются сложные предлоги *в отношении кого-либо, в отношении к кому-либо, по отношению к кому-либо*. В примерах ниже они вводят участник, семантически близкий к аспекту (29) или адресату (30–31):

- (29) *Он очень правдивый в отношении всего этого непоправимого пути...*
- (30) *Маленький, заботливый и осторожный в отношении к человеку, который ему важен и нужен.*
- (31) *Несколько горячих голов... требовали, чтобы маэстро, столь осторожный в отношении своей славы... (А. В. Чайнов)*

Отметим, что сложный предлог *по отношению к* почти не встречается в форме без *к*, тогда как *в отношении* употребляется в современном языке более часто, чем *в отношении к*. В русском языке предыдущих эпох, скорее всего, *в отношении к* было более частотно:

- (32) *Надо думать, однакож, не все так беззаботны в отношении к добру своему... (Д. В. Григорович, «Переселенцы»)*
- (33) *Эпическая поэзия есть по преимуществу поэзия объективная, внешняя, как в отношении к самой себе, так и к поэту и его читателю.* (В. Г. Белинский «Разделение поэзии на роды и виды»)

Оба предлога, скорее всего, проходят процесс грамматикализации, связанного с фонологической редукцией. Возможно, это будет сопровождаться изменением управления некоторых семантических классов прилагательных — задаваемые такими предлогами участники из факультативных станут обязательными.

Еще один сложный предлог, часто употребляющийся с прилагательными — *в сравнении с / по сравнению с*. Роль предложной группы, которую он задает, близка к роли объекта сравнения в сравнительных конструкциях:

- (34) а. *...окажется, что алюминиевый провод, хотя и более толстый по сравнению с медным, будет все же легче...*  
б. *Уровень детализации поражает, режим прост в использовании и гораздо более интересный по сравнению с прошлыми играми.*

Вопрос о том, можно ли в случае сложных предлогов говорить о некоторой гиперроли (и если да — какой) мы оставим для будущих исследований.

#### 4. Лексические валентности и семантика прилагательных — результаты

Подводя итоги обсуждению проблемы семантических участников и способу их выражения, скажем следующее. Во-первых, как мы постарались показать, семантика (точнее — принадлежность к некоторому семантическому классу) прилагательного определяет набор его участников и во многом способ их выражения. Во-вторых — существуют участники, характерные для прилагательных как класса стативных предикатов в целом: контрагент, репер квалификации, аспект, и, возможно, некоторые другие. Также начато исследование достаточно регулярного способа кодирования семантических участников, задаваемых сложными предлогами.

Несмотря на то, что отдельные прилагательные зачастую выбиваются из стройных семантических классов, имеющих схожее управление, в целом набор семантических ролей прилагательного определяется скорее его семантикой, чем синтаксической структурой составляющей, в которую оно входит. Это отличает участники прилагательных от внешнего и внутреннего участника глаголов, роли которых (агенс и пациенс соответственно) в большой степени определяются синтаксической структурой.

#### Литература

1. *Apresjan Ju. D.* (1974) *Lexical Semantics*. [Leksicheskaia semantika] Nauka: Moscow. 367 p.
2. *Arutyunova N. D.* (1988) *Types of language meanings: Evaluation. Event. Fact*. [Tipy jazykovykh znachenij: Ocenka. Sobytie. Fakt] Nauka: Moscow. 341 p.
3. *Bennis, H.* (2004) *Unergative Adjectives and Psych Verbs, The Unaccusativity Puzzle: Studies on the Syntax-Lexicon Interface* (Oxford Studies in Theoretical Linguistics 5). A. Alexiadou, E. Anagnostopoulou & M. Everaert (eds.) Oxford: Oxford University Press. pp. 84–114.
4. *Boguslavskaja O. Ju.* (2005) *The goal in the semantics of goalless adjectives*. [Cel' v semantike necelevykh prilagatel'nyh] *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* [Komp'juternaja lingvistika i intellektual'nye tehnologii. Trudy Mezhdunarodnoj konferencii "Dialog"] Moscow. pp. 54–58.
5. *Bulygina T. V., Shmelev A. D.* (1997) *The linguistic conceptualization of the world (on the basis of Russian grammar). Jazykovaja konceptualizacija mira (na materiale russkoj grammatiki)*. Shkola «Jazyki russkoj kul'tury»: Moscow. 576 p.
6. *Dixon, R. M.W.* (1977) *Where have all the adjectives gone?* // *Studies in language* 1, pp 19–80.
7. *Flickinger, D. and J. Nerbonne* (1992) 'Inheritance and Complementation: A Case Study of Easy Adjectives and Related Nouns', *Computational Linguistics* 18(3), 269–310.

8. *Givon, T.* (1970) Notes on the Semantic structure of English Adjectives. *Language*, Vol 46, 4, 816–837.
9. *Haritonchik, Z. A.* (1990) An adjective: problems of classification [Imja prilagatel'noe: problemy klassifikacii] *Theory of grammar: lexico-grammatical classes and ranks of words* [Teorija grammatiki: leksiko-grammaticheskie klassy i razrjady slov] *Berezina F. M., Kubrjakovoj E. S.* (eds.) INION RAN: Moscow. pp. 94–119.
10. *Iomdin L. L.* (1981) Symmetric predicates in Russian [Simmetrichnye predikaty v russkom jazyke] *Problems of structural linguistics 1979* [Problemy strukturnoj lingvistiki] 1979. Nauka: Moscow. pp. 89–104.
11. *Jackendoff, Ray* (1972) *Semantic Interpretation in Generative Grammar*. MIT Press.
12. *Karpova O. S., Reznikova T. I., Arhangel'skij T. A., Kjuseva M. V., Rakhilina E. V., Ry-zhova D. A., Tagabileva M. G.* (2010) The database on Russian polysemous adjectives and adverbs [Baza dannyh po mnogoznachnym kachestvennym prilagatel'nym i narechijam russkogo jazyka] *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* [Komp'juternaja lingvistika i intellektual'nye tehnologii. Trudy Mezhdunarodnoj konferencii "Dialog"] Moscow. pp. 163–169.
13. *Kibrik A. E.* (1992) *Essays on general and applied problems of linguistics* [Ocherki po obshhim i prikladnym voprosam jazykoznanija] MSU: Moscow. 356 p.
14. *Kjuseva M. V., Reznikova T. I., Ryzhova D. A.* (2013) A typologically oriented database of qualitative features [Tipologicheskaja baza dannyh adjektivnoj leksiki] *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* [Komp'juternaja lingvistika i intellektual'nye tehnologii. Trudy Mezhdunarodnoj konferencii "Dialog"] Moscow. pp. 367–376
15. *Kustova G. I.* (2006) Arguments and constructions of adjectives [Valentnosti i konstrukcii prilagatel'nyh] *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* [Komp'juternaja lingvistika i intellektual'nye tehnologii. Trudy Mezhdunarodnoj konferencii "Dialog"] Bekasovo. pp. 323–328.
16. *Kustova G. I.* (2009) The semantic database of verbal adjectives: structure and types of information [Elektronnyj semanticheskij slovar' glagol'nyh prilagatel'nyh: struktura i tipy informacii] *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* [Komp'juternaja lingvistika i intellektual'nye tehnologii. Trudy Mezhdunarodnoj konferencii "Dialog"] Moscow. pp. 271–277.
17. *Landau, I.* (2006) *Ways of Being Rude*. ms, Ben Gurion University, 50 p. at: <http://www.bgu.ac.il/~idanl/files/Rude%20Adjectives%20July06.pdf>
18. *Leisi E.* (1953) *Der Wortinhalt: Seine Struktur in Deutschen und Englischen*. Heidelberg.
19. *Lyashevskaya, Olga, and Egor Kashkin* (2016) Welcome to the club: Designing the inventory of semantic roles for adjectives. *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"* [Komp'juternaja lingvistika i intellektual'nye tehnologii. Trudy Mezhdunarodnoj konferencii "Dialog"] Moscow. pp. 440–454
20. *Mel'chuk I. A.* (1998) *Course on general morphology* [Kurs obshhej morfologii] v.2. *Wiener Slavistischer Almanach, Sonderband: Moscow-Vienna*. 544 p.

21. *Meltzer-Asscher, Aya* (2011) *Adjectives and Argument Structure*. Thesis submitted for the degree “Doctor of philosophy”. Tel-Aviv University, 236 p.
22. *Morzycki, Marcin* (2015) *Modification*. CUP, Cambridge, 346 p.
23. *Nissenbaum, Jon & Bernhard Schwarz* (2010) Parasitic degree phrases. *Natural Language Semantics*. March 2011 , Volume 19, Issue 1, pp. 1–38.
24. *Rosch, E.* (1978) *Principles of Categorization* // Rosch, E. & Lloyd, B. B. (eds), *Cognition and Categorization*, Lawrence Erlbaum Associates, Publishers, (Hillsdale), pp. 27–48.
25. *Shramm A. N.* (1979) *Essays on semantics of quality adjectives: on the data of modern Russian language*. [Ocherki po semantike kachestvennyh prilagatel'nyh: na materiale sovremennogo russkogo jazyka] Publishing house of the Leningrad State University: Leningrad. 135 p.
26. *Stowell, Tim* (1991) *The Alignment of Arguments in Adjective Phrases*. In: Rothstein, Suzan, (ed.), *Perspectives on Phrase Structure: Heads and Licensing, Syntax and Semantics* 25, Academic Press, 105–135.
27. *Volf E. M.* (1978) *Grammar and semantics of adjectives: on the data of Ibero-Romance languages*. [Grammatika i semantika prilagatel'nogo: Na materiale ibero-romanskih jazykov. Nauka: Moscow. 199 p.

# ПРИНЦИПЫ ОПРЕДЕЛЕНИЯ СТЕПЕНИ ЛИНГВОСПЕЦИФИЧНОСТИ КОННЕКТОРОВ<sup>1</sup>

**Инькова О. Ю.** (Olga.Inkova@unige.ch)

ИПИ ФИЦ ИУ РАН, Москва, Россия;  
Женевский университет, Женева, Швейцария

## PRINCIPLES OF HOW TO DETERMINE THE DEGREE OF LANGUAGE-SPECIFICITY OF CONNECTIVES

**Inkova O. Yu.** (Olga.Inkova@unige.ch)

Institute of Informatics Problems, FRC CSC RAS, Moscow,  
Russia; University of Geneva, Geneva, Switzerland

The paper deals with the question of how to determine the degree of language-specificity of connectives. As connectives are non-descriptive linguistic units and their morphological nature may vary, the author suggests to apply functional semantic criteria and distinguishes five types of interlingual isomorphism violation in Russian/French and Russian/Italian.

The author proposes to establish five categories of language-specific connectives:

1. The relation  $R$  expressed by the connective  $K_A$  of language  $A$  is conveyed in language  $B$  by a linguistic unit  $U_B$  not belonging to the functional class of connectives (for instance, the relation of simultaneousness expressed by *pri ètom* in Russian is frequently expressed in French or Italian by means of a present gerund);
2. The relation  $R$  expressed by the connective  $K_A$  of language  $A$  is conveyed in language  $B$  by the signal  $K_B$  of a relation  $R'$  which is semantically contiguous, but nevertheless possesses a more specific equivalent in language  $A$  (for instance, the *pri tom* of illocutive simultaneousness is rendered in French with *en outre*, *en plus*, *de plus* disposing in Russian of the more precise equivalents *k tomu že* and *krome togo*);
3. The connective  $K_A$  of language  $A$  possesses in language  $B$  a set of equivalents  $K_A^1$ , all occurring with the same frequency (it's the case of Russian *vpročem*, which is mainly translated with *d'ailleurs* and *du reste*);

---

<sup>1</sup> Исследование выполнено в ИПИ ФИЦ ИУ РАН в рамках проекта «Логическая структура текста: контрастивный анализ показателей логико-семантических отношений в русском, французском и итальянском языках», поддержанного РФФИ (грант № 16-18-10004).



4. The connective  $K_A$  of language A has in language B a systemic equivalent  $K_B$  which is the most frequent in translation, but doesn't reproduce all the values and uses of  $K_A$  (for instance, the connective *a to* of negative alternative has the French equivalent *sinon*, but the Russian connective is polysemous: it can express both the relations of cause and alternative, as well as a relation of addition — what is not the case of *sinon*);
5. There is in language B a mostly used translation variant  $K_B$  of the connective  $K_A$  of language A, but this 'equivalent'  $K_B$  doesn't possess the combinatory variety of  $K_A$  (thus the Russian connective *hotya* introducing a subordinate clause combines in the main sentence with the conjunction *no*, but *hotya ... no* is translated in French either by a single *bien que*, the equivalent of *hotya*, or by a single *mais*, the equivalent of *no*).

The author also describes features helping to determine the degree of language-specificity of connectives that would take into account the translation direction (from Russian to French/Italian and vice versa) and might be subject to a quantitative analysis. These features are: 1) frequency of occurrence of a "zero" functionally equivalent fragment (respectively, translation stimulus); 2) type of functionally equivalent fragment or translation stimulus (congruent or divergent); 3) number of possible translation patterns (respectively, translation stimuli). In conclusion, the author will comment on statistical data confirming the described types of interlingual isomorphism violation in case of the functional class of connectives. The author points out that statistics requires adequate interpretation and should be accompanied by a semantic analysis, as the latter also serves as a tool for the data's review. The study was carried out by using the Russian-French and Russian-Italian supracorpora database of connectives.

**Key words:** database of cross-linguistic equivalences, Russian, French, Italian, contrastive linguistics, corpus linguistics, language-specific connectives, statistical data

## 1. Введение

Исследования лингвоспецифичных языковых явлений ведутся в отечественной лингвистике начиная с 1990-х годов (см., например, [Зализняк, Левонтина 1996], [Зализняк, Левонтина, Шмелев 2012], [Зализняк 2015], [Шмелев 2015]. С одной стороны, они исходят из гипотезы Сепира-Уорфа, а с другой, опираются на работы А. Вежицкой, которой принадлежит и сам термин «лингвоспецифичный» ("language-specific (word-)meanings" [Wierzbicka 1992, 1996]). В [Зализняк 2015: 687] лингвоспецифичной называется «лексическ[ая] единиц[а], заключающ[ая] в себе *уникальную концептуальную конфигурацию*, а именно такую, что во всех существующих ее переводных эквивалентах не хватает какого-то семантического компонента или неустранимым образом присутствует 'лишний' компонент». Поэтому лингвоспецифичные единицы являются труднопереводимыми (или вообще непереводаемыми) на другие языки, а сам признак «труднопереводимости» является важнейшим отличительным свойством лингвоспецифичных слов.

Если до последнего времени исследования лингвоспецифичности проводилось на основе сопоставительного семантического анализа языковых единиц, то с развитием корпусной лингвистики, особенно, с созданием параллельных

корпусов, открываются новые возможности исследований межъязыковых соответствий, прежде всего, с привлечением статистических данных. Еще один шаг на этом пути был сделан благодаря созданию т.н. «надкорпусных баз данных» (НБД) ([Кружков 2015]). Они представляют собой своего рода надстройку над параллельным корпусом, позволяющей фиксировать реально встретившиеся соответствия языковых единиц и аннотировать их по заданному набору признаков. Отмечу два основных преимущества НБД перед параллельными корпусами: 1) для каждой анализируемой языковой единицы в тексте перевода уже выделен ее функционально-эквивалентный фрагмент (ФЭФ) (термин введен в работе [Добровольский и др. 2005]); 2) анализируемой единице и ее ФЭФ приписывается значение формально охарактеризованных признаков, релевантных для их анализа.

В результате работы с русско-французской НБД лингвоспецифичных единиц в работе [Зализняк 2016: 860] были выделены следующие пять типов отсутствия межъязыкового семантического изоморфизма:

1. Асимметричное членение концептуальной области (например, рус. *рука* — фр. *main* и *bras*).
2. Тип САМОВАР (переводной эквивалент отсутствует; используется заимствованное слово или описательное определение).
3. Тип БАБУШКА (имеется один преимущественный вариант перевода, но он неточный).
4. Тип РОДНОЙ (переводной эквивалент отсутствует; имеется несколько приблизительно равновероятных вариантов перевода, все неточные).
5. Тип РАЗЛУКА (имеется один преимущественный переводной эквивалент, имеющий более точное соответствие в русском языке: *разлука* — *séparation* — *расставание*).

Автором предлагается также ряд характерных признаков, указывающих на вероятную лингвоспецифичность языковой единицы. Эти признаки могут быть программно генерированы в НБД с учетом направления перевода и подаются количественному анализу.

При переводе с русского языка (Р—Ф):

1. наличие большого количества ФЭФ, в том числе имеющих приблизительно равную частотность;
2. наличие многокомпонентных ФЭФ;
3. слово остается без перевода (нулевой ФЭФ).

При переводе на русский язык (Ф—Р):

4. большое количество стимулов перевода (СП), т.е. тех языковых единиц оригинала, ‘реакцией’ на употребление которых является употребление русской языковой единицы;
5. наличие многокомпонентных СП;
6. отсутствие какой-либо единицы, которая ‘стимулирует’ появление анализируемой русской языковой единицы (нулевой СП).

В данной статье предлагается проверить, насколько данные пять категорий и перечисленные признаки лингвоспецифичности, разработанные, прежде

всего, для однозначных слов, применимы к анализу функционального класса коннекторов. Такого рода исследование представляет интерес для лексической типологии, лексикографии, теории и практики перевода, а также других областей, связанных с сопоставительным изучением языков. Предлагаемый анализ проводится с использованием русско-французской и русско-итальянской НБД коннекторов общим объемом более 3,5 млн словоупотреблений. Ее концепция, принципы аннотирования и отличия от других существующих НБД изложены в работах [Инькова, Кружков 2016; Зацман и др. 2016].

## 2. Параметры квантитативного анализа степени лингвоспецифичности коннекторов

Хорошо известно, что коннекторы, как и многие дискурсивные слова, принадлежат к лингвоспецифичным единицам. Даже такие, казалось бы, семантически пустые коннекторы как англ. *and*, нем. *und*, фр. *et*, рус. *и*, имеют в каждом языке особенности функционирования, которые не позволяют поставить между ними знак равенства. Достаточно привести хорошо известный пример расхождения в употреблении англ. *and*, нем. *und*, фр. *et*, с одной стороны, и русского *и*, с другой, в такой фразе: нем. *Ihre Mutter ist Lehrerin und ihr Vater ist Anwalt*, англ. *Her mother is a teacher and her father is a lawyer*; фр. *Sa mère est enseignante et son père est avocat*; рус. *Ее мать — учительница, а<sup>?</sup>и отец — адвокат*. Этот простой пример дает также возможность подчеркнуть три важных положения, которые должны учитываться при анализе лингвоспецифичности. Первое: само свойство лингвоспецифичности является сопоставительным, проявляющимся при анализе межъязыковых соответствий двух или более конкретных языков (это обстоятельство подчеркивалось, в частности, в работе [Зализняк 2015]). Иначе говоря, то, что является лингвоспецифичным для одной пары языков, может не быть таковым в другой. Второе: языковые единицы различаются степенью лингвоспецифичности: она очевидным образом будет выше у рус. *а*, не имеющего системного эквивалента во французском, английском, немецком и многих других языках, и минимальной у *или*, имеющего системные эквиваленты в большинстве европейских языков. Наконец, «труднопереводимость» не является детерминирующим признаком лингвоспецифичности. Как было показано в [Dobrovolskij, Röppel 2016] на примере *дело в том*, некоторые дискурсивные единицы, не представляя никакого труда для перевода, являются тем не менее лингвоспецифичными, т.е. обладают такой конфигурацией семантических черт, которые не позволяют найти им точного эквивалента в других языках.

Что касается пяти выявленных в [Зализняк 2016] типов отсутствия семантического межъязыкового изоморфизма, разработанного в первую очередь для однозначных слов, то их невозможно механически перенести на функциональный класс коннекторов в силу их семантических и функциональных особенностей. Коннекторы относятся к классу недескриптивных языковых единиц. Их функция — сигнализировать, что между двумя связываемыми ими фрагментами текста существует логико-семантическое отношение (ср.

категорию «сигналативы» в [Иорданская, Мельчук 2007: 347–358]). Говорящий может, однако, никак не выразить это отношение (например, причинное отношение в *Петя не пришел. Он заболел*) или, наоборот, выразить его неоднократно. Так, в высказывании:

Выяснилось, что в бюро иностранцев ни о каком Воланде, *а равно также и* Фаланде, маге, ровно ничего не слышали. [М. А. Булгаков. Мастер и Маргарита (ч. 1) (1929–1940)]

употреблен коннектор *а равно также и*, хотя можно было бы ограничиться более экономным *а также*. Нарушение межъязыкового изоморфизма может быть связано как раз с тем, что словарный эквивалент языковой единицы в другом языке не может входить в состав таких образований, что прежде всего связано с разнообразной морфологической природой коннекторов. Как известно, в функциональный класс коннекторов входят не только союзы, сочинительные и подчинительные, но и т.н. аналоги союзов наречного типа, а также предикативные конструкции (*из этого следует, дело в том, что* и др.), которые имеют другие синтаксические и сочетаемостные свойства. На этом основании в случае коннекторов мне представляется более адекватным говорить о нарушении функционально-семантического изоморфизма.

С другой стороны, в отличие от лингвоспецифичной лексики, включающей, как правило, однозначные лексемы, многие коннекторы многозначны, в том смысле, что в зависимости от смыслового соотношения соединяемых ими фрагментов текста они могут выражать разные отношения или устанавливать одно отношение, но на разных семантических уровнях. И наоборот, одно отношение может иметь целую парадигму показателей, различия между которыми часто заключаются не столько в семантике, сколько в тех синтаксических условиях, в которых они могут употребляться.

Таким образом, становится понятным, что проблема лингвоспецифичности коннекторов должна решаться на основе функционально-семантических критериев, а предложенные типы нарушения межъязыкового изоморфизма должны быть несколько переформулированы, а некоторые из них не находят себе места при анализе коннекторов. Так, тип 1. «Ассиметричное членение концептуальной области» на первый взгляд присутствует в такой тройке: рус. *скорее... чем и лучше... чем* и ит. *piuttosto che*.

<b>Piuttosto morire, che</b> bere quella medicina cattiva.	<b>Лучше</b> умереть, <b>чем</b> глотать такое ужасное лекарство!
— У меня <b>скорее</b> лапы отсохнут, <b>чем</b> я прикоснусь к чужому, — напыжившись воскликнул кот.	— Mi si paralizzino le zampe <b>piuttosto che</b> toccare la roba altrui, — esclamò il gatto, con sussiego.

Однако в итальянском языке существует также компаратив *meglio* 'лучше', который просто не обладает способностью русского *лучше* выступать в качестве «компонента двухместного союза» (РГ-80, II: 490) для выражения замещения по предпочтению. Поэтому речь здесь идет не столько о различном концептуальном членении, сколько о функциональных возможностях той или иной

языковой единицы. В парах русский — французский и русский — итальянский тип 1. пока не был обнаружен.

Тип 2. САМОВАР в том виде, в котором он сформулирован для однозначной лексики, также не находит себе места при анализе коннекторов. Использование заимствованного коннектора представляется мне мало вероятным явлением. По крайней мере, в паре языков русский — французский таких случаев, насколько мне известно, нет. Употребление в итальянском языке латинского *ergo* как показателя следственных отношений воспринимается как иноязычное вкрапление, имеющее явную книжную стилистическую окраску, поскольку в итальянском языке существуют стилистически нейтральные коннекторы, выражающие это отношение (*dunque, quindi*). Использование описательного определения для выражения логико-семантических отношений также не характерно. Типу 2. будет примерно соответствовать в типологии лингвоспецифичности коннекторов тип 1. Вообще говоря, установить какое-либо прямое соответствие между типами, предлагаемыми Анной А. Зализняк для лингвоспецифичной лексики, и типами, предлагаемыми нами в данной работе для анализа коннекторов, не представляется возможным в силу указанных выше причин.

Предлагается выделять следующие типы отсутствия функционально-семантического межъязыкового изоморфизма для класса коннекторов:

**Тип 1.** Отношение  $R$ , выражаемое коннектором  $K_A$  в языке  $A$ , выражается в языке  $B$  языковой единицей  $U_B$ , не принадлежащей к функциональному классу коннекторов. Это может быть синтаксическая конструкция, морфологическая форма, лексическая единица.

Так, отношение сопутствования (определение данного отношения см. в [Инькова в печати]), которое на пропозициональном уровне сигнализирует, что одно действие, представленное как главное, сопровождается другим действием, представленным как участвующее в его реализации, выражается в русском языке коннектором *при этом*, а во французском языке может при общем субъекте обоих действий выражаться формой деепричастия настоящего времени предиката, соответствующего сопровождающему действию.

<p>Маруся дарила сестре платья и кофты.  <b>При этом</b> говорилось: — Синюю кофту бери, а зеленую я еще поношу...</p>	<p>Maroussia donnait à sa cousine ses robes et ses blouses <b>en ayant soin d'ajouter</b>:          «Prends la blouse bleue, la verte je la porterai encore un peu... »</p>
--	---

**Тип 2.** Отношение  $R$ , выражаемое коннектором  $K_A$  в языке  $A$ , выражается в языке  $B$  показателями  $K_B$  близкого по семантике отношения  $R'$ , имеющего, однако, более точное соответствие в языке  $A$ . Например, коннектор *притом* не имеет системного эквивалента во французском языке. Он выражает отношения сопутствования на иллокутивном уровне, сигнализируя, что вводимый им речевой акт 'сопутствует предыдущему', добавляет по поводу сказанного новую информацию, часто с аргументативной функцией (см. [Инькова в печати]), как во втором примере. *Притом* может иметь нулевой эквивалент или переводиться несколькими равночастотными эквивалентами (*en outre, en plus,*

*de plus*), которые имеют более точные соответствия в русском языке (*кроме того, к тому же*).

Такое сочетание авантюриности, <b>притом</b> часто бульварной, с идеей, с проблемным диалогом, [...] воспринималось как грубое и ничем не оправданное нарушение «жанровой эстетики».	Cet alliage de l'aventure, souvent boulevardière, avec l'idée, avec le dialogue idéologique, [...] était perçu comme une infraction injustifiée à l'« esthétique de genre ».
Чтоб кончить все это разом, ей оставалось одно: заметив признаки рождающейся любви в Штольце, не дать ей пищи и хода и уехать поскорей. Но она уже потеряла время: это случилось давно, <b>притом</b> надо было ей предвидеть, что чувство разыграется у него в страсть; да это и не Обломов: от него никуда не уедешь.	Pour en finir d'un coup avec tout cela il ne lui restait qu'une chose : ayant deviné chez Stolz un amour naissant, partir au plus vite, sans lui fournir d'aliment. Mais il était trop tard : Stolz couvait son amour depuis quelque temps déjà; <b>de plus</b> il lui aurait fallu prévoir que ce sentiment se développerait chez lui en une passion. D'ailleurs, ce n'était pas Oblomov : on ne pouvait le fuir.

**Тип 3.** Коннектор  $K_A$  языка А имеет несколько равночастотных переводных эквивалентов-коннекторов  $K_A^1$ , адекватно отражающих его семантику. Так, *впрочем* в НБД из 285 случаев переводится *d'ailleurs* в 97 случаях и *du reste* — в 75. Ниже приводятся четыре перевода одного фрагмента русского текста, в котором распределение этих двух ФЭФ 50 % на 50 %.

Нужно знать, что одно значительное лицо недавно сделался значительным лицом, а до того времени он был незначительным лицом. <b>Впрочем</b> , место его и теперь не почиталось значительным в сравнении с другими, еще значительнейшими.	Il faut savoir que le monsieur important n'était devenu un monsieur important que tout récemment, et que, jusqu'alors, il n'avait rien eu d'important. <b>Du reste</b> , sa place, même à présent, n'était considérée pas comme importante si on la comparait à d'autres, bien plus importantes qu'elle.
	Il y a lieu d'indiquer que ce personnage n'était devenu important que depuis peu; son poste, <b>d'ailleurs</b> , n'était pas considéré comme très important, comparé à d'autres, plus importants encore.
	Il faut dire que ledit personnage n'était devenu important que depuis peu; <b>du reste</b> , par rapport à d'autres plus considérables, la place qu'il occupait n'était pas tenue pour bien importante.
	Il faut savoir qu'un certain personnage important était devenu depuis peu un personnage important bien qu'il fût jusque-là un personnage sans importance. <b>D'ailleurs</b> , même maintenant, sa place ne serait pas considérée comme importante par rapport à d'autres beaucoup plus importantes encore.

**Тип 4.** Коннектор  $K_A$  языка А имеет в языке В системный и наиболее частотный переводной эквивалент коннектором  $K_B$ , но он не охватывает всех случаев употребления коннектора  $K_A$ . Например, русскому *a to* в значении отрицательной альтернативы соответствует фр. *sinon*. Однако, в отличие от *sinon*, *a to* может выражать целый ряд других отношений (причинное, разделительное, соединительное и др.), сводимых тем не менее к единому семантическому ядру, что позволяет говорить о едином союзе *a to* ([Inkova-Manzotti 2005]). В этих случаях *a to* имеет другие ФЭФ.

**Тип 5.** В языке В имеется один преимущественный вариант перевода коннектора  $K_A$  языка А коннектором  $K_B$ , который, однако, не обладает сочетаемостными свойствами коннектора  $K_A$ . Речь в данном типе идет о возможности многократного маркирования выражаемого коннектором отношения, благодаря использованию нескольких его показателей. Так, *хотя* имеет преимущественный эквивалент *bien que*, адекватно передающий его семантику. Но *хотя* часто входит в состав двухместных образований, в частности, с союзом *но*. *Bien que* не обладает этим свойством. В результате двухместные образования *хотя... но* переводятся эквивалентами их составляющих: либо *bien que*, либо *mais*.

Хотя ты уверял меня, что она простила, <b>но</b> , знаешь... я не был покоен!	<b>Bien que</b> tu m'aies affirmé qu'elle m'avait pardonné... je n'étais pas tranquille!
Она очень обрадовалась Штольцу; <b>хотя</b> глаза ее не зажглись блеском, щеки не запылали румянцем, <b>но</b> по всему лицу разлился ровный, покойный свет и явилась улыбка.	Elle fut ravie de voir Stolz. Ses yeux ne brillèrent pas, ses joues ne devinrent pas vermeilles, <b>mais</b> tout son visage s'irradia d'une lumière étale et paisible. Elle sourit.

Что касается предлагаемых в [Зализняк 2016] параметров для применения количественных методов, то признаки 1., 3.–4. и 6. могут быть использованы без изменений, а 2. и 5. должны быть адаптированы к анализу коннекторов. Поскольку коннекторы являются в своем большинстве многокомпонентными единицами, то критерий многокомпонентности ФЭФ или СП не является релевантным: у однокомпонентного коннектора может быть многокомпонентный ФЭФ с идентичной семантикой (*например* — *par exemple*). Эти критерии предлагается заменить на критерии принадлежности ФЭФ и СП к классу коннекторов. Если коннектору соответствует коннектор, то ФЭФ и СП считаются конгруэнтными, если коннектору соответствует языковая единица, не принадлежащая к классу коннекторов, то ФЭФ и СП считаются дивергентными (термины заимствованы из [Johansson 2007]). Предлагаются, следовательно, такие параметры для применения квантитативных методов для анализа коннекторов.

При переводе с русского языка (Р–Ф):

1. наличие большого количества ФЭФ, в том числе имеющих приблизительно равную частотность;
2. доля дивергентных ФЭФ;
3. слово остается без перевода (нулевой ФЭФ).

При переводе на русский язык (Ф–Р):

4. большое количество стимулов перевода (СП);
5. доля дивергентных СП;
6. нулевой СП.

Ниже приводятся статистические данные по четырем коннекторам, полученные в результате сплошной выборки случаев их употребления в текстах на русском языке, загруженных в НБД. Они иллюстрируют первые четыре

типа нарушения межъязыкового изоморфизма и параметры их количественного анализа. Статистика дается для направления Р-Ф, но подтверждается и в направлении Ф-Р (объемы статьи, к сожалению, не позволяют их привести и прокомментировать).

Коннектор	Всего аннотаций	Zero	Zero%	Dvrg	Dvrg%	Models	Models Ratio
<i>при этом</i>	76	17	22,37	25	32,89	35	46,05
<i>притом</i>	58	12	20,69	1	1,72	16	27,59
<i>впрочем</i>	285	12	4,21	11	3,86	36	12,63
<i>а то</i>	109	14	12,84	2	1,83	24	22,02

Первые два коннектора — *при этом* и *притом* — принадлежат к типам нарушения межъязыкового изоморфизма, где в сопоставляемом языке нет специализированного коннектора для соответствующего логико-семантического отношения: *при этом* — Тип 1, *притом* — Тип 2. Различия между двумя типами отражаются в количественных показателях. Если количество нулевых эквивалентов у них примерно одинаковое (чуть больше 20%), то доля дивергентных ФЭФ достигает у *при этом* почти 33%, поскольку выражаемое им отношение сопутствования на пропозициональном уровне часто передается морфологической формой деепричастия настоящего времени. Однако, *при этом* может выражать и отношение сопутствования между двумя речевыми актами, для которого переводчики используют коннекторы близкой семантики добавления к сказанному, а таких коннекторов во французском языке целая парадигма (*du reste, pour le reste, par ailleurs, de surcroît, de plus, en plus, en outre* и др.). Этим объясняется большое количество моделей перевода *при этом*. Поскольку *притом* может выражать только отношение сопутствования на уровне речевых актов, доля дивергентных переводов у него не достигает 2%, но так как он не имеет системного эквивалента в паре русский — французский, то количество моделей переводов у него достаточно высоко (27, 59%).

*Впрочем* относится к Типу 3. Все три показателя у него невысокие. Если и фиксируется значительное количество моделей перевода, то все они принадлежат к одному семантическому классу, а их количество в два раза меньше, чем у *притом*, и в почти в четыре, чем у *при этом*. Помимо *du reste* и *d'ailleurs*, *впрочем* может быть переведен *au demeurant* (26 случаев) с такой же семантикой или семантически близкими противительно-уступительными коннекторами: *cependant* (13), *mais* (7), *pourtant* (6) и т.д. Принадлежащее к Типу 4. *а то* не только имеет в два раза больше, чем *впрочем*, моделей перевода, приближаясь по этому параметру к *притом*, но они, кроме того, не являются синонимичными: помимо *sinon*, на которое приходится более 50%, в качестве ФЭФ могут выступать причинные коннекторы (*car, parce que*), противительные (*tandis que*) или разделительные (*ou*). При интерпретации этих данных нужно, однако, иметь в виду, что даже при наличии системного эквивалента с минимальным различием в семантике, как например рус. *или* и фр. *ou*, этот эквивалент используется менее, чем в 80% случаев (для *или* в 76,22%, для *если* на долю *si* приходится 64,29%).



Что касается Типа 5., то двухместные образования *хотя... но* передаются в 50% одним *bien que*, в 18,18% одним *mais*, остальное приходится на другие коннекторы уступительной семантики.

### 3. Заключение

Предлагаемые пять типов нарушения межъязыкового семантико-функционального изоморфизма в функциональном классе коннекторов находят, таким образом, подтверждение при количественном анализе по шести признакам, указывающим на вероятную лингвоспецифичность коннектора. Не вызывает сомнения, что статистические данные требуют адекватной интерпретации и должны сопровождаться собственно семантическим анализом, являясь одновременно средством его проверки, что и было сделано в настоящем исследовании. Не вызывает сомнения и тот факт, что для надежности статистических данных необходимо увеличить объем доступных для обработки языковых фактов. Однако представляется очевидным, что использование квантитативных методов анализа с привлечением верифицируемых и программно генерируемых данных повышает надежность лингвистического описания.

### Литература

1. *Dobrovol'skiy D. O., Kretov A. A., Sharov S. A.* (2005), *Korpus parallel'nykh tekstov: Arkhitektura i vozmozhnosti ispol'zovaniya* [Corpus of parallel texts: Architecture and applications]. *Natsional'nyy korpus russkogo yazyka: 2003–2005* [Russian National Corpus: 2003–2005], Moscow, Indrik, pp. 263–296.
2. *Dobrovol'skiy D., Pöppel L.* (2016), *The Discursive Construction дело в том, что and its Parallels in other Languages: a Contrastive Corpus Study*, *Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference "Dialogue"* (2016), Vol. 15 (22), Moscow, RGGU, pp. 134–145.
3. *Inkova-Manzotti O.* (2005), *Encore sur la conjonction russe a to* [Again about the Russian conjunction *a to*], *Revue des études slaves* [Review of the Slavic Studies], Vol. 76/4, pp. 485–497.
4. *Inkova O.* (in press), *Konnektory s formantom pri: korpusnoe issledovanie* [The connectors of Russian with the element *pri*: corpus-based study], *Russian Linguistics*, 29 pp.
5. *Inkova O., Kruzchkov M.* (2016), *Nadkorpusnye russko-frantsuzskie bazy danykh glagol'nykh form i konnektorov* [Supracorpora databases of Russian and French verbal forms and connectors], *O. Inkova, A. Trovesi eds, Lingue slave a confronto* [Slavic languages in comparison], Bergamo, Bergamo University Press, pp. 365–392.
6. *Iordanskaya L. N., Mel'chuk I. A.* (2007), *Smysl i cochetaemost' v slovare* [The sense and the combinatory in dictionary], *Moskva, Yazyki slavyanskoy kul'tury*.
7. *Johansson S.* (2007), *Seeing through Multilingual Corpora. On the use of corpora in contrastive studies*, Amsterdam, John Benjamins.

8. *Kruzhkov M. G.* (2015), *Informacionnye resursy kontrastivnykh lingvisticheskikh issledovaniy: elektronnye korpusa tekstov* [Information resources for contrastive studies: electronic text corpora], *Sistemy i sredstva informatiki* [Systems and Means of Informatics], Vol. 25, N2, pp. 140–159.
9. RG-80 — *Russkaya grammatika* [Russian grammar]. Shvedova N. Yu. (ed.). Moscow, Nauka, 1980.
10. *Shmelev A. D.* (2015), *Russkie lingvospetsifichnye leksicheskie edinitsy v parallel'nykh korpusakh: vozmozhnosti issledovaniya i "podvodnye kamni"* [Russian language-specific lexical units in parallel corpora: prospects of investigation and "pitfalls"], *Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference "Dialogue"* (2015), Vol. 14 (21), Moscow, RGGU, pp. 584–594.
11. *Wierzbicka A.* (1992). *Semantics, culture, and cognition. Universal human concepts in culture-specific configurations*, N. Y., Oxford, Oxford University Press.
12. *Wierzbicka A.* (1996). *Semantics: primes and universals*, Oxford, Oxford University Press.
13. *Zaliznyak Anna A., Levontina I. B., Shmelev A. D.* (2012), *Konstanty i peremennye russkojazykovoj kartiny mira* [Constants and Variables of the Russian language picture of the world], Moskva, Yazyki slavyanskoy kul'tury.
14. *Zaliznyak, Anna A.* (2015), *Lingvospetsifichnye edinicy russkogo yazyka v svete kontrastivnogo korpusnogo analiza* [Russian language-specific words as an object of contrastive corpus analysis], *Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference "Dialogue"* (2015), Vol. 14 (21), Moscow, RGGU, pp. 683–695.
15. *Zaliznyak, Anna A.* (2016), *Baza dannykh mezh'yazykovykh ekvivalentsiy kak instrument lingvisticheskogo analiza* [A Database of cross-linguistic equivalences as an instrument of linguistic analysis]. *Computational Linguistics and Intellectual Technologies. Proceeding of the Annual Conference "Dialogue"* (2016), Vol. 15 (22), Moscow, RGGU, pp. 763–775.
16. *Zaliznyak, Anna A., Levontina I. B.* (1996), *Otrazhenie natsional'nogo kharaktera v leksike russkogo yazyka* [The reflection of the national character in the vocabulary of the Russian language], *Russian Linguistics*, Vol. 20, pp. 237–264.
17. *Zatsman I. M., Inkova O. Yu., Kruzhkov M. G., Popkova H. A.* (2016), *Predstavlenie krossyazykovykh znaniy o konnektorakh v nadkorpusnykh bazakh dannykh* [Representation of cross-lingual knowledge about connectors in supracorpora databases], *Informatika i ee primeneniya* [Informatics and Applications], Vol. 10/1, pp. 106–118.

# КАК НАМ БЫТЬ С КОНСТРУКЦИЯМИ ТИПА *КАК БЫТЬ?*<sup>1</sup>

**Иомдин Л. Л.** (iomdin@iitp.ru)

Институт проблем передачи информации  
им. А. А. Харкевича РАН, Москва, Россия

В статье рассматривается ряд сложных микросинтаксических единиц русского языка, образованных служебными словами и словами широкой семантики, такими как местоимения, союзы, частицы и связочные глаголы. Одна микросинтаксическая единица, *как быть*, описывается весьма подробно. Рассматриваются проблемы адекватного представления таких единиц в лингвистических ресурсах и в задачах автоматической обработки текстов. Особое внимание уделяется двум типам ресурсов: Микросинтаксическому словарю и аннотированному корпусу русских текстов «СинТагРус», в который встроена микросинтаксическая разметка.

**Ключевые слова:** микросинтаксис, синтаксические фраземы, микросинтаксический словарь, корпус с микросинтаксической разметкой

## WHAT TO DO ABOUT CONSTRUCTIONS LIKE *WHAT TO DO?*

**Iomdin L. L.** (iomdin@iitp.ru)

Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Moscow, Russia

The paper discusses a number of challenging microsyntactic units of Russian formed with functional and broad semantics words, like pronouns, particles, conjunctions and auxiliary verbs. One microsyntactic unit, “*kak byt*” (≈ “what to do”), is described in full detail. Issues of adequate presentation of such units in linguistic resources and their treatment in natural language processing tasks are considered. Special emphasis is laid on two types of resources: the Microsyntactic dictionary, developed as a specific lexicographic resource that can be used in theoretical linguistics and natural language processing, and a deeply annotated corpus of Russian, *Syntagrus*, supplied with microsyntactic tagging. The Microsyntactic dictionary consists of lexical entries compiled according to a uniform pattern that comprises 9 zones: (1) name of microsyntactic unit; (2) type of unit (syntactic idiom or nonstandard construction); (3) lexical composition of the unit

---

<sup>1</sup> Работа выполнена при поддержке Российского гуманитарного научного фонда (грант № 15-04-00562). Автор выражает фонду искреннюю признательность.

(individual words of the syntactic idiom or classes of words forming the non-standard construction, if any); (4) analytical lexicographic definition of the unit; (5) government pattern; (6) syntactic properties of the unit presented in terms of dependency grammar, including the identification of syntactic relations that hold between the unit's constituent elements, and relations that incorporate this unit into the sentence; (7) synonyms and analogues; (8) comments and (9) illustrations. The microsyntactic markup of Syn-TagRus shows what units are present, how they fit in the structure and which of the polysemous entities occur in the tagged sentence.

**Keywords:** Microsyntax, syntactic idioms, microsyntactic dictionary, corpus tagged for microsyntactic elements

## 1. Вводные замечания. Микросинтаксические конструкции, образованные служебными словами

Среди конструкций, так или иначе относящихся к сфере фразеологии, особое место занимают такие, которые составлены из слов весьма широкой или размытой семантики. Часто это служебные слова — местоимения, союзы, предлоги, частицы, связочные глаголы. Эти конструкции отличаются несколькими существенными особенностями, из которых мы упомянем четыре:

- (1) повышенный уровень некомпозициональности (смысл конструкции может быть настолько сложным и специфичным, что его невозможно вывести из смысла входящих в нее элементов даже посредством цепочки семантических переходов);
- (2) повышенный уровень неоднозначности (такие фразеологические конструкции могут иметь более одного значения сами по себе и иметь омонимичные свободные словосочетания);
- (3) повышенная употребительность; образно говоря, у конструкций, собранных из служебных слов, не было бы шанса утвердиться в языке, если бы они не использовались его носителями на каждом шагу;
- (4) повышенный уровень синтаксической сложности (во многих из таких конструкций нелегко определить типы синтаксических связей между их составными элементами или даже сам факт наличия таких связей, а в некоторых случаях бывает трудно установить даже частеречный статус этих элементов).

Примерами этих конструкций являются такие микросинтаксические единицы, как составные частицы *пока что*, *только что*, *разве что*, многозначные наречные речения *все равно* или *как раз*, союз *как будто*, многозначная единица *как бы* (частица или союз), конструкции типа *вот оно что*, *ни за что* и многие другие.

Одни из этих конструкций несомненно содержат несколько отдельных слов, другие (скажем, частицы *как бы* или *что ли*) в процессе языковых изменений движутся по направлению к единому слову подобно тому, как это произошло с союзом/наречием *якобы* (которое, согласно данным НКРЯ, могло писаться раздельно вплоть до начала XX века) и подобно партикулам Т. М. Николаевой

(2008), таким как *либо* и *неужели*, исторически образовавшимся из нескольких отдельных формантов.

В данной группе конструкций интересную подгруппу образуют единицы, в составе которых присутствуют связочные глаголы, прежде всего *быть* и *стать*. Даже при самом беглом взгляде на единицы *так и быть*, *была не была*, *будь что будет*, *стало быть*, *не стало (такого-то)*, *будешь (чай, кофе и т.п.)* становится ясно, что в их значение входят семантические компоненты, далеко выходящие за пределы семантики связочности.

Так, в предложениях

- (1) — *Удостоверений мы обычно не даём, — ответил кот, насупившись, — но для вас, **так и быть**, сделаем исключение* (М. А. Булгаков)

и

- (2) *Эй! вы там! Подайте мне корону. **Так уж и быть**, остаюсь на престоле* (Е. Л. Шварц)

вводный оборот *так и быть* обозначает **согласие говорящего выполнить просьбу слушающего после уговоров последнего и собственных колебаний** (то, что в «Золушке» Шварца король притворяется, что к нему поступила просьба не отрекаться и что он колебался, выполнить ли ее, только подтверждает верность толкования).

В предложениях

- (3) *В театре я не находила такой роли и поэтому однажды — **была не была!** — решила сниматься в кино* (Г. Александров),
- (4) *Не хочется попадаться немцам в руки. Ну ладно, думаю, **была не была!** Высунулся на улицу* (Ф. Искандер),
- (5) *Оставалась надежда только на «**была не была!**»* (Е. Евтушенко)

вводный оборот *была не была*, удивительным образом составленный из двух глаголов *быть* в женском роде, которые, несмотря на очевидную перспективность, имеют форму прошедшего времени, означает, что **говорящий готов совершить действие, допускающее два противоположных исхода и чреватое для него большим риском, нередко смертельным** (недаром этот оборот часто используется в военной прозе).

А конструкция *будь что будет*, в некоторых случаях означающая почти то же самое, что и *была не была*, а именно готовность идти на серьезный риск, как в примерах

- (6) *Несколько дней Глаша не была у себя в подвале, ночевала в прачечной, на гладильном столе, и как-то под утро, вдруг проснувшись, решила: **будь что будет**, пойду в полицию.* (Л. Юзефович),
- (7) *Схватился за голову, сел в машину. Эх, **будь что будет!** Поехал на малой скорости домой.* (Н. Мордюкова)

в целом значительно отличается от *была не была*, во-первых, выражая не столько готовность к риску, сколько **граничащее с фатализмом безразличие к результату действия или бездействия говорящего**, и, во-вторых, необязательной перспективностью: можно что-то сделать или решить, а потом сказать *будь что будет*. Интересно встречающееся порой столкновение этих конструкций в одном тексте, как в любовном стихотворении С. Гандлевского, где легкомысленность усиливается до безрассудства:

- (8) *...И в окне зацвела темнота. Будь что будет, вернее, была не была —  
Выцветай, как под утро фонарь на столбе, Доцветай, как сирень на столе  
доцветала... И сиреневый сор я смахнул со стола.*

В предложении

- (9) *Так вот её-то внучонок, мой, стало быть, правнучек, прибежал ко мне  
(П. П. Бажов)*

вводный дискурсивный оборот *стало быть* обозначает, что говорящий усматривает **логическое следствие** между двумя ситуациями, эксплицитно или имплицитно содержащимися в тексте.

Любопытно, что у этого оборота есть два однословных синонима *следовательно* и *значит*, которые не вполне взаимозаменяемы. К примеру, *стало быть* в

- (10) *Ну вот, стало быть, я иконку на грудь пришил и побежал...  
(М. А. Булгаков)*

можно заменить на *значит*, но вряд ли — *на следовательно*.

*Стало быть* состоит из двух глаголов с жестким порядком следования и с жестко фиксированными формами — *стать* в среднем роде прош. времени и *быть* в инфинитиве.

В противовес такой жесткости словоформ конструкция *не стало*, управляющая подлежащим исключительно в родительном падеже и имеющая смысл ‘умер’, допускает колебание формы времени глагола *стать*; ср.

- (11) *Через полгода у него внезапно обнаружили рак, и вскоре его не стало  
(А. Городницкий)*

и

- (12) *Завтра, как только ее не станет, я тоже умру. (Е. Шварц).*

Добавим, что в этой конструкции *стать* может даже оказаться в инфинитиве в контексте модального глагола:

- (13) *Ей было не до нее, мама тогда была совсем плоха, и однажды Ольга вдруг  
ясно увидела, что мамы может не стать (Г. Щербакова).*

А вот причастные и деепричастные варианты этой конструкции начисто исключены: *\*Этой больной, не ставший (не став) сегодня утром...*

Наконец, разговорная конструкция с глаголом *быть*, обозначающим в первом приближении **потребление еды и питья**, обладает рядом уникальных сочетаемостных, семантических и прагматических свойств.

В частности, она содержит глагол *быть* исключительно в личной форме и будущем времени; ср.

(14) *Я буду грушу и шарлотку,*

но не \**Я был грушу и шарлотку*, \**Я могу <хотел бы> быть грушу и шарлотку* и не \**Будучи только грушу и шарлотку, ты останешься голодным.*<sup>2</sup>

Во вторых, этот глагол, в отличие от любых базовых лексических значений *быть* — экзистенциального, связочного, локативного — допускает в нашей конструкции прямое дополнение в винительном падеже (ср. пример (14)) или — при отрицании — в родительном падеже (ср. (15)),

(15) *Я сказал: «Не буду каши, буду есть сегодня мед»* (Б. Симкин).

которое может заполняться словом со значением пищи, питья, а также, неожиданным образом, со значением курева: *Будешь сигаретку?* (Par extension, здесь могут появляться и слова со значением меры и количества, а также местоимения (*Буду рюмку коньяку; Буду пару пирожков; Вы что будете; Вы что-нибудь будете?* — *Ничего не буду, спасибо*). Расширение этого семантического класса невозможно: нельзя сказать что-либо вроде \**Я буду инъекцию аспирина*.

В третьих — и это представляется самым любопытным — смысловой компонент 'потреблять (еду или питье)' не исчерпывает значения нашей конструкции. Она употребляется тогда, когда потребитель оказывается в той или иной **актуальной ситуации обслуживания**: ему предлагают еду (или питье, или закурить), а он принимает или не принимает это предложение. Предлагать — вербально и невербально — может кто угодно: официант в ресторане, хозяйка, кормящая семью обедом, или приятель. При этом неважно, каков статус этого потребителя: это может быть обычный человек, маленький ребенок, не умеющий говорить, или даже животное: можно сказать *Мальш вряд ли будет эту молочную смесь* или *Кот ни за что не будет Вискас, даже не предлагай ему*. (Разумеется, в последних двух случаях вряд ли будет уместен глагол в первом или во втором лице). Невозможно представить высказывания типа *Смотри, бабушка завтракает. #Сейчас она будет чай с молоком. #Я скоро приду домой и буду пельмени* (отсутствует обслуживающее лицо) или даже *#Я скоро приду домой, жена мне приготовит обед, и я буду пельмени* (отсутствует актуальность).<sup>3</sup>

<sup>2</sup> Данное ограничение можно было бы объяснить эллипсисом полнозначного глагола, пусть и лексически неопределенного: *буду грушу и сок < буду есть грушу и пить сок*, если бы значения таких выражений были тождественными. Однако, как мы увидим ниже, это далеко не так.

<sup>3</sup> Эта конструкция исследовалась автором в некоторых публичных выступлениях и лекциях. Результаты исследования были учтены (и несколько переосмыслены) в статье Апресян 2010 (с соответствующей ссылкой на эту неопубликованную работу) и в написанной Ю. Д. Апресяном словарной статье глагола БЫТЬ Активного словаря русского языка. Сам автор публикует, однако, эти результаты впервые.

Приведенные примеры убедительно демонстрируют, насколько прихотливо и непредсказуемо могут быть устроены фраземные конструкции, сформированные из служебных слов, и какого детального описания они требуют. Разумеется, эта цель становится еще более сложной при решении компьютерно-лингвистических задач, в частности, задачи глубокого семантического анализа.

## 2. Микросинтаксический словарь и конструкция «как быть»

Предпримем теперь попытку описать одну из таких конструкций с несколько большей степенью подробности. Это описание выполняется в рамках создаваемого под руководством автора Микросинтаксического словаря русского языка, который мыслится как отдельный лингвистический ресурс, призванный вывести на новый уровень описание синтаксической фразеологии. Важным дополнением к этому ресурсу служит производимая в настоящее время микросинтаксическая разметка глубоко аннотированного корпуса текстов СинТагРус (см. Маракасова — Иомдин 2016, Iomdin 2016.)

Микросинтаксическая разметка корпуса будет кратко охарактеризована в следующем разделе.

Статьи Микросинтаксического словаря составляются по следующей схеме, в значительной мере опирающейся на опыт создания нового Активного словаря русского языка под ред. Ю. Д. Апресяна (2014):

- а) имя микросинтаксической единицы;
- б) тип микросинтаксической единицы (синтаксическая фразема или нестандартная синтаксическая конструкция)<sup>4</sup>;
- в) лексический состав единицы (конкретные слова синтаксической фраземы или классы слов, входящих в состав нестандартной синтаксической конструкции, если таковые имеются);
- г) лексикографическое толкование единицы (по возможности, аналитическое);
- д) модель управления (если она имеется);
- е) синтаксические свойства единицы, формулируемые в терминах грамматики зависимостей и включающие идентификацию синтаксических отношений между элементами единицы и синтаксических отношений, встраивающих данную единицу в предложение;
- ж) синонимы и аналоги;
- з) комментарии;
- и) иллюстрации.

---

<sup>4</sup> Синтаксическая фразема — это лексикализованная единица, образованная из конкретных слов (таковы все рассмотренные в разделе 1 единицы); нестандартная синтаксическая конструкция не привязана к конкретным словам (например, конструкции с повторяющимися элементами, такие как *X-овать не X-овал* 'не X-овал, делал что-либо более слабое, чем X', ср. *Дружить не дружили, но были приятелями*). Подробнее см. Иомдин 2013, 2014.



Приведем (в несколько сокращенном варианте, в неформальном виде и с необходимыми комментариями) статью Микросинтаксического словаря, представляющую синтаксическую фразу *как быть*. Эта синтаксическая фраза выступает в таких предложениях, как *Как нам быть с завтрашним семинаром?*, *Никто не знает, как быть в таких ситуациях* и т.д.

**Имя:** КАК БЫТЬ

**Тип:** синтаксическая фраза

**Лексический состав:** БЫТЬ, КАК<sup>5</sup>

**Лексикографическое толкование:**

*Как X-у быть с Y-ом в P?*

**Пресуппозиция:**

Лицо X является участником ситуации P, в которой имеет место факт, предмет или лицо Y. Говорящий считает, что Y создает для X-а трудности в P'.

**Ассерция:**

Говорящий хочет знать, что лицо X должно сделать, чтобы Y перестал создавать для X-а трудности в ситуации P'.

**Модель управления.**

В семантической структуре микросинтаксической единицы *как быть* имеется три переменных, соответствующих (1) валентности ситуации P, (2) валентности агенса X и (3) валентности объекта («помехи») Y.

В высказываниях, содержащих *как быть*, могут быть насыщены все три эти валентности или часть их, а может быть не насыщена ни одна:

- (16) *Как преподавателю (X) быть на экзамене (P) со студентами (Y), которые пользуются шпаргалками?* (выражены X, P и Y);
- (17) *Посоветуйте, как быть в такой ситуации (P)* (не выражено ни X, ни Y);
- (18) *Но как мне (X) быть с моей грудной клеткой И с тем, что всякой костности косней (Y)?* — Б. Л. Пастернак (не выражено P);
- (19) *Не знаю даже, как и быть* (не выражено ни P, ни X, ни Y).

Тем самым все валентности *как быть* являются факультативными.

Примечание: В полной семантической структуре, содержащей эту единицу, непременно должна фигурировать еще одна переменная, соответствующая слушающему, или контрагенту X2, от которого ожидается, что он сообщит X-у запрашиваемую им информацию. Однако эта переменная характеризует любые вопросительные высказывания и не должна входить в валентную структуру конкретных языковых единиц, в том числе и синтаксической фразы *как быть*, несмотря на её ингерентную вопросительность.

<sup>5</sup> КАК2 — местоименное наречие, в отличие от союза КАК1. Идентификация лексических единиц производится по комбинаторному словарю лингвистического процессора ЭТАП-3 (Iomdin et al. 2012), на основе которого строится и аннотированный корпус СинТагРус (Дяченко и др. 2015).

## 2.1. Синтаксические свойства

Обязательные элементы фраземы — глагол БЫТЬ, принимающий форму инфинитива, и вопросительное наречие КАК2. Последнее подчиняется инфинитиву БЫТЬ по обстоятельственному синтаксическому отношению (СинтО)<sup>6</sup>.

Валентность субъекта X выражается именной группой в дательном падеже. В синтаксической структуре эта именная группа подчиняется глаголу БЫТЬ по дательно-субъектному СинтО.

Валентность объекта Y выражается группой, вводимой предлогом С, который в синтаксической структуре подчиняется глаголу БЫТЬ по 1-му комплетивному СинтО (т.е. отношению, которое вводит первое дополнение). Тем самым имплицитно предполагается, что в пределах данной микросинтаксической единицы БЫТЬ имеет особое лексическое значение, не являющееся ни экзистенциальным, ни связочным, ни локативным (ср. аналогичное решение, принятое для БЫТЬ в рассмотренных в разделе 1 конструкциях типа *буду кофе и булочку*). Эта лексическая единица семантически и синтаксически близка к глаголу ПОСТУПАТЬ (хотя, разумеется, не тождественна ей).

Валентность ситуации Р выражается достаточно разнообразно. Помимо групп, вводимых локативными предлогами В, НА (ср. примеры (17) и (18) выше), а также ПРИ, как в

(20) *Петр Ильич — тот в письмах к благотельнице фон Мекк высказывал свои соображения о музыке и искусстве, изливал душу, а как быть при телефонах?* (В. Орлов)

она выражается еще и придаточными условия и цели, вводимыми союзами ЕСЛИ, КОГДА и ЧТОБЫ:

(21) *Но как же быть, если отказать высказанным Бердяевым суждениям в принудительной силе?* (Л. И. Шестов);

(22) *Долго сидела она с Лидией Александровной... и советовалась о жизни: как ей быть, когда Николай придет?* (Ю. Трифонов);

(23) *Как быть с таким человеком, чтобы дом показался ему именно домом, а не дворцом?* (Н. Г. Чернышевский).

Валентность Р представляется разумным в синтаксической структуре подчинять глаголу БЫТЬ по 2-му комплетивному отношению.

<sup>6</sup> В описании единиц Микросинтаксического словаря используется синтаксический формализм (в том числе и арсенал синтаксических отношений), разработанный для лингвистического процессора ЭТАП-3 (см. о нем, например, Iomdin et al. 2012). Тем самым обеспечивается совместимость этого словаря со всеми компонентами процессора.

## 2.2. Синонимы и аналоги

Точного синонима у синтаксической фраземы *как быть* нет. Аналогом является полусвободное сочетание *что делать*: *Как мне быть с опоздавшими? ≈ Что мне делать с опоздавшими?*

## 2.3. Комментарии

1. Микросинтаксическая единица *как быть* абсолютно не допускает отрицания при глаголе: *\*Как мне не быть с этой проблемой?* Выражение *Как не быть* — это реализация совершенно другой конструкции, имеющая значение, близкое к 'естественно быть', ср.

(24) *А девушки там рядом есть? — Как не быть...*  
(А. Стругацкий, Б. Стругацкий),

где *быть* имеет экзистенциальное значение, и

(25) *В своей поэзии Бунин сумел сделать много прекрасного. Как не быть ему благодарным?* (В. Ходасевич),

где *быть* выступает в связочном значении: и в (24), и в (25) фигурируют обычные лексические значения *быть*, а не то фразеологическое ≈ 'поступать', которое имеет место в данной синтаксической фраземе. Добавим, что скорее всего, мы имеем тут дело с другой синтаксической фраземой *как не + X-овать* (≈ 'естественно, что имеет место X'), которая формирует риторический вопрос.

2. Синтаксическая фразема *как быть*, в отличие от только что рассмотренной *как не + X-овать*, в общем случае формирует обыкновенный, риторический вопрос. Тем не менее из материала этого вопроса практически невозможно породить стандартный ответ, заменив в нем вопросительное слово. Трудно представить себе диалог типа *Как быть? — Быть так-то и так-то*. Редким исключением является диалог типа *Как быть? — (А) никак не быть*.

3. Единица *как быть*, как ясно из примеров в зоне модели управления, перенимает обычное для инфинитива управление дательным падежом субъекта. Что касается управления предлогом *с*, вводящим объект, то такое управление свойственно глаголу *быть* исключительно в составе данной микросинтаксической фраземы. Вне этой фраземы стоящий при *быть* предлог *с*, вводящий именную группу в творительном падеже, выступает в самостоятельном значении, чаще всего в значении совместности, а само *быть* утрачивает связь с идеей поведения, ср.

(26) *Где нам быть с братом?*,

(27) *Дети могли быть с няней*

и даже

(28) *Как нам быть с детьми в таком маленьком номере?*

4. Глагол *быть* в составе данной синтаксической фраземы может выступать в качестве вершины главного или придаточного предложения, но иногда может и подчиняться глаголу во втором лице из небольшого списка; ср.

(29) *Как вы посоветуете <порекомендуете> быть в такой ситуации?*

(30) *Как прикажете быть с соблюдением авторских прав?*

В (30) имеет место своеобразная формула, в которой обычный смысл глагола *приказывать* выветривается, так что *как прикажете быть* означает примерно ‘как вы ответите на вопрос „как быть“’.

Изредка вершиной предложения, в которое входит *как быть*, может выступать глагол *было* в прошедшем времени, обращая всю ситуацию в прошлое, ср.

(31) *Что же. Как было быть князю. Он мужчина еще не старьй, в поре, про которую сказано: «Не добро быть человеку одну...» (Ф. Н. Плевако).*

(32) *Вот как было быть С.А., когда этот граф днем писал одно, а по ночам делал другое, а потом наутро в дневнике плохими словами осыпал жену за то, что она, видите ли, вызывает у него «дурные» мысли (из отзыва о книге Софьи Андреевны Толстой).*

5. Два основных словесных элемента фраземы *как быть* почти не имеют собственных синтаксических зависимых, не являющихся зависимыми всей фраземы в целом. Исключения составляет частица *же* при *как* и частица *–то* при *быть*, вместе или порознь: последовательности *как же быть*, *как быть–то*, *как же быть–то* весьма частотны. Спорадически встречаются и другие зависимые, ср. наречие *правильно* в (33)

(33) *И как правильно быть, если мать продала нас цыганам, нет, хуже — бросила собакам? (А. И. Солженицын).*

Такие случаи представляются достаточно экзотичными.

6. Как говорилось в зоне толкования, смысл единицы *как быть* в ассертивной части содержит компонент ‘хочет знать’. Отсюда вытекает, что придаточное предложение с *как быть* может подчиняться далеко не любому предикату главного предложения, способному управлять инфинитивным оборотом с вопросительным словом. Нормально сказать *Спроси, как мне быть*; *Не скажу даже, как вам быть с билетом*, но ненормально выразиться *\*Мне все равно <безразлично, неважно>, как ему быть*, *\*Я забыл, как мне быть*, поскольку при этом возникает смысловое противоречие.

## 2.4. Иллюстрации и примеры

Микросинтаксическая единица *как быть* весьма частотна и встречается в текстах всех жанров: в художественной литературе, нейтрально-деловой и научно-популярной прозе, поэзии, в разговорной речи. В научно-технической литературе и в официальных документах единица встречается реже.

### **Примеры из основной части Национального корпуса русского языка и корпуса СинТагРус**

*Если уже вы, то есть, не поможете в нашей просьбе, то уж не знаем, как и быть: просто хоть в петлю полезай. (Н. В. Гоголь).*

*Проблема двуязычия легче разрешима, скажем, в Латвии или Армении, а как быть уйгуру или дунганину в Киргизии?*

*А как же быть с единой энергетической системой, связью или, скажем, добычей драгоценных металлов?*

Другие примеры из НКРЯ см. выше.

### **Примеры из поэтического корпуса НКРЯ**

*Как нам быть с тем ужасом, который был бегом времени  
когда-то наречён? (А. Ахматова)*

*Как мне быть с мечтаньями твоими, Устремленными далеко ввысь?  
(М. Светлов).*

*Как же нам быть с этим знаком, Ставить его или нет? Надо, чтоб сами  
глаголы Людям давали ответ (Э. Успенский).*

## **3. Микросинтаксическая разметка корпуса СинТагРус**

Известно, что для решения разных компьютерно-лингвистических задач, например, для разрешения многозначности лексических единиц, полезно использовать лексически размеченные корпуса текстов, в которых многозначным словам приписывается определенное лексическое значение, заданное каким-либо лексикографическим ресурсом.

Однако корпуса текстов, хотя бы частично размеченные словесными значениями, очень полезны и при решении задач разрешения неоднозначности в рамках теоретической семантики и лексикографии без непосредственной привязки к проблематике компьютерной лингвистики (см., например, Б. Иомдин 2014). Между тем, если корпуса, размеченные значениями обычных слов, уже имеются в нашем распоряжении, то подобных ресурсов, в которых бы аннотировались какие-либо фразеологические единицы, насколько нам известно, до последнего времени не существовало.

В рамках исследования микросинтаксиса мы решили восполнить эту лауну и ввели в число размечаемых свойств корпуса СинТагРус указания о присутствующих в предложениях корпуса микросинтаксических единицах.

На снимке экрана (см. рис. 1) приводится фрагмент СинТагРуса, в котором встречается рассматриваемая в данной статье единица *как быть*. Для каждой фразы приводятся границы этой единицы и ее имя, а, заглянув в синтаксическую структуру предложения, можно увидеть, как единица встраивается в эту структуру. Эту информацию предполагается использовать как для разрешения

многозначности лексических единиц, так и для задач глубокого семантического анализа.

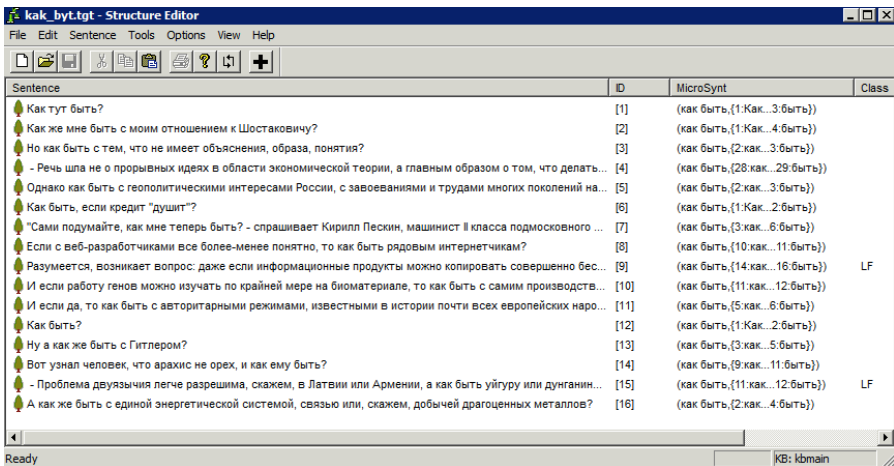


Рис. 1. Микросинтаксическая разметка корпуса СинTagРус

Следует отметить, что в случае с микросинтаксической единицей *как быть* материал СинTagРус'a обнаруживает редкое единообразие. Во всех предложениях, в которых содержится синтаксическая конфигурация, состоящая из инфинитива *быть* и подчиненного ему местоименного наречия *как*, присутствует именно указанная единица, а примеров типа (23)–(25) не встречается вовсе. Такая ситуация нетипична для микросинтаксической разметки корпуса: в других случаях мы чаще имеем дело с многозначностью самой единицы или случайным соположением входящих в нее словесных элементов (т. н. false positives разметки). Данное обстоятельство можно считать косвенным свидетельством фразеологической силы рассмотренной единицы.

## Литература

1. *Активный словарь русского языка* (2014). Т. I. Авторы: Апресян В. Ю., Апресян Ю. Д., Бабаева Е. Э., Богуславская О. Ю., Галактионова И. В., Гловинская М. Я., Иомдин Б. Л., Крылова Т. В., Левонтина И. Б., Лопухина А. А., Птенцова А. В., Санников А. В., Урысон Е. В. Под ред. акад. Ю. Д. Апресяна. М.: Языки славянской культуры. 408 с.
2. *Апресян Ю. Д.* (2010). Принципы лексикографического описания многозначных слов. // *Исследования по лингвистике и семиотике. Сборник статей к юбилею Вяч. Вс. Иванова. М., Языки славянских культур*, с. 217–233.
3. *Дяченко П. В., Иомдин Л. Л., Лазурский А. В., Митюшин Л. Г., Подлеская О. Ю., Сизов В. Г., Фролова Т. И., Цинман Л. Л.* (2015). Современное состояние глубоко аннотированного корпуса текстов русского языка (СинТагРус) // *Национальный корпус русского языка. 10 лет проекту. Труды Института русского языка им. В. В. Виноградова. М. Вып. 6. С. 272–299.*
4. *Иомдин Б. Л.* (2014). Многозначные слова в контексте и вне контекста. Вопросы языкознания, № 4. С. 87–103.
5. *Иомдин Л. Л.* (2013). ЧИТАТЬ НЕ ЧИТАЛ, НО...: об одной русской конструкции с повторяющимися словесными элементами // *Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции «Диалог» (2013). М.: Изд-во РГГУ. Вып. 12. Т. 1. С. 272–284.*
6. *Иомдин Л. Л.* (2014). *Хорошо меня там не было: синтаксис и семантика одного класса русских разговорных конструкций* // Сб. статей «Grammaticalization and Lexicalization in the Slavic Languages». По материалам Международного симпозиума «Грамматикализация и лексикализация в славянских языках», 11–14 ноября 2011 г. München-Berlin-Washington/D.C.: Verlag Otto Sagner. Band 55. P. 423–436.
7. *Маракасова А. А., Иомдин Л. Л.* (2016). Микросинтаксическая разметка в корпусе русских текстов СинТагРус // *Информационные технологии и системы 2016 (ИТиС'2016). Сборник трудов 40-ой междисциплинарной школы-конференции ИППИ РАН. Репино, Санкт-Петербург. С. 445–449.*
8. *Николаева Т. М.* (2008). *Непарадигматическая лингвистика (История «блуждающих частиц»)*. М.: Языки славянских культур. — 689 с.
9. *Iomdin Leonid* (2016). *Microsyntactic Phenomena as a Computational Linguistics Issue.* // *Grammar and Lexicon: Interactions and Interfaces. Proceedings of the Workshop. Osaka, Japan. 2016, pp. 8–18.* (<http://aclweb.org/anthology/W/W16/W16-38.pdf>). ISBN 978-4-87974-706-8
10. *Iomdin L., Petrochenkov V., Sizov V., Tsinman L.* (2012). ETAP parser: state of the art // *Computational Linguistics and Intellectual Technologies. International Conference (Dialog'2012). Moscow: RGGU Publishers. Issue 11(18). P. 830–843. ISSN 2221–7932.*

## References

1. *Active dictionary of Russian. [Aktivnyj slovar' russkogo jazyka]*. (2014). Vol I. Authors: Apresjan V. Yu., Apresjan Yu. D., Babaeva E. E., Boguslavskaya O. Yu., Galaktionova I. V., Glovinskaya M. Ya., Iomdin B. L., Krylova T. V., Levontina I. B., Lopukhina A. A., Ptentsova A. V., Sannikov A. V., Uryson E. V. Edited by Yu. D. Apresjan. Moscow: Jazyki slvjanskoj kultury [ ]. 408 p. (In Russian).
2. *Apresjan Yu. D.* (2010). The principles of lexicographic descriptions of polysemous words [Principy leksikograficheskogo opisanija mnogoznachnykh slov. // Issledovanija po lingvistike i semiotike Sbornik statej k jubileju Vjach. Vs. Ivanova. Moscow: Jazyki slavjanskix kultur, p. 217–233. (In Russian).
3. *Djachenko P. V., Iomdin L. L., Lazursky A. V., Mitjushin L. G., Podlesskaja O. Yu., Sizov V. G., Frolova T. I., Tsinman L. L.* (2015). The state-of-the art of a deeply annotated corpus of Russian texts (SynTagRus). [Sovremennoe sostojanie gluboko annotirovannogo korpusa tekstov russkogo jazyka (SynTagRus)] // Natsionalnyj korpus russkogo jazyka. 10 let proektu. Trydy Instituta russkogo jazyka im. V. V. Vinogradova. Moscow. Vyp. 6. p. 272–299. (In Russian).
4. *Iomdin B. L.* (2014). Polysemous words in context and out of context. [Mnogozhachnye slova v kontekste I vne konteksta. Voprosy jazykoznanija, N 4. p. 87–103. (In Russian).
5. *Iomdin L. L.* (2013). CHITAT' NE CHITAL, NO: On a Russian construction with repeated word elements. [CHITAT' NE CHITAL, NO: ob odnoj russoj konstrukcii s povtorjajushchimisja slovesnymi elementami] // Kompjuternaja lingvistika I intellektualnye tekhnologii. Po materialam ezhegodnoj Mezhdunarodnoj konferencii “Dialog” (2013). Moscow: Izdatelstvo RGGU. vyp. 12. T. 1. p. 272–284. (In Russian).
6. *Iomdin L. L.* (2014). *Good thing I wasn't there*: Syntax and Semantics of a class of Russian colloquial constructions. [Khorosho menja tam ne bylo: sintaksis i semantika odnogo klassa russkikh razgovornyx konstrukcij] // In: Grammaticalization and Lexicalization in the Slavic Languages», November 11–14, 2011. München-Berlin-Washington/D.C.: Verlag Otto Sagner. Band 55. P. 423–436. (In Russian).
7. *Iomdin Leonid* (2016). Microsyntactic Phenomena as a Computational Linguistics Issue. // Grammar and Lexicon: Interactions and Interfaces. Proceedings of the Workshop. Osaka, Japan. 2016, pp. 8–18. (<http://aclweb.org/anthology/W/W16/W16-38.pdf>).
8. *Iomdin L., Petrochenkov V., Sizov V., Tsinman L.* (2012). ETAP parser: state of the art // Computational Linguistics and Intellectual Technologies. International Conference (Dialog'2012). Moscow: RGGU Publishers. Issue 11(18). P. 830–843.
9. *Marakasova A. A., Iomdin L. L.* (2016). Microsyntactic tagging in the SynTagRus corpus of Russian texts. [Mikrositaksicheskaja razmetka v korpuse russkikh tekstov SynTagRus] // Informatsionnye tekhnologii i sistemy 2016 (ITiS'2016). Sbornik trudov 40-j mezhdistsiplinarnoj shkoly-konferencii IPPI RAN. Repino, Saint Petersburg. p. 445–449. (In Russian).
10. *Nikolaeva T. M.* (2008). (Nonparadigmatic linguistics. (The history of “straying particles”) [Neparadigmaticheskaja lingvistika (Istorija “bluzhdajushikh chastic”)]]. Moscow: Jazyki slavjanskikh kultur. — 689 p. (In Russian).



# КОАРТИКУЛЯЦИОННЫЕ ИЗМЕНЕНИЯ НОСОВЫХ СОГЛАСНЫХ КАК ПОКАЗАТЕЛЬ НАЛИЧИЯ ПРОСОДИЧЕСКОГО ШВА И ПОРЯДКА ПРИМЕНЕНИЯ ФОНОЛОГИЧЕСКИХ ПРАВИЛ В РУССКОМ ЯЗЫКЕ<sup>1</sup>

**Князев С. В.** (svknia@gmail.com),

**Утешева А. С.** (Utesheva102@gmail.com)

МГУ им. М. В. Ломоносова<sup>(1,2)</sup>,

Национальный исследовательский университет  
«Высшая школа экономики»<sup>(1)</sup>; Москва, Россия

**Ключевые слова:** фонетика, коартикуляция, носовые согласные, просодическое членение, паузальный маркер, порядок применения фонологических правил

## COARTICULATION OF NASALS AS A CUE FOR DETECTING PROSODIC BREAKS AND RULE ORDER IN STANDARD RUSSIAN

**Knyazev S. V.** (svknia@gmail.com),

**Utesheva A. S.** (Utesheva102@gmail.com)

Moscow State University<sup>(1,2)</sup>, Higher School of Economics<sup>(1)</sup>;  
Moscow, Russia

The paper reports some results of the research, aimed at finding out 1) whether place coarticulation occurs in clusters of [labial or dental nasal + labiodental obstruent] within the phonological word and in external sandhi position in Modern Standard Russian, and 2) whether it may serve as a cue for detecting the presence of prosodic breaks and the phonological rules order.

The results obtained show that the F2 value of nasal before labiodental obstruent is significantly higher for bilabial obstruent and significantly lower for coronal one in comparison with their F2 values in the position before gomorganic stops. This type of place coarticulation is found only within the phonological word and is not available in an external sandhi position; thus the absence of this type of coarticulation may serve as a cue for detecting the presence of prosodic break.

---

<sup>1</sup> Исследование проведено при поддержке гранта РФФИ 15-06-06103.

In the case of clusters with final palatalized labiodental obstruent, the F2 value of bilabial nasal is found to be noticeably higher than those of the coronal one because a palatalization coarticulation, which exists in Modern Standard Russian for bilabials but not coronals before labiodentals. Thus, we argue that the phonological rule of palatalization operates before the rule for place assimilation in Standard Russian.

**Key words:** phonetics, nasals, coarticulation, prosodic breaks' markers, rules ordering

Носовые согласные в разных языках могут подвергаться различным ассимилятивным изменениям, в том числе по признаку места образования, причем у носовых этот тип ассимиляции фиксируется чаще, чем у взрывных согласных [Cho 1990], так что не существует языков, в которых ассимилировались бы только взрывные согласные, но при этом описаны такие, где ассимиляции подвергаются только носовые [Mohan 1993: 74]. Ярким примером в этом отношении может служить испанский язык, где носовые согласные подвергаются варьированию по месту образования перед последующими губными, межзубными и заднеязычными [Barkhudarova 2011], а также английский, где носовые ассимилируются по месту образования и на стыке слов [Renwick et al. 2013].

Зачастую данное явление связывают с тем, что чаще нейтрализуется контраст, который менее ощутим: «speakers seem to be more willing to neutralize a contrast that is less perceptible» [Huang 2001], этот тип ассимиляции известен как «перцептивно неощутимая ассимиляция» (perceptually tolerated articulatory simplification) [Huang 2001].

**Цель** настоящего исследования заключалась в том, чтобы на основании экспериментально-фонетического анализа реализации губно-губного и переднеязычного носовых в позиции перед губно-зубным согласным (внутри и на стыках фонетических слов) определить, существуют ли коартикуляционные изменения носовых согласных по месту образования в русском языке.

## ЭКСПЕРИМЕНТ 1

В ходе первого эксперимента тестировалась **гипотеза** о том, что в русском языке наблюдается уподобление твердых губных и переднеязычных носовых согласных по месту образования в позиции перед твердым губно-зубным щелевым.

**Материалом** эксперимента служили слова с билабиальным и переднеязычным носовым согласным в позиции после ударного гласного перед лабиодентальным глухим щелевым, то есть с сочетанием  $V_1 C_2 C_4$  (EMF) и  $V_1 C_3 C_4$  (ENF), где  $V_1$  — гласный,  $C_2$  — носовой губно-губной,  $C_3$  — носовой переднеязычный — зубной,  $C_4$  — губно-зубной фрикативный согласный (*Стэнфорд / Стэмфорд*). Основанием для сравнения служили слова, в которых носовой согласный находился перед гомоганным взрывным (*тентом / темпом*). Все тестовые слова удовлетворяли следующим условиям:

- идентичный предшествующий гласный ([э]);
- идентичная позиция исследуемого сочетания относительно ударного гласного;
- идентичное количество слогов в слове;
- положение в идентичной фразовой позиции;
- одинаковая степень освоенности слов.

Варьировалось только место образования носового и последующего шумного согласного.

Каждое слово было помещено во фразу в сильной (под фразовым акцентом) и слабой фразовых позициях. Список фраз приведен ниже в Приложении.

В эксперименте в качестве **информантов** приняли участие 17 носителей русского языка (все женщины в возрасте от 17 до 20 лет). Дикторы представляли различные регионы России (Москва, Брянск, Курск, Нижегородская область, Белгород, Липецк, Орловская область, Тюменская область, Тула, Саратов).

В ходе **эксперимента** записывалось чтение дикторами подготовленного текста (см. Приложение) с тестовыми словами. Запись была произведена при помощи программы Praat. Анализ наличия / отсутствия коартикуляционных изменений осуществлялся путем сравнения качества носового в парах слов *тентом* / *Стэнфорд*, *темпом* / *Стэмфорд*. Для установления характера носового было определено значение  $F_2$  в начале и в конце его артикуляции, а также значение  $F_2$  в конце предшествующего гласного (в начале его последней четверти).

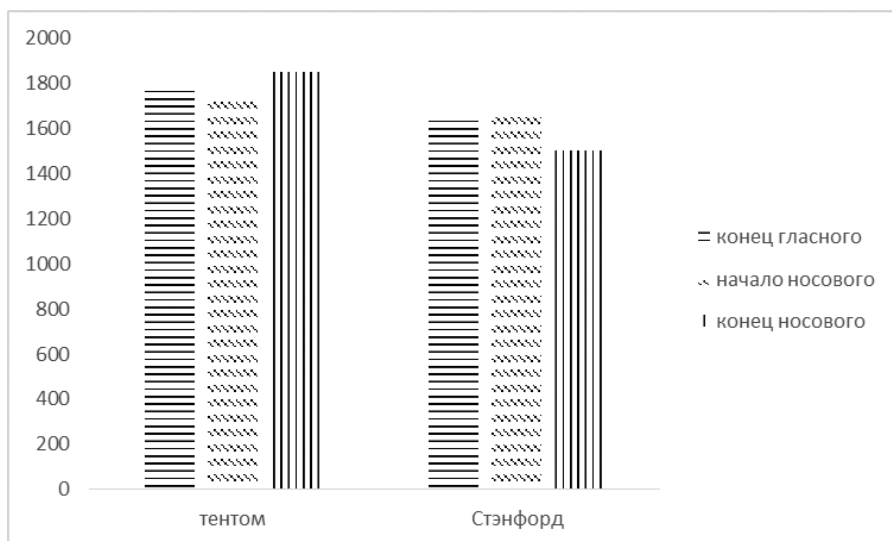
**Результаты исследования** приведены ниже в Таблице 1 и в обобщенном виде на рисунках 1–3.

В ходе эксперимента было установлено, что тип фразовой позиции (сильная или слабая) не влияет на характер коартикуляционных изменений носового согласного в русском языке, поэтому результаты, полученные для разных типов фразовых позиций, объединены.

На основании приведенных на рисунке 1 данных можно сделать вывод о том, что носовой в данной паре слов реализуется по-разному: в слове *тентом* вторая форманта носового повышается к последующему согласному, а в *Стэнфорд*, наоборот, понижается (разница в конце носового в исследованных словах составляет 140 Гц). Помимо этого значение второй форманты в начале носового в слове *Стэнфорд* ниже, чем в слове *тентом* (приблизительно на 80 Гц), а значение конца второй форманты гласного в слове *тентом* выше, чем в слове *Стэнфорд* (примерно на 130 Гц). Эти факты свидетельствуют о том, что в русском языке в словах с сочетанием «переднеязычный носовой + губно-зубной согласный» происходит коартикуляция носового по месту образования. Следует отметить, что значение второй форманты в конце гласного у всех информантов варьировалось достаточно широко, но сама закономерность (в слове *тентом* значение  $F_2$  в конце носового выше, чем в слове *Стэнфорд*) сохранялась у всех участников эксперимента.

**Таблица 1.** Значения  $F_2$  в конце гласного (вверху), начале и конце носового (внизу) в парах слов *тентом / Стэнфорд, темпом / Стэмфорд*

	<i>тентом</i>	<i>Стэнфорд</i>	<i>темпом</i>	<i>Стэмфорд</i>
1. МК (Нижегородская обл.)	1840 1712–1801	1703 1650–1494	1424 1371–1453	1578 1431–1607
2. ПК (Тюменская обл.)	1842 1674–1759	1655 1592–1370	1375 1360–1377	1428 1140–1203
3. ЛЖ (Тула)	1836 1847–1970	1683 1714–1400	1493 1387–1435	1597 1472–1644
4. ВЛ (Саратов)	1840 1770–1830	1717 1646–1512	1526 1466–1473	1698 1549–1562
5. АГ (Московская обл.)	1673 1630–1750	1775 1729–1588	1313 1259–1193	1484 1250–1350
6. ДК (Белгород)	1798 1780–1866	1553 1559–1359	1358 1295–1361	1459 1309–1439
7. СВ (Липецк)	1736 1728–1882	1608 1424–1268	1418 1253–1261	1430 1333–1401
8. МК (Москва)	1620 1475–1746	1400 1429–1316	1289 1111–1196	1384 1350–1450
9. АГ (Москва)	1815 1758–1902	1692 1800–1693	1549 1337–1626	1671 1335–1414
10. ВУ (Москва)	1683 1583–1890	1480 1480–1325	1217 1100–1173	1245 1167–1275
11. БС (Московская обл.)	1950 1930–2016	1747 1852–1683	1411 1321–1467	1318 1275–1505
12. КК (Московская обл.)	1879 1685–1710	1765 1896–1811	1490 1332–1124	1441 1277–1313
13. АГ (Москва)	1785 1707–1908	1678 1650–1532	1342 1234–1256	1451 1324–1396
14. ЖП (Брянск)	1849 1776–1920	1789 1747–1692	1423 1320–1418	1577 1440–1562
15. МБ (Орловская обл.)	1521 1469–1543	1552 1477–1493	1431 1272–1355	прочитано «Стэнфорд»
16. ЛИ (Москва)	1865 1738–1974	1682 1337–1465	1250 1147–1222	1407 1269–1280
17. СЛ (Волгоград)	1765 1720–1815	1608 1456–1550	1308 1156–1289	1421 1203–1298

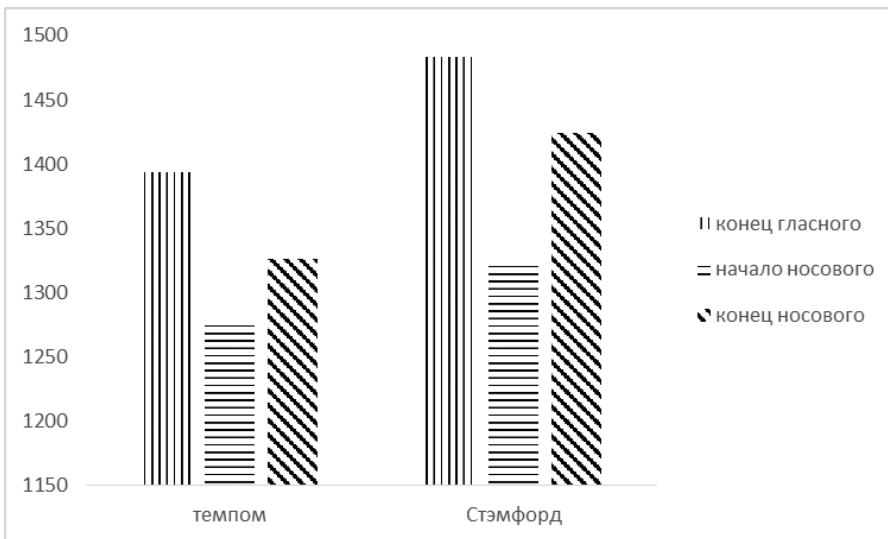


**Рисунок 1.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах *тентом* и *Стэнфорд* (Гц). Усреднено по всем произнесениям в сильной и слабой фразовых позициях

На основании приведенных на рисунке 1 данных можно сделать вывод о том, что носовой в данной паре слов реализуется по-разному: в слове *тентом* вторая форманта носового повышается к последующему согласному, а в *Стэнфорд*, наоборот, понижается (разница в конце носового в исследованных словах составляет 140 Гц). Помимо этого значение второй форманты в начале носового в слове *Стэнфорд* ниже, чем в слове *тентом* (приблизительно на 80 Гц), а значение конца второй форманты гласного в слове *тентом* выше, чем в слове *Стэнфорд* (примерно на 130 Гц). Эти факты свидетельствуют о том, что в русском языке в словах с сочетанием «переднеязычный носовой + губно-зубной согласный» происходит коартикуляция носового по месту образования. Следует отметить, что значение второй форманты в конце гласного у всех информантов варьировалось достаточно широко, но сама закономерность (в слове *тентом* значение  $F_2$  в конце носового выше, чем в слове *Стэнфорд*) сохранялась у всех участников эксперимента.

На основании данных, приведенных на рисунке 2, можно сделать вывод о том, что носовой согласный и в данной паре слов реализуется по-разному: значение  $F_2$  в начале второй форманты носового в слове *тентом* равно примерно 1270 Гц и незначительно увеличивается к последующему согласному (у некоторых информантов остается неизменным), в то время как в слове *Стэнфорд* значение второй форманты в начале носового равно 1330 Гц, а увеличение этого показателя к согласному составляет около 100 Гц. Эти факты в совокупности свидетельствуют о том, что в русском языке в словах с сочетанием «носовой билабиальный + лабиодентальный согласный» происходит уподобление по месту образования. И в этом случае значение второй форманты носового у участников

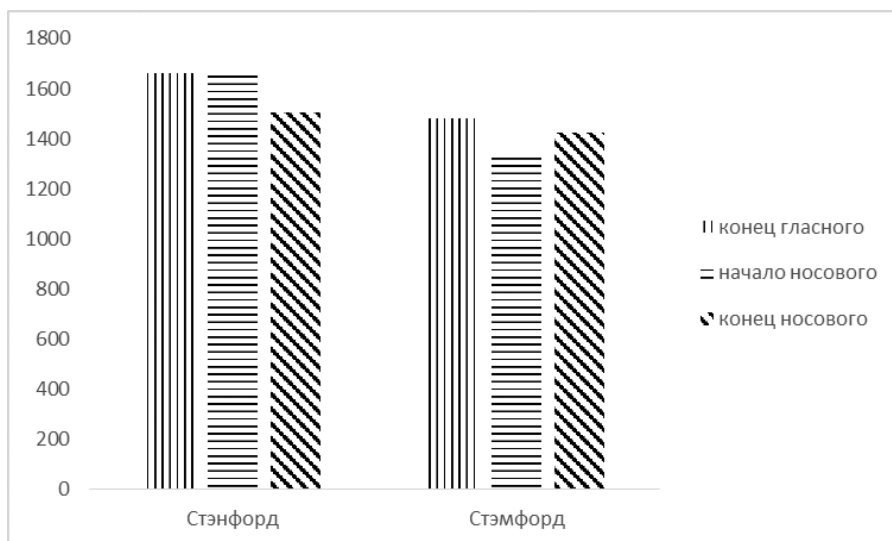
эксперимента варьировалось достаточно широко, но общая закономерность (носовой в слове *Стэмфорд* имеет более высокие значения  $F_2$ , чем в слове *темпом*, и эти значения увеличиваются к последующему согласному) сохраняется у всех дикторов. Кроме того, значение  $F_2$  в конце гласного в слове *Стэмфорд* больше, чем в слове *темпом* (примерно на 100 Гц), что также свидетельствует о том, что в сочетаниях MF носовой по месту образования уподобляется губно-зубному.



**Рисунок 2.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах *темпом* и *Стэмфорд* (Гц). Усреднено по всем произнесениям в сильной и слабой фразовых позициях

Анализ приведенных на рисунке 3 данных позволяет утверждать, что разница в значениях второй форманты в начале носового в паре слов *Стэнфорд/ Стэмфорд* достаточно значительная (примерно 330 Гц), но при переходе к последующему согласному она заметно уменьшается и составляет уже всего около 80 Гц. Этот факт свидетельствует о том, что в сочетаниях /мф/ и /нф/ носовые в конце артикуляции сближаются друг с другом по месту образования и совпадают в губно-зубном носовом согласном. Существенное отличие значений  $F_2$  в начале и конце носового согласного в позиции перед лабиодентальным спирантом свидетельствует, по всей вероятности, в пользу точки зрения о том, что описанный тип уподобления является не фонологическим правилом, а ко-артикуляционным процессом, с другой стороны, тот факт, что изменение это, вероятно, приводит к нейтрализации губных и зубных носовых по месту артикуляции<sup>2</sup>, позволяет отнести его к числу фонологических правил.

<sup>2</sup> Данное предположение нуждается, впрочем, в дополнительном экспериментальном исследовании.



**Рисунок 3.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах *Стэнфорд* и *Стэмфорд* (Гц). Усреднено по всем произнесениям в сильной и слабой фразовых позициях

Следует отметить, что формантные значения конца гласного, начала и конца носового в разных словах были различны: самые высокие показатели  $F_2$  у всех информантов зафиксированы в слове *тенгом*, далее в порядке убывания — *Стэнфорд*, *Стэмфорд*, *темпом*.

## ЭКСПЕРИМЕНТ 2

В ходе данного эксперимента тестировалась **гипотеза**, в соответствии с которой наличие / отсутствие просодической границы между фонетическими словами влияет на характер уподобления носовых согласных по месту артикуляции в русском языке.

**Материалом** исследования служили 8 предложений с тестовыми словосочетаниями (см. ниже) без физической паузы между ними. Внутри каждого словосочетания на границе двух фонетических слов были проанализированы следующие консонантные сочетания:  $V_1C_1$  (EM) +  $C_2$  (F) и  $V_1C_3$  (EN) +  $C_2$  (F), где  $V_1$  — гласный,  $C_2$  — лабиодентальный щелевой,  $C_1$  — губной носовой,  $C_3$  — переднеязычный-зубной носовой (*Тотем Фомы*, *Тотем формально* / *Шопен фонил*, *Шопен «Фантазия»*). Основанием для сравнения служили слова, в которых эти консонантные сочетания находились внутри слова (*Стэнфорд* / *Стэмфорд*) и словосочетания, в которых сочетания носового с гомоганным взрывным были разделены границей между двумя фонетическими словами (*Тотем по садовому кольцу*, *Шопен «Таинственный лес»* / *Тотем покинули*, *Шопен Тагира*).

Тестовые словосочетания удовлетворяли следующим условиям:

- идентичный гласный перед носовыми согласными (всегда [э]);
- идентичное место ударения: в первом фонетическом слове всегда перед [м] и [н], во втором фонетическом слове всегда на втором слоге;
- идентичное количество слогов в слове с носовым согласным;
- одинаковая степень освоенности слов.

Варьируемым параметром была просодическая организация словосочетания: позиция на границе двух синтагм или внутри синтагмы.

Предложения были помещены в связный текст (см. Приложение).

В эксперименте в качестве **информантов** приняли участие 19 носителей русского языка (18 женщин и 1 мужчина в возрасте от 18 до 25 лет). Дикторы представляли различные регионы России.

Предложения в составе связного текста были зачитаны дикторами вслух. Запись производилась с помощью программы Praat. Анализ влияния просодического оформления на реализацию носового осуществлялся путем сравнения качества носового в тестовых словосочетаниях и в словах *Стэмфорд / Стэнфорд*. Для этого было определено значение  $F_2$  в начале и в конце артикуляции согласного и в конце предшествующего гласного.

В ходе анализа результатов из рассмотрения были исключены те прочтения, в которых словосочетания были реализованы с физической паузой. Кроме того, были зафиксированы случаи, в которых значение  $F_2$  оказалось невозможным установить; они также не учитывались.

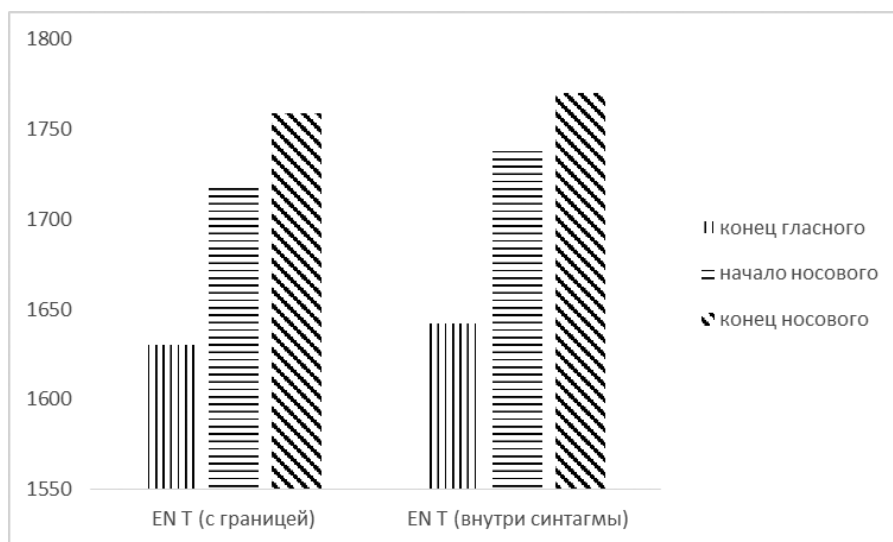
**Результаты исследования** приведены в Таблице 2 и в обобщенном виде на рисунках 4–7.

**Таблица 2.** Значения  $F_2$  в конце гласного (вверху), начале и конце носового (внизу) в тестовых словах при наличии и отсутствии просодического шва

	EM P граница	EM F граница	EM P – граница	EM F – граница	EN T граница	EN F граница	EN T – граница	EN F – граница
1	1447 1348–1236	1400 1243–1350	1500 1261–1222	1660 1557–1558	1807 1610–1785	1700 1814–1750	1703 1703–1763	1593 1508–1473
2	1408 1259–1209	1450 1020–1064	1283 1169–1100	1543 1393–1330	1600 1545–1601	1678 1633–1573	1633 1588–1678	1588 1543–1503
3	1370 1207–1257	1391 1326–1376	1282 1250–1186	1611 1530–1539	1992 1700–1811	2022 2009–1856	2056 1885–1950	
4	1291 1170–1260	1238 1066–1157	1304 1205–1042	1457 1325–1287	1629 1648–1629	1697 1720–1874	1678 1678–1705	1448 1501–1495
5	1284 1307–1338	1348 1305–1349	1393 1278–1085	1509 1430–1370	1747 1800–1847	2059 1724–1720	1974 1728–1720	1690 1662–1484
6	1268 1195–1216	1217 1197–1176	1238 1157–1103	1466 1286–1298	1756 1899–1702	1743 1702–1720	1829 1664–1601	1620 1530–1509
7	1320 1400–1407	1403 1268–1348	1339 1248–1411	1503 1348–1435	1774 1462–1590	1774 1666–1647	1938 1811–1810	1683 1676–1588
8	1357 1127–1167	1339 1176–1194	1295 1217–1195	1411 1286–1298	1841 1732–1872	1894 1916–1894	1890 1719–1938	1785 1670–1588



	EM P граница	EM F граница	EM P – граница	EM F – граница	EN T граница	EN F граница	EN T – граница	EN F – граница
9	1588 1556–1575	1457 1463–1476	1566 1489–1424	1632 1600–1610	1981 1883–1903	1950 1872–1850	1916 2069–1872	1850 1767–1650
10	1297	1413 1300–1304	1303 1272–1194	1413 1384–1336	1828 1807–1775	1938 1741–1714	1820 1850–1807	1719 1632–1566
11	1308 1060–1129	1320 1173–1042	1479 1248–1309	1300 1217–1308	1828 1723–1866	1872 1719–1828	1714 1700–1807	1670 1608–1513
12	1413 1217–1260	1588 1282–1326	1508 1260–1179	1544 1581–1547	2050 1916–1960	1916 1741–1800	1850 2100–2200	1772 1676–1807
13	1413 1129–1185	1284 1260–1212	1348 1261–1232	1506 1391–1390	1676 1600–1611	1758 1741–1763	1749 1686–1719	1544 1479–1457
14	1404 1260–1346	1370 1085–1286	1402 1282–1258	1544 1341–1348	1828	1818 1753–1763	1832 1772–1807	1700 1670–1482
15	1326 1306–1343	1282 1107–1151	1250 1260–1274	1348 1238–1326	1776 1766–1823	1840 1719–1646	1735 1789–1810	1600 1588–1502
16	1304 1000–1005	810–855	1304 1314–1109	1501 1435–1370	1903 1678–1856	1938 1741–1696	1937 1916–1960	1807 1566–1632
17	1519 1444–1313	914–974	1400 1271–1300	1697 1435–1348	1938 1807–1706	1850 1676–1700	1828 1807–1894	1648 1603–1693
18	1435 1338–1287	1544 1354–1403	1457 1305–1340	1697 1544–1690	1828 1785–1820	1785 1828–1785	1785 1789–1820	1700 1566–1523
19	1340 1139–1199	1289 1184–1154	1304 1154–1139	1428 1200–1259	1618 1643–1513	1633	1658 1673–1559	1580 1543–1513



**Рисунок 4.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словосочетании «зубной носовой + гоморганный взрывной» внутри и на стыке синтагмы. Усреднено по всем произнесениям

Приведенные в Таблице 2 данные позволяют утверждать, что носовой согласный в позиции перед гоморганным взрывным в тестовых словах внутри синтагмы и на границе двух синтагм реализуется одинаково:

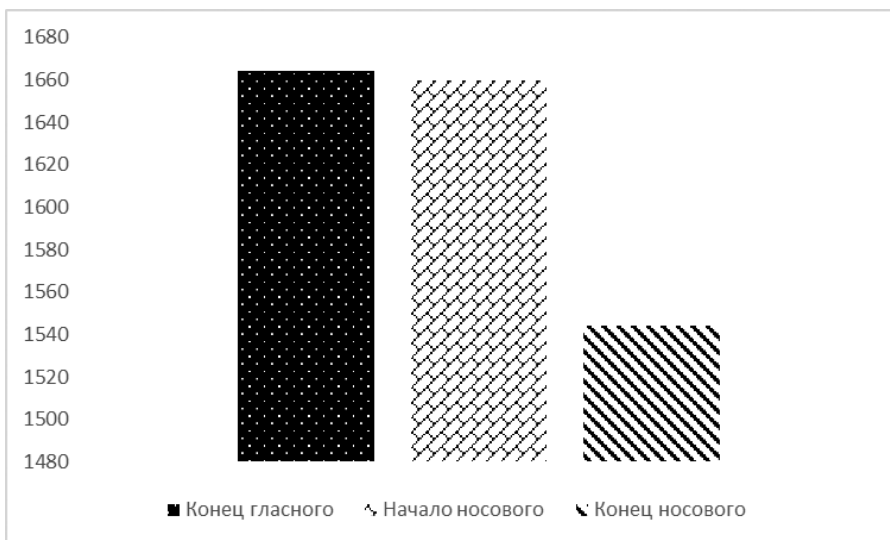
- перед губно-губным [м] значение  $F_2$  гласного резко понижается и становится равным 1350 Гц, перед [н] составляет 1800 Гц;
- $F_2$  носового либо незначительно повышается по направлению к последующему согласному, либо остается неизменной на протяжении всей артикуляции (значение  $F_2$  в начале и конце [м]: 1260–1280 Гц, [н]: 1735–1790 Гц (см. рис. 4).

Как видно из рисунка 5, реализация переднеязычного носового перед губно-зубным щелевым в позиции внутри синтагмы характеризуется тем, что:

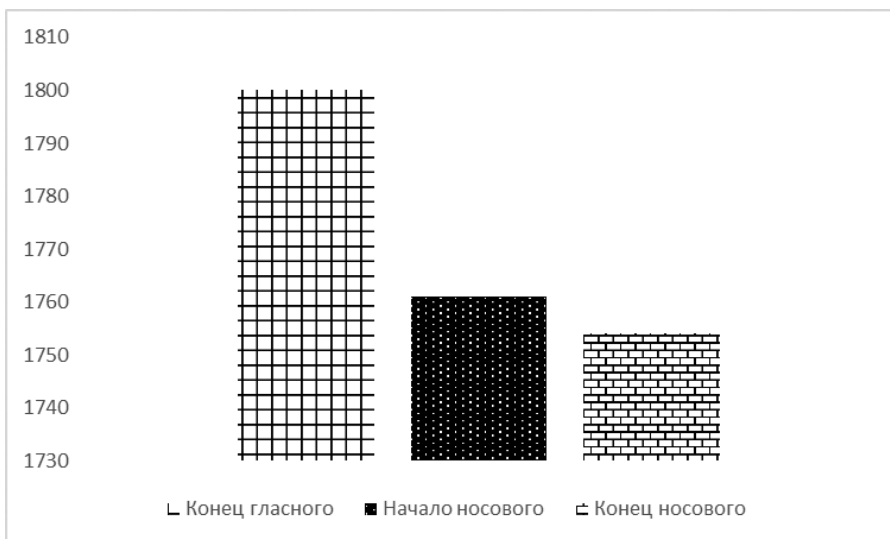
- значение  $F_2$  носового от начала к концу его артикуляции уменьшается (в то время как в позиции перед гоморганным взрывным оно увеличивается или остается неизменным, см. также рис. 4). Значение  $F_2$  в начале и конце носового в данной позиции меньше, чем в позиции перед гоморганным взрывным (разница составляет приблизительно 100 Гц в начале и 200 Гц в конце артикуляции носового);
- $F_2$  предшествующего гласного изменяется по отношению к альвеолярному носовому, но в меньшей степени, чем в положении перед гоморганным взрывным (разница составляет приблизительно 170 Гц).

На Рис. 6 можно видеть, что реализация переднеязычного носового в позиции перед губно-зубным щелевым согласным на границе двух синтагм характеризуется тем, что:

- значение  $F_2$  носового на протяжении всей его артикуляции не меняется или меняется незначительно и колеблется в районе 1750 Гц. Данное значение больше (на 100 Гц в начале артикуляции и на 200 Гц в конце артикуляции), чем значение второй форманты носового в том же консонантном сочетании внутри синтагмы. Те же значения фиксируются и в сочетании «альвеолярный носовой + гоморганный шумный», что означает, что в данном случае носовой реализуется без коартикуляции;
- предшествующий гласный не коартикулирует по отношению к носовому: значение его  $F_2$  от середины к концу артикуляции не изменяется и составляет приблизительно 1800 Гц.

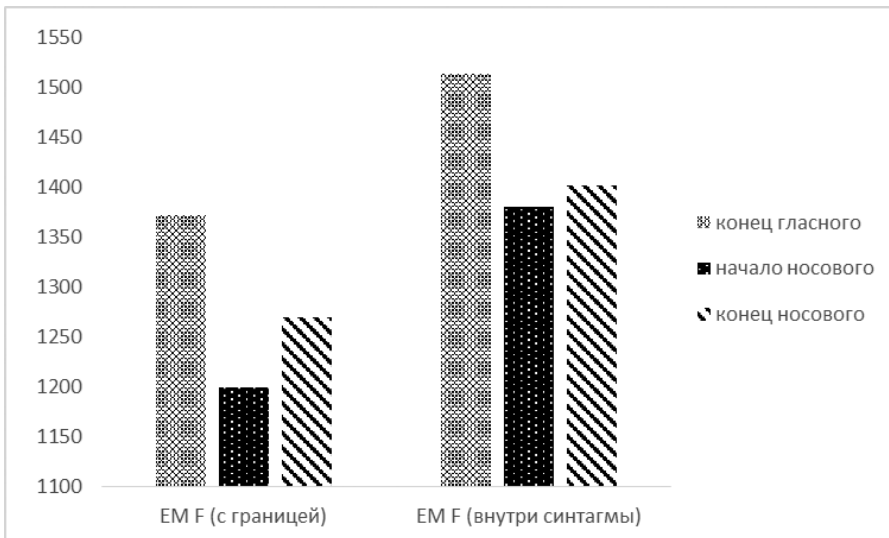


**Рисунок 5.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словосочетании *Шопен фонил* (Гц). Усреднено по всем произнесениям



**Рисунок 6.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словосочетании *Шопен «Фантазия»* на границе двух синтагм (Гц). Усреднено по всем произнесениям

Сходная картина наблюдается и в случае с билабиальным носовым (см. Рис. 7):



**Рисунок 7.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словосочетании «билабиальный носовой + лабиодентальный спирант» внутри и на стыке синтагмы. Усреднено по всем произнесениям

На основании полученных в ходе эксперимента данных можно сформулировать **вывод** о том, что уподобление носового перед лабиодентальным щелевым по месту образования фиксируется в случае отсутствия между этими согласными просодической границы, в случае же ее наличия коартикуляция отсутствует.

### ЭКСПЕРИМЕНТ 3

**Целью** эксперимента 3 было сопоставление данных о коартикуляционных изменениях губных и зубных носовых согласных по месту образования в позиции перед твердым губно-зубным спирантом с характером их реализации в положении перед мягким губно-зубным. Исходная **гипотеза** заключалась в том, что характер уподобления твердых губных и переднеязычных носовых по месту образования в позиции перед губно-зубным щелевым не зависит от твердости/мягкости губно-зубного согласного.

**Материалом** для исследования служили 8 тестовых слов с билабиальным и переднеязычным носовым согласным в позиции перед лабиодентальным мягким щелевым, то есть с сочетанием  $V_1 C_2 C_4$  (VMF') и  $V_1 C_3 C_4$  (VNF'), где  $V_1$  — гласный,  $C_2$  — носовой губно-губной,  $C_3$  — носовой переднеязычный — зубной,  $C_4$  — губно-зубной мягкий фрикативный согласный. Слова попарно

распределялись по группам в соответствии с сегментными позициями внутри фонетического слова, которые обуславливались следующими факторами:

- идентичная позиция относительно ударного гласного;
- в каждой паре идентичные/сходные гласные перед носовыми согласными;
- идентичное количество слогов в слове;
- идентичная фразовая позиция.

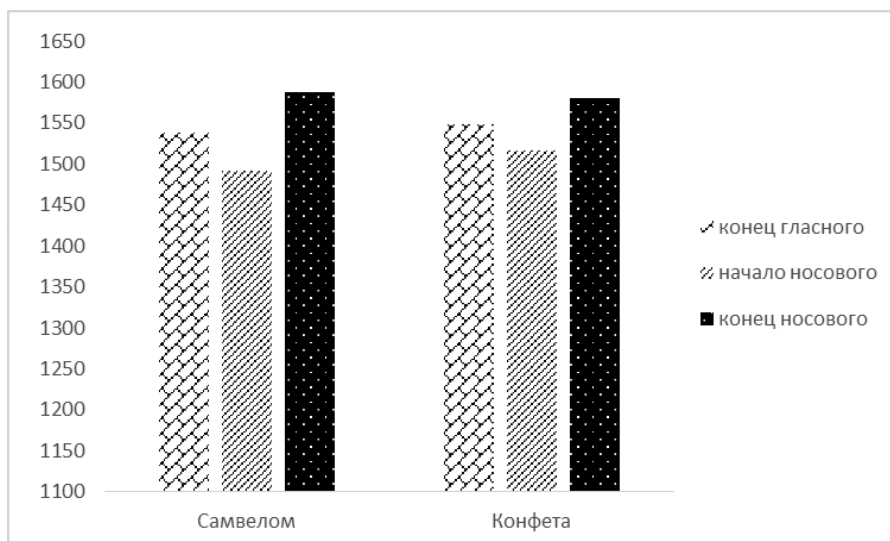
Список тестовых слов: *Самвелом / конфета; эмфизема / анфилада; Бенфика / Земфира; нимфетка / инфекция*. Тестовые слова были включены во фразы, а фразы в связный текст (см. Приложение).

В эксперименте в качестве **информантов** приняли участие 15 носителей русского языка в возрасте от 17 до 20 лет, представляющие различные регионы России.

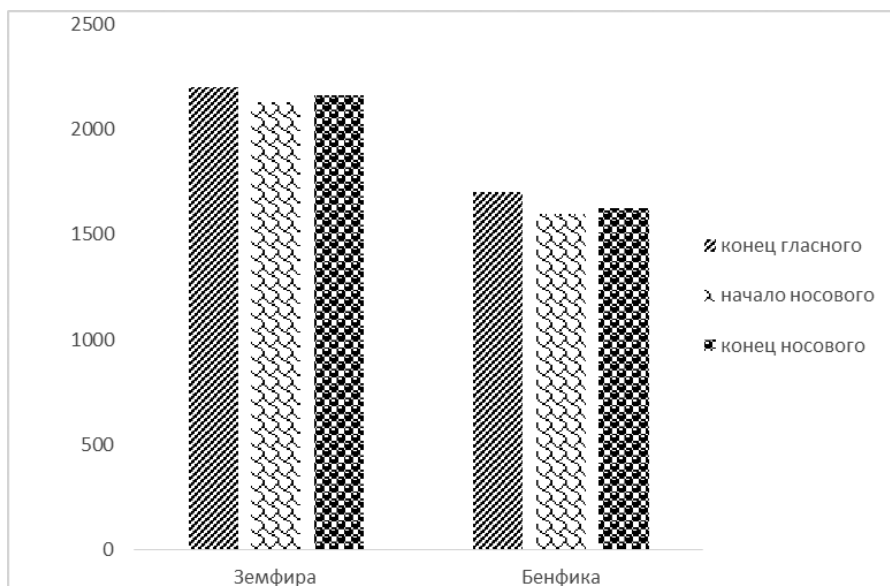
**Результаты исследования** приведены в Таблице 3 и в обобщенном виде на рисунках 8–13.

**Таблица 3.** Значения  $F_2$  в конце гласного (вверху), начале и конце носового (внизу) в тестовых словах в сильной фразовой позиции

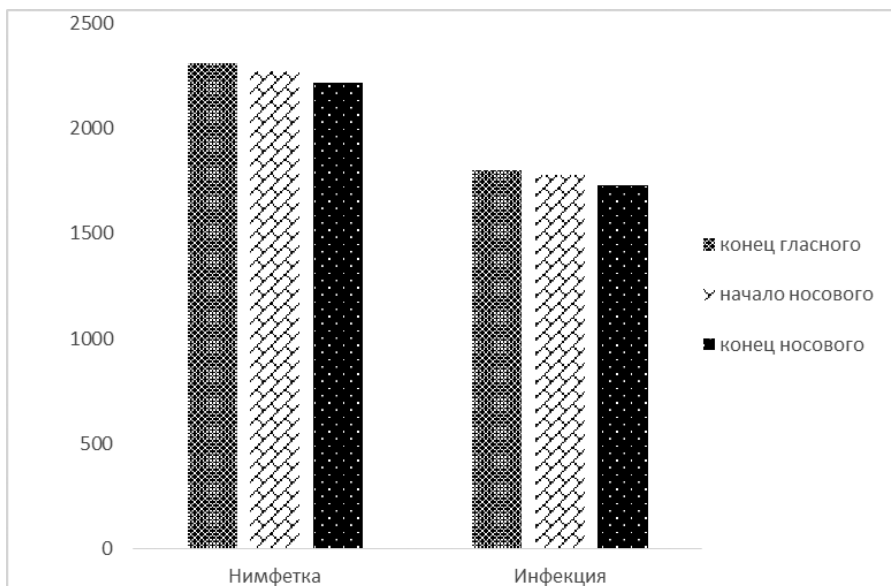
	VM <sup>(o)</sup> V' <i>Самвелом</i>	VN <sup>(o)</sup> F' <i>конфетой</i>	VM <sup>(o)</sup> F' <i>Земфира</i>	VN <sup>(o)</sup> F' <i>Бенфика</i>	VM <sup>(o)</sup> F' <i>нимфетка</i>	VN <sup>(o)</sup> F' <i>инфекция</i>	VM <sup>(o)</sup> F' <i>эмфизема</i>	VN <sup>(o)</sup> F' <i>анфилада</i>
1	1613 1523–1544	1566 1566–1588	2251 2220–2170	1700 1610–1654	2505 2374–2314	1413 1610–1566	Прочитано с твердым	1350 1523–1623
2	1623 1523–1606	1570 1480–1700	2100 2178–2222	2040 2083–1900	2380 2294–2011	1878 1601–1434	2079 2017–2035	1566 1523–1654
3	1891 1925–1936	1619 1593–1732	2087 2059–2071	1654 1606–1615	2283 2107–1846	2006 1952–1872	1763 1719–1807	1654 1563–1562
4	1494 1528–1829	1568 1545–1505	2287 2199–2185	? 1600–1501	2322 2305–2321	1790 1865–1791	2246 2329–2368	1676 1404–1402
5	1460 1450–1631	1318 1343–1268	1748 1614–1698	1539 1520–1646	2134 2171–2162	1504 1512–1670	1800 1512–1638	1400 1313–1300
6	1329 1353–1494	1538 1554–1615	2178 1905–1874	1482 1416–1464	1922 2085–2056	1626 1619–1759	2166 2228–2272	1508 1507–1662
7	1615 1513–1451	1461 1440–1520	2273 2192–2106	1707 1683–1618	2012 2001–2265	1807 1892–1816	2201 2243–2183	1415 1528–1676
8	1477 1330–1542	1339 1314–1380	2001 2142–2178	1667 1640–1612	2338 2486–2200	2022 1983–1978	2367 2357–2317	1424 1430–1536
9	1735 1647–1624	1652 ?	2646 2519–2287	1754 1716–1681	2003 1836–1986	1934 1880–1847	2068 2031–1985	1790 1790–1798
10	1659 1572–1722	1862 1890–2131	2719 2743–2712	2008 1941–2041	2700 2460–2444	2033 ?	2251 2149–2432	1781 1741–1680
11	1500 1431–1501	1346 1358–1455	2015 2226–2169	1701 1529–1610	2528 2597–2440	1676 1526–1231	1877 1817–2016	1544 1513–1468
12	1588 1686–1775	1660 1619–1641	2385 2189–2473	1918 1894–2042	2580 2259–2060	2005 1950–1917	1738 1629–1621	1710 1688–1640
13	1565 1471–1370	1713 1639–1618	2467 2680–2625	1975 1876–1888	2455 2110–2079	1900 1938–2000	2035 1627–1546	1730 1680–1719
14	1116 1235–1370	1410 1210–1400	1771 1738–1700	1491 1566–1505	2489 2045–2117	1478 1451–1395	1545 1631–1740	1345 1422–1376
15	1413 1363–1318	1607 1691–1593	1938 2219–2011	2008 2040–2091	2025 2031–2121	2091 2064–2013	1956 2156–2174	1840 1741–1763



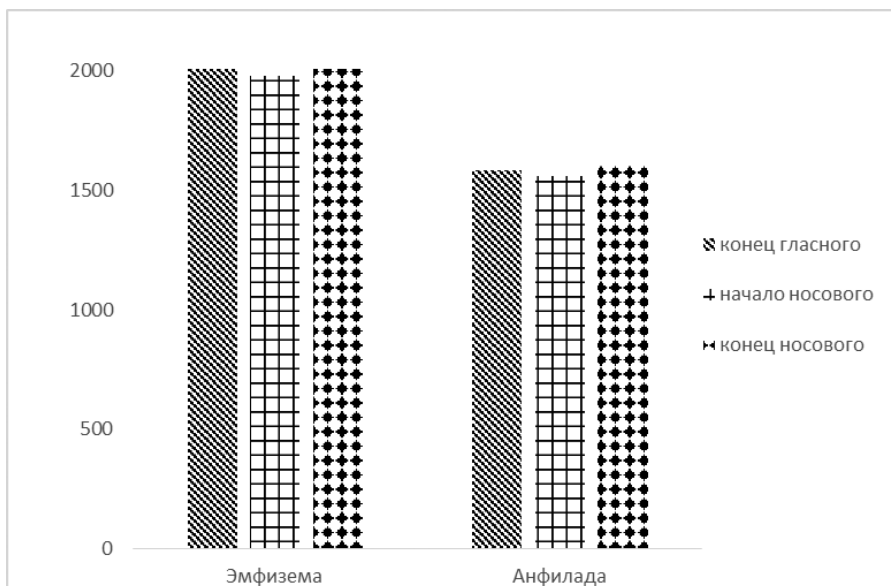
**Рисунок 8.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах *Самвелом* и *конфета* в сильной фразовой позиции (Гц). Усреднено по всем произнесениям



**Рисунок 9.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах *Земфира* и *Бенфика* в сильной фразовой позиции (Гц). Усреднено по всем произнесениям



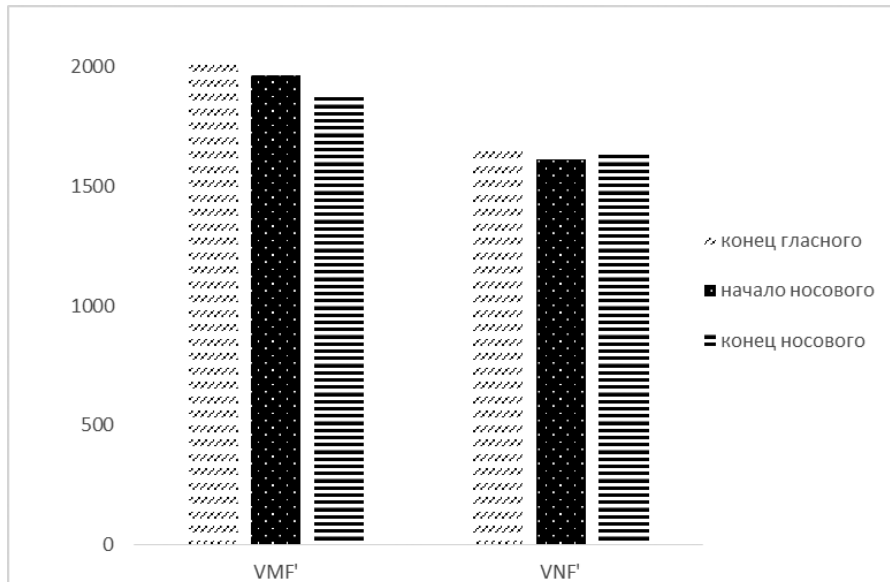
**Рисунок 10.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах *нимфетка* и *инфекция* в сильной фразовой позиции (Гц). Усреднено по всем произнесениям



**Рисунок 11.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах *эмфизема* и *анфилада* в сильной фразовой позиции (Гц). Усреднено по всем произнесениям

Попарное сравнение реализации губного и переднеязычного носового в положении перед мягким губно-зубным спيرانтом (Рис. 8–11) позволяет сделать вывод о кардинальном отличии этой позиции от позиции перед твердым губно-зубным, исследованной в эксперименте 1. Если в сочетании с твердым лабиодентальным значения  $F_2$  билабиального и коронального согласного изменялись по направлению к некоторой промежуточной точке (почти совпадая в области около 1400–1500 Гц), между их значениями в позиции перед гоморганным взрывным (около 1250 для губных и около 1750 для зубных), то кластеры с конечным мягким устроены принципиально иначе. В этом последнем случае значение  $F_2$  переднеязычного согласного фиксируется почти на том же уровне, что и в позиции перед соответствующим твердым согласным (около 1600 Гц), в то время как для губного носового отмечается резкое увеличение  $F_2$  до значений, превышающих значения  $F_2$  зубного носового перед гоморганным — до 1900 Гц в случае усреднения по всем тестовым словам и даже до 2200 в случае усреднения без учета наименее освоенных слов (*Самвел, эмфизема*) — см. Рис. 12, 13.

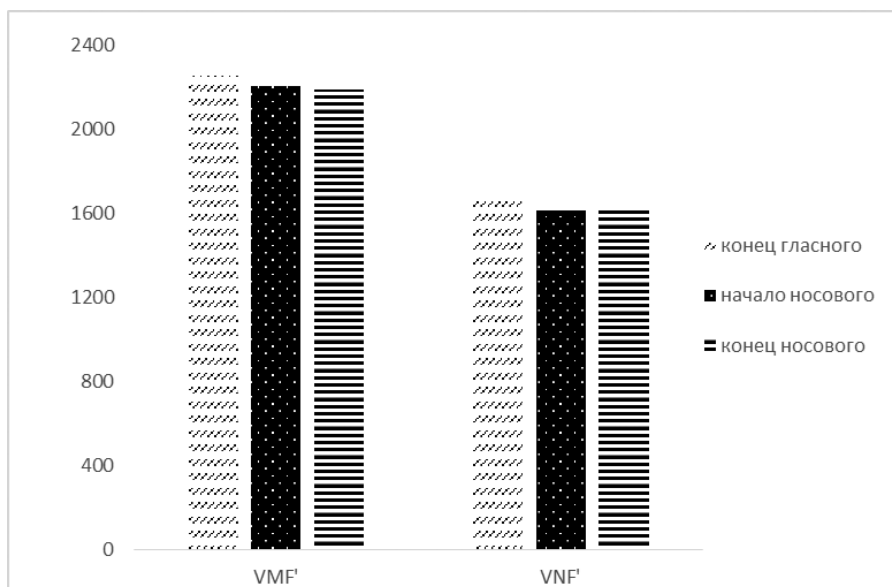
Описанные факты, по-видимому, могут быть объяснены тем, что в позиции перед мягким губно-зубным спيرانтом те губно-зубные носовые, которые восходят к билабиальным, подвергаются ассимилятивному смягчению, а те, которые восходят к зубным, — нет, что позволяет поставить вопрос о порядке применения фонетических правил в процессе построения звуковой оболочки высказывания в русском языке.



**Рисунок 12.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах с сочетаниями VMF' и VNF' (Гц).

Усреднено по всем словам и всем произнесениям





**Рисунок 13.** Значение  $F_2$  в конце гласного, начале и конце носового согласного в словах с сочетаниями VMF', VNF' без слов *Самвелом* и *эмфизема* (Гц). Усреднено по всем произнесениям

Современные представления о модели звукового поведения человека подробно описаны С. В. Кодзасовым и О. Ф. Кривновой [Kodzasov, Krivnova 2001: 505–509]. В этой модели фонологические (лингвистические) правила действуют на этапе контекстных преобразований глубинных символических репрезентаций и до действия моторной программы, порождающей явления коартикуляции (совмещения во времени неантогонистических элементов соседних артикуляционных жестов), которые не относятся к лингвистическим правилам. Этих правил в любом языке довольно значительное количество, поэтому одним из наиболее существенных вопросов в фонологии является вопрос о порядке применения этих правил. В классическом варианте порождающей фонологии [Chomsky, Halle 1968] было предложено считать, что этот порядок является фиксированным линейным, то есть правила действуют одно за одним в строго определенной последовательности. В дальнейшем — в развитие идей Н. Хомского и М. Халле — были предложены и другие решения относительно порядка применения правил:

- применение правил без заранее фиксированного порядка [Koutsoudas et al 1974]: правила сосуществуют, не будучи упорядоченными, при этом применяется то правило, для которого имеются условия применения, причем одно и то же правило может применяться неоднократно;
- немаркированный порядок [Kiparsky 1968, 1971]: правила стремятся расположиться в таком порядке, который дает наиболее полное их исполь-

зование, или в таком порядке, при котором достигается их наибольшая прозрачность<sup>3</sup>;

- локальный порядок [Anderson 1974]: иерархия правил определяется внутри каждой конкретной пары правил, а сами эти пары друг относительно друга не упорядочиваются;
- одновременное применение всех правил.

Возможно, однако, что в разных языках порядок применения правил может быть различным. Вопрос же о порядке осуществления коартикуляционных изменений никогда в фонетике не ставился.

Существует точка зрения, согласно которой в русском языке не имеется никаких конкретных данных для того, чтобы судить о порядке применения фонологических правил при построении высказывания на этом языке: «Вопрос о порядке правил имеет смысл лишь в случае их взаимодействия — когда в результате действия одного правила возникают или устраняются условия применения другого. Если правила не взаимодействуют, то их порядок безразличен, и они могут быть применены одновременно. Таковы, в частности, русские правила редукции гласных, ассимиляции согласных по глухости-звонкости, оглушения в конце слова, ассимиляции согласных по мягкости, выпадения согласных в консонантных сочетаниях» [Kodzasov, Krivnova 1981: 49–50]. Тем не менее, анализ фактов русского языка позволил сформулировать вывод о том, что по крайней мере для некоторых звуковых чередований последовательность применения правил такова: 1) внутрисловная ассимиляция по голосу, мягкости и месту образования, 2) конечное оглушение, 3) стирание межсловных границ, 4) ассимиляция по голосу и месту образования <на стыках слов>, 5) звуковой эллипсис [Кныазев 2004].

Полученные в ходе настоящего исследования данные позволяют уточнить положение о внутрисловных чередованиях: очевидным образом, правило изменения носовых по месту артикуляции действует в современном русском языке после правила уподобления по твердости/мягкости, так как один и тот же по месту образования (губно-зубной) носовой сонант в позиции перед мягким губно-зубным спирантом является палатализованным в том случае, если на глубинном фонологическом уровне он охарактеризован как билабиальный, и непалатализованным, если фонологически является переднеязычным. При этом представляется, что для принятия окончательного решения о статусе изменения носовых по месту образования (фонологическое правило или коартикуляция) данных пока ещё недостаточно: с одной стороны, это изменение может приводить к нейтрализации губных и зубных носовых по месту артикуляции (что характерно, в первую очередь, для фонологических правил), с другой стороны, это изменение осуществляется после действия фонологического правила ассимиляции по твердости/мягкости и может проявляться в разной степени, что характерно для фонетических коартикуляционных процессов.

---

<sup>3</sup> Прозрачными являются правила, условия применения и характер действия которых ясно видны из поверхностных репрезентаций.

## Выводы

На основании полученных в ходе настоящего исследования результатов могут быть сформулированы следующие выводы:

- в русском языке наблюдается уподобление твердых губных и переднеязычных носовых согласных по месту образования в позиции перед твердым губно-зубным щелевым, которое может приводить к их нейтрализации по признаку места артикуляции;
- это изменение осуществляется после действия правила ассимиляции по твердости/мягкости и проявляется в речи в разной степени;
- данное явление отсутствует на стыках фонетических слов, относящихся к разным синтагмам и является надежным показателем наличия просодического шва между словами.

## Приложение

### Эксперимент 1

Список фраз 1:

1. Джон из Стэнфордского университета;
2. Гордится, что живет в Стэнфорд-Холле;
3. Матч должен был состояться на домашней арене Челси — Стэмфорд Бридж;
4. Остановились они в маленьком городке Стэмфорд на юге Великобритании;
5. Когда тент открыли, стало все-таки посуше;
6. Надо было обзавестись тентом, на случай непогоды во время матча;
7. Во втором тайме все были поражены быстрым темпом и мощностью атак;
8. Команда проиграла достойному сопернику, поскольку темп игры был недостаточно высок.

Текст 1:

Джон из Стэнфордского Университета, который гордился тем, что живет в Стэнфорд-Холле, а не общежитии, решил отправиться с друзьями в Англию, чтобы побывать на одном из самых захватывающих матчей сезона Челси — Реал. Сам матч должен был состояться на домашней арене Челси — Стэмфорд Бридж, на стадионе, который впечатляет своим величием. Подготовка к поездке в страну туманного Альбиона была утомительна: надо было собрать вещи, обзавестись тентом, на случай непогоды во время матча, купить билеты и забронировать отель. Джон и его друзья отправились в путь за три дня до игры и остановились они в маленьком городке Стэмфорд на юге Великобритании. В день игры пошел сильный дождь, но когда тент открыли, стало все-таки посуше. Матч был поистине захватывающим и зрелищным. Во втором тайме все были поражены быстрым темпом и мощностью атак Реала, которому все-таки удалось забить гол на последней минуте. После матча тренер Челси отметил, что его команда проиграла достойному сопернику, поскольку темп игры был недостаточно высок.

## Эксперимент 2

### Список фраз 2:

1. Шопен «Таинственный лес»
2. Признаться честно, Шопен Тагира нам понравился больше
3. Распорядитель объявил: Шопен «Фантазия»
4. Когда Шопен фонил, впечатление сразу портилось
5. Отправились с друзьями в ресторан Тотем по садовому кольцу
6. Когда Тотем покинули, стало намного веселее
7. Тотем формально специализируется скорее на напитках, чем на еде
8. Если хотите хорошо провести время, то Тотем Фомы — идеальное место

### Текст 2:

Сегодня мы с друзьями отправились в ресторан «Тотем» по садовому кольцу праздновать день рождения нашего друга. Это уютное местечко принадлежит широко известному кинокритику Александру Фомину. Нас было трое: я, моя давняя знакомая Света и именинник Тагир. Место мы выбрали такое, где каждое воскресенье вживую играет Шопен, «Таинственный лес», «Болеро». Надо сказать, что «Тотем» формально специализируется скорее на напитках, чем на еде, потому что здесь огромная винная карта и скудное меню. В самом ресторане нам понравилось абсолютно все: приятная атмосфера, быстрое обслуживание. Если Вы хотите вкусно поесть и хорошо провести время, то Тотем Фомы — идеальное место. Больше всего мы, конечно, ждали, когда же зазвучат произведения классика. Примерно в 7 вечера распорядитель объявил Шопен, «Фантазия». За роялем сидел молодой исполнитель, явно робеющий перед публикой. Вроде бы его исполнение было вполне достойным, но когда Шопен фонил, впечатление сразу портилось. Так как Тагир был сегодня именинник, ему тоже разрешили сыграть небольшой кусочек из произведения его любимого композитора. Признаться честно, Шопен Тагира нам понравился больше. В общем, уже под конец вечера мы начали скучать и часов в 12 ушли от туда. После того, как Тотем покинули, стало намного веселее: сначала мы катались на катере, потом гуляли всю ночь и только утром вернулись домой.

## Эксперимент 3

### Список фраз 3:

1. Я делю комнату в общежитии с Самвелом
2. Матч Реал — Бенфика
3. Анфилада учебных комнат
4. Чай с конфетой
5. Мы были на концерте Земфиры
6. Подхватили какую-то инфекцию
7. Рассказывал о характере маленькой нимфетки
8. Что такое эмфизема легких

### Текст 3

Я делю комнату в общежитии с Самвелом уже года. Живем мы очень дружно и никогда не мешаем друг другу: когда один хочет посмотреть матч

*Реал — Бенфика*, другой идет делать уроки вниз, где располагается анфилада учебных комнат. На четвертом курсе учеба занимает много времени: с утра обычно мы не успеваем даже позавтракать, только если чай с конфетой. Иногда, конечно, нам удастся развлечься, но редко. Например, в субботу мы были на концерте Земфиры — хоть какое-то развлечение. Последнюю неделю, правда, мы на занятия не ходим: подхватили какую-то инфекцию и не можем никак вылечиться. Целый день лежим и читаем: я Набокова, «Лолиту», а Самвел Аронович свои книжки по биологии, потом делимся впечатлениями. Сегодня я ему рассказывал о характере маленькой нимфетки, а он мне о том, что такое эмфизема легких. Так и живем.

## Литература

1. *Anderson, S. R.* (1974) *The organization of phonology*. N.Y., S.-Fr., L. Academic Press.
2. *Barkhudarova E. L.* (2011) Paradigmatics and syntagmatics of sound units in the context of teaching Russian pronunciation [Paradigmatika i sintagmatika zvukovykh yedenits v kontekste obucheniya russkomu proiznosheniyu] // *Vestnik MGU, ser. 9, Philology*. № 4.
3. *Cho, Young-mee Yu* (1990). *Parameters of Consonantal Assimilation*. Doctoral dissertation, Stanford University.
4. *Chomsky, Noam & Morris Halle* (1968). *The Sound Pattern of English*. Harper & Row, Publishers. New York, Evanston, and London.
5. *Huang, Tsan* (2001). The interplay of perception and phonology in tone 3 sandhi in Chinese Putonghua. In *Ohio State University Working Papers in Linguistics* 55: 23–42.
6. *Kiparsky, P.* (1968) Linguistic universals and linguistic change // *Universals in linguistic theory*. Ed. by E. Bach and R. T. Harms. N.Y.
7. *Kiparsky, P.* (1971) *Historical linguistics // A survey of linguistic science*. Ed. by W. O. Dingwall. University of Maryland Press.
8. *Knyazev S. V.* (2004) Towards the phonological rules' hierarchy in Russian [Ob yerarkhii fonologicheskikh pravil v russkom yazyke] // *Semiotics, Linguistics, Poetics [Semiotika, lingvistika, poetika]*. Moscow.
9. *Koutsoudas A., Sanders G. and C. Noll* (1974) On the application of phonological rules // *Language*, 50.
10. *Kodzasov S. V., Krivnova O. F.* (1981) *Contemporary American phonology [Sovremennaya amerikanskaya fonologiya]*. Moscow.
11. *Kodzasov S. V., Krivnova O. F.* (2001) *General phonetics [Obshchaya fonetika]*. Moscow.
12. *Margaret E. L. Renwick, Ladan Baghai-Ravary, Rosalind Temple, John S. Coleman* (2013) Assimilation of word-final nasals to following word-initial place of articulation in UK English. *Phonetics Laboratory, University of Oxford*.
13. *Mohanan, K. P.* (1993) *Fields of attraction in phonology*. Chicago: University of Chicago Press, 61–116.
14. *Trask, R. L.* (1996) *A Dictionary of Phonetics and Phonology*. Routledge. London and New York.

# ТИПОЛОГИЯ РУССКИХ ПОХОДОК

**Крейдли́н Г. Е.** (gekr@iitp.ru)

Российский государственный гуманитарный университет, Москва, Россия

## TYPOLOGY OF RUSSIAN GAITS

**Kreydlin G. E.** (gekr@iitp.ru)

Russian State University for the Humanities, Moscow, Russia

The paper continues a series of research on gaits as a particular type of nonverbal complex sign units. It presents the detailed description of various types of gaits peculiar to Russian people. The research is conducted in the framework of the feature-based approach, which is a main tool for constructing a semiotic conceptualization of the somatic objects or human body in general. Using this approach we analyze physical features of gaits, such as width of pace, speed of walking, type of contact with the surface, noise, which is produced while walking, shape of the body or particular somatic objects (i.e. legs, feet), etc. We describe how the particular manner of walking reflects various characteristics of the human being. These are social characteristics, such as gender and age; traits of character and psychological features, such as being nervous, depressed or happy; various types of dysfunctions, such as being halting. Such parallels between the specific features of the gait and elements of the emotional and psychological states of the person highlight the connections between various systems of the human being.

**Key words:** body, gait, gesture, Russian, semiotic conceptualization, walking

### Введение. Постановка задачи

Темой моей предыдущей статьи о походках [4: 127–142] были концепт «походка» и значение и употребление слова *походка* применительно к человеку. Было показано, что слово *походка* имеет два значения — вопреки тому, что о нём говорится в объяснительных и иных словарях — и два типа употребления.

В одном из них *походка* обозначает свойство человека, а именно ‘то, как человек обычно ходит’, а в другом — его актуальное состояние, то есть ‘то, как человек ходит сейчас или идёт сейчас’. Походка, понимаемая как свойство, постоянна, это — характеристика антропоморфного положения тела, которое в одних чертах наследуется с рождения, а в других приобретается человеком на протяжении его жизни. Походка-состояние переменна и зависит от многих внешних и внутренних факторов. В этом значении слово *походка* обозначает

способ ходьбы человека, в котором отражается его актуальное физическое и психологическое состояние.

В той же статье были представлены отличительные свойства походок как элементов класса знаковых телодвижений и как форм поведения, указаны соматические объекты, релевантные для реализации хотя бы одной походки. Например, при *размашистой походке* человек ходит более широким шагом, чем обычно, энергично, с силой размахивает руками, а *угрожающая* или *решительная походки* предполагают среди прочего определённую мимику — чуть выдвинутое вперёд лицо и подбородок, прямой строгий взгляд, напряжённость рта, отсутствие улыбки. Были отмечены некоторые признаки, отличающие одни походки от других, и продемонстрирована зависимость походок от одежды и обуви их исполнителей. Наконец, в упомянутой статье были приведены примеры признаков, по которым можно и уместно квалифицировать и оценивать походки.

\*\*\*

В настоящей работе строится типология русских походок<sup>1</sup>. Я не утверждаю, однако, что на основе выделенных характеристиках походок можно будет в результате построить «настоящую» внутриязыковую типологию походок с пересекающимися классами, но я полагаю, что даже простая каталогизация множества походок, проведённая с использованием ряда содержательных признаков, способствует системной организации и семиотической концептуализации класса походок.

В русском языке, если судить по данным НКРЯ [7], походки часто описываются с помощью атрибутивных сочетаний со словом *походка* в синтаксической роли вершинного элемента, хотя существует также много других разнообразных, языковых способов репрезентации походок. Например, они могут передаваться в тексте глаголами, семантическая структура которых содержит пропозицию, характеризующую способ движения (я имею в виду такие глаголы, как *ползёт*, *семенит*, *переваливается*, *плывёт*, *подплывает*). Походки могут также описываться синтаксическими группами, состоящими из глаголов движения и наречия, а также с помощью наречного, деепричастного или иного оборота, ср. *ходит важно*; *идёт вразвалку*; *еле ступая*, *с низко опущенной головой*; *идёт, виляя бёдрами*; *еле движется*, *озираясь по сторонам*), либо сочетаниями со словами *шаг* (*шаги*), *шагать*, *ступать* и характеристиками шагов или способа передвижения (*идёт быстрым шагом*, *передвигается быстрыми/мелкими шагами*, *шагает широко*, см.:

- (1) *Мужичонок бледен <...> Он широко шагает, нагибает туловище вперёд и, видимо, спешит* (А. П. Чехов. Он понял).

Примеры употребления атрибутивных сочетаний со словом *походка* позволяют выделить те типы походок, в которых отражены определённые признаки

<sup>1</sup> Настоящая статья, как и предыдущая, посвящённые походкам, обе написаны при поддержке РФНФ (грант 16-34-00023).

походок. Приводимые группы признаков сопровождаются примерами из НКРЯ и из текстов художественной литературы XIX–XXI веков:

## 1. Физические признаки походок

Из физических признаков походок выделяю:

(а) характеристику шага (его размер, форму, амплитуду). **Примеры:** *широкая, крупная, мелкая, частая, размеренная, семенящая*, ср.:

(2) <...> *Маша худа и тонка, стянута в корсет, и походка у неё мелкая* <...> (А. П. Чехов. *Бабье царство*),

(б) скорость и ритм передвижения, амплитуда переноса тела при движении  
**Примеры:** *быстрая, медленная, медлительная, плавная, раскачивающаяся, стремительная, тихая, (не)торопливая, спорая*.

(в) тип контакта ног с поверхностью передвижения. **Примеры:** *летающая, летучая, парящая, скользящая, пружинистая, мягкая, лёгкая, тяжёлая*, ср.:

(3) <...> *Каждый вечер тяжёлой походкой приходила дочь, её округлый лобик был собран страдальческими бровями под самые волосы* (О. Славникова. *Стрекоза, увеличенная до размеров собаки*).

(г) наличие / отсутствие звука от контакта ног с поверхностью передвижения. **Примеры:** *бесшумная, шумная; шаркающая, тихая, беззвучная*

(д) форму тела или форму отдельных телесных объектов, принимаемых при походке (часто вместе с оценкой). Синтаксическая и семантическая сферы действия прилагательного здесь обычно расходятся, то есть синтаксически оно присоединяется к слову *походка*, а семантически относится к форме тела или каких-то его частей при движении. **Примеры:** *подпрыгивающая, вальяжная, прямая, косая, расхлябанная, (не)ровная*, ср.:

(4) Он <...> *ровной походкой направился в свою каюту*  
(А. С. Новиков-Прибой. *Цусима*),

*танцующая, прихрамывающая, косолапая, суетливая,*

(5) <...> *Походка у неё была, как у богомолки, быстрая и суетливая*  
(А. П. Чехов. *Мужики*),

(е) сравнение походки человека с некоторыми чертами походки животного. **Примеры:** *медвежья, лисья, кошачья, павлинья, утиная, утичья*, ср.:

(6) *Походка утичья, с перевальцем и виляньем турнюра, есть признак добродушия, податливости и иногда тупости*  
(А. П. Чехов. *Художество*),

*журавлиная, лебединая* и др.



## 2. Социальные признаки исполнителя походки

### 2.1. Пол/гендер

Выдающийся французский социолог и этнограф Марсель Мосс как-то заметил, что все техники тела разделены между полами [5]. Не удивительно, что и такая важная «техника тела», как походка, тоже бывает мужской и женской.

Большее место в жизни общества занимают, конечно, *женские походки*, поскольку с ними связаны суждения о женском изяществе и женской красоте<sup>2</sup>. Женщин походке даже специально обучают — и до школы, и в школе, и далее. Особенно это касается женщин, которым овладение походкой важно для реализации в профессии (манекенщица, балерина, актриса, стилист и др.). Ср., например, предложение из одного телефильма:

(7) *На первом занятии манекенщицы отрабатывали походку.*

Рисунок движений ног и тела может быть своеобразным и сугубо индивидуальным, присущим только одному человеку, но принято считать, что *красивая грациозная женская походка* создаётся неподвижными бёдрами и плечами, плечи при этом должны быть приподняты, а голова не должна качаться из стороны в сторону. *Красивая походка* — лёгкая, невесомая, летучая, то есть ступни ноги должны лишь едва касаться поверхностей земли или пола, однако эта походка не *быстрая*. Ср.:

(8) *Женская походка, когда женщина идет быстро, крайне редко бывает красивой* (Е. Белкина. От любви до ненависти).

М. Мосс обращает внимание на то, что по положению рук при ходьбе можно понять кое-что о воспитании девушки [5]. Он пишет с изрядной долей иронии, что девушка, которая ходит, сжав кулаки, скорее всего, воспитывалась в монастыре.

От женской походки мужчины ожидают обольстительности и сексуальности, ср.:

(9) *У Лёльки слегка сильный голос, воспаленные, чуть навывкате козы глаза и обольстительная развинченая походка* (О. В. Волков. Из воспоминаний старого тенишевца).

В книге Элсона Беста «The Maori» [1], которую вспоминает М Мосс, приводится свидетельство, касающееся сексуальной и правильной походки у женщин народа маори, проживающего на территории Новой Зеландии. Говорится, что походка у женщин маори — это чётко очерченное и одновременно соразмерное покачивание бедрами, которое, хотя нам и кажется некрасивым, вызывает восхищение у маори. Матери там специально обучают дочерей этому способу хождения, который получил специальное название — *аниои*. [1: 408–409]. Аниои — это приобретенный, а не наследуемый способ ходьбы.

<sup>2</sup> См. об этом [3].

## 2.2. Возраст

Хорошо известна и визуально чётко определима так называемая *старческая походка*. В результате телесных изменений, связанных с разными заболеваниями, прежде всего, сердечно-сосудистыми и неврологическими, у пожилых и старых людей могут возникнуть затруднения в удержании равновесия при ходьбе. Поэтому они во время ходьбы наклоняют тело вперёд, иногда весьма значительно, опускают плечи и перемещаются по поверхности мелкими, шаркающими, осторожными шагами с неуверенными, плохо сбалансированными движениями рук и ног. У них нарушены локомоторные функции, что ведёт, в конечном счёте, к ограничениям в размере шага, в скорости передвижения и в форме телесных объектов, участвующих в походке. Иными словами, *старческая походка* — это обозначение определённого вида телесной патологии.

Пока ребёнок учится ходить, о *детской походке* обычно не говорят. Затем отмечают, например, что детская походка «приближается к взрослой на седьмом году жизни» (интернет). Но если детской походкой движутся взрослые, то это тоже вид телесной патологии. Соответствующие изменения в манере ходьбы проявляются у взрослых людей в осторожных, мелких шагах, аномальном поджимании пальцев ног, косолапости.

Детскую походку взрослых отличают и другие расстройства, обычно вызывающие боль при ходьбе в разных частях ног, в пояснице или в области таза. Все они обусловлены уменьшением гибкости суставов, возросшей ригидностью скелета и ослаблением мышц.

## 2.3. Социальный статус

С появлением в Европе так называемых *деловых женщин* и стереотипно связываемым с ними внешним обликом и стилем поведения появляется и типичная *деловая женская походка*. Вспомним, что говорила секретарша Верочка из фильма «Служебный роман»: «Чем отличается деловая женщина от просто женщины? Походкой!» См. также: предложение:

(10) *Они умолкали и рассеивались, когда туда деловой походкой вривалась какая-нибудь делающая карьеру комсомолка* (Н. Мандельштам. Воспоминания).

Деловая походка, впрочем, судя по примерам, в большей мере присуща мужчинам, см.:

(11) *Иногда директор, проходя мимо деловой походкой, кратко приказывал <...>* (Ю. Дружников. Виза в позавчера).

(12) *<...> Лукин видел, как меняется его походка, приобретая черты деловой устремлённости* (Н. Дежнев. В концертном исполнении).

Однако женская походка, в том числе и деловая, в большей степени, чем мужская, намеренная. Иногда женщины осознанно пытаются двигаться так, а не иначе, по примеру некоторых известных киноактрис. Известно, например, что многие француженки по мере роста популярности голливудских фильмов

стали передвигаться, подражая героиням этих фильмов, сильно виляя бёдрами и одновременно стараясь идти максимально прямо.

А *мужские походы* больше, чем женские, связаны с родом занятий и профессией. Отсюда *генеральская, морская, кавалерийская, начальственная, борцовская, спортивная* (последняя бывает и женской). Тем не менее, атрибутивные сочетания, выражающие женские и нейтральные по полу социальные и профессиональные походы, тоже встречаются. Помимо *спортивной*, это *балетная, монашеская*, ср.:

(13) Она шла *монашеской походкой*, опустив глаза долу <...> (М. Булгаков. Театральный роман),

*хозяйская, воровская, бандитская* и некоторые другие виды походов. Есть также походы *арестантская* и *дембельская*, подчёркивающие временный социальный статус исполнителя похода, см.:

(14) <...> *Размеренною арестантскою походкой*, он пересекал <...> двор (О. Славникова. Стрекоза, увеличенная до размеров собаки);

(15) На палубу *дембельской походкой* выходит Дэвид Бауи: «Я спасу!» (Митьки. Митя и Гудзя).

С приобретением высокого положения в обществе походка у человека обычно меняется: она становится <более> *медленной, размеренной, важной, солидной*.

По тому, как человек движется, нередко можно понять, к какому социальному классу или слою он принадлежит. Сказанное относится, в частности, к скорости передвижения мужчин из сферы бизнеса. В [2] можно прочесть следующее замечание:

«Мужчины с низким статусом двигаются медленнее, так как не стремятся к экономическому успеху либо в силу определённых философских установок, либо полагая, что этот успех для них маловероятен.

Мужчины с высоким статусом осознают свою значимость и стараются вести себя спокойно и неторопливо (они вообще редко ходят по улице пешком на большие расстояния, а если ходят, то в сопровождении провожатых).

Мужчины со средним статусом часто достаточно амбициозны и проявляют отчётливое стремление к продвижению по социальной лестнице. В эту категорию попадают молодые люди, стремящиеся сделать карьеру».

Интересно, что какой-либо статистически значимой корреляции между скоростью движения и социальным статусом женщин из той же сферы исследователи не обнаружили.

### 3. Физические и психологические признаки исполнителя похода

В русском языке есть слова, которые обозначают телесные дефекты или дисфункции, проявляющиеся при ходьбе. Это, например, глаголы *хромать, прихрамывать, припадать, волочить <ногу>, ковлять*, имена *хромой*,

*хромота, ковыляние, прихрамывание.* Они характеризуют походку — реально аномальную или лишь имитирующую некую телесную дисфункцию, ср.:

(16) *Легкая, чуть прихрамывающая походка точно бросает Бальмонта вперед, в пространство* (А. Белый. Бальмонт).

(17) *Он <...> почти молящим взором скользил за Иваном Матвеечем, когда тот хромающей походкой пробежал мимо него* (Л. М. Леонов. Русский лес).

Без расширенного контекста в примере из «Русского леса» Леонида Леонова нельзя понять, является ли *хромающая походка* описанием реального дефекта ног у Ивана Матвееча (походка-свойство) или это описание актуального движения, не являющегося телесным дефектом (походка-состояние).

Едва ли про человека, про которого нам известно, что он с рождения хромает, мы выскажем оценочное суждение типа *\*У него плохая (некрасивая, странная, неправильная... походка)*, но вполне допустимо квалифицирующее суждение типа

(18) *Британский гость <...> проследовал своей хромающей походкой к выходу из парадных залов* (Р. Штильмарк. Наследник из Калькутты).

Очень многие прилагательные передают впечатление о походке стороннего наблюдателя и отражают связь этого впечатления с характером исполнителя походки. При языковом выражении восприятия походки нередко передаётся её оценка. С походкой люди связывают самые разные черты её исполнителя: и позитивные, и негативные, и физические, и психологические, и часто переносят признаки походки на физические и психологические особенности самого человека или его актуальное настроение (хотя тут, разумеется, возможны и ошибки). В НКРЯ встречаем примеры с такими признаками походов, как *беззаботная, благородная, вальяжная, ленивая, небрежная, нерешительная, неуверенная, осторожная, подхалимская, элегантная* и многими другими.

Естественно, что физически не очень здоровый человек, человек уставший, утомлённый, грустный или подавленный чем-то движется иначе, чем человек бодрый, радостный, с хорошим настроением. Однако выделить в точности те признаки походки, которые соответствуют в норме данному типу личности, склонности или эмоции, не удаётся. Телесные движения и жесты, определяющие узуальный или актуальный способы ходьбы, могут быть самыми разными. Я попросил шесть студентов (4 девушек и 2 юношей) представить себе, что они чем-то взбешены, и попросил их пройти вдоль комнаты (аудитории) по очереди (ни один из них не видел при этом ни одного другого). Четверо (3 девушки и 1 юноша) прошли быстрыми и относительно мелкими шагами, а двое, напротив, шли медленно, я бы сказал даже нарочито медленно. И когда я спросил их, почему они так шли, то услышал почти идентичные ответы типа «Я представил (а) себе некую ситуацию, при которой взбесился (взбесилась) бы, но стал себя как бы успокаивать и думать, что будет (вариант: что делать) дальше». Иными словами, способ актуальной ходьбы зависит и от темперамента исполнителя, и от его отношения к произошедшему событию и связанным с ним людям, и от осознания необходимости каких-то дальнейших действий и приведения

себя в надлежащее для этого состояние, и ещё от очень многого другого. Поэтому, когда читаешь безоговорочные суждения психологов о соотношении данной походки и конкретных внутренних свойств человека, такие, как:

- (19) *Походка, почти как глаза, отражает душу человека* (Интернет, сайт «Походка: как сделать её красивой и уверенной»);
- (20) *Психологи давно уже поняли (??? — знак мой Г. К.), что по походке можно определить характер* (Интернет, сайт «О чём может сказать наша походка»),

относишься к этим и подобным суждениям с недоверием.

\*\*\*

Походка не только способствует социализации личности, поскольку от умения держаться, сидеть и ходить, как и от умения вести себя в обществе и общаться с другими людьми, существенно зависит карьера человека, она может быть симптоматическим невербальным знаком, потому что даёт ключ к таким чувствам, как скорбь, страдание, гнев, гордость или радость. Так, язык говорит нам о том, что *тяжёлая походка* характерна для людей, испытывающих отрицательные эмоции (хотя тяжелой походкой могут обладать и другие люди). *Тяжёлой походкой* идут и люди, уставшие от изнурительной работы или больные, и люди скорбящие, ср.:

- (21) *Я шёл по улице **тяжелой походкой** прачки* (М. Дяченко, С. Дяченко. Магам можно всё);
- (22) *Тяжёлая походка выдавала болезни: от мужа я знала, что владыка перенес два обширных инфаркта <...>* (Е. Чиждова. Лавра), с
- (23) *Ни на кого не глядя, уставившись в землю, <...> шли **тяжёлой походкой** за гробом двое подростков* (А. Варламов. Купавна).

*Тяжёлая походка* — это также походка страдающего человека и человека грустного, ср. неправильное \**Он шёл радостной тяжёлой походкой*. Чувство радости хорошо согласуется как раз с *лёгкой походкой*, ср.:

- (24) *И мы видим, как она, эта женщина, уже совершенно свободная, **легкой походкой** идет по улице, видим, как она весело играет со своим ребенком, какое счастливое у нее лицо* (В. Розов. Удивление перед жизнью).

Отмечено, что если человек страдает, то его *походка вялая, угнетённая*, а если он полон сил и энергии, если у него высокая самооценка, он решителен и целеустремлён, то у него *уверенная походка*, ср.:

- (25) *Представим профессора, который полон сил и энергии, хочет созидать и двигать вперед науку, у него горят глаза и **уверенная походка*** (Б. Якеменко. Путь к России).

Трудно определить, что такое *правильная походка* и *неправильная походка*, в частности, что означают фразы выдающегося отечественного физиолога В. М. Бехтерева:

(26) *Хорошей, правильной походкой она никогда не обладала*

и современной писательницы Веры Белоусовой:

(27) *Неправильная походка, — подумал Мышкин, глядя ей вслед и ощущая странную пустоту.*

Далее, однако, в тексте из повести писательницы «Жил на свете рыцарь бедный» разъясняется, что имеется в виду под *неправильной походкой* — «Зачем она сутулится?» *И в ту же секунду, точно подслушав его мысли, девушка расправила плечи и пошла совсем иначе — упругой и элегантной походкой.* Неправильная походка здесь вызвана не движением ног, а сутулостью, то есть нарушением осанки, выражающейся в том, что у человека избыточно согнутая, полукруглая спина.

Походка может отражать не только актуальное эмоциональное состояние человека (*печальная, виноватая, смеющаяся* и др. *походки*), но и такие прагматические характеристики, как отношение её к наблюдателю, ср. (*не*)*знакомая походка, неслышная походка*, или к разворачивающейся ситуации, ср. *обычная, особенная, необыкновенная походка.*

И саму походку, и общую стилистику актуального телесного поведения человека характеризуют такие сочетания, как *игривая, театральная, сдержанная, суетливая* или *развязная, самодовольная походка*. Из наиболее частых метафорических характеристик физических или психологических свойств исполнителя походки отмечу метафоры веса, материала и свойств материала, ср. такие часто встречающиеся эпитеты при слове *походка*, как *лёгкая, тяжелая, увесистая, грузная, невесомая* и *деревянная, упругая, гибкая, эластичная* и т. п.

#### 4. Заключение

Анализ реальных жестов-походок и того, как они представлены русскими атрибутивными сочетаниями, показал, что попытка построить на основе этих сочетаний нетривиальную семантическую классификацию походок в её буквальном, точном смысле не удаётся: классы разбиения получаются пересекающимися. Между тем на примере рассмотренных сочетаний можно увидеть богатый и разнообразный семантический спектр этого вида жестов, охватывающий как характеристики самих походок, так и характеристики их исполнителей. В последнем случае обычно наблюдается рассогласование семантической и синтаксической сфер действия прилагательного.

О походках писали многие, но лингвистов среди них почти нет. Зато большое внимание походкам уделяли деятели искусства — театральные режиссёры и кинорежиссёры, актёры, писатели. На протяжении двух работ о походках — первой, посвящённой концепту походки, а также семантике и синтактике слова

*походка* — и второй, то есть данной работе, в которой речь шла о типологии русских походок, мы приводили много примеров тонких и пронизательных высказываний о походках. В заключение приведём ещё несколько таких высказываний.

В своём «Руководстве для желающих жениться», представляющем собой сборник юмористических рассказов-советов, как вести себя при сватовстве и женитьбе, А. П. Чехов пишет: «Походка торопливая, с оглядками, говорит о ветрености и легкомыслии. Походка ленивая бывает у женщин, сердце которых уже занято, — тут ты не пообедает. Походка утичья, с перевальцем и виляньем турнюра, есть признак добродушия, податливости и иногда тупости. Походка горделивая, лебединая бывает у этих дам и содержанок. Чем спесивее походка, тем, значит, старше и богаче содержатель. Такая походка у девиц означает самомнение и ограниченность. Если барыня не идёт, а плывёт, как пава, то поворачивай оглобли: она накормит, утешит, но непременно возьмёт под башмак».

Известна следующая байка, которую рассказывают о замечательном актёре Алексее Баталове (текст байки в каком-то варианте есть в интернете, но я знаю её давно). Баталова пригласили на главную роль в фильме «Дама с собачкой» по рассказу Чехова. Консультантом фильма была старая дворянка, хорошо помнившая дореволюционную эпоху. Увидев, как ходит Баталов, старушка заявила: «Он косолапит. Так русские интеллигенты той эпохи не ходили». Эти слова расстроили артиста, и он принялся работать над походкой, тем более что режиссёр тоже постоянно делал ему по этому поводу замечания. В конце концов, Баталову это надоело, а походка никак не хотела измениться. Однажды они приехали в Крым, в Ялту, и стали обходить места будущей съёмки. И вдруг какой-то старичок на лодке, увидев одетого для съёмки Баталова, закричал: «Смотрите-ка, шляпа точь-в-точь, как до революции носили! Такая была, помнится, у Антона Павловича Чехова. И, между прочим, молодой человек, — обратился он к Баталову, — у вас походка чрезвычайно напоминает походку Чехова. Он тоже чуть косолапил».

Существует также много пословиц о походках, причём самых разных народов. Вот, например, несколько пословиц, выделяющих разные аспекты, которые связаны с походкой и которые были отмечены нами выше. Это японская пословица: *Когда легко на сердце — и походка легка*; турецкая пословица: *Ворона, попытавшись изобразить походку куропатки, ходить разучилась*; русские пословицы: *Хоть к обедне не поспеть, а походки не потерять*; *Журавлиная походка не нашей стати*; *Ясного сокола видать по полёту, а молодца — по походке*; таджикская пословица: *Если ворона будет подражать походке аиста, то и свою походку потеряет*. И они показывают, что люди давно подметили тесную связь походки с личностью её исполнителя и связь характера походки с характером человека.

## References

1. *Best E.* (1979). *The Maori*. AMS Press, New York.
2. *Butovskaya M. L., Levashova V. V.* (2004). Speed and body language of the pedestrians in the modern cities: ethological analysis [Skorost' dvizheniya i yazyk tela peshehodov v usloviyah sovremennogo goroda: etologicheskiy analiz] // *Archeology, ethnography and anthropology of Eurasia [Arheologiya, etnografiya i antropologiya Evrazii]*, № 3 (19), pp. 147–156.
3. *Kreydlin G. E.* (2005). *Men and women in nonverbal communication [Muzhchiny i zhenshchiny v neverbal'noy kommunikatsii]*. Yazyki slavyanskoy kul'tury, Moscow.
4. *Kreydlin G. E.* (2016). A gait as a concept and a word [Pohodka: concept i slovo] // *Ekologiya yazyka i kommunikativnaya praktika [Language ecology and communicative practice]*, № 2, pp. 127–142.
5. *Moss M.* (1996). *Societies. Exchange. Personality [Obshchestva. Obmen. Lichnost']*, Nauka, Moscow.
6. [www.osipov.vinchi.ru](http://www.osipov.vinchi.ru)
7. [www.ruscorpora.ru](http://www.ruscorpora.ru)



# ФОНЕТИЧЕСКИЕ ХАРАКТЕРИСТИКИ ДЫХАТЕЛЬНЫХ ПАУЗ С РАЗНОЙ ТЕКСТОВОЙ ЛОКАЛИЗАЦИЕЙ<sup>1</sup>

**Кривнова О. Ф.** (okrivnova@mail.ru)

Московский государственный университет  
им. М. В. Ломоносова, Москва, Россия

**Ключевые слова:** фонетика, устная речь, просодическое членение, интонационно-смысловая дыхательная пауза, текстовая локализация, инструментальный анализ

## PHONETIC CHARACTERISTICS OF BREATHING PAUSES WITH DIFFERENT TEXT LOCALIZATION

**Krivnova O. F.** (okrivnova@mail.ru)

Lomonosov Moscow State University, Moscow, Russia

In spite of its fundamental role in the processes of speech production, speech breathing is studied experimentally quite insufficiently. The main task of this paper was to obtain phonetic-acoustic data on the breathing pauses (BP) with different text localization in oral Russian texts. In the introduction we formulate the problems being discussed with the main goal to analyze the correlation of BP's acoustics with the boundaries of principal text units—paragraphs, sentences, clauses, taking into account the interspeaker variability in reading of the same text. In our previous studies it was supposed that the distinctive function of BP as prosodic markers of macrosegmentation is realized by the stable differences in their general phonetic patterns and such acoustic parameters as duration, intensity and noise spectrum of the inhalation phase. The quantitative analysis of these differences comprises the main part of the paper. The second section describes the material and the methods of research. In the third section the analysis of the general acoustic-physiological pattern of BP with different text localization is given, which allows to set the regular objective differences between them. The fourth section contains the data of BP's duration in connection with their text localization. It is shown that BP's text localization is the main factor to control its duration. In the fifth section the similar information is presented for the average BP's intensity. The sixth section contains an analysis of the spectral characteristics of the inhalation noise in BP.

---

<sup>1</sup> Исследование проведено при поддержке гранта РФФИ 15-06-06103

It is shown that it is the most invariant feature for all BP types, almost independent of its text localization. The conclusion summarizes the study's results emphasizing the fact that the presence of BP in any oral text can serve as a sufficient indication of its structure, but interspeaker variability shows that BP's realization is not a necessary feature of the text boundary.

**Key words:** phonetics, spoken language, prosodic phrasing, breathing pause, text localization, interspeaker variability, instrumental analysis

## 1. Введение

Механизм речевого дыхания, несмотря на свою фундаментальную роль в процессах речеобразования, в экспериментальном плане до сих пор изучен мало. Наиболее обстоятельные результаты в этой области были получены американским фонетистом П. Ладефогедом. Основные результаты обобщены в (Ladefoged 1967), где речевое дыхание рассматривается в разных аспектах — физиологическом, аэродинамическом, акустическом, и во взаимодействии с другими артикуляционными процессами. Нельзя не отметить, к сожалению, что физиологическая сторона речевого дыхания по-прежнему мало доступна для прямого анализа в естественных условиях. Однако в последнее время появились исследования, в которых используются компьютерные методы детектирования и анализа дыхательных пауз в устной речи, как при чтении, так и в спонтанном режиме. Так, в работе (Johansson et al. 2005) идентификация ДП основана на спектральных эталонах и алгоритме динамического программирования. Авторы отмечают, что алгоритм работает лучше на мужской речи по сравнению с женской, а в целом, результаты сильно зависят от степени зашумленности и интенсивности сигнала. В работе (Braunschweiler, Chen 2013) ДП идентифицируются с использованием скрытой Марковской модели с 3-мя состояниями, которая обучается на речевом корпусе и учитывает 3 акустических параметра: ЧОТ, среднюю амплитуду и спектр. Также показано, что добавление акустической модели ДП в систему синтеза речи на немецком языке значительно повышает ее естественность, особенно при озвучивании больших текстов, в частности аудиокниг. Оказалось также, что введение информации о локализации ДП в модель языка при распознавании речи, основанном на Марковских моделях, улучшает точность распознавания (Weilhammer, Schiel 1999).

На материале русского языка подобных исследований не проводилось. В связи с этим задача настоящей работы заключалась в том, чтобы получить более детальные акустико-фонетические сведения о ДП с разной текстовой локализацией для русской устной речи. Нас интересовала как общая акустико-физиологическая картина ДП, так и их конкретные акустические характеристики: длительность, интенсивность и спектр шумовой части ДП (вдоха).

Вопрос о текстовой локализации ДП подробно обсуждался в (Кривнова 2016), поэтому ниже указаны кратко лишь наиболее важные результаты проведенного нами исследования, а основное внимание будет уделено анализу фонетических характеристик ДП учетом их текстовой локализации и ряда других факторов.

Анализ текстовой локализации ДП на материале прочтения одного и того же текста разными дикторами позволяет сделать следующие выводы:

1. Главным фактором, который влияет на организацию дыхания в репродуцированной речи, является стратегия интонационного паузирования диктора, для которой характерна тенденция к реализации темпоральных пауз после каждой клаузы в предложении. Однако эта достаточно регулярная тенденция взаимодействует с когнитивными характеристиками дикторов: некоторые дикторы в определенных синтаксических условиях «пропускают» конечные границы произносимых клауз, в то время как другие часто реализуют темпоральные паузы в определенных точках внутри произносимой клаузы.
2. В организации речевого дыхания находит отражение иерархическая структура текстовых единиц, основание которой образуют отдельные клаузы.
3. Текстовые фрагменты упорядочиваются по убыванию вероятности последующего вдоха следующим образом (в скобках дается частота реализации вдоха в среднем по 10 дикторам): абзац (100%) > самостоятельное предложение внутри абзаца (94%) > клауза внутри предложения (65%) > компонент внутри клаузы (34%).
4. Полученные результаты подтверждают мнение многих исследователей о центральной роли клаузы и текстовых границ более высокого ранга в процессах порождения, понимания и озвучивания русских текстов, и в том числе в организации речевого дыхания. Аналогичные результаты были получены и на материале других языков, причем не только для локализации ДП, но и для глубины вдоха и объема вдыхаемого воздуха в интервале ДП (Winkworth et al. 1994; 1995; Rochet-Capellan, Fuchs 2013). Для изучения физиологических параметров дыхания современные исследователи чаще всего используют миниатюрный респираторный плейтисмограф, регистрирующий сигналы дыхания на уровне грудной клетки и диафрагмы с одновременным использованием ротового микрофона. Данная технология подробно описывается в (Włodarczak et al. 2015). Авторы планируют применить ее для создания речевого корпуса разговорной речи с разметкой на речевые и дыхательные фрагменты и с разграничением разных фаз в ДП. Аналогичная техника уже использовалась на материале немецкого языка для исследования связи дыхательного ритма со сменой ролей в диалоге (Rochet-Capellan, Fuchs 2016). Полученные результаты довольно противоречивы: на протяжении диалога авторы не обнаружили отчетливой координации в дыхательной активности участников, однако при этом выявили ряд признаков, которые показывают, что локально смена реплик связана с вдохами говорящего на стадии подготовки реплики. А в целом, как и в других, более «плавных» формах устного дискурса (чтении и спонтанном монологе), в диалоге имеется тенденция к координации речевого дыхания с важными дискурсивно-текстовыми событиями и их таймированием.

## 2. Материал и методика настоящего исследования

В качестве материала использовался корпус прочтений небольшого рассказа о посещении научного учреждения<sup>2</sup>. Текст был прочитан «с листа»<sup>3</sup> десятью дикторами, носителями русского языка с высшим образованием, но без специальной дикторской подготовки; средняя длительность озвученного текста 3–3,5 минуты. Материал записывался на компьютер (SR 22050 Гц, 16-bit, Mono) в условиях тихой комнаты с использованием высокочувствительного микрофона, что позволило в большинстве случаев выявить ДП в каждом прочтении текста, при некотором локальном усилении сигнала на темпоральной паузе.

Материал был отобран из большего массива, включавшего 30 прочтений текста разными дикторами (суммарный объем исходного речевого массива около 400 Мб). Для дальнейшего анализа было выбрано 10 наилучших прочтений, среди которых удачно оказалось 5 мужских и 5 женских — далее они обозначаются соответственно -m-i и f-i, где i меняется от 1 до 10 и обозначает место, которое занял диктор в десятке лучших текстовых прочтений.

Дыхательное заполнение пауз определялось на слух и визуально по осциллограммам с использованием анализатора Speech Analyzer SA SIL, версия 1.5-2002 и SA 3.1-2007. Акустические характеристики ДП и других речевых отрезков измерялись с помощью того же анализатора с учетом предварительной «ручной» сегментации речевого материала. Статистическая обработка инструментальных данных производилась средствами Excel MS-2000.

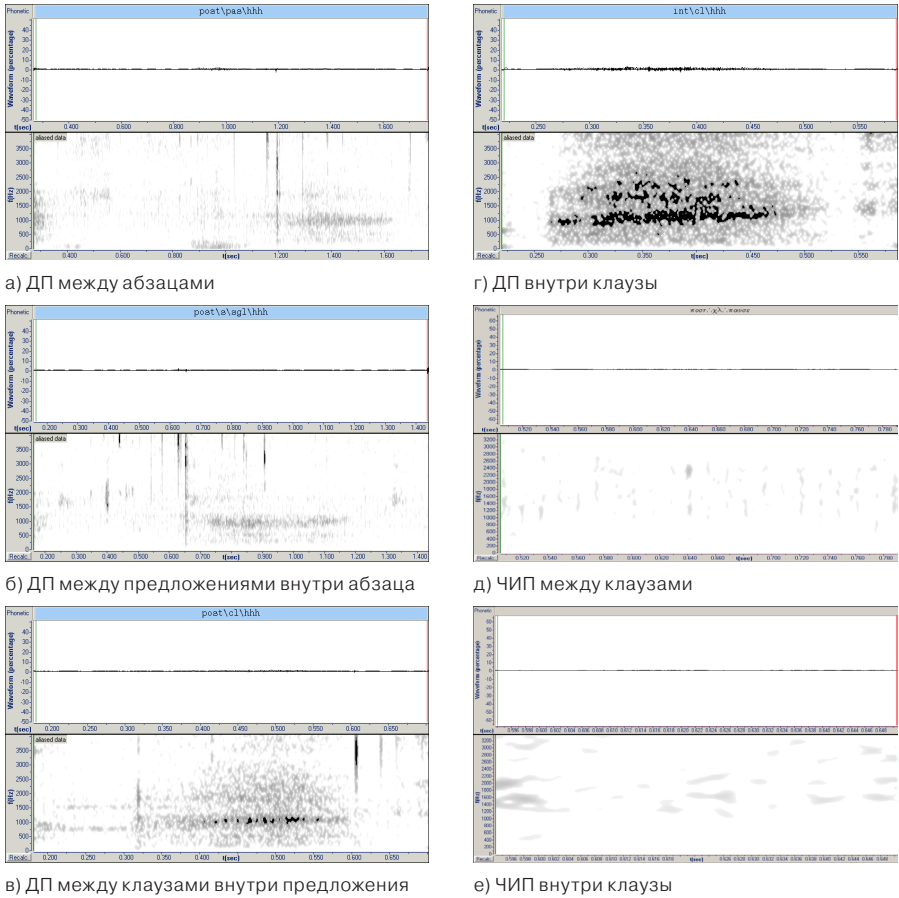
## 3. Общая фонетическая картина дыхательных пауз с разной текстовой локализацией

На рис. 1 представлены осциллограммы и спектрограммы ДП с разной текстовой локализацией в прочтении экспериментального текста диктором-женщиной f-2. Отметим, что по экспертному рейтингу это второе наилучшее прочтение в экспериментальном массиве. Аналогичные данные для наилучшего мужского прочтения того же текста приведены в работе (Кривнова 2010). Для сравнения на рис. 1 даны также иллюстрации для «чистых» интонационных пауз (ЧИП) с тождественной текстовой локализацией.

---

<sup>2</sup> Текст был взят из методической разработки по составлению текстовых массивов [(Штерн 1984), а в качестве основы для него использовался отрывок из книги С. Иванова «Схватка с роботом». М., 1977.

<sup>3</sup> Материал для эксперимента был любезно предоставлен московской компанией Stel Computer Systems, ведущей разработку систем автоматического распознавания речи для русского языка.



**Рис. 1.** Диктор f-2. Акустическая картина дыхательных пауз (ДП) — слева. ЧИП — чистая интонационная пауза (без элементов вдоха) — справа. В иллюстративных целях рисунки масштабированы по оси Y в соотношении 1х2, а затем приведены к одной высоте, а также приведены к одной длине по оси X. Паузы были выделены курсорами без захвата соседних участков речевого сигнала.

Даже простой визуально-слуховой анализ материала, аналогичного представленному на рис. 1, приводит к заключению, что ДП с разной текстовой локализацией имеют различное акустико-физиологическое наполнение, которое регулярно воспроизводится в прочтениях разных дикторов. А именно:

1. В ДП между абзацами отчетливо выделяются три фазы: сначала носовой вдох, которому может предшествовать краткий выдох, затем достаточно резкий ротовой шумный вдох, которому обычно предшествуют явления чмокания, сглатывания и под. (это т.н. артикуляционный

шум). На спектрограммах ДП они видны хорошо и реализуются подобно взрывам смычных согласных. За ротовым вдохом идет короткий относительно «немой» участок, на котором тоже может присутствовать артикуляционный шум. Спектрограммы обнаруживают заметные различия в интенсивности и спектре шума на носовой и ротовой фазах вдоха в ДП. В (Garman 1990), со ссылкой на исследования речевого дыхания (Butcher 1981), также говорится о трехфазной структуре ДП с выделением 3 участков *pre-inhalation*, *inhalation*, *post-inhalation*, с разными акустическими характеристиками. Отмечается, что временное соотношение фаз ДП различно для спонтанной речи и чтения, при этом фаза ротового вдоха практически постоянна и равна 400 мс, что, возможно, объясняется чисто физиологическими факторами.

Предполагается, что ДП большей длительности оказываются таковыми за счет двух других фаз — предшествующей и следующей за вдохом. Именно эти фазы увеличиваются по длительности в спонтанной речи, что позволяет, по мнению автора, ввести в интегральную оценку плавности речи такой показатель, как *доля собственно вдоха в общей длительности ДП*. Однако, как показывает наш материал, этот показатель зависит не только от формы речи, но и от текстовой локализации ДП (см. ниже).

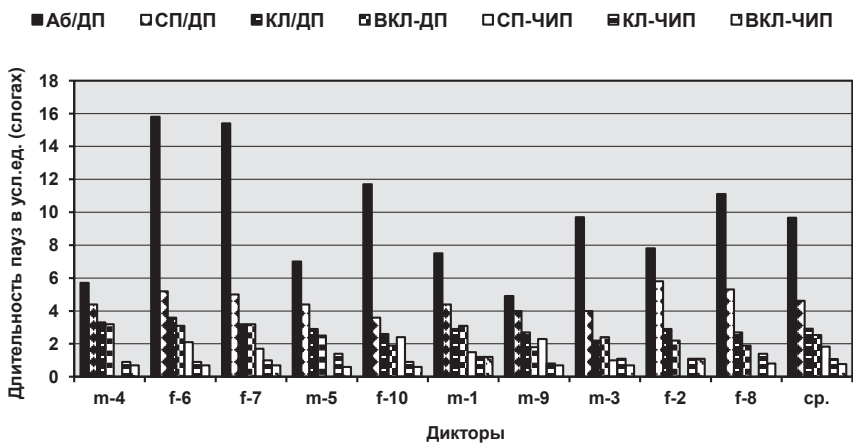
2. ДП между самостоятельными предложениями внутри абзаца характеризуются фонетической картиной, сходной с ДП между абзацами. Возможно, есть некоторые различия во временных характеристиках носовой и ротовой фаз ДП, что, в свою очередь, может быть связано с различиями в общей длительности ДП между и внутри абзаца.
3. В ДП между клаузами внутри предложения, как правило, отчетливо выражена только шумная фаза ротового вдоха, а явления чмоканья в ее начале менее заметны и встречаются реже, чем в ДП более высокого текстового уровня. Кроме того, шум на ротовой фазе вдоха имеет существенно бóльшую интенсивность.
4. ДП внутри клаузы демонстрируют дальнейшее нарастание явлений, отмеченных для ДП после клаузы внутри предложения. Так, темпоральная пауза практически полностью заполнена шумом ротового вдоха, инициальных явлений чмоканья не наблюдается, шум вдоха очень интенсивен и отчетливо фиксируется на спектре ДП, аналогично шумным фрикативным согласным.

Что касается гендерных различий, то при сходстве общей фонетической картины ДП определенные различия наблюдаются в выраженности и интенсивности шума вдоха, особенно в ротовой фазе. В среднем, дикторы-мужчины в своих текстовых прочтениях дышат более шумно, чем дикторы-женщины.

#### 4. Длительность дыхательных пауз с разной текстовой локализацией

Мысль о том, что длительность интонационных пауз, включая ДП, отражает иерархическую структуру текста, высказывалась уже давно (Волконский 1913; Светозарова 1982). Однако конкретных данных в литературных источниках немного.

Нас интересовала прежде всего средняя длительность ДП с различной текстовой локализацией в прочтениях одного текста разными дикторами, а также соотношение длительностей ДП и ЧИП с аналогичной текстовой локализацией. Междикторская и внутрдикторская вариативность длительности текстовых пауз в настоящей работе не была предметом специального анализа, так как общее количество темпоральных пауз с разной текстовой локализацией в прочтении каждого отдельного диктора было сравнительно невелико (30–60). Некоторое представление о вариативности можно получить из показателей стандартного отклонения от средних значений паузальных длительностей (табл.1). Длительность пауз определялась как в абсолютных единицах (мсек), так и в условных, нормированных относительно темпа (средней длительности слога) в прочтении конкретного диктора. Подробные данные о длительности пауз разного типа представлены ниже в таблице 1 и в более наглядной форме на рис. 2 для нормированных длительностей ДП и ЧИП.



**Рис. 2.** Нормированная длительность ДП и ЧИП с разной текстовой локализацией в прочтениях разных дикторов, в среднем и по отдельным дикторам. Обозначения текстовых границ: Аб — между абзацами, СП — между предложениями внутри абзаца, КЛ — между клаузами внутри самостоятельного предложения, ВКЛ — внутри клаузы.

**Таблица 1.** Абсолютная (мс) и нормированная (в слогах) длительность ДП и ЧИП с различной текстовой локализацией. Данные упорядочены по темпу (средней длительности слога) в прочтении диктора (в сторону возрастания); по строкам m-5 — m-1 проходит примерная граница между быстрым и замедленным темпом произнесения относительно нейтрального.

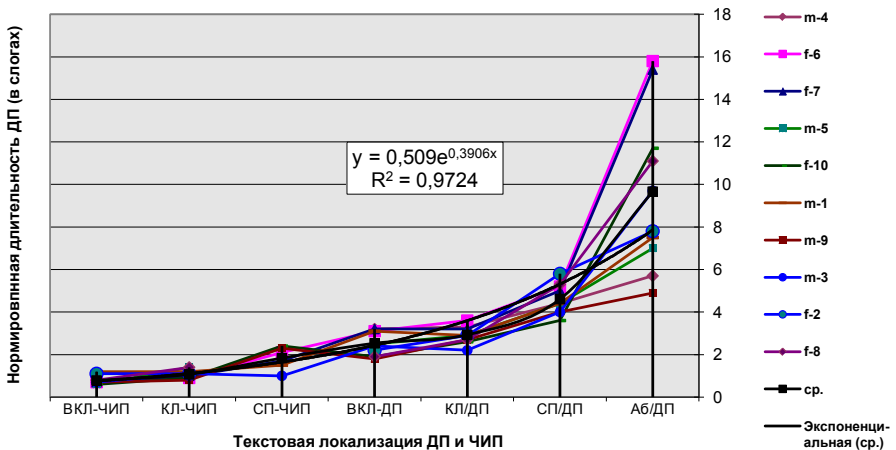
Дикторы	Ср. длительность слога	Аб-ДП		СП-ДП		КЛ-ДП		ВКЛ-ДП		СП-ЧИП		КЛ-ЧИП		ВКЛ ЧИП	
		мс	сл	мс	сл	мс	сл	мс	сл	мс	сл	мс	сл	мс	сл
m-4	0,141	0,796	5,7	0,624	4,4	0,462	3,3	0,455	3,2	–	–	0,126	0,9	0,099	0,7
		0,376	2,7	0,193	1,4	0,123	0,9	0,055	0,4	–	–	0,084	0,6	0,046	0,3
f-6	0,142	2,236	15,8	0,741	5,2	0,516	3,6	0,444	3,1	0,295	2,1	0,132	0,9	0,105	0,7
		0,475	3,3	0,193	1,4	0,102	0,7	0,048	0,3	–	–	0,114	0,8	0,079	0,6
f-7	0,146	2,247	15,4	0,727	5,0	0,446	3,2	0,467	3,2	0,248	1,7	0,145	1	0,106	0,7
		1,209	8,3	0,156	1,1	0,121	0,8	0,166	1,1	0,158	1,1	0,113	0,8	0,079	0,5
m-5	0,153	1,076	7,0	0,670	4,4	0,448	2,9	0,381	2,5	–	–	0,212	1,4	0,092	0,6
		0,210	1,4	0,165	1,1	0,107	0,7	0,033	0,2	–	–	0,143	0,9	0,070	0,5
f-10	0,165	1,941	11,7	0,592	3,6	0,434	2,6	0,309	1,9	0,398	2,4	0,148	0,9	0,104	0,6
		0,502	2,3	0,129	0,9	0,089	0,5	0,033	0,2	–	–	0,096	0,6	0,058	0,4
m-1	0,167	1,251	7,5	0,739	4,4	0,488	2,9	0,523	3,1	0,256	1,5	0,196	1,2	0,196	1,2
		0,249	1,5	0,253	1,5	0,206	1,2	0,265	1,6	0,110	0,7	0,098	0,6	0,097	0,6
m-9	0,174	0,858	4,9	0,695	4,0	0,463	2,7	0,321	1,8	0,398	2,3	0,132	0,8	0,117	0,7
		0,133	0,8	0,293	1,1	0,122	0,7	0,128	0,7	–	–	0,084	0,1	0,112	0,6
m-3	0,176	1,711	9,7	0,703	4,0	0,382	2,2	0,421	2,4	0,178	1,0	0,189	1,1	0,123	0,7
		0,430	2,4	0,168	1,0	0,081	0,5	0,166	0,9	0,069	0,4	0,095	0,5	0,070	0,4
f-2	0,177	1,381	7,8	1,025	5,8	0,511	2,9	0,397	2,2	–	–	0,187	1,1	0,187	1,1
		0,245	1,4	0,350	2,0	0,122	0,7	0,128	0,7	–	–	0,059	0,0	0,060	0,3
f-8	0,187	2,068	11,1	0,998	5,3	0,495	2,7	0,354	1,9	–	–	0,261	1,4	0,144	0,8
		0,133	0,2	0,520	2,8	0,159	0,9	0,091	1,5	–	–	0,165	0,9	0,082	0,4
Ср. длительность паузы		1,557	9,7	0,751	4,6	0,465	2,9	0,407	2,5	0,296	1,8	0,173	1,1	0,127	0,8
т. отклонение		0,557	3,8	0,147	0,7	0,040	0,4	0,068	0,6	0,088	0,5	0,044	0,2	0,037	0,2

Анализ полученных данных приводит к следующим выводам:

1. Дыхательные паузы имеют большую длительность, чем чистые интонационные паузы независимо от текстовой категории последних. Это выполняется как в среднем, так и по отдельным дикторам. Однако для пауз одинаковой текстовой категории это соотношение выдерживается более отчетливо. Реализация вдоха увеличивает длительность соответствующей ЧИП условно на 2–2,5 слога (~ 300–400 мс).
2. Текстовые фрагменты упорядочиваются по убыванию длительности следующей ДП следующим образом: в скобках дается средняя нормированная длительность ДП: абзац(~10 слогов) > самостоятельное предложение внутри абзаца(~5 слогов) > клауза внутри предложения (~3 слога) > компонент внутри клаузы (~3 слога). Таким образом, предположение о зависимости длительности ДП от ее текстовой категории в целом подтверждается, хотя в этом отношении ДП между клаузами и внутри клауз



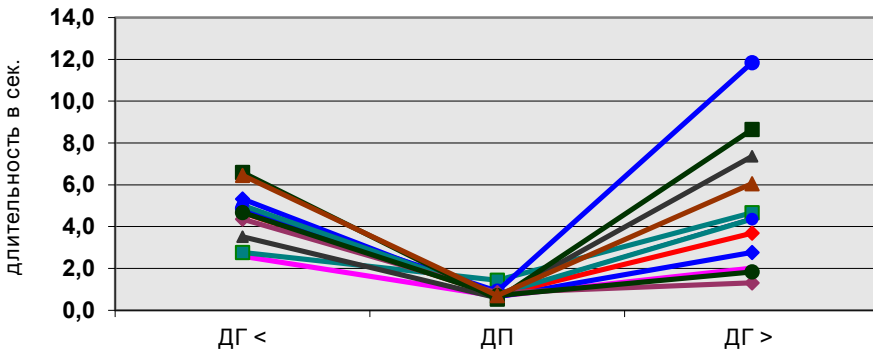
как самые короткие среди ДП, по длительности могут пересекаться, причем в различной степени у разных дикторов. То же самое наблюдается и для ЧИП. На рис. 3 приводится также формула, которая позволяет количественно оценивать данную зависимость. R2-показатель, который отражает степень соответствия наблюдаемых и расчетных данных по длительности ДП в зависимости от ее текстовой категории, близок к 1, что свидетельствует о хорошем согласовании данных.



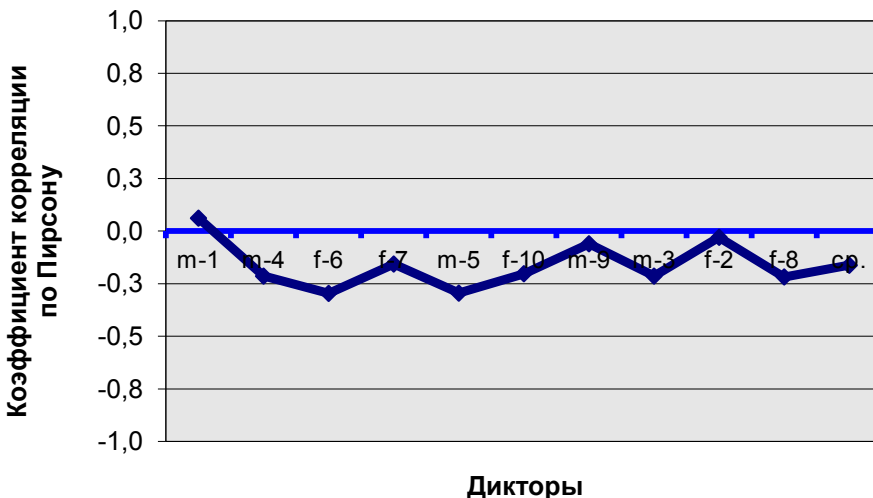
**Рис. 3.** Зависимость длительности ДП от ее текстовой локализации в среднем и по разным дикторам. Условные обозначения текстовых границ см. выше на рис. 2.

- Дополнительный анализ показал, что текстовая локализация ДП является главным фактором, определяющим ее длительность. На нашем материале не обнаружено (рис. 4) значимых корреляций между длительностью ДП и таким фактором, как длина предшествующей и последующей дыхательных групп (ДГ), который часто указывается как контролирующий длительность ДП при порождении речи (Congrad et al.1983; Henderson et al. 1965; Mitchell et al. 1996).
- Не обнаружено и корреляции с длительностью предшествующей дыхательной паузы, см. рис. 5, которую можно было бы ожидать, исходя из принципа экономии дыхательных усилий, на что указывается в работе (Златоустова 1968).

<sup>4</sup> В (Rochet-capellan, Fuchs 2013) отмечается также, что длительность ДП в спонтанной речи зависит не только от длины последующей ДГ, но и от синтаксического типа первой клаузы в ней.



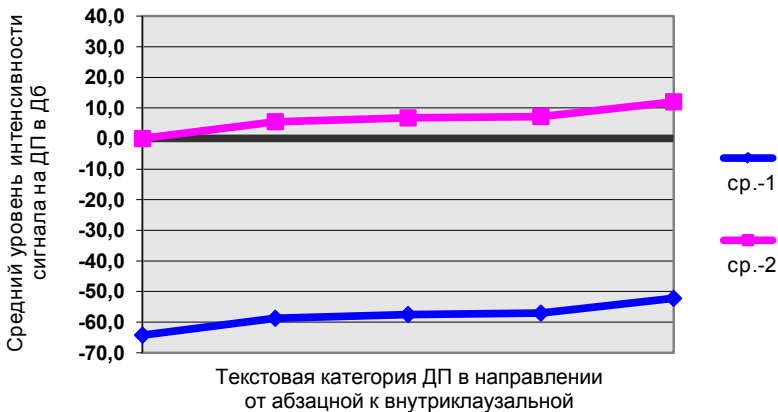
**Рис. 4.** Зависимость длительности ДП между самостоятельными предложениями внутри абзаца от длины соседних ДГ в сек. (диктор m-1). Условные обозначения: ДГ1 < дыхательная группа перед ДП на границе между предложениями, ДГ2 > — дыхательная группа после ДП на границе между предложениями.



**Рис. 5.** Корреляция между длительностями соседних ДП (в среднем и по разным дикторам). Коэффициенты корреляции (по Пирсону) близки к 0, что не позволяет говорить о какой-либо корреляционной зависимости между длительностями соседних ДП.

## 5. Интенсивность на отрезке ДП с разной текстовой локализацией

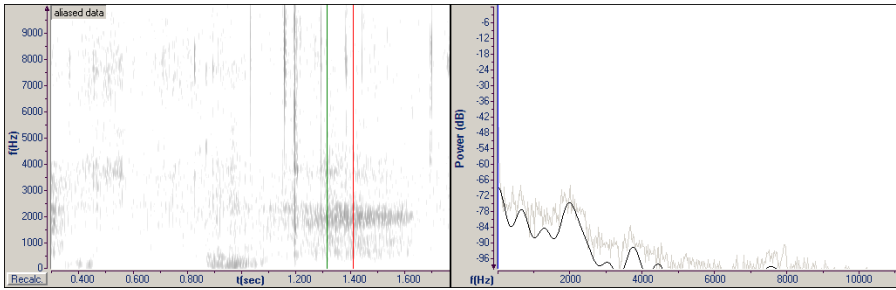
Особенности общей картины ДП, о которых говорилось выше, позволяют предположить, что интенсивность акустического наполнения ДП также может отражать их текстовую локализацию. Как видно из рис. 6, это предположение подтверждается данными. Понижение ранга текстовой границы, совмещенной с вдохом, в направлении от абзацной ДП к внутриклаузуальной, сопровождается увеличением *среднего* уровня интенсивности сигнала на интервале ДП, правда, весьма незначительным — до 10 дБ при усреднении по 10 дикторам. Более детальные данные показывают, что это происходит как за счет сокращения общей длительности ДП, так и за счет усиления шумовой части вдоха.



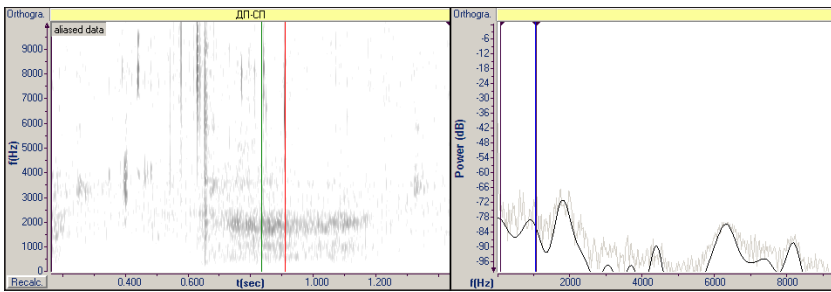
**Рис. 6.** Средний уровень интенсивности акустического наполнения ДП относительно максимального уровня записи (ср-1) и относительно уровня фона в начале записи (ср-2).

## 6. Спектральные характеристики шума вдоха на отрезке ДП с разной текстовой локализацией

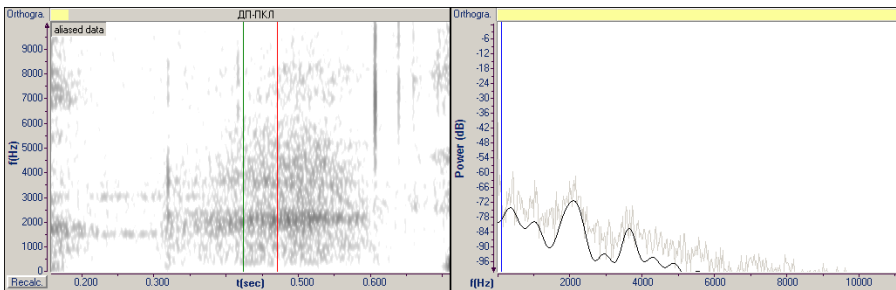
На спектрограммах рис. 1, видно, что ДП содержат участок шума приблизительно одинаковой длительности и разной интенсивности, спектр мощности которого находится в низкочастотной области с выраженным максимумом (формантой) на частоте 2000 Гц. В целом, такая спектральная картина характерна для фрикативных согласных заднего места образования, что естественно для шума речевого вдоха. Более точные спектральные данные можно получить из спектральных срезов на участке шума. Выборочные иллюстрации на рис. 7 подтверждают наблюдения об определенной инвариантности спектральных характеристик шума для ДП с разной текстовой локализацией. По спектрограммам видно также, что при уменьшении ранга текстовой локализации ДП от абзацной к внутриклаузуальной увеличивается частотный диапазон шума за счет его верхней границы.



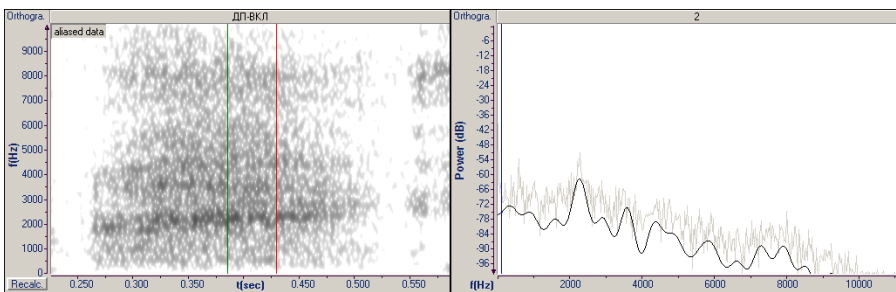
а) ДП-Абзац



б) ДП-СП



в) ДП-ПКЛ



г) ДП-ВКЛ

**Рис. 7.** Спектральные характеристики шума вдоха на ДП с разной текстовой локализацией (слева динамические спектрограммы, справа спектральные срезы, полученные на временных фреймах, выделенных курсорами).

## 7. Заключение

Устойчивые различия в общей фонетической картине и физических характеристиках ДП с разной текстовой локализацией создают возможность детектирования ранжированных границ между смысловыми отрезками текста как в естественном устном дискурсе, так и в задачах автоматической обработки речи, по крайней мере в режиме чтения. Отчетливая зависимость между акустическими и текстовыми показателями наблюдается для длительности: чем выше ранг текстовой границы, тем больше длительность ДП. Аналогичная зависимость, но менее яркая, характерна для среднего уровня сигнала на интервале ДП. Спектр шума в ДП («ротового» вдоха) — наиболее инвариантная характеристика для всех ДП, практически независимая от текстовой локализации. Этот акустический параметр может использоваться как достаточно надежный детектор ДП, что продемонстрировано на материале английского языка (Johansson et al. 2005). Корреляция ширины частотного диапазона шума ДП с разной текстовой локализацией требует отдельного исследования на более представительном речевом материале.

В итоге можно сказать, что, реализация вдоха в интонационной паузе является достаточным признаком наличия смысловой текстовой границы, а различия в длительности и интенсивности шума в ДП с разной текстовой локализацией несут информацию о смысловой и синтаксической структуре текста.

При этом нужно иметь в виду, что в общем случае *отсутствие* ДП *не* является абсолютно надежным признаком отсутствия текстовой границы. Так, даже в режиме чтения некоторые дикторы в определенных текстовых условиях не делают вдохов между самостоятельными предложениями внутри абзаца (Кривнова 2016).

## References

1. Garman M. (1990), Psycholinguistics. Cambridge Univ. Press.
2. Braunschweiler N., Chen L. (2013), Automatic detection of inhalation breath pauses for improved pause modelling in HMM-TTS // 8th ISCA Speech Synthesis Workshop. Barcelona, Spain.
3. Butcher A. (1981), Aspects of the speech pause: Phonetic correlates and communicative functions // Arbeitsberichte. N15. Kill.
4. Johansson A., Frid J., Horne M. (2005), Using cepstral coefficients for inhalation pause detection in spontaneous speech. Department of Linguistics and Phonetics. Lund University, Lund, Sweden. VINNOVA Language Technology Program.
5. Krivnova O. F., Chardin I. S. (1999), Pausing for automatic speech synthesis [Паузирование при автоматическом синтезе речи], Theory and practice of speech research (ARSO-99). Proceedings of the conference [Теория и практика речевых исследований (ARSO-99). Материалы конференции]. М.
6. Krivnova O. F. (2010), General phonetic pattern of breathing pauses in a spoken Russian text. (reading mode) [Общая фонетическая картина дышател'ных пауз в звучащем русском тексте (в режиме чтения)], Phonetics and grammar: present, past and future: to the 50th anniversary of the scientific activities

- of S. K. Pozharitskoj [Fonetika i grammatika: nastojashcheje, proshedsheje i budushcheje: k 50-letiju nauchnoj dejatel'nosti S. K. Pozharitsoj]. Ser. Issues of Russian linguistics. V. XIII. MGU, M., pp. 61–70.
7. Krivnova O. F. (2016), Prosodic phrasing in spoken text: localization of breathing pauses [Prosodicheskoje chlenenije v zvuchashchem tekste: lokalizatsija dyhatel'nyh pazv], Computer linguistics and intellectual technologies. Proceedings of the annual international conference "Dialogue [Komp'juternaja lingvistika i intellektual'nyje tehnikigii. Materialy jezhegodnoj mezhdunarodnoj konferentsii 'Dialog'] M., RGGU, v.14,t.1, pp. 340–354.
  8. Ladefoged P. (1967), Three Areas of Experimental Phonetics, Oxford UP.
  9. Rochet-capellan A., Fuchs S. (2013), The interplay of linguistic structure and breathing in German spontaneous speech // *Interspeech-2013*, 2014–2018.
  10. Rochet-Capellan A., Fuchs S. (2016), Take a breath and take the turn: how breathing meets turns in spontaneous dialogue // *Philos Trans R Soc Lond B Biol Sci*. 2014 Dec 19; 369(1658).
  11. Shtern A. S. (1984), Articulation tables. Methodical recommendations for developing listening skills and testing of auditory function [Artikuljatsionnyje tablitsy. Metodicheskaja razrabotka dlja razvitija navykov audirovanija i testirovanija sluhovoj funktsii], L.
  12. Svetozarova N. D. (1982), Intonation system of the Russian language [Intonatsionnaja sistema russkogo jazyka] L.
  13. Volkonsky S. (1913), Expressive Word [Vyzritel'noje slovo]. St. Petersburg.
  14. Winkworth A. L., Davis P. J., Adams R. D., & Ellis, E. (1994), Variability and Consistency in Speech Breathing During Reading: Lung Volumes, Speech Intensity, and Linguistic Factors // *J. of Speech, Language, and Hearing Research*, 37 (6), pp. 535–556.
  15. Winkworth A. L., Davis P. J., Adams R. D., Ellis E. (1995), Breathing Patterns During Spontaneous Speech // *J. of Speech, Language, and Hearing Research*, 38(1), 124–144.
  16. Weilhammer K., Schiel F. (1999), Investigation of language structure by means of language models incorporating breathing and articulatory noise // *CPhS-14*, pp. 1649–1652.
  17. Włodarczak M., Heldner M., Edlund J. (2015). Breathing in Conversation: an Unwritten History// *Proceedings of the 2nd European and the 5th Nordic Symposium on Multimodal Communication* pp. 107–112.
  18. Zlatoustova L. V. (1968), Some remarks on speech breathing [Nekotoryje zamechanija o rechevom dyhanii], Researches on speech information [Issledovanija po rechevoj informatsii], M.

# LINGUISTIC TENDENCIES IN ENGLISH TO RUSSIAN TRANSLATION: THE CASE OF CONNECTIVES<sup>1</sup>

**Kunilovskaya M. A.** (mkunilovskaya@gmail.com)

University of Tyumen, Tyumen, Russia

Lexical elements, which signal text cohesion (*though, moreover, etc*), are a convenient and attractive research object in Translation Studies. Unlike other words, connectives are structurally optional, more context-independent and, therefore, more revealing in terms of motivations behind translators' linguistic choices. Their frequencies are used to establish differences between translations and non-translations and are interpreted as a linguistic indicator of several tendencies in translation process such as explicitation, simplification and convergence. Particular patterns of translator's choices result in different degrees of 'being a translation' and can be related to translation quality and translational norms. We set out to reveal tendencies in translational behavior at different competence levels by describing frequency distributions of a single functional type of discourse markers (connectives) in English to Russian translation of mass-media texts. To this end, we compared data from a parallel translational learner corpus and a corpus of professional translations to customized selections from English and Russian national corpora. Using independent predefined lists of connectives for each language, we explored cross-linguistic differences and their influence over students' and professional translations. We hypothesized three possible tendencies: translations follow source language pattern (interference); translations follow target language pattern (normalization) or translations demonstrate independent idiosyncratic (over)use of connectives (explicitation). The observations were done with regard to the overall frequencies of the list items, their semantic groups and individual frequencies. Beside obvious pedagogical implications, the findings are useful in understanding cognitive processes behind translation, applicable in detecting translational errors for a given language pair and assessing textual quality of translations.

**Keywords:** translation universals, translational learner corpora, parallel corpora, connectives, cohesion, interference, frequency distribution, translation quality assessment

---

<sup>1</sup> This work was partially supported by the Russian Foundation for Basic Research (Grant 17-06-00107).

# ЧАСТОТНОЕ РАСПРЕДЕЛЕНИЕ КОННЕКТОРОВ В ПЕРЕВОДЕ С АНГЛИЙСКОГО ЯЗЫКА НА РУССКИЙ И ЗАКОНОМЕРНОСТИ ЯЗЫКОВОГО ВЫБОРА ПЕРЕВОДЧИКОВ

**Куниловская М. А.** (mkunilovskaya@gmail.com)

Тюменский государственный университет, Тюмень, Россия

## 1. Theoretical framework and motivation

Levels of explicit text connectedness have been on linguistic research agenda in computational and corpus linguistics for many years. Indeed, connectives most readily lend themselves for objective comparative and contrastive analysis, because they are essentially functional and often structurally optional items. Translators are free to choose between explicit and implicit renditions, and their lexical choices are indicative of the adopted global textual and translational strategy.

Most research in this vein resorts to the theory of translation universals (Baker 1993) for explanation of differences found between translations and original texts in the target language (non-translations). According to Baker's definition, translation universals are *'features which typically occur in translated texts rather than original utterances and which are not the result of interference from specific linguistic systems'* (Baker 1993). Since then it was demonstrated that statistically significant distinctions of translations as a subsystem of target language (TL) could also be non-universal, that is, typical for a given direction of translation in a given language pair. Hence, according to Castagnoli (2009) the analysis of connectives use in sources and targets *'cannot be severed from considerations about differences in their use across languages'* (p. 67).

Statistical findings on the use of connectives in translations (Chen 2006; Cartoni et al. 2011) usually corroborate the *explicitation hypothesis*, according to which translations tend to be more explicit than their source texts (ST) or comparable non-translations (Blum-Kulka 1986).

Explicitation is one way in which translations establish themselves as 'third code'. Other possibilities include combining features of source language (SL) and target language (TL). Translations, usually simultaneously, bear features 1) inherited from SL, and 2) features resulting from attempts to produce a text acceptable under TL norms. The textual features arising from the first process are referred to as *interference and transfer*; in the second case they speak of *normalization and convergence (levelling-out)*. The former two refer to the tendency to transfer phenomena pertaining to the make-up of the source text to the target text (Toury 1995). For the purposes of this study, normalization is defined in accordance with Baker (1996) as the *'tendency to exaggerate features of the target language and to conform to its typical patterns'* (p. 183). Convergence is defined as a *'relatively higher level of homogeneity of translated texts with regard to their own scores on given measures of universal features'* (Laviosa 2002, p. 72).



While translations at large can be different from non-translations according to a chosen parameter of variation, learner translations represent its own variety of translationese.

## 2. Aim and research questions

This research aims to describe quantitative differences between learner and professional translations of mass-media texts from English into Russian as to the use of connectives and establish translational tendencies at work at different levels of competence. We approach cohesive specificity of translations from cross-linguistic perspective by supplementing intra-linguistic comparisons with SL data.

The system of research questions can be laid down as follows:

1. What is the relation between the two types of translation into Russian (learners and professionals) in terms of explicit text cohesion?
2. Do translations and non-translations feature differences with regard to the frequency distribution of individual connectives and their semantic groups?
3. Is there a cross-linguistic difference in the use of connectives in the reference corpora? How can patterns of overuse and underuse be explained from cross-linguistic perspective (interference, convergence)?

The rest of the paper is structured as follows. Section 3 delineates the group of lexical items focused in this research as an operator of linguistic differences. Section 4 describes our corpus resources and methodology. Section 5 has the results presented for all relevant comparisons in terms of overall frequency of connectives, its distribution across semantic categories and individual frequencies. The final section brings together all findings and draws conclusions.

## 3. Connectives: delineation of the class

### 3.1. Definition

For the purposes of this research we define connectives as a functional class of lexical items, which indicate the type of relation that holds between parts of the text.

We see connectives as a subclass of a broader category of discourse markers (DM) that specializes in signalling text structure (as opposed to conveying speaker's attitudes or arranging speaker-hearer interaction). They guide and facilitate recipient's interpretation of the text structure and help inferring textual relations by overtly naming the ones intended by the author.

### 3.2. Preliminary search lists

To proceed we need to arrive at a motivated and practicable list of connectives. We limit research to structurally independent, truly optional items and did not consider conjunctions and syntactically undetachable cohesive elements (*with regard to*, *относительно того, что*).

To produce search lists, we proceeded from the inventories of connectives given in grammar references and academic papers as either connectives, DM, linking adverbials, particles or parenthesis responsible for managing text structure. For English we relied upon Biber et al. (1999), Fraser (2006), Liu (2008), Meyer and Webber (2013), Swan (1992). The Russian list was derived from Berson et al. (1984), Kogut (2014), Novikova (2008), Priyatkina (2015), Russian Grammar (Shvedova 1980).

To filter these lists we introduced a number of selection criteria to ensure uniformity of the lists and comparability of results for English and Russian. The lists were reviewed and refined iteratively against corpora until there were no further changes to them.

### 3.3. Extraction experiments

To improve precision of extraction and operability of data we lemmatized and tagged all items from the search lists through *redefining their default POS annotations*. Standard POS-tagging was performed by *TreeTagger* with the supplied English model (Schmid 1995). For Russian we used the model based on MULTEXT-East tagset (Sharoff et al. 2008) and trained by Sharoff<sup>2</sup>.

In this approach to disambiguate different uses of the same forms we coded positional and punctuational occurrences of the items we want to extract. The resulting lists of connectives include 105 items for English and 95 items for Russian.

### 3.4. Semantic classification

Our approach to semantic classification is borrowed from Mayer and Webber (2013) and consists in assigning polyfunctional items to the majority class. In the alternative approach (Castagnoli 2009; Liu 2008) researchers resorted to manual disambiguation and included polyfunctional items in several groups. Schiffrin argues that it is typical for DM to combine meanings of different semantic groups, but they still have some dominant functionality and can be assigned to a particular group (Schiffrin 1987).

In terms of granularity we wanted reasonably few classes and roughly followed the taxonomy suggested by Fraser (2006) and consistent with Biber's classification (1999) as well as the systems suggested by Russian linguists (Berson 1984; Kogut 2014; Inkova 2015). Most researchers agree on the the four broad meanings of elements in cohesive function:

1. elaborative (*in other words, which is to say, а именно, следует отметить*);
2. inferential/causative (*it follows that, hence, в результате этого, соответственно*);
3. contrastive (*instead, by way of contrast|comparison, всё же, НО на самом деле*);
4. sequential/transitional (*lastly, briefly, on the whole, к слову, напомним, во-вторых*)

Following suggestions in literature, dictionary definitions (see references) and analysis of contexts in our samples we arrived at the classification given in Appendix 1<sup>3</sup>.

---

<sup>2</sup> <http://corpus.leeds.ac.uk/mocky/msd-ru.html>

<sup>3</sup> [http://www.rus-ltc.org/download/Appendix\\_Semgroups.pdf](http://www.rus-ltc.org/download/Appendix_Semgroups.pdf)

#### 4. Corpus resources and implications for statistical analysis

This research is based on three pairs of corpora for English-Russian language pair:

- 1) learner translations and their sources,
- 2) professional translations and their sources and
- 3) two reference corpora for the respective languages.

We took efforts to make these corpora comparable by controlling for register at sampling. Both parallel collections are unidirectional and contain translations from English into Russian.

Learner translations ('learners') are a sample of Russian Learner Translator Corpus<sup>4</sup>, limited to essays, informational and educational texts. Their sources ('leST') are published in various electronic mass media between 2001 and 2016. Translations are done by senior students majoring in translation in 5 Russian universities.

Professional translations ('pro') into Russian are collected from a range of established electronic media, such as *Nezavisimaya Gazeta*, *InoSMI* and *Polit.RU* (8 sources carrying translator's name or endorsement of the translation by editorial board). The source texts ('proST') for these translations were published in 2008–2016 by established English media.

The English reference corpus is a sample of the British National Corpus<sup>5</sup> (BNC). We collected all texts tagged as *written*, *newspaper* in David Lee's scheme regardless of the type of the paper and topical domain. All texts in our sample were created between 1984–1992.

The Russian reference corpus is a sample from the Russian National Corpus (RNC)<sup>6</sup>, which fits the following sampling frame: *article*, intended for large adult non-specialist readership, created after 2003 and marked as neutral of style. Table 1. has the basic statistics for the resulting datasets after pre-processing and linguistic annotation.

**Table 1.** Basic parameters of the research datasets

	leST	learners	proST	pro	BNC	RNC
Size (tokens)	214,234	197,583	359,252	330,899	5,735,951	3,130,183
No. of texts	208	208	200	200	—	1,562
No. of sentences	9,739	9,904	14,977	15,648	246,029	166,383

Preliminary statistical tests showed that we have reasonably homogeneous samples with regard to sentence length, an important, but admittedly crude, indicator of register. Furthermore, they suggest that we can reliably normalize frequencies over number of sentences rather than tokens. This approach is considered more reasonable for text-level features (Bolton 2002) and, from the cross-linguistic perspective, compensates for the more analytical ('wordier') nature of English in comparison to Russian<sup>7</sup>.

<sup>4</sup> <http://www.rus-ltc.org/>

<sup>5</sup> <http://www.natcorp.ox.ac.uk/>

<sup>6</sup> <http://www.ruscorpora.ru/>

<sup>7</sup> Normalization over all text tokens would unfairly push relative frequencies down for English compared to Russian

## 5. Frequency distribution of connectives

### 5.1. Overall frequencies across texts: levels of cohesive explicitness of texts

In frequency analysis of connectives, we used two approaches. First, we characterized and compared corpora based on the overall frequencies of the list items per text. Second, we collect individual frequencies for each item and their semantic groups across each corpus to locate overused and underused connectives that can be targeted in parallel data.

In Table 2, we report standard measures of central tendency for the relative frequencies along with the raw counts. We normalized cumulative per-document frequencies by the number of sentences in each text to the base of 100 and then averaged over all texts.

**Table 2.** Descriptive statistics for overall connectives frequency across all texts by corpora

Corpora	leST	learners	proST	pro	BNC	RNC
No. of CONN	976	1,494	1,306	2,115	12,439	22,821
Mean	9.76	16.06	9.22	14.17	6.72	13.65
Median	8.33	14.29	8.08	12.72	6.52	12.50
<i>SD</i>	9.00	11.60	6.22	7.93	3.13	7.78

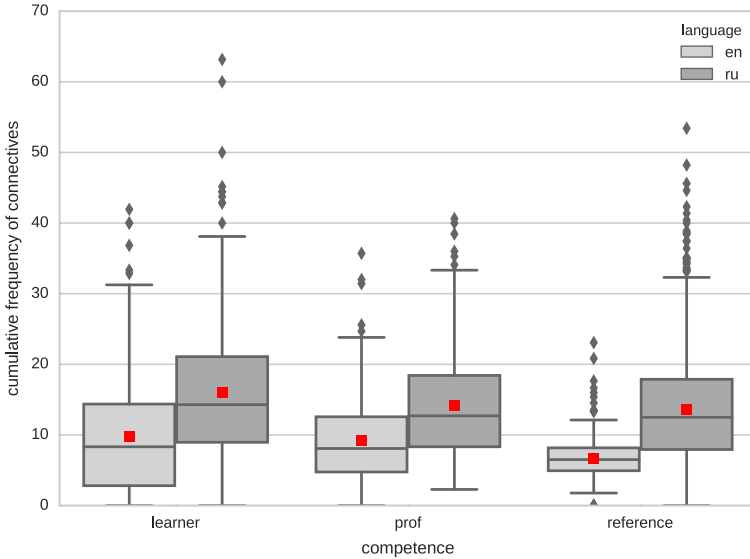
Comparisons of frequency distributions and dispersions of connectives are visualized in a boxplot (Fig. 1). The boxes in the plot represent frequencies in the middle 50% of observations (=texts).

The boxplot reveals a contrast between BNC and the other corpora both in terms of average frequency of connectives and variation in this parameter: BNC seems to have less connectives that are spread more evenly around the mean.

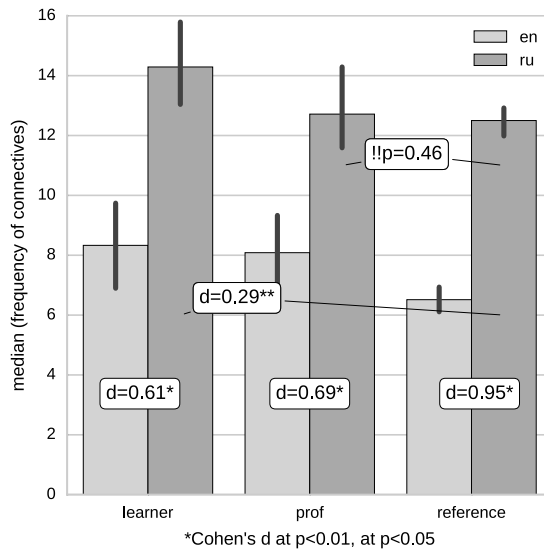
In Fig. 1 professional translations look very similar to the reference Russian corpus, while learner translations seem to be the most explicitly cohesive.

For non-parametric data like ours, median is a better measure of central tendency. In Fig. 2, we have a bar chart which compares medians of the same corpora pairs. Again, all Russian corpora demonstrate greater cohesive explicitness, while BNC is unexpectedly contrasted to the other two English corpora. Fig. 2 also carries results of Cohen's *d* effect size test and associated two-tailed Wilcoxon rank tests, paired or unpaired depending on the possibility to match texts. The differences between both translational collections as compared to respective sources as well as between the two reference corpora were statistically significant at  $p < 0.01$ . The effect size for the first two pairs was 0.61 and 0.69. It is above medium, if we accept the usual cut-offs (Cohen 1977). The difference between the reference corpora amounts to 0.95, indicating sharp contrast between them.

Interestingly, there was a significant difference between all English sources in this research (created 2001–2016) and the BNC sample (created 1984–1992). A two-tailed Wilcoxon rank sum test returned  $p\text{-value}=5.946e\text{-}05$ , Cohen's  $d=0.497$ . We can tentatively suggest that these differences should be due to diachronic distance of more than 15 years between these collections. We can explain lack of variation in BNC by lack of conventional internal structure in the corpus.



**Fig 1.** Frequency distributions of connectives across texts at different competence levels



**Fig 2.** Medians for overall connectives frequencies and inferential statistics

Learner translations are different from the reference data (at 95% confidence level) but the effect size of the differences was small ( $d=0.29$ ). We do not have enough evidence that professional translations were distinct from the reference corpus.

## 5.2. Overused and underused connectives

At the next stage we looked at distributions of individual connectives and their semantic types. Unlike the previous experiment, where we assessed the levels of explicit cohesiveness of texts, we focused on the qualitative make up of connectives contributing to this textual feature.

To obtain lists of over- and underused items we ran a two-tailed Wilcoxon rank sum test on vectors of normalized frequencies for each of 95 connectives on our search list obtained from every text in both translational corpora and from the reference corpus. The lists of connectives with the respective frequency statistics and tests results were sorted by 1) the sign of difference between normalized frequencies with respect to the translational corpora ('+' or '-') and by 2) test result—p-value for Wilcoxon ranksum test, cutting off at the significance level of 95%. Table 3 has the connectives with the largest discrepancy in frequencies between translations and RNC.

**Table 3.** Translationally distinctive connectives (at  $p < 0.05$ )

	<b>learners</b>	<b>professional</b>
<b>Overused</b>	например, <i>вместо этого/того</i> , при этом, <i>однако</i> , затем	<i>вместо этого/того</i> , также, <i>однако</i> , в конце концов
<b>Total</b>	5	4
<b>Underused</b>	поэтому, <i>причем</i> , <i>так</i> , <i>то есть</i> , <i>кстати</i> , <i>впрочем</i> , <i>ведь</i> , <i>кроме того/этого</i> , <i>прежде всего</i> , <i>в частности</i> , наконец, значит, вместе с тем/этим, <i>в связи с тем/этим</i> , во-первых	<i>Так</i> , <i>прежде всего</i> , <i>кроме того/этого</i> , <i>кстати</i> , <i>причем</i> , например, <i>то есть</i> , <i>в частности</i> , <i>ведь</i> , <i>в связи с тем/этим</i> , <i>впрочем</i>
<b>Total</b>	15	10

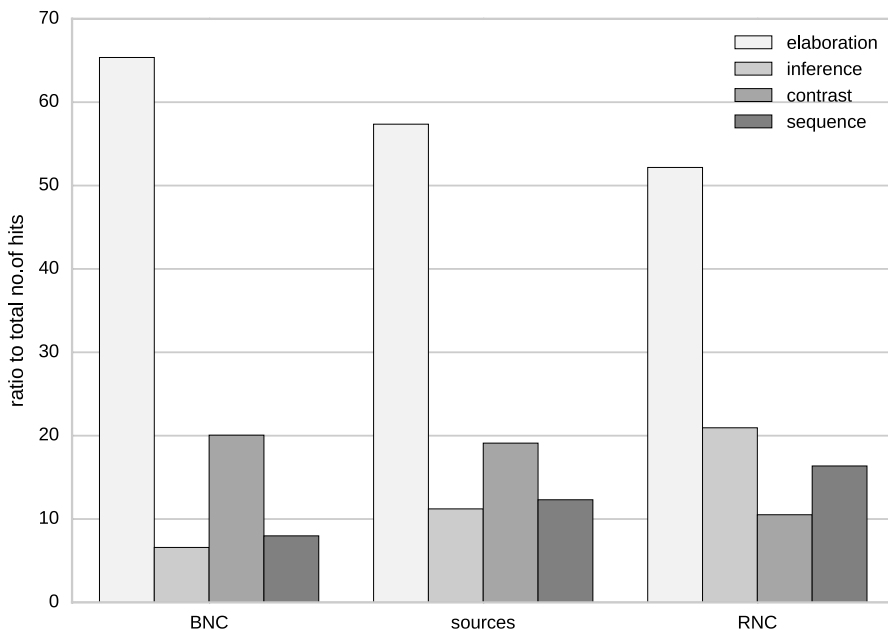
Most of the translationally distinct connectives are shared by both lists (italicized in Table 3) picking up on common features of English to Russian translationese.

All underused items revealed in this research came from top frequency quartile of the reference corpus list of connectives, while overused items included connectives from the third frequency band (*вместо того/этого*). It means that out-of-English translation into Russian primed the use of less typical connectives, reflecting commonalities of translators' language use, and at the same time it de-prioratizes other markers, narrowing the variety of expression. This is in line with findings by Cartoni et al. (2011) who showed that different SL influenced the distribution of connectives in French as a TL and that there was no 'unique translationese' (p.85).

## 5.3. Semantic groups

A closer look at the semantic make up of overused connectives reveals that 2 out of 4 are contrastive, while the most numerous underused connectives signal inference

and elaboration. In an attempt to explain translationally distinctive frequencies by cross-linguistic differences in preferred types of explicit textual relations, we calculated the ratio of each semantic group in the reference corpora. The relation of English and Russian non-translated corpora are visualized in Fig. 3.



**Fig. 3.** Ratio of semantic types of connectives in the reference corpora

In our data contrastive connectives made up 10.5% in RNC and twice as much in BNC (20.1%) or in sources combined (19.1%) (see Table 4). Markers of inference were noticeably more frequent in Russian (20.9%) than in English (BNC: 6.6%, sources: 11.2%). Both languages had elaboration as the most frequent type (52.2% and 65.4% for Russian and English respectively). Note that these frequencies were not immediately related to the size of the search lists for each group in each language.

We can hypothesize that the overuse of contrastive connectives in translations was a case of English text-structuring strategies ‘shining through’. Underused items in learner translations belonged to the semantic types that were less marked in English (inference and sequential), or that were most varied in expression in the TL (elaboration). This tendency can be indicative of simplification and convergence, which are associated with the reduction of lexical variety and overexploitation of most typical TL patterns.

Contrastive inspection of overall frequencies with regard to the ratio of semantic groups were indicative of the shift towards adding markers of inference and sequence, while contrastive relations lost some of their relative prominence in SL. These shifts are marked with solid upward and downward arrows ▲▼ in Table 4.

**Table 4.** Aggregated frequencies for semantic groups of connectives across the corpora

Semantic type	No. of items		leST	learners	proST	pro	BNC	RNC
	EN	RU						
Elaboration	35	41	550 56.4%	756 50.6%▼	759 56.4%	1,037 49.0%▼	8,129 65.4%	11,906 52.2%
Inference	18	15	124 12.7%	221 14.8%▲↓	132 12.7%	313 14.8%▲↓	820 6.6%	4,780 20.9%
Contrast	21	15	198 20.3%	271 18.1%▼↑	238 20.3%	399 18.9%▼↑	2495 20.1%	2,400 10.5%
Sequential	30	24	104 10.7%	246 16.5%▲≈	177 10.7%	366 17.3%▲≈	993 8.0%	3,735 16.4%
Total (100%)	105	95	976	1,494	1,306	2,115	12,437	22,821

As compared with the Russian reference corpus translations demonstrated an increase in contrastive markers and a decrease in inferential markers (shown with plain arrows ↓↑ in Table 4). Below there are some translations that exemplify the phenomena behind higher frequency of connectives in translations. In (1) the translator explicates the underlying inferential discourse relation. Translation (2) offers reinterpretation of the situation, which leads to introduction of markers absent in the source.

- |  |   |
|--|---|
| (1) <i>The Nobels are a great way to get people interested in science, they'll say, <b>and</b> it's good that we have them.</i>  | <i>Говорят, что Нобелевская премия действительно побуждает людей интересоваться наукой, <b>и поэтому</b> здорово, что она есть.</i>   |
| (2) <i>Anyone who, as a child, possessed a Junior Conjuror's Set will have learned two simple lessons about magic. Audiences long to be deceived and are invariably disappointed when they are told how the trick is done.</i> | <i>Каждый у кого был набор юного фокусника <b>в конечном итоге</b> поняли два основных правила магии — не бойтесь обманывать зрителя, <b>ведь</b> они ожидают увидеть чудо, и ни в коем случае не рассказывайте секрет фокуса — их это сильно разочарует.</i> |

#### 5.4. Parallel approach

To disentangle discrepancies in the observed contrastive and comparative frequencies we focused on the overused items from Table 3 and their tentative English counterparts, as well as on hypothetically added inferential and sequential markers. We limited bitext analysis to the translationally distinctive items from each group.

The increase in contrastive markers in comparison with reference is generated largely by three items (*тем не менее, однако, вместо этого*). Most of *однако* occurrences in our data were prompted by sentence-initial *but* (42%) and *however* (36%), while the rest corresponded to other English contrastive markers (*nevertheless, yet,*



*instead, though*) or otherwise explicitly expressed contrast. In examples like (3) the use of *однако* compounds the expressed contrastive effect. About 9% of occurrences explicate an implied contrast (like in (4)).

- |     |   |   |
|-----|---|---|
| (3) | <i>The scene can be interpreted otherwise.</i>                                    | <i>Однако эту сцену можно понять по-другому.</i>  |
| (4) | <i>On the face of it Asia stacks up not at all well against this achievement.</i> | <i>Однако Азиатские страны, кажется, не стремятся соответствовать в этом отношении европейским.</i> |

*Тем не менее* is widely used as a generic rendition for *though, yet, nonetheless, however, although, still, all the same*. English conjunctions *rather (than)* and *instead (of)* are typically mapped into *вместо этого*. A noticeable opposite tendency in this group is represented by comparative frequencies of *впрочем*. While in the reference data it is the most frequent contrastive connective, it is atypical for translations.

The group of inferential markers has more items that are less used in translations than in non-translated Russian. Notably underused items are *поэтому, так, следовательно, значит* and *потому*. The only spike reaching above the reference corpus line is в *результате*.

Only 4 out of 15 source sentences, which got в *результате* in translation, contained the predictable prompt *as a result*. The respective logical relation between propositions was expressed in a variety of ways (*consequently, we'll end up, it led to, hence, the result is*), which converged in translation.

Among sequential connectives, *в общем, в конце концов* and *итак* had higher frequencies in both types of translation. The group of elaborative connectives in translations showed a lot of similarity to non-translations, except for *также*. It was twice more frequent in learner translations (4.17 items per hundred sentences against 2.35 in non-translations, with professional translations counts in between (3.27)).

## 6. Discussion and conclusions

In this research we tested whether connectives were a reliable feature of Russian out-of-English translationese and whether it can be used to distinguish between translations produced by people at different competence levels. We found statistically significant contrasts in overall frequencies between learner translations and Russian reference corpus, while professional translators seem to be able to adapt to TL norm better.

We also revealed cross-linguistic differences in levels of explicit text cohesion, with English being more implicit, which is also the case in other language pairs with English (Castagnoli, 2009, p. 69). In our experiments languages were dissimilar in the preferred type of textual relations marked in texts, given our lists. The English sample had more markers of contrast, while the Russian texts in our research relied more heavily on signals of inferential relations.

The analysis of the individual frequencies revealed a few translationally distinctive connectives. Parallel text analysis along with information on the frequencies of semantic groups, suggested that the overuse was due to higher frequencies of the English triggers for these items. With some connectives there was a tendency

to be used as a generic counterpart for a range of more specific expressions. We have also detected a mild trend towards explicitation. Thus, the overuse was a composite effect of interference, convergence and explicitation.

## References

1. *Baker M.* (1993), *Corpus Linguistics and Translation Studies: Implications and Applications, Text and Technology: In honour of John Sinclair*, Amsterdam: J. Benjamins, pp. 232–250.
2. *Baker M.* (1996), *Corpus-based translation studies: The challenges that lie ahead. Terminology, LSP and Translation*, J. Sager (Ed.), Benjamins Translation Library, Vol. 18, pp. 175–186). DOI: <http://dx.doi.org/10.1075/btl.18.17bak>
3. *Berzon V. E., Blehman M. S., Piotrovskij R. G.* (1984), Relations, units and unities of supersentential language level [Svjazi, edinicy i edinstva sverhfrazovogo urovnja jazyka], *Tartu State University Bulletin [Uchenye zapiski Tartuskogo gosudarstvennogo universiteta]*, *Papers on linguistics [Trudy po lingvostatistike]*, pp. 16–34.
4. *Blum-Kulka Sh.* (1986), Shifts of cohesion and coherence in translation. *Interlingual and intercultural communication: Discourse and cognition in translation and second language acquisition studies*, pp. 17–35.
5. *Bolton K., Nelson G. & Hung J.* (2002), A corpus-based study of connectors in student writing: Research from the International Corpus of English in Hong Kong (ICE-HK), *International Journal of Corpus Linguistics*, 7 (2), pp. 165–182.
6. *Cartoni B., Zufferey S., Meyer T., Popescu-Belis, A.* (2011), How comparable are parallel corpora? Measuring the distribution of general vocabulary and connectives, *Proceedings of the 4th Workshop on Building and Using Comparable Corpora: Comparable Corpora and the Web*, Association for Computational Linguistics, pp. 78–86.
7. *Castagnoli S.* (2009), *Regularities and variations in learner translations: a corpus-based study of conjunctive explicitation* (Doctoral thesis, University of Pisa, Italy), available at: [http://etd.adm.unipi.it/theses/available/etd-04252009-135411/unrestricted/Castagnoli\\_thesis\\_body.PDF](http://etd.adm.unipi.it/theses/available/etd-04252009-135411/unrestricted/Castagnoli_thesis_body.PDF)
8. *Chen, W.* (2006), *Explication Through the Use of Connectives in Translated Chinese: A Corpus-based Study*, PhD thesis, University of Manchester
9. *Cohen, J.* (1977), *Statistical power analysis for the behavioural sciences* (Rev. ed.), New York: Academic Press.
10. *Fraser B.* (2006), *Towards a theory of discourse markers, Approaches to discourse particles*, 1, pp. 189–204.
11. *Inkova O.* (2015), On the delimitation of functional class of connectors [O graničah funkcional'nogo klasa konektorov], *Trends in Slavic Studies*, pp. 657–668.
12. *Kogut S. V.* (2014), Discourse markers in Russian and German academic papers on geology [Diskursivnye markery v russkojazyčnyh i nemeckojazyčnyh geologičeskikh naučnyh stat'jah], *Tomsk State University Herald [Vestnik Tomskogo gosudarstvennogo universiteta]*, No 380, pp. 18–23.

13. *Laviosa S.* (2002), *Corpus-based translation studies: theory, findings, applications*, Rodopi, Vol. 17, Amsterdam-New York.
14. *Liu D.* (2008), Linking adverbials: An across-register corpus study and its implications, *International Journal of Corpus Linguistics*, 13(4), pp. 491–518.
15. *Meyer T., Webber B.* (2013), Implicitation of discourse connectives in (machine) translation, *Proceedings of the 1st DiscoMT Workshop at ACL 2013*, pp. 19–26.
16. *Novikova N. I.* (2008), Connectives as cohesive devices in an asyndetic composite sentence [Konnektory kak svjazujushhie sredstva v bessojuznom slozhnom predlozhenii], *Herald of the Voronezh State Architecture University [Nauchnyj Vestnik Voronezhskij gosudarstvennyj arhitekturno-stroitel'nyj universitet]*, *Advanced Linguistic and Pedagogical Research Series [Ser.: Sovremennye lingvističeskie i metodiko-didaktičeskie issledovanija]*, pp. 92–100.
17. *Priyatkina A. F., Starodumova E. A.* (2015), Text clips in A Dictionary of Functional Words [Tekstovye skrepy v «Slovarě sluzhebnyh slov»], *Siberian Philological Journal [Sibirskij filologičeskij žurnal]*, Novosibirsk, 2, pp. 134–141.
18. *Schiffrin D.* (1987), *Discourse markers*, Cambridge: Cambridge University Press.
19. *Schmid H.* (2008), *Treetagger a language independent part-of-speech tagger*, Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.
20. *Sharoff S., Kopotev M., Erjavec T., Feldman A., & Divjak D.* (2008), *Designing and Evaluating a Russian. LREC 2008*, pp. 279–285.
21. *Toury G.* (1995), *Descriptive Translation Studies — and beyond*, Amsterdam, New York, John Benjamins Publishing Company.

### List of dictionaries and grammar references

1. *Apresjan V. Ju., Pekelis O. E.* (2011), *Russian Corpus-based grammar [Russkaja korpusnaja grammatika]*, available at: <http://rusgram.ru/%D1%81%D0%BE%D1%8E%D0%B7#111>
2. *Biber D., Johansson S., Leech G., Conrad S., Finegan E., Quirk R.* (1999), *Longman grammar of spoken and written English*, MIT Press.
3. *Burtseva V. V.* (2010), *A Russian Dictionary of Adverbs and Function Words [Slovar' narečij i sluzhebnyh slov russkogo jazyka]*, *Russian Language-Media [Russkij jazyk-Media]*.
4. *Efremova T. F.* (2000), *New Morphological Russian Dictionary [Novyj slovar' russkogo jazyka. Tolkovo-slovoobrazovatel'nyj]*, *Russian Language [Russkij jazyk]*, available at: <http://dic.academic.ru/dic.nsf/efremova/>
5. *Macmillian Dictionaries*, available at: <http://www.macmillandictionary.com>
6. *Oxford Dictionaries*, available at: <https://en.oxforddictionaries.com>
7. *Priyatkina A. F., Starodumova E. A., Sergeeva G. N.* (2001), *A Russian Dictionary of Functional Words [Slovar' sluzhebnyh slov russkogo jazyka]*.
8. *Rogozhnikova, R. P.* (2003), *A Dictionary of word equivalents [Tolkovyj slovar' sochetanij, jekvivalentnyh slovu]*, Moscow, Astrel.
9. *Russian Grammar (1980)*, Ed. N. Ju. Shvedova, Moscow, Science [Nauka].
10. *Swan M.* (1992), *Practical English usage*, New-York.

# ТИПЫ КОНСТРУКЦИЙ С АДВЕРБИАЛАМИ<sup>1</sup>

**Кустова Г. И.** (galinak03@gmail.com)

Институт русского языка им. В. В. Виноградова РАН;  
Московский педагогический государственный университет;  
Москва, Россия

Адвербиалы рассматриваются как свернутые (редуцированные) предикации, которые являются промежуточной фазой процесса грамматикализации (образования производных предлогов от номинализаций). Выделяются два типа адвербиалов:

- (1) **Включенные** адвербиалы дополняют основную пропозицию, конкретизируют какие-то ее аспекты. Они могут относиться к двум подтипам: (а) **корреферентные** (субъект адвербиала совпадает с субъектом главной предикации) — соответствуют деепричастиям (причем только таким, которые в грамматике называются второстепенными сказуемыми, т. е. не-обстоятельственным): *Стоял на остановке в ожидании автобуса*, ср.: *ожидая автобуса*; (б) **некорреферентные** (субъект адвербиала не совпадает с субъектом главной предикации) — соответствуют пассивным причастиям: *Находился в гостинице под охраной милиции*, ср.: *охраняемый милицией*.
- (2) **Обстоятельственные** адвербиалы (*в защиту кого* — цель; *под давлением / под действием чего* — причина; *при обыске* — время; *при желании* — условие) соотносятся с обстоятельными деепричастными оборотами или с обстоятельными придаточными.

В статье рассматриваются семантические и синтаксические свойства адвербиалов, типы конструкций, в которых они употребляются. Основное внимание уделяется включенным адвербиалам, которые не имеют отдельного статуса в грамматиках и вообще не выделяются как особая группа.

**Ключевые слова:** адвербиал, номинализация, грамматикализация, валентность

## THE TYPES OF ADVERBIALS

**Kustova G. I.** (galinak03@gmail.com)

Vinogradov Russian Language Institute of the Russian Academy  
of Sciences; Moscow State Pedagogical University;  
Moscow, Russia

---

<sup>1</sup> Исследование выполнено в рамках проекта РФФ № 16-18-02003, осуществляемого в МПГУ. Примеры извлечены из Национального корпуса русского языка (www.ruscorgpora.ru).

The article discusses adverbial expressions (adverbials), cf. *v ozhidanii* ('in expectation'), *pod okhranoj* ('guarded by smb'), *po priglaseniju* ('by invitation'), *v blagodarnost'* ('being thankful for') as one of the stages of the process of verbal nouns grammaticalization — namely, as a stage of their transformation into derivative prepositions.

An adverbial is regarded as a reduced predications related to the main predication. There are two types of adverbials: included adverbials and circumstantial ones. Circumstantial adverbials correlate with adverbial subordinate clauses. The article discusses semantic-syntactic properties of included adverbials.

Among included adverbials two types can be distinguished:

- 1) coreferential (having the same subject) adverbials that correlate with an adverbial participle: the subject of the adverbial is identical to the subject of the main predication; and the subject (agent) of the adverbial is not expressed and can't be expressed syntactically, cf. (*Passazhiry khodili po Perronu v ozhidanii poezda* 'Passengers were walking along the platform **waiting for the train to arrive**');
- 2) non-coreferential (having different subjects) adverbials that correlate with a passive participle: the subject of the adverbial is not identical to the subject of the main predication; and the subject (agent) of the adverbial is necessarily expressed (*Prishel v sopovozhdenii advokata* 'came **accompanied by** a lawyer'). In contrast to such adverbials, nominalizations can express all their arguments.

**Key words:** adverbial, nominalization, grammaticalization, valency

## Адвербиалы в процессе грамматикализации

Материалом настоящей работы являются выражения, образованные на базе номинализаций и имеющие в предложении обстоятельственное значение: *с согласия кого; под присмотром кого; в подтверждение чего; в сопровождении кого; при участии кого; при проверке кого-чего; без ущерба для кого-чего* и т. п.

Подобные выражения не обязательно восходят к отглагольным существительным (ср. *без ущерба для*), — важно, что они обозначают ситуацию и имеют набор валентностей, соответствующих партиципнтам этой ситуации. Мы будем называть такие выражения адвербиалами. Термин адвербиал, или адвербиальное выражение, используется в литературе (ср. [Филипенко 2003]), особенно англоязычной (ср. ([Grennbaum 1969]; [Heny 1983]; [Haspelmath 1995]), для обозначения различных обстоятельственных единиц (в том числе наречий и деепричастий). Но мы будем использовать его в более узком смысле, который он имеет в работах [Богуславский 1985; 1996; 2008]. В этих работах адвербиалами, или адвербиальными дериватами (термин из концептуального аппарата модели «Смысл ↔ Текст», см. [Мельчук 1974 / 1999]), называются выражения типа *в доказательство, в награду, в ответ на, в подарок, по привычке, по совету, под предлогом* и под., которые, в рамках модели «Смысл ↔ Текст», включаются в словарную статью глагола (или существительного, ср. *под предлогом* — *предлог*) в качестве его синтаксических дериватов. Работа [Богуславский 2008] посвящена тому, как адвербиалы выражают валентности исходного глагола.

В настоящей работе адвербиалы будут рассматриваться под другим углом.

Мы будем рассматривать адвербиалы как одну из ступеней процесса грамматикализации, а именно — превращения исходной номинализации в производный предлог. Адвербиал — «срединная» фаза этого процесса, он занимает промежуточное положение между «обычной» номинализацией и производным предлогом. Поэтому прежде всего необходимо обсудить, чем отличается этот «промежуточный» элемент от исходного и конечного.

Сначала сравним адвербиал и «обычную» номинализацию, от которой он образован и которую мы будем называть управляемая номинализация (см. [Кустова 2013]).

Управляемые номинализации, см. ниже примеры (а), являются в предложении подлежащими и дополнениями и входят в парадигму исходного отглагольного существительного в качестве образующих ее форм. Адвербиалы — это обороты (обычно предложно-падежные), которые являются адъюнктами главной предикации (связаны с ней атрибутивными отношениями), ср. примеры (б):

- (1) а) *Начал с согласия, а закончил отказом*  
б) *Дети участвуют в мероприятии с согласия родителей*
- (2) а) *Не вижу ничего странного в сопровождении нашим сотрудником такой важной персоны*  
б) *Пришел на допрос в сопровождении адвоката*
- (3) а) *Вы разобрались с учетом готовой продукции?*  
б) *Водители Ларионов и Гутько обещали сэкономить тонно-километров вдвое и бой тары произвести с учетом интересов* (М. М. Жванецкий).

Для нас важно подчеркнуть не производность, а, напротив, отдельность, отделённость адвербиалов от управляемых номинализаций. Поскольку синхрония — это срез диахронии, лингвистическое описание застаёт адвербиалы в определенной фазе эволюции от знаменательной единицы к служебной, на определенной ступени грамматикализации, когда уже появился набор новых свойств, но еще не полностью утрачены старые. Адвербиалы в этом смысле не уникальны, а, напротив, вполне типичны, ср. наречия типа *весной, вечером*, которые еще присоединяют прилагательные — *ранней весной*; ср. также так наз. свободные релятивы типа *кто угодно, что хочешь* и под. (см., в частности, [Лютикова 2015], [Кустова 2016]). То, что адвербиалы не принято рассматривать как особое явление, во многом объясняется идеологией традиционной лексикографии. Если бы большинство существующих толковых словарей не основывалось на «номинативном» принципе (когда слово представлено начальной формой — номинативом или инфинитивом), то адвербиалы должны были бы фиксироваться в словарях как отдельные единицы. Разумеется, они семантически связаны с производящей номинализацией, но уже обособились, «оторвались» от ее парадигмы, — точно так же, как сама номинализация не входит в парадигму глагола, хотя связана с ним семантически и сохраняет его валентности (выражаемые иногда другими, а иногда теми же средствами). Кроме того, адвербиалы можно рассматривать как конструкции (в идеологии Грамматики

конструкций, ср. [Croft, Cruse 2004]; [Goldberg 1995]; [Goldberg 2006], такой подход вполне допустим), и тогда окажется, что адвербиалы, как и другие конструкции, характеризуются той или иной степенью некомпозициональности значения и определенными запретами и ограничениями (подробнее см. ниже).

Соотношение адвербиалов как промежуточной ступени грамматикализации и производных предлогов как конечной фазы этого процесса не столь прозрачно, как соотношение адвербиалов и номинализаций.

Производные предлоги (далее — ПП) — разнородная группа, значительную часть которой составляют отыменные предлоги. Основным поставщиком отыменных ПП является параметрическая лексика — *причина, время, случай, цель, условия* и под. (в основном, это не параметры предметов, а параметры ситуаций), ср. *по причине, во время, в случае, по случаю, с целью* и под. Другой источник ПП — конструкции с номинализациями (и другими существительными с ситуативной семантикой), т. е. адвербиалы, ср. *в согласии с, в сопровождении, с помощью, в содружестве с, в зависимости от, в соответствии с, за исключением, в интересах* и под.

Разные авторы ([Русская грамматика 1980], [Бурцева 2005], [Лепнев 2010], [Черкасова 1976], [Шереметьева 2008]) приводят разные списки ПП. При этом все авторы отмечают, что остаются неясными и частеречный статус ПП, и критерии разграничения адвербиалов и ПП.

Некоторые адвербиалы окончательно вошли в списки ПП: *по сравнению с, в интересах кого-чего*; другие отнесены к предлогам лишь некоторыми авторами: *за неимением, под давлением, с оглядкой на, с учетом* ([Лепнев 2010]), *в надежде на* ([Шереметьева 2008]), *в расчете на* ([Лепнев 2010], [Шереметьева 2008]). Легко показать, что «общепризнанные» производные предлоги и адвербиалы имеют сходные синтаксические свойства. Так, предлог *на основании чего* входит в список ПП в [Русская грамматика 1980], а адвербиал *с учетом чего* — не входит. При этом оба они участвуют в аналогичных синтаксических преобразованиях, ср. (4а) и (4б):

- (4) а) *На основании полученных нами данных был определён химический состав воды в реке Москва в июле 2003 г.* [«Геоинформатика», 2004];  
**Основываясь на полученных данных, мы определили химический состав воды;**  
*Определение химического состава воды **основывалось** на полученных данных.*
- б) *Работодатель определяет размер вознаграждения **с учётом** фактически совершённых исполнителем действий* [«Адвокат», 2004.12.01];  
*Работодатель определяет размер вознаграждения, **учитывая** фактически совершённые исполнителем действия;*  
*При определении размера вознаграждения **учитываются** фактически совершённые исполнителем действия.*

Ни объем статьи, ни ее задачи не позволяют нам углубляться в проблему разграничения адвербиалов и ПП и рассматривать соображения, по которым разные авторы включают или не включают те или иные единицы в состав ПП. Мы исходим

из того, что проблема разграничения адвербиалов и ПП не может быть решена в полном объеме (что признается всеми специалистами), поэтому не существует «идеальной» и однозначной классификации ПП. В данной работе мы включаем единицы типа *в согласии с, в сопровождении кого, с помощью кого / чего* и под., которые некоторыми авторами трактуются как ПП, в состав адвербиалов.

Задача настоящей статьи — привлечь внимание к адвербиалам, которые, во-первых, оказываются в тени ПП и рассматриваются как фон (или, хуже того, — как материал, который надо отбросить, формируя списки ПП), а во-вторых, как уже было сказано, в силу «номинативности» существующих словарей не считаются особыми лексическими единицами. Далее мы коротко рассмотрим два основных типа адвербиалов — включенные и обстоятельственные, — а затем более подробно остановимся на свойствах включенных адвербиалов.

## Типы и свойства адвербиалов

Синтаксическая структура с адвербиалом содержит две пропозиции. Вообще говоря, эти пропозиции могут быть представлены в виде двух самостоятельных предикаций: *X пришел в сопровождении Y-а* — ‘X пришел, X-а сопровождал Y’. Однако в предложении их статус неодинаков: одна из них — сказуемое — имеет статус главной предикации, другая — адвербиал — имеет более низкий статус, т. к. адвербиал — это редуцированная (свернутая, подчиненная) предикация. Она связана определенными смысловыми отношениями с главной предикацией — сказуемым. Адвербиалы аналогичны другим редуцированным предикациям — причастным или деепричастным оборотам (которые обычно тоже могут быть «развернуты» в полноценные предикации) или придаточным (т. е. подчиненным) клаузам. Вопрос о статусе субъекта адвербиала и о его соотношении с PRO подчиненных клауз мы в данной работе не рассматриваем).

Обычно у адвербиалов (в зависимости от исходной номинализации) имеются две семантические валентности — условно — субъектная и объектная. Субъект, как правило, — лицо (бывает и ситуация, но такие случаи здесь не рассматриваются). Объект может быть личным или (чаще) пропозициональным: *Брат пришел в надежде на встречу* (‘Брат [Сб] надеется на встречу [Об]’); *Школьники проводили опыты под наблюдением учителя* (‘Учитель (Сб) наблюдал за проведением опытов (Об)’). Для нашей темы значимым является субъект адвербиала.

Для выделения подтипов адвербиалов необходимо учитывать:

- какова семантика адвербиала;
- какой конструкции он соответствует (во что он может быть преобразован);
- выразим или невыразим субъект адвербиала и как он соотносится с субъектом главной предикации.

В зависимости от семантики, возможностей преобразования и соотношения между субъектами главной предикации и адвербиала (см. [Кустова 2013]), адвербиалы можно разделить на две основные группы — включенные и обстоятельственные.



## Включенные адвербиалы (ВА)

Включенные адвербиалы (ВА) дополняют основную пропозицию, конкретизируют какие-то ее аспекты. Можно сказать, что предложение с ВА представляет собой расширенный вариант реализации ситуации, обозначенной главным предикатом: *Стоял на остановке в ожидании автобуса* (ср. *стоял и ждал*); *Пришел на переговоры в сопровождении адвоката* (ср. *пришел с адвокатом*). Именно в силу такой семантической спаянности адвербиала с главной пропозицией мы и назвали этот тип адвербиалов включенными.

Различаются два класса ВА — кореферентные и некорреферентные. Они различаются не только по формальным свойствам (см. ниже), но и по семантике. Кореферентные ВА, ср. *пришел в надежде на встречу*, обозначают пропозициональные установки субъекта (обычно — Агенса) — ‘учитывать, опираться на, базироваться на, рассчитывать на, иметь в виду, принимать во внимание, надеяться’ и под. Т.е. это ментальное (в широком смысле) состояние агенса, сопровождающее какое-либо действие.

В группу некорреферентных ВА, ср. *пришел в сопровождении адвоката*, входят действия другого субъекта, который «сопровождает» действие главной предикации (либо участвует в нем) в качестве второстепенного субъекта (*P при участии Y-a*) или, наоборот, в качестве доминирующего субъекта (*P под присмотром, под надзором, под руководством Y-a*).

### А) Кореферентные ВА

Субъект основной пропозиции и семантический субъект адвербиала кореферентны (т.е. у главной пропозиции и адвербиала — один и тот же субъект): *X переработал статью с учетом замечаний рецензентов* = ‘X переработал; X учел’. Адвербиал с кореферентным субъектом обычно соответствует деепричастию, но не любому, а такому, которое в традиционной грамматике называется второстепенным действием или второстепенным сказуемым и которое может быть трансформировано в «основное» сказуемое: *стоял на остановке в ожидании автобуса* — *стоял, ожидая автобуса* — *стоял и ждал автобуса*; ср. также: *без оглядки на* = *не оглядываясь на*; *с опорой на* = *опираясь на*; *в надежде на* (или: *в надежде + Инф*) = *надеясь на* (такие деепричастия противопоставлены обстоятельству деепричастиям, которые могут быть трансформированы в придаточные времени, причины, условия, уступки, ср.: *будьте осторожны при спуске с эскалатора* → *когда спускаетесь с эскалатора*, деепричастия-второстепенные сказуемые обычно в придаточные не преобразуются, ср.: *пересмотрел выводы, основываясь на данных эксперимента* — \**пересмотрел выводы, когда основывался на данных эксперимента*, — либо меняется исходное значение, ср.: *стоял на остановке, ожидая автобуса* — ?*стоял на остановке, когда ждал автобуса* — такое предложение допустимо, но оно не соответствует исходному).

При этом субъект адвербиала, как и субъект деепричастия, системно не выражается: *Петя стоял на остановке в ожидании \*ИМ автобуса*; *Брат пришел на вечер в \*ЕГО надежде встретить Машу*. Системный запрет на выражение субъекта кореферентного адвербиала аналогичен запрету на выражение субъекта деепричастия или инфинитива (в соответствующих случаях). В принципе, можно было бы

считать, что субъект адвербиала не запрещен, а выражен нулевым местоимением (так наз. PRO), которое контролируется субъектом глагола-сказуемого, — однако в конечном счете это ситуацию не меняет: для нас важно, что при кореферентном адвербиале субъект лексически выражен быть не может, тогда как при других типах адвербиалов, напротив, может. Этот признак кореферентных адвербиалов важно учитывать в программах автоматического анализа текста.

## Б) Некореферентные ВА

Субъект главной предикации и адвербиала не кореферентны — такой адвербиал обычно соответствует пассивному причастию (если оно может быть образовано от однокоренного глагола; в некоторых языках, например, в испанском, встречаются и деепричастия, которые допускают некорреферентность с субъектом главного глагола, см. [Богуславский 2008]): *X пришел на переговоры в сопровождении адвоката* — *сопровождаемый адвокатом* ('X пришел, Y его сопровождал'), ср. также: *X сделал P без разрешения / с согласия / под наблюдением / под руководством / при участии / в присутствии Y-а / на виду (на глазах) у Y-а*. У некорреферентных адвербиалов субъект выражается, а выражение объекта (если он есть в семантической структуре адвербиала), напротив, запрещено: *Петр пришел в сопровождении \*ПЕТРА адвокатом*.

Некорреферентные адвербиалы удобны в том отношении, что не всегда имеется переходный глагол (как в случае *X, сопровождаемый Y-ом*), от которого можно образовать пассивное причастие, ср.: *в присутствии Y-а* — \**присутствуемый Y-ом*.

Важным свойством включенных адвербиалов и важным отличием их от обстоятельственных адвербиалов является то, что ВА не могут быть преобразованы в придаточные.

**Терминологическое замечание.** Термины «кореферентные» и «некорреферентные» адвербиалы, возможно, не являются оптимальными (например, некорреферентный субъект адвербиала теоретически может быть кореферентен какой-то не-субъектной именной группе в главной клаузе — хотя практически в таких случаях требуются дополнительные средства: *Мы помогаем клиентам при участии самих клиентов*, ср.: ?*Мы помогаем клиентам при участии клиентов*, — и звучат такие предложения не очень естественно, если только не считать их языковой игрой).

Альтернативой могли бы служить термины «односубъектный» и «разносубъектный» (same-subject constructions — different-subject constructions), которые используются в англоязычной литературе, ср. [Haspelmath 1995]; [Adverbial Constructions 1998]; [Language Typology 2007] (в отечественных работах по отношению к деепричастиям используется термин «односубъектность», ср. [Онипенко, Биккулова 2016]). Однако эти термины имеют тот же недостаток — они сами по себе не однозначны и нуждаются в дополнительных комментариях.

Мы надеемся, что термины «кореферентные» и «некорреферентные» адвербиалы понятны и могут быть использованы с учетом сделанной оговорки, что речь идет о кореферентности или некорреферентности субъектов адвербиала и главного глагола.

## Обстоятельственные адвербиалы

Обстоятельственные адвербиалы имеют, как следует из их названия, обстоятельную семантику. Они соотносятся с деепричастными оборотами с обстоятельным значением или с соответствующими обстоятельными придаточными (цели, причины, времени, условия), ср.: *При проверке сочинения учитель нашел ошибку — проверяя сочинение / когда проверял сочинение.*

Среди обстоятельных адвербиалов также есть кореферентные (у них семантический субъект совпадает с субъектом главной предикации) и некорреферентные (субъекты не совпадают). К кореферентным относятся, например, адвербиалы со значением цели: *в благодарность, в доказательство, в дополнение, в защиту, во избежание, во исполнение, в обоснование* ('чтобы обосновать'), *в наименование, в отместку* ('с целью отомстить'), *в поддержку* ('чтобы поддержать'), *в подкрепление, в подтверждение; на потеху* и др. К некорреферентным относятся, например, адвербиалы со значением причины, основания, побудительного мотива: *под влиянием, под гнетом, под давлением, под действием, под угрозой; по вине, по приглашению, по приказу, по просьбе, по требованию, по настоянию* и под.

Адвербиалы со значением времени нейтральны относительно признака кореферентность-некорреферентность; они могут быть и кореферентными: *При выборе товара обращайтесь внимание на срок годности — выбирая товар / когда выбираете товар* — 'когда X выбирает...', X должен обращать внимание...'; и некорреферентными: *При проверке [таможенником] чемоданов пассажиры заметно нервничали* — 'У проверял, X нервничал'.

Теоретически можно представить себе ситуацию, когда от одной номинализации образуются кореферентный и некорреферентный включенные адвербиалы. Однако практически такой случай нам не встретился (и это можно объяснить семантическими ограничениями на включенные адвербиалы). При этом встречается ситуация, когда от одной номинализации образуется и включенный, и обстоятельный адвербиал (см. ниже). Поэтому при обсуждении включенных адвербиалов мы будем в качестве фона привлекать однокоренные обстоятельные.

Если обстоятельные адвербиалы еще имеют шанс попасть в поле зрения исследователей, — хотя бы в качестве синонимов обстоятельных деепричастий и придаточных, — то включенные адвербиалы практически не изучаются.

Далее мы рассмотрим свойства включенных адвербиалов на примере нескольких представителей этой группы.

## Включенные адвербиалы

### Включенные адвербиалы с кореферентным субъектом (с оглядкой на кого-что; в расчете на что; с учетом чего)

Первая особенность кореферентных адвербиалов — более широкие синтаксические возможности по сравнению с их ближайшим аналогом, с которым

их естественно сравнивать, — с деепричастиями. Рассмотрим, как деепричастия и кореферентные адвербиалы взаимодействуют с главной предикацией и чем может выражаться сама главная предикация.

Кореферентные ВА, как и соответствующие им деепричастия, в норме присоединяются к глагольным конструкциям с агентивным подлежащим и, соответственно, агентивным глаголом:

- (5) *Кукушкина с оглядкой на дверь прошептала: — Вы, небось, к товарищу Сталину хотите попасть?* [Анатолий Приставкин. Кукушата, или жалобная песнь для успокоения сердца (1992)];
- (6) *Годы спустя, когда Гроссман напишет роман «Сталинград», редактор «Нового мира» Твардовский пошлет рукопись члену редколлегии Шолохову в расчёте на поддержку* [С. Б. Рассадин. Книга прощаний (2004–2008)];
- (7) *Бюджет рассчитали с учётом завышенных расценок* [«Дело», 2002].

При этом по отношению к главной клаузе деепричастие и соответствующий ему адвербиал ведут себя по-разному. Рассмотрим два случая: присоединение к пассивной конструкции и присоединение к номинализации.

Семантический субъект деепричастия кореферентен не просто субъекту главной предикации, но субъекту-подлежащему, поэтому в контексте деепричастной конструкции пассив в главной предикации (когда в позиции подлежащего оказывается семантический объект) нежелателен:

- (8) *Основываясь на полученных данных, был определен химический состав воды.*  
Тем более запрещен некорреферентный неодоушевленный субъект:
- (9) *Основываясь на полученных данных, эксперимент прошел успешно* (ср. хрестоматийный пример из А. П. Чехова *Подъезжая к станции, у меня слетела шляпа*).

Деепричастие не может присоединяться непосредственно к номинализации: *\*определение химического состава, основываясь на полученных данных*. Речь идет о случаях, когда в предложении нет агентивного глагола, ср.: *\*У нас до сих пор нет анализа обстановки, основываясь на полученных данных* (подобные конструкции могут встречаться в канцелярском языке, но их нельзя признать грамматичными).

Адвербиал снимает эти запреты, и предложение остается правильным — лишь бы сохранялась кореферентность субъектов на семантическом уровне.

Адвербиал свободно присоединяется к сказуемому в пассиве (с неодоушевленным подлежащим и синтаксически не выраженным субъектом-агентом):

- (10) *Каким бы дорогим и редким ни был этот автомобиль, он все равно **сделан с оглядкой на то, что его нужно продать*** [«Автопилот», 2002.10.15];
- (11) *Ядерные арсеналы России и США времён холодной войны **создавались в расчёте на тотальное уничтожение друг друга*** [«Известия», 2003.02.19];

- (12) *С учётом накопленного опыта были разработаны требования к рациональным режимам труда и отдыха военнослужащих* [«Военная мысль», 2003.07.14];
- (13) *Эти взгляды должны постоянно уточняться с учётом жизненных реалий.*
- Кроме того, кореферентные адвербиалы присоединяются и к номинализации, т. к. у нее сохраняется агентивный семантический субъект:
- (14) *Планирование акций с оглядкой на выпуски новостей* (‘планируют с оглядкой на...’);
- (15) *Проведение девальвации в расчете на займы* (‘провели в расчете на...’);
- (16) *Необходимость решения ядерной проблемы в КНДР с учётом северокорейских интересов* [«Дипломатический вестник», 2004].

Адвербиал может даже присоединяться к конструкции с неодушевленным субъектом и неагентивным глаголом. Иногда такие предложения выглядят не полностью грамматичным, ср.:

- (17) *Идея введения «краевого знака соответствия» вызрела в недрах местного министерства финансов с оглядкой на удачный московский эксперимент* [Виктор Мясников. Водка (2000)].

Однако если агенса, которого нет в семантической структуре главной предикации, можно восстановить, реконструировать (и он окажется кореферентным с субъектом адвербиала), то такие конструкции вполне корректны:

- (18) *Тогда я достал чемодан. И раскрыл его. Сверху лежал приличный двубортный костюм. В расчёте на интервью, симпозиумы, лекции, торжественные приёмы. Полагаю, он сгодился бы и для Нобелевской церемонии* [Сергей Довлатов. Чемодан (1986)].

Конструкция \**в чемодане лежал костюм, рассчитывая на...* из-за формального запрета невозможна, а конструкция *В чемодане лежал костюм в расчете на интервью, симпозиумы, лекции...* — возможна, т. к. восстанавливается цепочка — ‘лежал костюм, который я положил в расчете надеть его...’. Ср. также:

- (19) *Журнал будет выходить каждую неделю по четвергам — в расчете на более спокойное, «выходное» чтение* [«Русский репортер», 2007] — ‘журнал будут выпускать в расчете на...; издатели рассчитывают, что...’.

В процессе функционирования в официальном («канцелярском») языке адвербиалы семантически «выветриваются» и теряют соотнесенность в пропозицией, ср.: *Выпускаются наборы для сервировки стола в расчёте на двух человек — наборы на двух человек.* В таком употреблении утрачивается связь с агентивной семантикой, что приводит к формированию собственно предлогов. Ср. также:

- (20) *Большинство крестьян, попавших в списки присяжных на основании «служебного» ценза, являлись выходцами из беднейших слоёв народа* [«Отечественные записки», 2003] — *попали в списки присяжных*

на основании «служебного» ценза = благодаря служебному цензу, в силу, из-за (адвербиал превращается в ПП со значением причины).

Другой пример — формирование у кореферентного включенного адвербиала обстоятельственного значения условия, что тоже способствует превращению в ПП:

(21) *Цена за единицу оборудования с учётом НДС* [Договор на поставку оборудования (2003)] — ‘если учитывать НДС’.

Итак, практически у всех кореферентных адвербиалов есть существенное отличие от деепричастий: субъект деепричастия требует кореферентности не просто субъекту главной предикации, но подлежащему. Адвербиал же предполагает только кореферентность семантическому субъекту главной предикации, в том числе синтаксически не выраженному. Поэтому адвербиал не только допускает пассивную конструкцию (от агентивного глагола) в главной предикации, но даже предпочитает ее (в официальном языке, для которого и характерен пассив, конструкции с пассивом занимают значительное место). Более того, адвербиал допускает присоединение к неагентивной структуре.

Вторая особенность кореферентных адвербиалов — «обратимость» статусов главной и зависимой пропозиций: адвербиал может превращаться в полноценное сказуемое, а главная предикация — в адвербиал или деепричастный оборот со значением времени. При этом общее значение исходного высказывания сохраняется:

(22) *Работодатель определяет размер вознаграждения с учётом фактически совершённых исполнителем действий* [«Адвокат», 2004.12.01] vs. *Определяя вознаграждение, работодатель учитывает действия исполнителя / При определении вознаграждения работодатель учитывает действия исполнителя;*

(23) *Твардовский пошлет рукопись члену редколлегии Шолохову в расчете на поддержку* vs. *Посылая рукопись Шолохову, Твардовский рассчитывал на поддержку.*

У других типов адвербиалов свои наборы синтаксических преобразований, и эти преобразования обычно приводят к существенным изменениям исходной структуры, например, требуют включения дополнительного глагола, ср. *Х отказался от проекта под угрозой разоблачения* — *Угроза разоблачения заставила Х-а отказаться*; или приводят к изменению коммуникативного статуса пропозиций: *Оппозиционеры провели митинг в поддержку политзаключенных* (у целевого адвербиала в поддержку — неутвердительный статус, см. [Падучева 1985]) vs. *Проведя митинг, оппозиционеры поддержали политзаключенных* (у сказуемого поддержали — утвердительный статус).

Иногда на базе одной и той же номинализации образуются и включенный, и обстоятельственный адвербиал. Так, на базе номинализации *учет* (*учитывать*) образуется не только включенный кореферентный адвербиал с *учетом*, для которого действует запрет на выражение субъекта, ср.: *Исследователь провел эксперимент с учетом всех факторов* — с *учетом \*им всех факторов*,

но и обстоятельственный адвербиал *при учете* со значением времени или условия. У него нет запрета на выражение каких-л. валентностей, т. е. выражаются обе валентности — субъект и объект, — и он не требует кореферентности субъектов главной предикации и адвербиала, хотя и допускает ее: *При учете всех факторов исследователем [X] он [X] получит достоверный результат* (кореферентность); *При учете [Росстатом — Y] всех показателей получится не столь радужная картина [X]* (некореферентность).

### Включенные адвербиалы с некореферентным субъектом (*при участии кого*)

Общее значение конструкций с адвербиалом *Р при участии Y-а*: ‘Y участвовал в главной предикации в качестве неосновного субъекта, наряду с другими субъектами’:

- (24) *При участии В. А. Дегтярева было создано первое в стране проектно-конструкторское бюро (ПКБ) стрелкового оружия* [«Вокруг света», 2004];
- (25) *Цель конкурса «FootballGirls», созданного при участии Центра болельщиков, — поддержка женского футбольного движения на трибунах* [«Огонек», 2014] — ‘X создал конкурс, и болельщики [Y] в этом тоже участвовали’.

На базе той же номинализации (*участие кого в чем*) образуется обстоятельственный адвербиал *при участии*:

- (26) *Извечный радетель за здоровье подведомственного народа Геннадий Онищенко не забывает предупредить об инфекционной угрозе при участии в митингах* [«Огонек», 2013].

Обстоятельственный адвербиал *при участии* может иметь значение условия или времени:

- (27) *При участии в заключении мирового соглашения третьих лиц мировое соглашение подписывается в том числе этими лицами* [«Арбитражный и гражданский процесс», 2003] — ‘если участвуют третьи лица’;
- (28) *Свои возможности эта система подтвердила при участии в боевых действиях на Ближнем Востоке* [«Воздушно-космическая оборона», 2001.04.15] — ‘когда участвовала’.

Если у включенного адвербиала *при участии* может быть выражен только субъект, но не может быть выражено мероприятие (‘в чем участвует Y’), ср.:

- (29) *Данный обряд совершался при участии всех женщин-родственниц невесты и подруг* [Сб] [«Народное творчество», 2004] — *Обряд совершается при участии подруг невесты \*в обряде / при участии \*в нем подруг невесты*;

то обстоятельственный адвербиал *при участии* не имеет ограничений на заполнение валентностей, т. е. они обе могут быть выражены, ср.: *При участии подруг [Сб] невесты в обряде* [мероприятие] *его структура изменяется*.

Таким образом, ВА имеют системные ограничения на выражение валентностей исходного предиката, обстоятельственные адвербиалы времени и условия — не имеют. Это важное синтаксическое отличие может быть использовано при автоматической обработке текста: если имеются обстоятельственный и включенный адвербиалы от одной и той же номинализации, они будут различаться синтаксическими связями.

\*\*\*

Итак, мы попытались показать, что хотя адвербиалы обычно не считаются самостоятельными единицами, однако они имеют ряд семантических и синтаксических свойств, которые позволяют их рассматривать как особые единицы. С одной стороны, адвербиалы участвуют в процессе грамматикализации, удаляясь от исходной номинализации и теряя свойства из исходного набора. С другой стороны, адвербиалы как особый тип редуцированных предикаций соотносятся с деепричастными и причастными оборотами, в сравнении с которыми адвербиалы имеют более широкий круг употреблений и меньше синтаксических ограничений, чем и объясняется их распространенность в книжном языке.

## Литература

1. *Богуславский И. М.* Исследования по синтаксической семантике. М.: Наука, 1985.
2. *Богуславский И. М.* Сфера действия лексических единиц. М.: Наука, 1996.
3. *Богуславский И. М.* Актантное поведение адвербиальных дериватов // Динамические модели: Слово. Предложение. Текст. Сб. статей в честь Е. В. Падучевой. М.: ЯСК, 2008. С. 110–128.
4. *Бурцева В. В.* Словарь наречий и служебных слов русского языка. М., 2005.
5. *Кустова Г. И.* Семантические механизмы формирования адвербиальных выражений на базе отглагольных существительных // Компьютерная лингвистика и интеллектуальные технологии. Труды международной конференции «Диалог-2013». Вып. 12 (19). М., 2013. С. 354–366.
6. *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.: ЛЕКСПУС, 2016. С. 302–319.
7. *Лепнев М. Г.* Русские производные предлоги. СПб., 2010.
8. *Лютикова Е. А.* Безвршинные относительные предложения в русском языке: эмпирические данные и теоретические проблемы // Rhema. Рема. 2015. № 3. С. 74–85.
9. *Мельчук И. А.* Опыт теории лингвистических моделей «Смысл ↔ Текст». М.: Наука, 1974. 2-е изд. — М.: ЯРК, 1999.
10. *Онипенко Н. К., Биккулова О. С.* Проблема деепричастной нормы и категория субъекта // Труды Института русского языка им. В. В. Виноградова.



- Вып. 10. Материалы международной научной конференции «Грамматические процессы и системы в синхронии и диахронии» (30 мая — 1 июня 2016 г.). СПб.: Нестор-История, 2016. С. 220–231.
11. Падучева Е. В. Высказывание и его соотносительность с действительностью (Референциальные аспекты семантики местоимений). М.: Наука, 1985.
  12. Русская грамматика. Т. I. / под ред. Н. Ю. Шведовой. М., 1980.
  13. Филипенко М. В. Семантика наречий и адвербиальных выражений. М., 2003.
  14. Черкасова Е. Т. Переход полнозначных слов в предлоги. М., 1967.
  15. Шереметьева Е. С. Отыменные релятивы современного русского языка. Семантико-синтаксические этюды. Владивосток, 2008.
  16. *Adverbial Constructions in the Languages of Europe*. Johan van der Auwera (ed.). Mouton de Gruyter. Berlin-New York, 1998.
  17. Croft W., Cruse D. A. *Cognitive linguistics*. Cambridge: Cambridge University Press, 2004.
  18. Goldberg A. E. *Constructions: A Construction Grammar approach to argument structure*. Chicago and London: The Univ. of Chicago Press, 1995.
  19. Goldberg A. E. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press, 2006.
  20. Grennbaum S. *Studies in English adverbial usage*. London, 1969.
  21. Haspelmath M. The converb as a cross-linguistically valid category // *Converbs in cross-linguistic perspective*, Martin Haspelmath, Ekkerhard König (eds). Mouton de Gruyter, Berlin-New York, 1995. Pp. 1–55.
  22. Heny F. W. *Sentence and Predicate Modifiers in English // Syntax and Semantics*. Vol. 2. N.Y., London, 1973.
  23. *Language Typology and Syntactic Description: Vol. 2*, Timothy Shopen (ed.). Cambridge: Cambridge University Press, 2007.

## References

1. *Adverbial Constructions in the Languages of Europe* (1998), Johan van der Auwera, ed., Mouton de Gruyter, Berlin—New York.
2. Boguslavskii I. M. (1985), *Research on syntactic semantics [Issledovaniya po sintaksicheskoi semantike]*, Moscow.
3. Boguslavskii I. M. (1996), *Scope of lexical units [Sfera deistviya leksicheskikh edinit]*, Moscow.
4. Boguslavskii I. M. (2008), *Argument behavior of adverbial derivatives [Aktantnoe povedenie adverbial'nykh derivatov]*, *Dynamic models: Word. Sentence. Text. [Dinamicheskie modeli: Slovo. Predlozhenie. Tekst]*, *Yazyki slavyanskoj kul'tury*, Moscow, pp. 110–128.
5. Burtseva V. V. (2005), *Dictionary of Russian adverbs and grammatical words [Slovar' narechii i sluzhebnykh slov russkogo yazyka]*, Moscow.
6. Cherkasova E. T. (1967), *The transformation of lexical words to prepositions [Perekhod polnoznachnykh slov v predlogi]*, Moscow.
7. Croft W., Cruse D. A. (2004), *Cognitive linguistics*, Cambridge: Cambridge University Press.

8. *Goldberg A. E.* (1995), *Constructions: A Construction Grammar approach to argument structure*, Chicago and London, The Univ. of Chicago Press.
9. *Goldberg A. E.* (2006), *Constructions at work: The nature of generalization in language*, Oxford: Oxford University Press.
10. *Filipenko M. V.* (2003), *The semantics of adverbs and adverbials [Semantika narechii i adverbial'nykh vyrazhenii]*, Moscow.
11. *Grennbaum S.* (1969), *Studies in English adverbial usage*, London.
12. *Haspelmath M.* (1995), *The converb as a cross-linguistically valid category, Converbs in cross-linguistic perspective*, Martin Haspelmath, Ekkerhard König, eds., Mouton de Gruyter, Berlin—New York, pp. 1–55.
13. *Henry F. W.* (1973), *Sentence and Predicate Modifiers in English, Syntax and Semantics*, vol. 2, N.Y., London.
14. *Kustova G. I.* (2013), *Adverbial expressions based on verbal nouns [Semanticheskie mekhanizmy formirovaniya adverbial'nykh vyrazhenii na baze otglagol'nykh sushchestvitel'nykh]*, *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue 2013” [Komp'yuternaya lingvistika i intellektual'nye tekhnologii. Trudy mezhdunarodnoi konferentsii “Dialog-2013”]*, Issue 12 (19), Moscow, pp. 354–366.
15. *Kustova G. I.* (2016), *Free relatives as reduced constructions: meaning and use [Svobodnye relyativy kak reducirovannye konstruksii: znachenie i upotreblenie]*, *Yazyk: poiski, fakty, gipotezy [Language: searches, facts, hypotheses]*, M. V. Lyapon, ed., LEKSRUS, Moscow, pp. 302–319.
16. *Language Typology and Syntactic Description: Volume 2.* (2007), Timothy Shopen, ed., Cambridge: Cambridge University Press.
17. *Lepnev M. G.* (2010), *Russian derivative prepositions [Russkie proizvodnye predlogi]*, St. Petersburg.
18. *Lyutikova E. A.* (2015), *Free relatives in Russian: empirical data and theoretical issues [Bezvershinnye otnositel'nye predlozheniya v russkom yazyke: empiricheskie dannye i teoreticheskie problemy]*, *Rhema*, № 3, pp. 74–85.
19. *Mel'chuk I. A.* (1974 / 1999), *An essay on the theory of linguistic models “Meaning ↔ Text” [Opyt teorii lingvisticheskikh modelei “Smysl ↔ Tekst”]*, Nauka, Moscow, 1974, 2-nd edition — YaRK, Moscow, 1999.
20. *Onipenko N. K., Bikkulova O. S.* (2016), *Russian adverbial participle — grammatical rule and category of subject [Problema deeprichastnoi normy i kategoriya sub'ekta]*, *Proceedings of the V. V. Vinogradov Russian Language Institute, 2016, No. 10. Materials of International Scientific Conference “Grammatical Processes and Systems in Synchrony and in Diachrony” (May 30—June 1, 2016) [Trudy Instituta russkogo yazyka im. V. V. Vinogradova. Vyp. 10. Materialy mezhdunarodnoi nauchnoi konferentsii “Grammaticheskie protsessy i sistemy v sinkhronii i diakhronii” (30 maya—1 iyunya 2016 g.)]*, Nestor-Istoria, St. Petersburg, pp. 220–231.
21. *Paducheva E. V.* (1985), *Utterance and its relationship to reality*, Moscow.
22. *Russian grammar* (1980), [Russkaya grammatika], v. I, N. Yu. Shvedova, ed., Moscow.
23. *Sheremet'eva E. S.* (2008), *Denominative relatives in modern Russian. Semantic and syntactic essays [Otymennye relyativy sovremennogo russkogo yazyka. Semantiko-sintaksicheskie etyudy]*, Vladivostok.

# 'TENSED' AND 'NON-TENSED' PREDICATIVES IN RUSSIAN<sup>1</sup>

**Letuchiy A. B.** (alexander.letuchiy@gmail.com)

National Research University Higher School of Economics,  
Russia, Moscow

In the article, I focus on tense marking in Russian constructions with predicatives, such as *xolodno* '(it is) cold' and *ploxo* '(it is) bad'. Statistical data from the Russian National Corpus show that the frequency of past tense forms (e.g., combinations with the form *bylo*) is much greater for some predicatives than for others. This difference results both from semantic and formal factors. On the one hand, some predicatives denote evaluation (e.g. *ploxo* 'bad'). Evaluation can be applied to events that have finished or have never been realized. What is relevant is that the evaluation is made at the moment of speech, and this is why the present tense (= the zero copula verb) is used. On the other hand, it is important that the present tense is unmarked with predicatives, while with verbs, it is marked with special verbal affixes. The unmarked present tense form of a predicative can get the temporal meaning from the embedded verb. Interestingly, this phenomenon is in a sense opposite to the well-known phenomenon of relative tense marking. While the latter presupposes that the tense assignment in the embedded event is anchored to the tense meaning of the main event, the tense value of the construction with evaluation predicatives is assigned by 'agreement' with the embedded verb.

**Keywords:** predicatives, tense, evaluation, emotion, copula verb, the Russian language

## ПРЕДИКАТИВЫ, МАРКИРОВАННЫЕ И НЕ МАРКИРОВАННЫЕ ПО ВРЕМЕНИ, В РУССКОМ ЯЗЫКЕ

**Летучий А. Б.** (alexander.letuchiy@gmail.com)

Национальный исследовательский университет  
Высшая школа экономики, Россия, Москва

**Ключевые слова:** предикативы, время, оценка, эмоции, глагол-связка, русский язык

---

<sup>1</sup> The study was supported by the grant #16-18-02071 from the Russian Science Foundation.

## 1. Introduction

In Russian, a special type of units called ‘predicatives’ exists. These units are mostly homonymous to short adjectives and / or adverbials and tend to be found in the predicate position. Here belong, for instance, *xolodno* ‘(it is) cold’, *ploxo* ‘(it is) bad’, *stranno* ‘(it is) strange’, though some go back to nouns (*vremja* ‘time’). The group of predicatives includes the following semantic classes:

- Modal predicatives: *nado* ‘(it is) necessary’, *nužno* ‘(it is) necessary’, *možno* ‘(it is) possible’, *vozmožno* ‘(it is) possible’
- Physical state: *tesno* ‘(it is) crowded’.
- Mental state: *interesno* ‘(it is) interesting’, *legko* ‘(it is) easy (e.g., to do something)’.
- Emotions: *radostno* ‘(it is) joyful’, *neprijatno* ‘(it is) unpleasant’.
- Characteristics of a situation in terms of its properties: *stranno* ‘(it is) strange’, *glupo* ‘(it is) silly’.
- Characteristics of time: *pozдно* ‘(it is) late (e.g., to do something)’, *rano* ‘(it is) (too) early (e.g., to do something)’, *vremja* ‘(it is) time (e.g., to do something)’.
- Ethical evaluation: *neetično* ‘(it is) immoral’, *žestoko* ‘(it is) cruel’, *blagorodno* ‘(it is) generous’.
- General evaluation: *ploxo* ‘(it is) bad’, *xorošo* ‘(it is) good’.

In the article, I consider tense marking of predicatives. I will show that the present tense in constructions with predicatives sometimes has a non-standard interpretation. The analysis is based on data taken from the Russian National Corpus ([www.ruscorpora.ru](http://www.ruscorpora.ru)), as well as some elicited examples.

Tense is not expressed morphologically in predicatives. It is marked in the form of the auxiliary verb ‘be’, which is used with the predicative as a finite element. When a predicative has to be marked for present tense, no overt predicate is used (Testelets 2009 and Letuchiy 2015 show that the question whether these cases are instances of zero copula or do not contain any verb at all is not trivial, and the answer is different for different types of constructions).

(1) *Мне холодно.*

It turns out, however, that in some contexts, the present tense (the zero copula) is used, though the meaning seems to require past or future.

(2) *Хотел сказать—но что-то остановило. И хорошо, что не сказал.*  
[Анатолий Азольский. Лопушок // «Новый Мир», 1998]

(3) *Да вот этот визит занял весь вечер и напрочь разрушил столь любимое им чувство одиночества. В конце концов, может, и хорошо, что разрушил ...* [Василь Быков. Бедные люди (1998)]

(4) *Оставшийся кирпич лежал под навесом и странно, что вовсе не продавался.* [В. А. Солоухин. Смех за левым плечом (1989)]

(5) *Вообще возбуждение возрастало, и неизвестно, во что бы всё это вылилось ...* [М. А. Булгаков. Мастер и Маргарита, часть 1 (1929–1940)]

Now I will consider these contexts in detail.

## 2. Present tense for present vs. past meanings

Examples (2) through (5) are heterogenous in that present tense can refer either to a moment in present or in past. Below I term the interpretation where the zero copula refers to a moment in present 'standard interpretation'.

For instance, in (5), the interpretation can be standard. While the whole context describes a situation in past, the predicative *neizvestno* remains unmarked for tense (or combined with a zero copula). This can result from the fact that the answer to the indirect question is unknown to the author at the moment of speech (in this case, the moment when the text is written or read).

This interpretation seems to be incompatible with example (3). A broader context apparently describes the hero's (and not the author's) attitude to the events:

- (3') *Жена Скварыша с внуками тоже вчера уехала на дачу, а он остался в городской квартире, сказал, есть дело. Хотя дела у него не было никакого, просто хотелось побыть одному, наедине со своею бедой. Да вот этот визит занял весь вечер и напроць разрушил столь любимое им чувство одиночества. В конце концов, может, и **хорошо**, что разрушил: одиночество тоже с каких-то пор перестало быть для него спасением, как не было им и многолюдье. Настало время, когда людей вокруг него становилось всё меньше, а порою они и вовсе исчезали, особенно из числа друзей, сослуживцев; одиночество разрасталось, заключая его в плотный глухой пузырь. Это было непривычно, пугало, и он не знал, где ему лучше: дома без людей или на людях вне дома. То, что чужих автомобилей внизу не было, несколько его успокоило, хотя, подумал он, машина могла уже и отъехать, пока он прощался с Краснянским. [Василь Быков. Бедные люди (1998)]*

The hero perceives the situation in past, so its attitude is also situated in past (we do not know anything about the hero's state in present).

Interestingly, all constructions with verbal predicates in (3) (*uexala, ostalsja, zanjal*) are marked for past, while the predicative is not accompanied with a past tense form of the copula verb *byt* 'be'. The insertion of *bylo* makes the example stylistically poor:

- (6) *?В конце концов, может, и **хорошо было**, что разрушил.*

The same interpretation seems to be valid for examples (2) and (4). In (4), for instance, the evaluation *stranno* 'strange' is made by the heroes at the reference time, rather than at the moment of speech. Below I refer to examples like (3), (4) and (5) using the term 'non-standard construction' and 'non-standard interpretation (of the present tense)'.

## 3. Statistical distribution: percentage of past tense

Statistics, based on the Russian National Corpus, shows that predicatives differ from each other in respect of tense expression. Table 1 below contains percentages of examples with past tense marking with several predicatives in the whole sample of examples for the same predicatives. Of course, the verb *byt* 'can be separated from

the predicative. However, here I counted only those contexts in which the form *bylo* is situated immediately before or after the predicative.<sup>2</sup> To count the percentage in the last column, the number of contexts with *bylo* was divided to the whole number of examples with the predicative and the complementizer *čto* or an interrogative pronoun (*kto* ‘who’, *kogda* ‘when’ and so on).<sup>3</sup>

Predicative	Total	<i>bylo</i> + PRAEDIC + complementizer / pronoun	PRAEDIC + <i>bylo</i> + complementizer / pronoun	Percentage of past
<i>непонятно</i> + <i>кто</i>   <i>как</i>   <i>что</i>   <i>где</i>   <i>когда</i>   <i>почему</i>   <i>сколько</i>	1,991	154	133	14.4
<i>странно</i> + <i>что</i>	1,777	118	79	11.1
<i>плохо</i> + <i>что</i>	559	7	3	1.8
<i>приятно</i> + <i>что</i>	980	237	49	29.2
<i>важно</i> + <i>что</i>	1,425	29	20	3.4
<i>обидно</i> + <i>что</i>	796	104	44	18.6
<i>хорошо</i> + <i>что</i>	6,745	38	18	0.8
<i>интересно</i> + <i>кто</i>   <i>как</i>   <i>что</i>   <i>где</i>   <i>когда</i>   <i>почему</i>   <i>сколько</i>	4,566	121	32	3.4

The percentage of contexts where the predicative is marked for past apparently differs from one predicative to another. Predicatives associated with emotions, such as *prijatno* and *obidno*, show the highest percentage of past. For those marking cognitive activity, *stranno* and *neponjatno*, the proportion is a little lower. Finally, generalized properties and evaluation predicatives, namely, *važno*, *ploxo* and *xorošo*, are rarely marked for past.

Of course, this difference can result from different factors:

- (1) rarity of past contexts. It may be the case that generalized properties and evaluation are mainly described as belonging to the present time.
- (2) difference in interpretation. Perhaps, the statistical difference results from the fact that the three predicatives of the evaluation group are formally distinct from other groups: they are used without past marking, even if they denote a situation in past.

<sup>2</sup> This condition is imposed on the search results to avoid contexts like *Bylo interesno i prijatno* ‘(It) was interesting and pleasant’, where *bylo* is syntactically connected with two predicatives: *interesno* and *prijatno*. Another type of contexts I would like to sort out in this way is *Bylo li prijatno* ‘whether it was pleasant’ with an interrogative / indirect question particle *li*: interrogative contexts can differ from affirmative ones in many respects.

<sup>3</sup> The site [www.ruscorpora.ru](http://www.ruscorpora.ru) was accessed on 08.02.1017.

In fact, both factors appear to be relevant. On the one hand, the evaluative component can be applied to a situation as a whole (without a temporal localization). A situation as such can be regarded as 'good' or 'bad' theoretically, even if this situation has never taken place or if we speak about a class of situations.

On the other hand, we have seen that even for evaluative predicatives, the specific interpretation is also possible: very often the same evaluative predicatives denote single definite events, as in (2)–(5) above. This is where the second factor (formal properties of predicatives) is relevant. In the next section I will describe the group of predicatives that tend to remain unmarked for tense.

#### 4. The relevant group of predicatives

As mentioned above, the relevant group of 'non-tensed' predicatives includes mainly predicatives of evaluation ('good', 'bad') and other predicatives that denote continuous properties of events ('important'). At first glance, the fact that they can be used without tense marking does not seem very surprising. To be characterized as 'good' or 'important', an event should not obligatorily be directly related to the present state of affairs—moreover, it can be the case that an event has not been realized.

On the other hand, two more things are to be explained:

- 1) why for other predicatives, the non-standard interpretation of present tense is impossible or strange;
- 2) why the standard marking of tense with the form *bylo* sometimes sounds awkward for this group of predicates, as in (7), contrary to non-standard construction (8), which sounds normally:

(7) \**Было хорошо, что Петя вовремя прибежал и помог нам.*

(8) *Хорошо, что Петя вовремя прибежал и нам помог.*

##### 4.1. Impossibility / awkwardness of the non-standard construction with other predicates

Why is the non-standard interpretation of the present tense unacceptable with predicatives of emotions and many other classes? I shall explain it on the example of emotional predicatives. The reason seems to be that emotions are linked to the time of event much tighter than evaluation meanings.

For instance, example (9) is non-canonical semantically, while (10) is standard:

(9) *Маме приятно, что гости вели себя вежливо.*

(10) *Маме было приятно, что гости ведут / вели себя вежливо.*

Of course, both (9) and (10) are perfectly grammatical. What is unusual about (9), is that the emotion is described in present, while the event took place in past. Contexts like this occur in the RNC, but in a special situation when the past situation is relevant for the moment of speech. For instance, this is the case with 'resultative events':

(11) *Очень приятно, что вы остались с нами.* [О подписке на второе полугодие 2007 года // «Наука и жизнь», 2007]

In (11), the stimulus situation is relevant for the moment of speech: the addressee is currently with the speaker. This is not the case in (9), where (in the absence of a broader context) the embedded situation does not seem to be relevant.

In (10), the emotion is described as occurring simultaneously to the stimulus. This variant seems to be more canonical.

#### 4.2. Awkwardness of the standard construction with evaluation predicates

Contrary to emotions, evaluation meanings are not linked to the time when the situation occurs. The description of a situation as ‘good’, ‘bad’, or ‘strange’ is not related to the individual instance of the situation. If we say something like ‘It was raining, and it is good that Peter took his umbrella’, this evaluation results from the general logic (‘If it rains, an umbrella is necessary’). Also contrary to emotions, evaluation meanings are not sensitive to the relevance of the situation: we can evaluate even a situation that is completely irrelevant to the moment of speech: in (12), for instance, the evaluation constitutes a part of the description of the author’s childhood. The events, which are described, are not necessarily relevant for the current time:

(12) *Да, я родился в Советском Союзе и горжусь этим. Все детство отдыхал в пионерлагерях и на базах с родителями. И самое интересное, мы все были счастливы—без пятизвездочных апартаментов, мнимой экзотики и обжираловки по системе «все включено». Хорошо, что тогда мерилом жизни не был ни гламурный вид, ни иномарка, ни офисная работа, ни обязательный отдых за рубежом ...* [Евгений РУМЯНЦЕВ. Турецкие врачи не разглядели у 9-летнего россиянина перелом руки // Комсомольская правда, 2013.07.11]

The fact that life used to be different when the speaker was child does not have much to do with the moment of speech. However, this fact is really important for the way of speaker’s life in that period.

Yet, the present tense is used because the evaluation meaning ‘good’ is applicable even to events which are no longer relevant. If it is changed to *prijatno*, it is necessary that the event should be relevant for the speech act time.

If a predicative like *xorošo* was used with *bylo*, it would imply that the evaluation predicative refers to past in the sense that the situation is evaluated in past. However, the evaluation like ‘good’ by default refers to the general state of affairs. The hypothetical meaning ‘it was good that ...’ presupposes that the evaluation changed from the reference time, which is highly non-canonical (‘in my childhood it was good, but it is no longer good’). To make the standard construction with *bylo* possible in (12), a subjective component should be included in the sentence. For instance, (12’) sounds good with *bylo*, because the PP *dlja menja* ‘for me’ is added: the situation might be good for the speaker in past, and then the speaker might have changed his opinion.

(12’) *Для меня было хорошо, что мерилом жизни тогда не был гламурный вид.*



In other words, evaluation predicatives tend to choose a non-standard interpretation, because evaluation tends to be regarded as objective and constant in time (though in fact, this is not necessarily the case).

### 4.3. Pragmatic motivation for not marking past: speech act markers

There is another group that typically does not take the past tense marker *bylo*: it includes the predicative *interesno*. We have not mentioned it, but note that the proportion of past for this predicative is as low as for evaluation predicatives. It seems, however, that it is semantically different from the evaluation group: the fact that something is interesting for someone is not a constant property, it rather changes from one period to the other.

If we have a look at the data, it turns out that *interesno* can be combined with *bylo*, but only if the dative-marked experiencer is expressed in the sentence (16). Constructions with *bylo* but without an experiencer sound a bit strange (14):

(13) *Интересно, кто этот человек—ходит, спрашивает всех о чём-то.*

(14) *?Какой-то человек ходил и спрашивал всех о чём-то. Было интересно, кто этот человек.*

(15) *Мне интересно, кто этот человек.*

(16) *Мне было интересно, кто этот человек.*

I think that for *interesno*, the rarity of past results from the discourse function of the predicative. When used without an explicit experiencer, *interesno* either marks the question or the relevance of the information. Since the predicative is used here as a discourse marker<sup>4</sup>, it is natural that usually it does not refer to the past: discourse meanings tend to be used deictically, with reference to the speech act time and location. In contrast, when the experiencer is mentioned, as in (15)–(16), the predicative *interesno* expresses the usual meaning of the cognitive state of the experiencer ('something is interesting for the experiencer'). Not surprisingly, this meaning allows transposition to past contexts, as well as emotional meanings like 'surprise'.

Contrary to evaluation predicates, *interesno* does not allow the non-standard interpretation of present tense. In general, the use of this predicative without an experiencer is rare if it refers to past.

Notably, for *interesno*, the possibility of the non-standard interpretation depends on the strategy of complementation. The standard interpretation is possible if an indirect question is used (13)–(16), but almost impossible if the verb hosts a *čto*-clause:

(17) *Интересно, что никто не заметил подмены документов.*

(18) *\*Было интересно, что никто не заметил подмены документов.*

<sup>4</sup> Note that, contrary to verbs like *kažetsja*, that are often used as parentheticals, the predicative *interesno* does not lose the properties of the head predicate in constructions like this: its function of a discourse marker does not lead to a change in the syntactic position.

## 5. Factors of the choice of the strategy

Although the standard and the non-standard strategy appear to be free variants, some factors are statistically relevant for their distribution. The probability of the standard marking increases if the sentence contains a nominal subject-like pronoun (*это, одно*), a dative-marked experiencer or the complement clause is an infinitive construction.

### 5.1. The pronoun *это*

Zimmerling (2009) shows that in constructions with predicatives, the pronoun *это* can be used pleonastically. It is coreferent with the clausal complement (note that the precise syntactic position of *это* and the clausal complement, their identity or non-identity are irrelevant for the present work).

What is relevant, is that the use of *это* makes the standard use of tense more probable.

(19) *И это было хорошо, что уроки кончались раньше работы.* [Андрей Битов. Сад (1960–1963)]

Example (19) with the past tense is acceptable, though in general, past sounds strange with the predicative *хорошо* 'good'. The reason seems to be that the clause includes *И это было хорошо*, while the complement clause *что уроки кончались раньше работы* is added afterwards, and the syntactic position of the complement clause is unclear. If the initial form does not include an embedded clause, marked for tense, the clause *И это было хорошо* cannot be assigned the tense value if the past tense is not expressed.

### 5.2. Infinitive vs. finite clause

Another relevant feature is the strategy of complementation, chosen in the particular construction with a predicative. The variant without tense-marking and the past interpretation in the main clause is only possible if the embedded clause is finite. If infinitive is chosen, the tense value must be marked on the predicative. For instance, in (20), with the predicative *приятно*, and (21), with the lexeme *хорошо*, omission of the tense form of *быть* will change the interpretation: the speaker's perception will refer to the moment of speech.

(20) *Мне было приятно слышать эти слова.*

(21) *Так хорошо было сидеть здесь, пить кофе, слушать море, смотреть в ласковые глаза Георгия Васильевича...* [Дмитрий Быков. Орфография (2002)]

Some rare examples with a non-standard construction with infinitive are not fully acceptable and semantically unclear.

(22) *На некоторых крышах видны были люди, но неизвестно, как к ним добраться.* [Анатолий Кузнецов. Бабий яр (1965–1970)]

This opposition reveals an important thing about the nature of the non-standard interpretation. The non-standard (past-like) interpretation of the present tense appears to result from the fact that tense is marked in the finite complement clause. In this case, the past interpretation of the predicative is clear from the interpretation of the embedded clause. If the embedded clause contains infinitive, which does not distinguish tense values, the non-standard interpretation is impossible.

This mechanism is rather close to the relative tense-marking, which was described by Barentsen (1996), Schlenker (2003) Khomitsevich (2007), Say (2016), Letuchiy (2016). The notion of relative tense-marking covers cases when the tense form of the embedded verb denotes the temporal localization of the event with respect to the main situation. For instance, in (23), the present tense form *rabotaet* 'works' does not mean that the subject works in a bank at the moment of speech. The point is that he worked there at the time when the main situation (*dumal* 'thought') takes place:

(23) *Я думал, что он работает в банке.*

In contrast, under the absolute interpretation, the tense form in the embedded clause describes the localization of the event with respect to the speech act: before, after or simultaneously with *it*.

(24) *Я вошёл, когда все обедали.*

The relative interpretation is characteristic of Russian complement clauses, while the absolute one tends to occur in adjunct clauses, though both tendencies know some exceptions (see, for instance, Paducheva 2014 on the non-standard behaviour of the adverbial subordinator *poka*).

The situation with predicatives is almost opposite to the relative tense marking. The tense value is marked in the embedded clause, while the time of the main event is understood from the embedded clause. The difference is that relative tense marking does not presuppose that the tense value of the embedded verb *coincides* with that of the main verb (the latter is just counted from the former). By contrast, in the case under analysis, the tense value of the clause with a predicative coincides with the value of the embedded clause.

In other words, we can treat the non-standard (past-like) interpretation of present tense with predicatives as a case of 'accommodation' of the tense value in the main clause to the one in the embedded clause. Under this analysis, it becomes clear why this non-standard interpretation is mainly observed in cases, where the main clause is headed by a predicative, rather than a verb. Past tense remains unmarked with predicatives, this is why accommodation of the present tense value in the main clause to the past tense value of the embedded clause does not create any grammatical conflict between present tense and past tense markers. In contrast, if the main predicate is verbal, it is explicitly marked for the present tense, which makes accommodation to the past tense value very difficult.

### 5.3. The expression of experiencer

The expression of the experiencer is also relevant for the choice of the standard / non-standard marking. If the experiencer is expressed, the standard variant becomes better. Not surprisingly, a significant part of examples of tense-marking with the predicative *xorošo* (11 out of 56) contain a dative-marked experiencer.

(25) *И мне было хорошо, что я видел счастливую женщину.* [Фазиль Искандер. Сюжет существования (1965)]

(26) *И еще ему было хорошо, что, обидев Егоровну, он узнал ее по-настоящему.* [Владимир Железников. Последний парад (1985)]

The relevance of the experiencer expression can result from the degree of syntactic complexity of the main clause (this factor has been mentioned before, where we discussed the presence of a pronominal element). However, the same link can be addressed in semantic terms. If the experiencer is expressed, the interpretation of the situation changes: while the predicative *stranno* or *xorošo* without an experiencer is interpreted as denoting a general property of the situation, if the experiencer is included, the main clause denotes experiencer's perception, and, thus, its meaning becomes more subjective. The experiencer's state is linked to the precise temporal localization, whereas the generalized property, as mentioned above, characterizes the situation as such, with no link to the particular time or space. If something is strange or good, this is true even the situation has never occurred, or, on the contrary, has occurred and ended much time ago.

## 6. Coordination with past tense descriptions

The fact that the construction under analysis is interpreted in a special way is evident from its behaviour in coordinate constructions. In examples (2) and (4), formally present tense predicative constructions are coordinated with past tense verbal constructions. The same is shown in the example below:

(27) *Семен смотрел то на меня, то на Илью, и неизвестно, кто вызывал у него большее удивление.* [Сергей Лукьяненко. Ночной дозор (1998)]

Recall that the non-standard interpretation is awkward or impossible if the experiencer is expressed. Not surprisingly, coordinate constructions of the same type also become strange in presence of the explicit experiencer, as in (28) and (30):

(28) *?Я его не любил, и мне приятно, что приходил он редко.* (example (28) becomes fully acceptable if *мне приятно* is changed to *xorošo*):

(29) *Я его не любил, и хорошо, что приходил он редко.*

(30) *?Семен смотрел то на меня, то на Илью, и мне неизвестно / непонятно, кто вызывал у него большее опасение.*

An interesting question is whether the acceptability of coordinate constructions like (27) shows that predicatives like *neizvestno* are grammaticalized in contexts

with a non-standard interpretation and are syntactically modal markers, and not matrix predicates. Currently I do not have any reliable data to answer this question. In any case, it is important that in examples like (27), predicatives retain their syntactic pattern: they host an embedded clause with *čto* and not, for instance, occur as parentheticals.

## Conclusions

In the paper, I claimed that Russian predicatives fall into two groups. This division is based on their tendency to be marked or unmarked for tense in past contexts. The first group includes predicatives of various semantic classes, for instance, predicatives of emotion, physical perception, and so on ('unpleasant', 'cold', 'hot', and so on). The other group includes the evaluation class and similar meanings ('good', 'bad', 'strange'). Importantly, not only does the second group allow the past tense value to be unmarked, but sometimes the hypothetical past tense marking even looks awkward.

The facts under analysis can be accounted for both by semantic and by formal factors. On the one hand, evaluation is semantically possible even for those situations that are finished, have not been realized or simply are not highly relevant for the moment of speech. Thus, even in past contexts, the evaluation predicative can be used in present (= with a zero copula), because evaluation is made by the speaker at the moment of speech. On the other hand, this phenomenon is mainly characteristic of constructions with predicatives, and not with verbs. I supposed that this results from the fact that present tense remains unmarked with predicatives, but is morphologically marked with verbs. This is why the tense value of the unmarked form of predicatives can be accommodated to the tense value of the embedded verb. The same is impossible for synthetic and explicitly marked present tense forms of verbs. I showed that this case of accommodation is in a sense parallel to relative tense marking in Russian subordinate clauses.

The problem which has been discussed has theoretical implications lying beyond the domain of predicatives. The situation under analysis shows that classes of matrix predicates (evaluation predicates, emotion predicates and so on) differ in their preference to combine with a specific instance of an event (emotion predicates) vs. with a generalized event (evaluation predicates).

The future research perspective can include testing Russian verbs of emotion for their combinability (in the present tense) with different tense forms of the embedded predicate. This can constitute an important empirical test for the type of event (e.g., concrete / abstract, realized / non-realized) that the main predicate is compatible with. For instance, as shown before, evaluation predicatives can characterize an event irrespectively of whether it was realized and whether it is regarded as an abstract event or as a concrete instance of the situation.

Another theoretical conclusion is that the range of strategies of tense marking depends on the part of speech (and, perhaps, other properties) of the main predicate. This makes us doubt the parallelism of constructions with lexical verbs vs. zero copulas. Other syntactic features, which are said to characterize the Russian complex sentence, are to be tested against constructions with predicatives.

## References

1. *Barentsen Adrian* (1995), Shifting points of orientation in Modern Russian. Tense selection in 'reported perception', Theo Janssen & Wim van der Wurff (eds.), Reported speech: form and functions of the verb, John Benjamins, Amsterdam & Philadelphia, pp. 15–55.
2. *Khomitsevich Olga* (2007), Dependencies across phases. From sequence of tense to restrictions on movement. Ph.D. diss, University of Utrecht.
3. *Letuchiy Alexander B.* (2015), Russian zero copulas and lexical verbs: Similar or different. *Lingue e linguaggio* 2015, 2, pp. 233–250.
4. *Letučij Alexander B.* (2016), Patterns of complementation in Russian: core and periphery [Markirovanie sentencial'nyx aktantov v russkom jazyke: centr i periferija], Working papers of Russian Language Institute [Trudy IRYa RAN], 2016.
5. *Padučeva Elena V.* (2014), Expletive negation and semantics of the subordinator *poka* [Ėkspletivnoe otricanie i semantika sojuza *poka*], Plungjan V. A., Daniél' M. A., Ljutikova E. A. & Tatevosov S. G., Fedorova O. V. (eds.), Language. Constants. Variables. In the memory of Alexander Evgen'jevič Kibrik [Jazyk. Konstany. Peremennye. Pamjati Aleksandra Evgen'eviča Kibrika], Aleteja, Moscow, pp. 339–350.
6. *Say Sergey S.* (2016) Vremja v russkix finitnyx sentencial'nyx aktantax: nejtralizacija i točka otsčeta. Working papers of Russian Language Institute [Trudy IRYa RAN], 2016.
7. *Schlenker Philippe* (2003), A Plea For Monsters, *Linguistics and Philosophy*, 26, pp. 29–120.
8. *Testelefs Yakov G.* (2007), Structure of sentences with an unexpressed copula in Russian [Struktura predloženijs nevyražennoj svjazkoj v russkom jazyke], G. I. Kustova & R. I. Rozina (eds.), Dynamic models. Word. Sentence. Text. [Dinamičeskie modeli. Slovo. Predloženie. Tekst], Jazyki slavjanskix kul'tur, Moscow, pp. 773–789.
9. *Zimmerling Anton V.* (2003), Predicatives and qualitative adverbs: word classes and directions of derivation [Predikativy i kačestvennyje narečija: klassy slov i napravlenija derivacii], *Rusistics on the threshold of the 21st century* [Rusistika na poroge XXI veka: problemy i perspektivy], Jazyki slavjanskoj kul'tury, Moscow, pp. 54–59.
10. *Zimmerling, Anton.* (2009) Dative Subjects and Semi-Expletive pronouns in Russian, G. Zybatow, U. Junghanns, D. Lenertová, P. Biskup (eds.), *Studies in Formal Slavic Phonology, Syntax, Semantics and Information Structure*, Frankfurt am Main; Berlin; Bern; Bruxelles; New York; Oxford; Wien, 2009.

# ИТАЛЬЯНСКОЕ *MAGARI* И ЕГО РУССКИЕ ПЕРЕВОДНЫЕ ЭКВИВАЛЕНТЫ: РАЗНЫЕ ДИСКУРСИВНЫЕ СТРАТЕГИИ<sup>1</sup>

**Левонтина И. Б.** (irina.levontina@mail.ru)

ИРЯ РАН, Москва, Россия

**Денисова Г.** (galina.denissova@unipi.it)

Пизанский Государственный Университет, Италия

## ITALIAN *MAGARI* AND ITS RUSSIAN EQUIVALENTS: DIFFERENT DISCOURSE STRATEGIES

**Levontina I. B.** (irina.levontina@mail.ru)

RLI RAS, Moscow, Russia

**Denissova G.** (galina.denissova@unipi.it)

Pisa State University, Italy

This paper is aimed to analyze the Russian equivalents of the Italian focus particle “magari”. This lexeme has attracted much attention among linguists from different countries due to its especially intriguing polyfunctionality that sometimes knows no parallel in other languages. Russian equivalents of “magari” (extracted from the Russian-Italian subcorpus of the Russian National Corpus) clearly demonstrate that “magari” corresponds in Russian to a wide range of lexemes/units with different modality—equipotential non exclusion of factuality, concessivity, weakening of the illocutionary force of imperative, optative modality. The set of functions held by “magari” (non factual, non factual concessive, imperative, optative) also recalls in Russian the semantic network developed by several irrealis markers of non factuality. Moreover, in some contexts in the Italian translations from Russian “magari” appears while there is no concrete equivalent in the source language. In other words, the connotative range of “magari” is mostly achieved in Russian by different semantic mechanisms. Cross-linguistic analysis helps in clarifying the set of interlinguistic Russian-Italian correspondences of the lexeme “magari” and in circumscribing the different constructions and contexts where it occurs.

**Key words:** cross-linguistic equivalence, magari, parallel corpora, Russian language, Italian language, discourse markers

---

<sup>1</sup> Работа выполнена при поддержке РФФИ (грант 16-06-00339).

1. Контрастивные корпусные исследования дискурсивных слов позволяют не только сравнить собственно значения слов и их переводных эквивалентов, но и обнаружить различия в дискурсивных стратегиях, характерных для сравниваемых языков. В частности, описано уже много случаев, когда в русском переводе частицы возникают там, где в иноязычном оригинале никакого дискурсивного показателя не было (феномен прагматической обязательности частицы; см. [Levontina, Shmelev 2007]). В настоящей работе сравнение *magari* с русскими переводными эквивалентами и использование этой частицы в переводах с русского позволяет сделать важные выводы, касающиеся дискурсивных стратегий русского языка в отличие от итальянского. Работа выполнена на материале русско-итальянского параллельного корпуса Национального корпуса русского языка ([www.ruscorpora.ru](http://www.ruscorpora.ru)).

2. Яркое и загадочное итальянское слово *magari* неоднократно становилось объектом внимания лингвистов, см. [Arcaini 1997; Pietrandrea 2007; Schiemann 2008; Masini, Pietrandrea 2010]. В работе [Masini, Pietrandrea 2010: 78] общая идея *magari* описывается как «commonly understood as an epistemic adverb, as a generic marker of non factuality, more precisely as a marker of “non exclusion of factuality”».

Приведем основные значения *magari*, которые выделяются в работе [Masini, Pietrandrea 2010]:

I. Значение возможности:

*Magari è a casa*  
'Возможно, он/она дома'

*Magari* более утвердительно (в сравнении с *forse* и *chissà*):

*Forse è venuto ieri, ha passato qui tutto il pomeriggio e magari si è fermato a dormire*  
'Может быть, он приходил вчера, провел здесь весь день, а, возможно, остался на ночь'

II. В уступительном контексте:

*Magari è intelligente, ma non è abbastanza preparato*  
'Может, он и умен, но недостаточно подготовлен'

III. В повелительном наклонении с аттенуативным значением:

*Magari parlagliene tu!*  
'Поговорил бы ты с ним!'

IV. Для выражения пожелания (при употреблении отдельной репликой):

*Vorrei tanto vedere questo film. Magari ne facessero ancora!*  
'Как бы мне хотелось посмотреть этот фильм! Вот бы его показали еще раз!'

V. В контексте восклицания/опатива (при употреблении отдельной репликой):

A: *Vuoi un po' di riposo?*  
'Отдохнуть хочешь?'



В: *Magari!*

‘Хорошо бы! / О если бы я мог!’

VI. В контексте предположения, соответствующего реальности с большей или меньшей долей вероятности:

*Non potrebbe essere uscito con un amico? Non sarà magari con suo fratello?*

‘Может, он пошел куда-то с другом? А может, с братом?’

Мы попытались сравнить русские переводы итальянских контекстов с использованием *magari*, с одной стороны, и русские контексты, в переводе которых фигурирует итальянское *magari* — с другой. Сравнение с контекстами употребления итальянского *magari* позволяет оценить важные дискурсивные характеристики русских текстов.

3. Самый частый и, казалось бы, самый простой случай, — это контексты, в которых выражена **идея возможности**.<sup>2</sup> В переводах с итальянского на русский и с русского на итальянский в большинстве случаев *magari* соответствуют слова *может, мог бы, может быть, возможно*:

- (1) — *Может, скоро будет вылетать очередная экспедиция на далекий холодный Юг нашего маленького глобуса?* (А. Курков “Закон улитки”, 2005). [*Magari presto sarebbe partita una delle solite spedizioni dirette al gelido estremo sud del nostro piccolo globo*] (А. Kurkov “I pinguini non vanno in vacanza”, tr.: В. Osimo).
- (2) — *Doveva essere un posto incredibile, magari ci viveva qualche animale strano* (Niccolò Ammaniti “Io non ho paura”, 2001). [*Нечто притягательное было в этом холме, казавшемся невероятным, может быть, там обитало какое-нибудь странное животное*] (Н. Амманити “Я не боюсь”, пер.: В. Николаев).
- (3) — *Ведь ты мог бы оказаться человеком, которого прислали, чтобы меня убить...* (А. Курков “Закон улитки”, 2005). [*Magari ti avevano mandato qui per farmi fuori...*] (А. Kurkov “I pinguini non vanno in vacanza”, tr.: В. Osimo)

При этом итальянская частица *magari* фигурирует в переводах русских текстов в случаях, когда в оригинале дискурсивные показатели предположительности отсутствуют:

- (4) *Я подумала: «Просто устали, спят. Или боятся брать трубку»* (А. Политковская. “Путинская Россия”, 2004). [*Pensai che stessero dormendo, stanchi com'erano. O che magari avessero paura dirispondere*] (А. Politkovskaja “La Russia di Putin”, tr.: С. Zonghetti).

<sup>2</sup> Причем речь может идти о разных видах возможности (алетической, деонтической, эпистемической), но здесь мы не будем на этом подробно останавливаться, так как для наших целей это несущественно.

Как мы видим, предположительность в русском оригинале формально не выражена (автор ограничивается использованием ментального предиката), однако итальянский переводчик вводит частицу. Такая ситуация достаточно типична.

Однако необходимо заметить, что в русском языке чрезвычайно дробная шкала предположительности: от почти полной уверенности: *навверняка*, как *пить дать, к гадалке не ходи, наверно*, в меньшей степени *вероятно, вероятнее всего, скорее всего* — до средних значений: *возможно, может быть, пожалуй* и маркеров робкого предположения: *а что если, а вдруг, случайно не, чем черт не шутит* и т. п.<sup>3</sup> В литературе отмечается и еще одно важное деление внутри вводных показателей достоверности. Ментальные предикаты делятся на предикаты **мнения** (типа *считать*) и предикаты **знания** (типа *знать, видеть, чувствовать*). Соответствующее деление имеется и во вводных словах. Вводные слова, выражающие мнение, — это *возможно, наверно*. А слова, выражающие знание, — это *кажется, явно, определенно, как будто, вроде бы* [Яковлева 1988; Падучева 1996].

Если в русском тексте используется слово типа *навверняка*, то оно, естественно, переводится на итальянский как *magari*. Но интересно, что и итальянское *magari* часто переводится русскими словами, определенно указывающими на высокую степень уверенности:

- (5) *Dato che lei pensa di conoscermi bene, saprà magari che io sono persona che non si fa pigliare per il culo* (A. Camilleri “Il cane di Terracotta”, 1996).  
[Раз уж вы думаете, что хорошо меня знаете, то **навверняка** знаете то, что не такой я человек, чтоб позволить держать себя за дурака] (А. Камиллери “Собака из терракоты”, пер.: А. Кондюрина)<sup>4</sup>.

В других же случаях переводчик выбирает менее радикальные варианты:

- (6) *Meglio della memoria vera perché quella, magari a prezzo di duro esercizio, impara a ricordare ma non a dimenticare* (U. Eco “Il pendolo di Foucault”, 1988).  
[Эта память лучше настоящей, потому что настоящая, **вероятно** благодаря суровой школе, научена запоминать, но не научена забывать] (У. Эко “Маятник Фуко”, пер. Е. Костюкович).
- (7) *Doveva conoscere il modo di fare entrare e uscire dalla cinta, e sapere che in cucina ci sarebbero state quelle frattaglie (e magari domani si sarebbe detto che, la porta restata aperta, un cane era entrato e se le era mangiate* (U. Eco “Il nome della rosa”, 1980).  
[Кроме того, он знает, что на кухне есть свежий ливер (а наутро **скорее всего** решили бы, что в незапертую дверь пробралась собака и украла его)] (У. Эко “Имя розы”, пер.: Е. Костюкович).

<sup>3</sup> О маркерах со значением «оценка достоверности» см. [Виноградов 1947: 739].

<sup>4</sup> С примерами из произведений Андреа Камиллери следует работать осторожно. Дело в том, что этот автор часто (и намеренно) использует диалектизмы и/или стилизует свой язык под диалект. Отсюда довольно частое присутствие *magari* в значении (а иногда и в форме) сицилийского (и шире — южноитальянского) *macari* в значении *perfino, anche* (даже, а также).

Здесь **переводчик вынужден оценивать степень вероятности**, исходя из собственного понимания ситуации. В итальянских контекстах нет никаких формальных маркеров, позволяющих различить, имел ли автор в виду *может быть* или *наверняка*.

Используется *magari* и при передаче маркеров робкого предположения и даже слабой вероятности:

- (8) *И вдруг лет через десять мы встретимся, — говорил он* (А. П. Чехов “Рассказы”, 1885–1903).  
[*E magari, tra una decina d'anni, un bel giorno ci incontreremo, diceva lui*] (А. Čechov “Racconti”, tr.: F. Malcovati)].
- (9) *Но сегодняшний Иван уже значительно отличался от Ивана вчерашнего, и первый путь показался ему сомнительным: чего доброго, они укоренятся в мысли, что он буйный сумасшедший* (М. Булгаков “Мастер и Маргарита”).  
[*Ma l'Ivan odierno era ormai molto diverso dall'Ivan del giorno innanzi, e la prima via gli sembrò dubbia: magari in loro si sarebbe radicata l'idea che lui era un pazzo furioso*] (М. Bulgakov “Il Maestro e Margherira”, tr.: V. Dridso).
- (10) — *Не Агнеса Шеридан, случайно?* (В. Набоков “Лолита”).  
[— *Agnes Sheridan, magari?*] (V. Nabokov “Lolita”, tr.: Giuli Arborio Mella).
- (11) — *Разве я не могу полюбить?* — промолвила она. — **Едва ли!** Только я напрасно назвал это несчастьем (И. С. Тургенев “Отцы и дети”).  
[— *Non posso amare forse? — Magari! Ho fatto male però a parlare di sventura*] (I. Turgenev “Padri e figli”, tr.: F. Verdinois, 2012).

Заметим в скобках, что русский язык тут различает среди маловероятных событий **желательные** (*чем черт не шутит*<sup>5</sup>) и **нежелательные** (*чего доброго*); при этом, *а вдруг, а что если* применимы к обеим ситуациям. А итальянское *magari*, хотя этимологически связано с идеей счастья<sup>6</sup>, не различает благоприятные и неблагоприятные возможности.

4. Можно продолжать приводить примеры, но уже ясно, что итальянское *magari* покрывает весь спектр предположительности, а в русском языке ему соответствует большой арсенал средств, среди которых нет нейтрального, которое бы просто выражало идею предположительности без детализации степени. Говоря по-русски, человек обязан, указывая на предположительность, определить степень вероятности события (и, возможно, некоторые другие его свойства). При этом следует отметить, что **обязательно выражаемые смыслы**,

<sup>5</sup> Как справедливо отметил один из рецензентов «Диалога», это выражение может использоваться и применительно и к нежелательной ситуации, однако такое нехарактерно: *Опасность, правда, невелика, пожалуй, сомнительна, не прямой родственник Аграфены, а племянник, но ведь чем черт не шутит?* (Ю. Трифионов).

<sup>6</sup> *Magari* — от греч. *makári, makárie, makários* [благословенный, счастливый].

те параметры ситуации, которые не могут быть не отражены, — это, может быть, главное, в чем проявляется лингвоспецифичность в области семантики.

Важно также иметь в виду, что *tagari* часто соответствуют и другие довольно специфические показатели предположительности.

Приведем всего два примера.

В следующей русской фразе используется конструкция *X не X* со значением: ‘*X* или, возможно, не *X*, но очень кто-то похожий на *X*-а (что-то очень похожее на *X*)’. В итальянском переводе этому соответствует просто *tagari*:

- (12) *Так слушайте же: вчера вечером я на Патриарших прудах встретился с таинственной личностью, иностранцем не иностранцем, который заранее знал о смерти Берлиоза и лично видел Понтия Пилата* (М. Булгаков “Мастер и Маргарита”).

[*Allora ascolti: ieri sera, agli stagni Patriaršie, ho incontrato un personaggio misterioso, magari straniero, che sapeva in anticipo della morte di Berlioz e aveva visto personalmente Ponzio Pilato*] (М. Bulgakov “Il Maestro e Margherira”, tr.: V. Dridso).

В переводе, тем самым, смысл упрощается: остается предположительность, а все смысловые тонкости сняты.

Очень интересен следующий пример:

- (15) *Magari, con gli occhi della mente, la vede nuotare fluida nel silenzio, immagina il suo corpo bagnato ricoprirsi di un riflesso come di stagnola mentre esce dall'acqua e si asciuga minuziosamente* (G. Faletti “Io uccido”, 2002).

[*Пожалуй, он даже представляет себе, как она плавала в тишине, представляет ее влажное, сверкающее, словно фольга, тело, когда она выходит из воды и тщательно вытирается*] (Дж. Фалетти “Я убиваю”, пер.: И. Константинова).

В русском переводе использовано слово *пожалуй* (точнее, сочетание *пожалуй даже*). *Пожалуй* имеет специфическое “риторическое” (в смысле статьи Л. Н. Иорданской о риторических союзах типа *раз*) значение; оно не указывает на предположительность, а отражает **колебания говорящего в процессе осуществления речевого акта**, после которых он допускает возможность соответствующей ситуации<sup>7</sup> [Иорданская 1988, 1992; Разлогова 1998; Падучева 2011].

При этом *пожалуй*, как кажется, лишь приблизительно отражает функцию *tagari*. Скорее всего, *tagari* в примере (15) маркирует переход повествования в мир мечты, фантазии или воспоминаний. Ср.:

- (16) *C'era solo da sperare che Gegè, in ricordo degli anni trascorsi l'uno al lato all'altro sullo stesso banco delle elementari, amicizia continuata poi magari quando s'erano fatti grandi, noi si fosse risolto, per interesse suo, a venderlo come carne da porco* (A. Camilleri “Il cane di Terracotta”, 1996). [*Можно было только надеяться, что*

<sup>7</sup> В этом отношении с *пожалуй* сходно слово *впрочем*, а также *хотя* во фразах типа: *Я не пойду... хотя ладно, идем*.

*Джедже, в память о годах, проведенных за одной с ним партией в начальной школе, — дружба их продолжалась и когда они выросли, — не собирался продать его на корм свиньям (А. Камиллери “Собака из терракоты”, пер. А.: Кондюрина)].*

Если в русский перевод вставить маркер возможности в часть, соответствующую итальянскому *amicizia continuata poi magari quando s'erano fatti grandi*, — получится, что выражается сомнение в продолжении дружбы. Но по-итальянски здесь просто дополнительный маркер того, что автор говорит не о фактах. Частица здесь относится к надежде, а продолжение дружбы — как раз факт, поэтому частица не переводится.

Показателен в этом отношении пример из набоковской “Лолиты”. В русском тексте герой предается мечтам, но ни грамматически, ни лексически это не маркируется<sup>8</sup>. Однако в итальянском тексте, как и в английском, используется ирреалис, а в итальянском к тому же фигурирует *magari*.

- (17) *...мне пришло в голову, что продолжительные роды, с основательным кесаревым сечением и разными другими осложнениями, в укротном родильном приюте, этак будущей весной, дадут мне возможность побыть наедине с моей Лолитой несколько недель сряду — и закармливать размянную нимфетку снотворными порошками (В. Набоков “Лолита”).*

[...mi frullò per la testa che un parto difficile in un sicuro reparto ostetrico, con un bel cesareo e altre complicazioni, mi avrebbe dato a primavera l'opportunità di restar solo con la mia Lolita magari per settimane — e di ingozzare l'inerte ninfetta di sonniferi] (V. Nabokov “Lolita”, tr.: Giulio Arborio Mella). Cp.: [... it occurred to me that a prolonged confinement, with a nice Cesarean operation and other complications in a safe maternity ward sometime next spring, would give me a chance to be alone with my Lolita for weeks, perhaps — and gorge the limp nymphet with sleeping pills].

Ситуация погружения в мир воспоминаний — типичный случай, когда в итальянском есть маркер *magari*, в то время как в соответствующем русском тексте маркер отсутствует:

- (18) *I pomeriggi trascorsi al Barchette del Duca in ambiente saturo di memorie rinascimentali, col tè delle cinque preso in compagnia della famiglia al completo — e la signora Olga rientrava dal parco molto spesso a quell'ora, le braccia piene di fiori — nonché più tardi, magari, su in biblioteca, godendo fino al buio della dotta conversazione del professor Ermanno, quei pomeriggi straordinari rappresentavano evidentemente per lui qualcosa di troppo prezioso perché non ne facesse materia anche con noi di continui discorsi e divagazioni (G. Bassani “Il giardino dei Finzi-Contini”, 1962).*

[Вечера, которые он проводил в «Лодочке герцога», в обстановке, наполненной духом Возрождения, чай, который он пил в пять часов в кругу семьи (даже синьора Ольга часто приходила в этот час из сада с огромным букетом цветов), ученая беседа с профессором Эрманно,

<sup>8</sup> Только приближительность срока отмечается при помощи слова *этак*.

*которой он наслаждался потом в библиотеке до темноты, — эти вечера имели для него, несомненно, особую ценность и служили ему темами для постоянных рассказов и лирических отступлений во время уроков*] (Дж. Бассани “Сад Финци-Контини”, пер.: И. Соболева).<sup>9</sup>

Нельзя сказать, что русский язык не умеет вычленять и маркировать дискурс мечты и воспоминаний, однако это происходит на совсем ином уровне — **интонационном**.

В книге [Янко 2008: 109] описывается интонационный контур, названный «интонацией ментальной деятельности, т.к. этот интонационный тип направлен на отражение средствами интонации разного рода информационных и мыслительных процессов».

В этой связи упоминаются припоминание, недоумение, погружение в мечты, а также и передача чужой речи (*Тетя сказала, надо чего-то там уколды делать*). Эта интонация описывается так: «Соответствующий акцент характеризуется подъемом тона и существенным удлинением ударного слога акцентоносителя ремы. Вся заударная область ровная (иногда с небольшим естественным падением)».

Поразительно, что, если проанализировать случаи, когда в итальянском тексте есть *magari*, а в соответствующем русском (при переводе в обе стороны) этой частицы нет, кажется, что в большинстве случаев это как раз будут разные типы **контекстов ментальной деятельности**<sup>10</sup>. Например, это может быть **планирование**:

- (19) *Оставалось побеседовать со знаменитостями. Ну, и кое-что послушать — так, для общего развития* (С. Довлатов “Филиал”, 1989).  
[*Mi restavano da fare due chiacchiere con le celebrità e magari anche da ascoltare qualcosa, così, per aggiornamento personale*] (S. Dovlatov “La filiale”, tr.: L. Salmon).
- (20) *И вот Леша скурил чинарик и решил ехать в редакцию, а если что, вернуться*. (Ю. Дружников “Ангелы на кончике иглы”, 1988).  
[*Ma ormai Aleksej l'aveva fumata e decise di andare in ufficio per poi magari tornare, se era il caso*]. (Ju. Druznikov “Angeli sulla punta di uno spillo”, tr.: F. Aceto).

Или просто **воображение**:

- (21) *Gente che si è sporta troppo, un capogiro, magari...* (D. Buzzati “Eppure battono alla porta”, 1942).  
[*Подойдешь слишком близко — и голова закружится...*] (Д. Быццати “И все же стучат в дверь”, пер.: И. Смагин).

<sup>9</sup> Заметим, что при описании погружения в мир воспоминаний для английского характерно использование *would*.

<sup>10</sup> Кстати, очень часто в русском языке соответствующие фразы естественно прочтешь именно с указанной интонацией.

Ср. также следующий пример:

- (22) *Ну а вы уж сами сделаете выводы, прикинув ситуацию на себя*  
(А. Политковская “Путинская Россия”, 2004).  
[*A voi trarre le conclusioni, magari cercando di mettervi nei nostri panni*]  
(А. Politkovskaja “La Russia di Putin”, tr.: С. Zonghetti).

Здесь появление *magari* можно объяснить тем, что возникает сама идея гипотезы, отступления от фактов. Дальше Политковская поясняет, о какой именно мыслимой ситуации идет речь: “Хотели бы лично вы жить в ТАКОЙ стране и исправно платить налоги на ПОДОБНУЮ армию?” — однако существенно, что в итальянском переводе *magari* появляется не в этой фразе, а раньше, как только возникает сама идея гипотетического построения.

5. Разумеется, в небольшой статье невозможно рассмотреть все критические точки, в которых сопоставление контекстов *magari* с их русскими соответствиями позволяет увидеть важные различия в языковой концептуализации действительности. Так, чрезвычайно показательны **уступительно-противительные** контексты, где итальянский не требует обязательного уточнения характера отношения между ситуациями, а для русского без такого уточнения фраза стала бы аномальной:

- (23) *Magari era lì che si stava aggiustando i pantaloni... alzava la testa un attimo, buttava un occhio verso il mare...*” (Alessandro Baricco “Il Novecento”, 1994).  
*Даже если он поправлял брюки... он вдруг поднимал голову, бросал взгляд на море... и видел ее...*” (Алессандро Барикко “Легенда о пианисте”, пер.: Наталья Колесова)].

Интересно использование *Magari!* как отдельной реплики и для подтверждения (*Еще бы!*), и для опровержения (*О если бы!*, *Едва ли!*), при этом такие употребления различаются интонационно.

Отдельного рассмотрения в будущем потребуют следующие случаи употребления *magari*.

- в контексте предположения, соответствующего реальности с большей или меньшей долей вероятности:
- (24) *Magari è a casa* — *Может, он дома / Вдруг (!) он дома / Хотя бы он был дома* (в зависимости от контекста и с различной интонацией),
- в повелительном наклонении с аттенуативным значением:
- (25) *Magari parlagliene tu!* — *Поговорил бы ты с ним / Сам бы с ним поговорил / Поговори-ка с ним сам!*,
- в контексте восклицания / оптатива:
- (26) *Magari ne facessero ancora* — *Побольше бы таких фильмов.*

Однако и уже рассмотренный материал, как кажется, позволяет сделать важные предположения о различии дискурсивных стратегий, типичных для

русского и итальянского языков. В русском языке чрезвычайно дробная шкала предположительности и степеней уверенности и высокая степень обязательности детализации по этой шкале. Со своей стороны, итальянский, по сравнению с русским, подразумевает в целом более эксплицитное маркирование предположительности и подобных модальностей. В частности, для итальянского характерно маркирование перехода к дискурсу ментальной деятельности (мира мечты, воспоминаний, фантазий и планов), который в русском языке если и маркируется, то (в устной речи) интонационно.

## References

1. *Arcaini E.* (1997), The connective particle *magari* from the comparative perspective [Le connecteur *magari* dans une perspective comparative]. Wotjak G. (Ed.), Studien zum romanisch-deutschen und innerromanischen Sprachvergleich. Peter Lang, Frankfurt am Main, pp. 59–76.
2. *De Mauro T.* (2000), Italian Dictionary [Il dizionario della lingua italiana], Paravia.
3. *De Mauro T., Mancini F., Vedovelli M., Voghera M.* (1993) Frequent Words in the Italian spoken [Lessico di frequenza dell'italiano parlato], ETAS Libri, Milano.
4. *Jordanskaja L.* (1988) The semantics of Russian conjunction *raz* (in the comparison with other Russian conjunctions) [Semantika russkogo sojusa *raz* (v sravnenii s nekotorymi drugimi russkimi sojuzami)], Russian Linguistics, 12, pp. 239–267.
5. *Jordanskaja L.* (1992), Performative verbs and “rhetoric” conjunctions [Performativnye glagoly i ritoricheskiye sojuzy] // Reuther T. (Ed.) Festschrift for Viktor Jul'evich Rozencvejk. Wien: WSA [Sonderband 33]. P. 29–41.
6. *Jakovleva E.* (1988), Pragmatics and problems of intensionality [Progmatika i problemy intensional'nosti], Moskva, pp. 278–302.
7. *Janko T.* (2008), Intonation strategies in Russian in the comparative aspect [Intonacionnye strategii russkoj rechi v sopostavitel'nom aspekte], Moskva.
8. *Masini F., Pietrandrea P.* (2010), *Magari*, Cognitive Linguistics, 21 (1), pp. 75–121.
9. *Paducheva E.* (1996), Semantics Studies. Semantics of Narrative [Semanticheskie issledovanija. Semantika narrativa], Moskva.
10. *Paducheva E.* (2011) Egocentric valencies and the reconstruction of a speaker [Egocentricheskije valentnosti i rekonstrukcija govoriaschego] // Voprosy jazykoznanija, №3, cc. 3–18. <http://lexicograph.ruslang.ru/TextPdf1/egocentrals.pdf>
11. *Pietrandrea P.* (2007), The grammatical nature of some epistemic-evidential adverbs in spoken Italian, Italian Journal of Linguistics, 1/2007, pp. 39–64.
12. *Razlogova E.* (1998), *Pozhaluj* or Duality [Pozhaluj, ili razdvoennost'], Discourse markers in Russian [Diskursivnye slova russkogo jazyka], Metatekst, Moskva, pp. 331–336.
13. *Schiemann A.* (2008), The polysemy of “magari” and “forse”. Corpus based italian C-ORAL-ROM analysis. [La polisemia di magari (e forse). Analisi corpus based su C-ORAL-ROM italiano]. Cresti E. (Ed.), Perspectives in the Italian Vocabulary Studies [Prospettive nello studio del lessico italiano]. Firenze University Press, Florence, vol. I, pp. 299–307.
14. *Vinogradov V.* (1947) Russian Language [Russkij jazyk. Grammaticheskoe uchenie o slove], Moskva.



# АННОТИРОВАНИЕ РУССКИХ МАНУАЛЬНЫХ ЖЕСТОВ: ТЕОРЕТИЧЕСКИЕ И ПРАКТИЧЕСКИЕ ВОПРОСЫ<sup>1</sup>

**Литвиненко А. О.** (allal1978@gmail.com)<sup>1</sup>,  
**Николаева Ю. В.** (julianikk@gmail.com)<sup>2,1</sup>,  
**Кибрик А. А.** (aakibrik@gmail.com)<sup>1,2</sup>

<sup>1</sup>Институт языкознания РАН, Москва, Россия;

<sup>2</sup>МГУ имени М. В. Ломоносова, Москва, Россия

Доклад относится к области мультимодальной лингвистики и посвящен аннотированию мануальных жестов в мультимедийном (мультимодальном) дискурсе на материале создаваемого русского корпуса «Рассказы и разговоры о грушах». Рассматривается задача системного вычленения коммуникативно значимых жестов из общего кинетического фона на основании кинетических признаков (усилие, скорость, траектория и т. д.). Вводится понятие жестикуляционного портрета говорящего, включающего такие характеристики, как склонность к подвижности/неподвижности и к частому использованию физиологически мотивированных движений (адапторов), а также набор привычных положений, служащих точками отсчета для значимых движений. Предлагается аннотационная схема для среды ELAN, позволяющая системно аннотировать иерархические жестовые единицы — жестовые цепочки, отдельные жесты и их фазы, — в том числе с учетом асинхронного движения рук. Аннотация ориентирована на сопоставление мануальных жестов с единицами других коммуникативных каналов. Приводятся количественные наблюдения о предпочтениях в области руки жестов.

**Ключевые слова:** мультимодальный дискурс, корпус, мануальный жест, аннотация, аннотационная схема

## ANNOTATION OF RUSSIAN MANUAL GESTURES: THEORETICAL AND PRACTICAL ISSUES

**Litvinenko A. O.** (allal1978@gmail.com)<sup>1</sup>,  
**Nikolaeva Ju. V.** (julianikk@gmail.com)<sup>2,1</sup>,  
**Kibrik A. A.** (aakibrik@gmail.com)<sup>1,2</sup>

<sup>1</sup>Institute of Linguistics RAS, Moscow, Russia;

<sup>2</sup>Lomonosov Moscow State University, Moscow, Russia

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ (проект № 14-18-03819).

This paper contributes to the field of multimodal linguistics and discusses annotation of manual gestures in multichannel (multimodal) discourse, using a corpus of Russian discussions of the Pear Film (Chafe ed. 1980); the corpus is currently under construction ([www.multidiscourse.ru](http://www.multidiscourse.ru)). The study is based on a fully annotated subcorpus of 3 recordings, total time about one hour, with three participants in each recording. One of the key issues in gesture annotation is the importance of systematic distinction between communicatively significant gestures and general kinetic background. To address this issue, we introduce the notion of a speaker's gesticulation portrait, involving properties such as (dis)inclination to stillness and to frequent use of physiologically motivated self-adaptors, as well as a set of habitual postures ("neutral hand positions"), serving as starting points for significant movements. An annotation scheme for the ELAN environment is proposed, allowing one to systematically annotate hierarchically organized gestural units as annotations in independent ELAN tiers: gesture chains, individual gestures, and gesture phases. Gesture phases are annotated separately for the left and the right hand; an individual gesture (sometimes called "gesture phrase") is treated as a combined unit that can be one-handed or two-handed. To divide an ongoing stream of gesticulation into a series of gestures, annotators identify the points where kinetic features change. A simultaneous change of two or more parameters (effort, velocity, trajectory, movement direction, hand shape and orientation, location in gesture space, etc.) marks a phase boundary; a change of several features points to a gesture boundary. A set of subordinate ELAN tiers is used to describe each gesture, including handedness, phase structure, multi-strokes, gesture overlaps, gesture repetitions, etc. As left and right hands movements are often asynchronous, a set of formal rules is proposed for such instances. On the basis of the annotated subcorpus evidence, quantitative observations were made concerning speakers' preferences in gesture handedness. Three out of nine participants prefer (>60%) two-handed gestures, while six participants prefer either right- or left-handed gestures most of the time. While all the speakers identified themselves as right-handed persons, two of them show preference for the left hand (>63%); the degree of right hand preference varies strongly (from 60% to 97%). This data suggests that adding the "dominant hand" parameter for two-handed gestures may be quite useful. The proposed annotation scheme is oriented towards comparing manual gestures with units of other communication channels.

**Key Words:** multimodal discourse, multimodal corpus, manual gesture, annotation, annotation scheme

## 1. Введение

Данная работа относится к области мультимодальной лингвистики — научного направления, признающего принципиальное равноправие различных формальных средств, используемых в естественной коммуникации (см. напр. Gibbon et al. eds. 2000, Kress 2002, Granström et al. eds. 2002, Scollon 2006, Кибрик 2010, Müller et al. eds. 2013). Наряду с вокальной (слуховой) модальностью, в рамках которой представлены вербальная структура и просодия, в коммуникации активно используется кинетическая (зрительная) модальность, включающая ряд каналов и компонентов: мануальные жесты, головные жесты, телесные жесты, мимика, движения глаз и т. д. (ср. Крейдлин 2002, Бутовская 2004). Авторы настоящей работы участвуют в коллективном проекте Института

языкознания РАН, целью которого является сбалансированное описание и исследование множества коммуникативных каналов, используемых в дискурсе. В этом докладе мы остановимся на одном из аспектов этого исследования, а именно на аннотировании мануальных жестов.

В рамках проекта создается мультимедийный (мультимодальный) корпус «Рассказы и разговоры о грушах». Корпус основан на записях естественной коммуникации в группах из четырех участников, обсуждающих «Фильм о грушах» У. Чейфа, см. Chafe ed. 1980; [www.linguistics.ucsb.edu/faculty/chafe/pearfilm.htm](http://www.linguistics.ucsb.edu/faculty/chafe/pearfilm.htm). Каждое обсуждение, или «запись», включает фрагменты монологов и общего разговора. Корпус состоит из 24 записей общей длительностью около 9 часов и объемом около 100 тыс. словоупотреблений. Каждая запись представлена в корпусе в виде комплекта медиафайлов (видео- и аудиофайлы, а также данные айтрекинга) и комплекта аннотаций. Подробнее см. [www.multidiscourse.ru](http://www.multidiscourse.ru).

Мультимедийный дискурс не может стать объектом научного анализа, пока он не подвергнут **аннотированию**. В корпусе реализуется аннотирование следующих основных компонентов: вокальный (вербально-просодический) компонент, жесты рук, жесты головы, окуломоторное поведение. Аннотирование каждого из них предполагает выделение релевантных объектов и разметку их свойств. Так, если мы рассматриваем звучащую речь, при аннотировании (транскрибировании) выделяются такие виды объектов, как элементарные дискурсивные единицы, паузы, дискурсивные акценты и др., и им приписываются свойства (тип ЭДЕ, длительность паузы, направление тона в акцентах и т.д.); см. Кибрик и Подлесская ред. 2009; [www.spokencorp.org.ru](http://www.spokencorp.org.ru). Аналогично, в потоке жестикуляции должны быть идентифицированы те или иные релевантные единицы и их свойства. С теоретической точки зрения, важно выделить конституирующие явления жестикуляции; с практической точки зрения, необходимо выбрать программную среду и разработать технические решения — аннотационную схему.

Данная статья посвящена аннотированию **мануальных жестов**, то есть коммуникативно значимых движений рук, сопровождающих речь. Подавляющее большинство мануальных знаков в коммуникации составляют иллюстративные жесты, имеющие спонтанную природу: для них не существует фиксированного словаря, заранее заданных форм и правил сочетания, и в этом отношении они контрастируют с языковыми знаками (Fricke, Freiburg, Chemnitz 2013). При этом иллюстративные жесты играют значительную роль в коммуникации, дополняя или уточняя содержание вербального сообщения, а иногда и противореча ему, в связи с этим изучение жестов существенно расширяет наше понимание языка.

## 2. Мануальные жесты как объект лингвистического анализа

Один из первых вопросов, с которым сталкивается любой исследователь жестикуляции, — это вычленение жестов из общей массы движений,

совершаемых участником коммуникации. Граница может проводиться как на функциональных, так и на формальных основаниях. Одним из базовых свойств жеста считается **коммуникативная информативность** в заданном контексте, причем одни исследователи предполагают ее намеренность (Kendon 2004: 13–14), другие настаивают на нецеленаправленной и/или неосознаваемой природе жеста (McNeill 2013). Важными свойствами жеста являются его контролируемость (говорящий волен не совершать жеста) и отсутствие явной практической цели (Гришина, в печати). В рамках формального подхода определяющим свойством жеста является **выраженное артикуляционное (двигательное) усилие** на общем фоне кинетического поведения (Müller, Bressemer, Ladewig 2013), и это усилие воспринимается и распознается адресатом как свидетельство коммуникативной значимости. Авторы настоящей работы считают наиболее конструктивным сочетание функционального и формального подходов. В любом случае важно уметь определять для конкретного говорящего кинетический «фон» (слабые и/или незначимые движения), по сравнению с которым жест будет «фигурой» (значимым движением).

Существенную часть этого фона составляют физиологически мотивированные движения — **адапторы** (т. н. self-adaptors, см. Ekman, Friesen 1969), связанные с поиском более удобного положения, снятием телесного дискомфорта и т. д. Их можно разделить на две большие группы. Первая — движения с явной практической целью: почесаться, поправить одежду и т. д. Вторая — мелкоамплитудные движения (обычно многократные), не имеющие явной цели: шевеление пальцами, потирание рук и т. п. Такие движения П. Экман называл также манипуляторами и трактовал как показатель беспокойства или просто привычного действия (Ekman 1999: 43). В первом случае адапторы отличаются от жестов наличие практической цели (ср. жест на рис. 1 и адаптор на рис. 2). Во втором случае адапторы характеризуются меньшей амплитудой, более простой и/или менее четкой траекторией, низким или отсутствующим ускорением, отсутствием явного направленного усилия.

Важной составляющей кинетического поведения являются **нейтральные положения** — те, к которым говорящий регулярно возвращается. Как правило, нейтральных положений у говорящего более одного, и они организованы иерархически. Абсолютное нейтральное положение показывает минимальную ситуативную вовлеченность в коммуникацию, например: спокойная поза, взгляд обращен не на собеседника, руки лежат на коленях. Нейтральное положение среднего уровня является отправной точкой для продолжительных фрагментов коммуникации: например, один из участников подался к собеседнику, опершись локтями на колени. Наконец, локальное нейтральное положение — то, к которому происходит возврат в пределах небольшого фрагмента дискурса, между отдельными жестами. Характерный пример — сложенные перед собой на весу руки, см. рис. 3. Такое положение вне контекста могло бы считаться маркированным, однако в контексте служит точкой отсчета и разграничивает жесты или их последовательности.



Рис. 1. Жест 'груши'



Рис. 2. Адаптор «почесать нос»



Рис. 3. Локальное нейтральное положение

Еще одно важное понятие, относительно которого оценивается всякий жест, — **неподвижность**: отсутствие значимого движения в течение заметного отрезка времени. Применительно к жестикуляции мы стремимся различать два разных типа неподвижности, в зависимости от того, в каком

положении — нейтральном или маркированном — произошла остановка. Задержку в нейтральном положении мы называем **покоем**; покой не является частью жестикуляции.

Таким образом, в основе предлагаемого подхода лежит деление кинетического поведения на жесты и кинетический фон, который включает периоды покоя и реализацию адапторов разного вида. Жесты идентифицируются относительно этого типичного фона. Адапторы на текущем этапе исследования не аннотируются как отдельные объекты, но их наличие может влиять на разметку граничащих с ними жестов.

### 3. Жестовые единицы различных уровней. Структура и границы

Внутренняя структура жеста первоначально исследовалась на материале жестовых языков глухих (Stokoe 1960, Battison 1974), в которых кинетическое поведение является единственным кодирующим средством. Позднее выработанные принципы были применены к спонтанной жестикуляции, сопровождающей речь. В настоящее время для аннотирования жестов используются разнообразные схемы, во многом зависящие от конкретных задач исследования; см. обзоры в Bressemer 2013a, b, Wagner et al. 2014: 223–224, Karpiński et al. 2015. В нашем подходе приоритет отдается интегральным признакам, неизменным на протяжении жеста.

Ключевым понятием предлагаемой аннотационной схемы является **жест** (gesture); в работах А. Кендона (Kendon 1980, 2004) и, вслед за ним, других авторов используется термин *gesture phrase* или *G-phrase*. Отдельные жесты, с одной стороны, делятся на фазы (phase, Kendon 1980), с другой — объединяются в **жестовые цепочки** (gesture chains; G-units по Kendon 1980, 2004). Под жестовой цепочкой понимается такая непрерывная последовательность жестов, в которой в любой момент времени хотя бы одна рука не находится в состоянии покоя. Нахождение обеих рук в положении покоя прерывает цепочку.

**Фазы** — структурные элементы жеста (Ladewig and Bressemer 2013: 1061). Главная фаза жеста — **мах** (stroke); значение жеста реализуется именно в махе. Это обязательная фаза, остальные являются второстепенными и могут отсутствовать. **Подготовка** (preparation) представляет собой переход к маху от нейтрального положения или от предыдущего жеста, а **ретракция** (retraction) — переход от маха к нейтральному положению. **Удержание** (hold) имеет место, когда рука останавливается в маркированном положении.

Основанием для деления жестикуляционного потока на жесты и их фазы служит изменение базовых признаков: скорость и направление движения, амплитуда, траектория, ориентация и конфигурация руки, место производства. Изменение двух и более из этих характеристик сигнализирует о вероятной границе фаз. Значимое изменение нескольких признаков позволяет говорить о том, что перед нами два разных жеста.

Отдельный случай — жесты с **многократным махом** (multi-stroke), в которых повторяющееся движение представляет собой единый мах (например, ‘стучать молотком’). Жест с многократным махом отличается постоянным

ритмом и общим паттерном скорости в пределах маха, между тем как в серии повторяющихся жестов скорость и ритм меняются на границе между жестами.

Важнейшей характеристикой целого жеста является **рукость**, то есть выполнение жеста правой рукой vs. левой рукой vs. двумя руками. В случае двуручных жестов актуально сопоставление движений рук: руки могут двигаться симметрично, идентично или же различным образом.

Жестовые единицы (цепочка, жест, фазы) устроены иерархично, однако на практике возникают сложности, связанные с неполной согласованностью движений двух рук. С одной стороны, руки в целом действуют как единый орган (см., например, понятие *growth point* в McNeill 2013; о синтаксическом и семантическом единстве жестовых и вербальных средств см. также Николаева 2004, 2009). С другой стороны, движения рук не всегда синхронны; границы фаз одного и того же жеста нередко не совпадают для левой и правой руки. Разработанная нами аннотационная схема предлагает стандартные решения для точного и последовательного отражения этих рассогласований.

#### 4. Схема аннотирования мануальных жестов

Предлагаемая аннотационная схема ориентирована на среду ELAN — одну из наиболее известных программ для аннотирования коммуникативного, в том числе кинетического, поведения (см. [tla.mpi.nl/tools/tla-tools/elan](http://tla.mpi.nl/tools/tla-tools/elan), Wittenburg et al. 2006). Это среда со встроенной возможностью многослойной кадровой аннотации медиафайлов.

Основой разметки являются 4 главных слоя, где с точностью до кадра (в нашем корпусе при частоте видео 100 к/с это дает шаг в 10 мс) размечаются временные границы следующих жестовых единиц:

- mChain — жестовые цепочки (общий слой для обеих рук)
- mGesture — жесты (общий слой для обеих рук)
- mLtPhases — фазы левой руки
- mRtPhases — фазы правой руки

Эти слои являются независимыми; аннотации в них привязаны непосредственно к временной шкале. Так как данная схема входит в разрабатываемую систему описания различных компонентов мультимедийного дискурса, причем при аннотировании других кинетических компонентов используется схожая структура слоев, все слои мануального компонента отмечены префиксом *m*.

Слою *mGesture* подчинена группа зависимых слоев, в которых аннотации привязаны к соответствующим аннотациям в слое жестов. Слои *mHandedness* и *mPhaseStructure* предназначены для указания руконости и фазовой структуры жеста. Слой *mTags* позволяет снабжать жест метками (тэгами) для указания других важных свойств жеста, причем тэги могут сочетаться между собой. В настоящее время мы пользуемся закрытым списком тэгов. Слой *mComments* предназначен для сопроводительных комментариев в свободной форме. На рис. 4 представлен образец аннотированной видеозаписи, а в таблице 1 приводится сводное описание аннотационной схемы.

ELAN 4.9.4 - pears04N-manu\_2017-04-10.eaf

File Edit Annotation Tier Type Search View Options Window Help

Comments Recognizers Metadata Controls

Grid Text Subtitles Lexicon

N-mGesture				
Nr	Annotation	Begin Time	End Time	Duration
>				
506	N-mg506	00:16:08.860	00:16:09...	00:00:01...
507	N-mg507	00:16:09.870	00:16:10...	00:00:00...
508	N-mg508	00:16:10.040	00:16:10...	00:00:00...
509	N-mg509	00:16:10.480	00:16:10...	00:00:00...
510	N-mg510	00:16:10.810	00:16:11...	00:00:00...
511	N-mg511	00:16:11.120	00:16:11...	00:00:00...
512	N-mg512	00:16:11.610	00:16:11...	00:00:00...
513	N-mg513	00:16:11.800	00:16:12...	00:00:00...
514	N-mg514	00:16:38.860	00:16:39...	00:00:00...
515	N-mg515	00:16:39.420	00:16:40...	00:00:00...
516	N-mg516	00:16:43.860	00:16:44...	00:00:00...

00:16:10.716 Selection: 00:00:00.000 - 00:00:00.000 0

10.000 00:16:10.500 00:16:11.000

- N-mChain [71]
- N-mGesture [533]
  - N-mHandedness [533]
    - Rt
    - Lt
    - Lt
    - Rt
  - N-mPhaseStructure
    - P S
    - P S
    - S H
    - P S
  - N-mTags [271]
    - rlap Lt P S overlap
    - Rt P overlap
    - Rt P S overlap, S Rebound, Lt P
  - N-mComments [89]
    - 10Ж
  - N-mLTPHases [561]
    - P
    - S
    - S
  - N-mRTPHases [419]
    - S
    - P
    - S

Рис. 4. Образец аннотированного видео в ELAN



Таблица 1. Схема аннотирования мануальных жестов в ELAN

Слои		Обозначения	Описание
mChain		(нумерация)	
mGesture		(нумерация)	
	mHandedness	Выбор из списка: Lt, Rt, Bh-sym, Bh-id, Bh-dif, Other	Вид жеста: леворучный, праворучный, двуручный симметричный / идентичный/с различной траекторией, другое. Значения этого параметра относятся преимущественно к маху и могут не совпадать в других фазах. Симметрия может быть как относительно сагиттальной (чаще всего), так и относительно другой оси либо плоскости. Симметрия относительно точки (во вращательных жестах) также попадает в эту категорию.
	mPhase-Structure	Формула из элементов P, S, H, R	Фазы в порядке появления: подготовка, мах, удержание, ретракция. Если граница между фазами не поддается дискретной идентификации, обозначения пишутся через дефис: P-S и т.д. Структура типа P-S характерна, например, для указательных жестов, а S-R — для постукиваний или хлопков.
	mTags	Multi-S	Жест с многократным махом
		Shuttle	Двухчастный жест (туда-обратно, P-S-R без возможности выделить границы)
		S rebound, R rebound	Отскок в конце маха или ретракции
		Long R	Длинная ретракция (в 2 или более этапа)
		Lt/Rt P/S/ H/R overlap	Наложение на текущий жест фазы другого жеста (например, Lt P overlap: наложение подготовки леворучного жеста на текущий — праворучный либо двуручный)
		Repeat	Повтор предыдущего жеста
		Unspec	Трудные случаи определения границ жеста, в том числе из-за наложения адапторов
	mComments		Комментарии
mLtPhases		Выбор из списка: P, S, H, R	Обязательны: S, H; P и R — в случае наложений.
mRtPhases			

Фазы жеста для левой и правой руки размечаются независимо и строго — так, как руки действуют «на самом деле». Границы жеста определяются для обеих рук по следующему правилу: жест начинается там, где начинается первая фаза (подготовка или мах) хотя бы для одной из рук (по более ранней из двух границ); правая граница ставится там, где начинается следующий жест.

Исключением является случай, когда подготовка следующего жеста накладывается на мах предыдущего; в этом случае, чтобы избежать укорачивания основной фазы жеста, граница ставится после окончания маха (см. рис. 5).

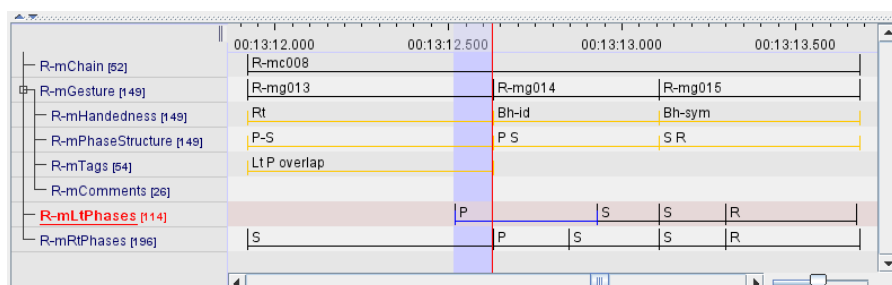


Рис. 5. Наложение подготовки следующего жеста (№14) на мах текущего (№13)

Окончание последнего жеста в цепочке ставится по правой границе последней из фаз (см. жест №15 на рис. 5). В случае наложения махов граница ставится там, где начинается второй мах (см. жесты №46 и №47 на рис. 6 ниже).

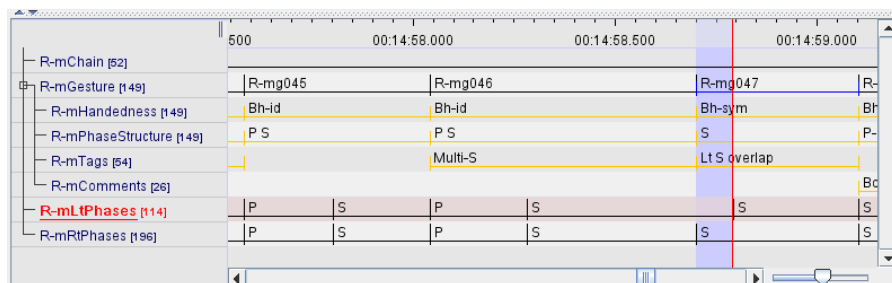


Рис. 6. Наложение махов

## 5. Опыт применения в корпусе.

### Индивидуальное варьирование

На данный момент (февраль 2017) мануальные жесты полностью аннотированы для трех записей общей продолжительностью 1 час, в каждой по три активных участника (Рассказчик, Комментатор и Пересказчик<sup>2</sup>). Все девять участников в этих трех записях — правши.

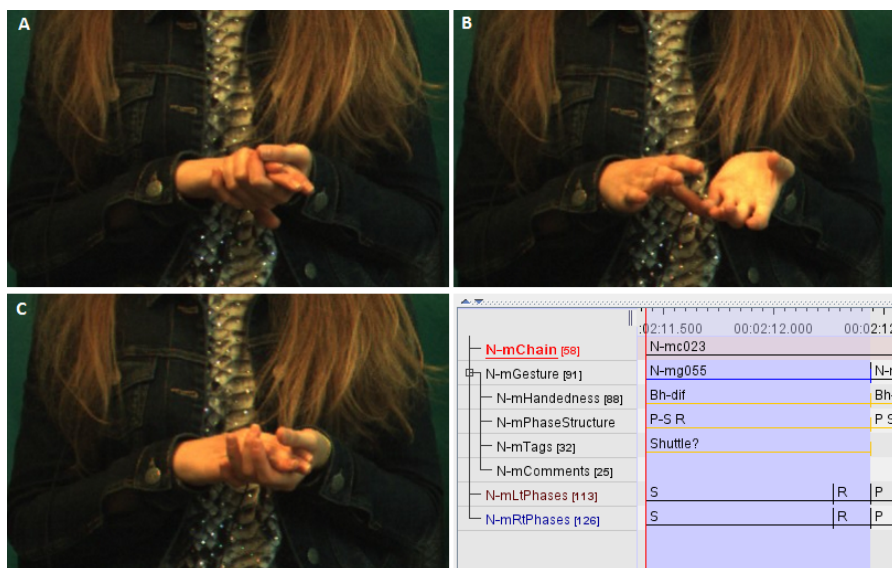
В настоящее время мы придерживаемся методологического принципа «аннотирование жестов без речи» (Bressemer 2013b: 1084). Цель этого подхода — минимизировать влияние содержания речи на интерпретацию жестов и определение

<sup>2</sup> В идентификаторах обозначены как N, C, R соответственно.

их границ. Полностью исключить такое «смысловое наведение» не удастся, поскольку аннотаторам все же известен предмет обсуждения — фильм о грушах. Однако отключение звука при разметке позволяет экспертам сосредоточиться на кинетических (визуальных) признаках. Кроме того, независимое аннотирование кинетического и вокального компонентов позволяет избежать порочного круга при изучении их взаимодействия (см., например, исследование временной координации жестовых и просодических единиц в Федорова и др. 2016).

Индивидуальные различия в области кинетического поведения весьма значительны. Одни говорящие между периодами жестикуляции могут долго сохранять неподвижность; другие часто меняют положение рук, но, тем не менее, существенное время пребывают в покое. Третьи постоянно переходят от жестикуляции к сериям адапторов и обратно, что может затруднять определение границ жестов и жестовых цепочек, даже если маховые фазы выделяются отчетливо.

Процедура аннотирования, как правило, начинается со сплошного просмотра видеофайла, что позволяет составить **жестикуляционный портрет** коммуниканта: каковы естественные для него нейтральные положения рук, типичная скорость и амплитуда движений, привычное количество адапторов и т.д. (ср. просодические портреты, аналогичным образом используемые при транскрибировании звучащей речи, — Кибрик 2009). Все это необходимо для определения того кинетического фона, на котором должны выделяться жесты. У коммуникантов, не склонных даже к кратким остановкам внутри цепочек, локальные нейтральные положения часто служат основным средством выявления границ жестов (см. рис. 7).



**Рис. 7.** Локальные нейтральные положения на границах жеста №55 (кадры А и С) в сравнении с началом маха (кадр В)

При аннотировании возможно как движение от крупных единиц (цепочек) к более мелким (жесты, фазы), так и выделение на первом этапе центральных единиц (жестов) с дальнейшим уточнением их границ и объединением в цепочки. При длящейся рассинхронизации рук правильно определить границы жестов зачастую удастся лишь после тщательной разметки всех фаз фрагмента (см. рис. 8); в подобных случаях целесообразно аннотирование «снизу вверх».

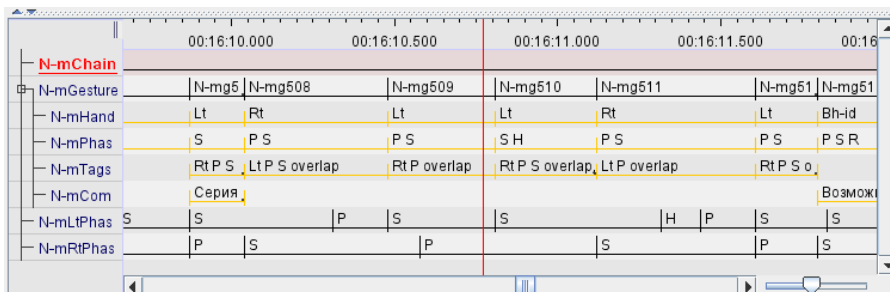


Рис. 8. Множественные наложения при длящейся рассинхронизации рук

Важным параметром жестикуляционного портрета являются предпочтения в использовании правой/левой руки при жестикуляции. Есть данные о том, что выбор доминирующей руки в конкретном жесте связан не столько с право-/леворукостью, сколько с тем, какое полушарие мозга задействовано в решении текущей когнитивной задачи (обзор см., например, в Lausberg, Sloetjes 2015). Предварительные количественные тенденции, основанные на уже аннотированном подкорпусе, резюмированы в таблицах 2 и 3.

Таблица 2. Преобладание двуручных vs. одноручных жестов

Участник	Двуручные		Одноручные		Всего
23С	66	67%	32	33%	98
23N	229	66%	120	34%	349
04С	71	60%	47	40%	118
22R	66	44%	83	56%	149
23R	93	42%	126	58%	219
22С	69	40%	102	60%	171
04R	158	32%	329	68%	487
22N	110	32%	239	68%	349
04N	162	27%	441	73%	603

**Таблица 3.** Распределение одноручных жестов по рукости

Участник	Праворучные		Леворучные		Всего
22N	231	97%	8	3%	239
22R	72	87%	11	13%	83
22С	79	77%	23	23%	102
04R	225	68%	104	32%	329
04С	32	68%	15	32%	47
23С	20	63%	12	38%	32
23R	47	37%	79	63%	126
04N	144	33%	297	67%	441
23N	40	33%	80	67%	120

Согласно этим данным, часть коммуникантов склонна к двуручным жестам, но большинство — к одноручным. Наиболее типичное соотношение праворучных и леворучных жестов — примерно 2:1, но склонность использовать правую руку в одноручных жестах существенно варьирует — от 60% до почти 100%. Наконец, хотя все участники идентифицировали себя как правшей, у трех из них леворучные жесты преобладают. При расширении аннотированной части корпуса (и, возможно, добавлении сведений о доминирующей руке в асимметричных двуручных жестах) эти данные позволят изучать возможную связь рукости жеста и, к примеру, его функционального типа или речевого контекста.

## 6. Заключение

В рамках создаваемого мультимедийного корпуса «Рассказы и разговоры о грушах» был разработан оригинальный теоретический подход к описанию мануальных жестов естественного русского дискурса, опирающийся на предшествующий опыт вокальной аннотации (Кибрик, Подлесская, ред. 2009). Основное внимание уделяется интегральным характеристикам, релевантным для жеста на всем его протяжении. На основе этого подхода предложена схема аннотации мануальных жестов, позволяющая последовательно идентифицировать жестовые единицы и их свойства. Разработаны специальные конвенции для частотных сложных случаев: взаимодействие жестов и адапторов, рассинхронизация рук, наложение жестов, разграничение жестов с многократным махом и серий повторяющихся жестов. Система аннотирования мануальных жестов не автономна, а увязана с аннотацией других кинетических компонентов, а также вокального поведения. Предлагаемая система нацелена на согласованность аннотаций разных каналов и компонентов между собой. Это отражает реальное положение дел в естественной коммуникации, где все каналы задействованы синхронно и направлены на передачу единого послания. Создаваемый на базе данной аннотационной схемы ресурс позволит проводить широкий круг исследований в области взаимодействия коммуникативных каналов.

## Литература

1. *Бутовская М. Л.* (2004), *Язык тела: природа и культура (эволюционные и кросс-культурные основы невербальной коммуникации человека)*, Москва: Научный мир.
2. *Гришина Е. А.* (в печати), *Русская жестикуляция с лингвистической точки зрения (корпусные исследования)*.
3. *Кибрик А. А.* (2009), *Просодические портреты говорящих*, Кибрик А. А., Подлеская В. И. (ред.). *Рассказы о свидениях: корпусное исследование устного русского дискурса*. Москва: ЯСК, 478–487.
4. *Кибрик А. А.* (2010), *Мультиmodalная лингвистика*, Ю. И. Александрова, В. Д. Соловьев (ред.) *Когнитивные исследования*, вып. IV. М.: ИПРАН, 134–152.
5. *Кибрик А. А., Подлеская В. И.* (ред.) (2009), *Рассказы о свидениях: корпусное исследование устного русского дискурса*, Москва: ЯСК.
6. *Крейдлин Г. Е.* (2002), *Невербальная семиотика*, Москва: НЛО.
7. *Николаева Ю. В.* (2004), *Функциональные и семантические особенности иллюстративных жестов в устной речи (на материале русского языка)*, *Вопросы языкознания*, №4, 48–64.
8. *Николаева Ю. В.* (2009), *Сегментация устного нарратива и изобразительные жесты: кинетические признаки границ и связей между сегментами дискурса*, *Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог 2009»*, Вып. 8 (15), 340–345.
9. *Фёдорова О. В., Кибрик А. А., Кортаев Н. А., Литвиненко А. О., Николаева Ю. В.* (2016), *Временная координация между жестовыми и речевыми единицами в мультиmodalной коммуникации*, *Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции Диалог (Москва, 1–4 июня 2016 г.)*, Вып. 15 (22), М.: Изд-во РГГУ, 159–170.

## References

1. *Battison R.* (1974), *Phonological deletion in American Sign Language*, *Sign Language Studies*, 5, pp. 1–19.
2. *Bressem J.* (2013 a), *Transcription systems for gestures, speech, prosody, postures, and gaze*, in Müller C., Cienki A., Fricke E., Ladewig S. H., McNeill D. & Teßendorf S. (eds.), *Body — Language — Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 1, Berlin: Mouton, pp. 1037–1060.
3. *Bressem J.* (2013 b), *A linguistic perspective on the notation of form features in gestures*, in Müller C., Cienki A., Fricke E., Ladewig S. H., McNeill D. & Teßendorf S. (eds.), *Body — Language — Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 1, Berlin: Mouton, pp. 1079–1098.
4. *Butovskaya M. L.* (2004), *Body language: nature and culture (evolutionary and cross-cultural foundations for non-verbal human communication)* [*Yazyk tela: priroda i kul'tura (evolyutsionnye i kross-kul'turnye osnovy neverbal'noy komunikatsii cheloveka)*], Moscow: Nauchny mir.

5. *Chafe W.* (ed.) (1980), *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*, Norwood, NJ: Ablex.
6. *Ekman P.* (1999), *Emotional and Conversational Nonverbal Signals*, in Messing, L. S. & Campbell, R. (Eds.), *Gesture, Speech, and Sign*, New York, NY: Oxford University Press, pp. 45–55.
7. *Ekman, F., Friesen W. V.* (1969), *The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding*, *Semiotica*, 1 (1), pp. 49–98.
8. *Fedorova O. V., Kibrik A. A., Korotaev N. A., Litvinenko A. O., Nikolaeva Ju. V.* (2016), *Temporal coordination between gestural and speech units in multimodal communication [Vremennaya koordinatsiya mezhdzhu zhestovymi i rechevymi edinitami v mul'timodal'noy kommunikatsii]*, *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2016" [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii]*, RGGU, Moscow, pp. 159–170.
9. *Fricke E., Freiburg i. Br., Chemnitz* (2013), *Towards a unified grammar of gesture and speech: A multimodal approach*, in Müller C., Cienki A., Fricke E., Ladewig S. H., McNeill D. & Teßendorf S. (eds.), *Body—Language—Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 1, Berlin: Mouton, pp. 733–754.
10. *Gibbon D., Mertins I., Moore R. K.* (eds.) (2000), *Handbook of multimodal and spoken dialogue systems: Resources, terminology and product evaluation*, Berlin: Springer.
11. *Granström B., House D., Karlsson I.* (eds.) (2002), *Multimodality in language and speech systems*, Dordrecht: Kluwer.
12. *Grishina E. A.* (in press), *Russian gesticulation in linguistic perspective (corpus studies) [Russkaya zhestikulyatsiya s lingvisticheskoy tochki zreniya (korporusnye issledovaniya)]*.
13. *Karpiński M., Jarołowicz-Nowikow E., Czoska A.* (2015), *Gesture annotation scheme development and application for entrainment analysis in task-oriented dialogues in diverse cultures*, *Proceedings of GESPIN 2015 Conference*, Nantes, France, pp. 161–166.
14. *Kendon A.* (1980), *Gesticulation and speech: Two aspects of the process of utterance*, in M. R. Key (ed.), *The relationship of verbal and nonverbal communication*, pp. 207–227.
15. *Kendon A.* (2004), *Gesture: Visible Action as Utterance*, New York: Cambridge University Press.
16. *Kibrik A. A.* (2009), *Speakers' prosodic portraits [Prosodicheskie portrety govoryashchikh]*, in Kibrik A. A., Podlesskaya V. I. (eds.) (2009), *Corpus of spoken Russian "Night Dream Stories" [Korpus ustnoy russkoy rechi "Rasskazy o snovideniyakh"]*, *Jazyki slavyanskikh kul'tur*, Moscow.
17. *Kibrik A. A.* (2010), *Multimodal linguistics [Mul'timodal'naya lingvistika]*, in Yu. I. Aleksandrov, V. D. Solov'yev (eds.), *Cognitive studies [Kognitivnyye issledovaniya]*, Vol. IV, Institute of psychology, Moscow, pp. 134–152.
18. *Kibrik A. A., Podlesskaya V. I.* (eds.) (2009), *Corpus of spoken Russian "Night Dream Stories" [Korpus ustnoy russkoy rechi "Rasskazy o snovideniyakh"]*, *Jazyki slavyanskikh kul'tur*, Moscow.

19. *Kress G.* (2002), The multimodal landscape of communication, *Medien Journal*, Vol. 4, pp. 4–19.
20. *Kreydlin G. E.* (2002), Nonverbal semiotics [Neverbal'naya semiotika], New literary review, Moscow.
21. *Ladewig S. H., Bressemer J.* (2013), A linguistic perspective on the notation of gesture phases, in Müller C., Cienki A., Fricke E., Ladewig S. H., McNeill D. & Teßendorf S. (eds.), *Body — Language — Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 1, Berlin: Mouton, pp. 1060–1078.
22. *Lausberg H., & Sloetjes H.* (2015), The revised NEUROGES-ELAN system. An objective and reliable interdisciplinary analysis tool for nonverbal behavior and gesture, *Behavior Research Methods*, DOI: 10.3758/s13428–015–0622-z, available at: <http://neuroges.neuroges-bast.info/publications>.
23. *Müller C., Cienki A., Fricke E., Ladewig S. H., McNeill D. & Teßendorf S.* (eds.) 2013, *Body — Language — Communication: An International Handbook on Multimodality in Human Interaction*, Berlin: Mouton.
24. *McNeill D.* (2013), Gesture as a window onto mind and brain, and the relationship to the linguistic relativity and ontogenesis, in Müller C., Cienki A., Fricke E., Ladewig S. H., McNeill D. & Teßendorf S. (eds.), *Body — Language — Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 1, Berlin: Mouton, pp. 28–54.
25. *Müller C., Bressemer J. and Ladewig S. H.* (2013), Towards a grammar of gestures: A form-based view, in Müller C., Cienki A., Fricke E., Ladewig S. H., McNeill D. & Teßendorf S. (eds.), *Body — Language — Communication: An International Handbook on Multimodality in Human Interaction*, Vol. 1, Berlin: Mouton, 707–733.
26. *Nikolaeva Yu. V.* (2004), Functional and semantic aspects of illustrative gestures in Russian spoken discourse [Funktsional'nye i semanticheskie osobennosti illyustrativnykh zhestov v ustnoy rechi (na materiale russkogo yazyka)], *Issues of General Linguistics [Voprosy yazykoznanania]*, No. 4, pp. 48–64.
27. *Nikolaeva Yu. V.* (2009), Segmentation of oral narrative discourse and illustrative gestures: visual clues as segment markers [Segmentatsiya ustnogo narrativa i izobrazitel'nye zhesty: kineticheskie priznaki granits i svyazey mezhdu segmentami diskursa], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2006” [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2009”]*, Bekasovo, pp. 340–345.
28. *Scollon R.* (2006), Multimodality and the language of politics, in K. Brown (ed.) *Encyclopedia of language and linguistics*, Elsevier, Vol. 9, pp. 386–387.
29. *Stokoe W.* (1960), *Sign Language Structure*, Buffalo, NY: Buffalo University Press.
30. *Wagner P., Malisz Z., Kopp S.* (2014), Gesture and speech in interaction: An overview, *Speech Communication*, 57, pp. 209–232.
31. *Wittenburg P., Brugman H., Russel A., Klassmann A., Sloetjes H.* (2006), ELAN: A professional framework for multimodality research, in *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*.



# КОМПЬЮТЕРНАЯ СИСТЕМА ОБУЧЕНИЯ ИНТОНАЦИОННЫМ КОНСТРУКЦИЯМ РУССКОЙ РЕЧИ

**Лобанов Б. М.** (Lobanov@newman.bas-net.by),  
**Житко В. А.** (zhitko.vladimir@gmail.com),  
**Харламов А. А.** (kharlamov@analyst.ru)

Объединённый институт проблем информатики  
НАН Беларуси, Минск, Беларусь

# A COMPUTER SYSTEM OF TEACHING INTONATION PATTERNS OF RUSSIAN SPEECH

**Lobanov B. M.** (Lobanov@newman.bas-net.by),  
**Zhitko V. A.** (zhitko.vladimir@gmail.com),  
**Kharlamov A. A.** (kharlamov@analyst.ru)

United Institute of Informatics Problems NAS Belarus, Minsk,  
Belarus

A software system is presented which is designed to train learners in producing the basic intonation patterns of Russian speech. The system is based on comparing the melodic portraits of a reference sentence and a sentence pronounced by the learner and involves active interaction of the learner with the system. While parametric representation of intonation features of the speech signal faces fundamental difficulties, the paper shows how they can be overcome. The basic algorithms of analyzing and comparing intonation features, used in the proposed learning system, are presented. The features of the acoustic database composed of reference sentences and used in the learning system are presented. The set of reference sentences represents intonation patterns of Russian speech (IP1 to IP7) and their basic varieties. The system's interface is presented and the results of system operation are illustrated.

**Keywords:** Speech intonation, speech analysis and generation, melodic portrait, systems of analysis and assessment of intonation, learning system, Russian intonation, Russian as the second language

## Введение

Интонация играет важную роль при восприятии речи человеком. С помощью интонации определяется коммуникативная направленность высказывания, логический смысл, осуществляется вычленение семантически связанных

отрезков речи и объединение речевых элементов внутри этих отрезков. Зачастую в зависимости от интонационного оформления логически идентичные высказывания могут иметь различное семантическое значение.

Интонация представляет собой важнейшее средство передачи как собственно языковой, так и социокультурной информации. В практике преподавания существует мнение, что иноязычный акцент особенно ярко проявляется в интонации, поэтому при обучении и преподавании иностранных языков ей следует уделять особенное внимание. Интонационный акцент часто возникает вследствие контакта различных языковых систем, например, при билингвизме, как результат языковой интерференции. Следует иметь в виду, что функции интонации в речи столь многообразны, что нарушения в этой области могут привести к серьезным смысловым изменениям, а также создать неверное впечатление о говорящем человеке.

Правильность воспроизведения интонации при говорении и адекватность её восприятия при слушании с *трудном поддаётся самоконтролю* учащимся (особенно, при отсутствии музыкального слуха). Существующие лингафонные курсы и оборудование обеспечивают только *СЛУХОВУЮ обратную связь* контроля правильности интонирования речи, что явно *недостаточно*.

Настоящая работа посвящена описанию прогресса в создании компьютерного тренажёра, обеспечивающего дополнительную *ВИЗУАЛЬНУЮ обратную связь*, а также *количественную ОЦЕНКУ* правильности интонирования русской речи в процессе обучения русскому языку как иностранному (РКИ).

В мировой практике освоение языка с помощью компьютера (Computer Assisted Language Learning — CALL) является широко распространённой междисциплинарной областью, которая включает также подобласть обучения произношению на иностранных языках с использованием речевых технологий [Catia Cucchiarini, 2008]. Появилось также описание компьютерных систем, предназначенных для обучения различным аспектам просодики английской речи, в том числе — интонации [Sixuan Zhao, 2010].

Для создания описываемой ниже компьютерной системы обучения ИК русской речи потребовалось решение ряда принципиальных проблем, к которым относятся:

1. ***Адекватное сопоставление произносимого и эталонного сигналов при условии наличия взаимной временной деформации и заранее неизвестного начала и конца произносимого сигнала.***

Решение этой проблемы стало возможным благодаря разработанному ранее [Lobanov, 1997] методу текущего динамического сопоставления (ТДС) двух сигналов. Его использование обеспечивает автоматическое определение конца и начала произнесённой фразы в процессе её сопоставления с эталонной фразой.

2. ***Сегментация анализируемого сигнала на участки, для которых понятие ЧОТ имеет смысл с точки зрения формирования интонационного контура фразы в целом: участки гласных и большинства сонорных согласных.***

Эта проблема решается путём нелинейного переноса сегментных меток с предварительно размеченной эталонной фразы на произнесённую фразу с использованием предложенной ранее технологии клонирования просодических характеристик речи [Lobanov, 2008]

3. **Высокоточное вычисление частоты основного тона (ЧОТ) произносимого и эталонного сигналов без подстройки для мужских и женских голосов в широком диапазоне ЧОТ –  $F_0 = \{40-1000 \text{ Гц}\}$ .** Задача решается путём использования известных методов выделения ЧОТ речевого сигнала. Успешному решению данной проблемы посвящено большое число работ (см. например, [Shimamura, 2001]).

4. **Интерполяция значений ЧОТ на тех участках, для которых определение ЧОТ является некорректным, т.е. на участках большинства согласных звуков.**

Эта задача решается путём использования известных математических формул интерполяции, определяющих способ нахождения промежуточных значений по имеющемуся дискретному набору известных значений.

5. **Вычисление меры подобия интонационных характеристик произнесённой и эталонной фраз в условиях их различной длительности и различных диапазонах ЧОТ.**

Эта задача решается путём использования представления интонационной кривой в виде **Универсального (унифицированного) Мелодического Портрета (УМП)**, подробно описываемого ниже в 1-м разделе. Собственно расчёт меры подобия двух УМП осуществляется по известным формулам путём вычисления выборочного коэффициента корреляции, либо через векторное расстояние между кривыми.

Данная работа посвящена описанию разработанной компьютерной системы, ориентированной на начальное обучение РКИ в рамках освоения учащимися интонационных конструкций русской речи. Понятие *интонационных конструкций (ИК1 — ИК7)*, предложено в 1960-х гг. [Vryzgunova, 1968] и эффективно используется во многих современных методических пособиях по обучению РКИ [Odintsova, 2011].

Работа построена следующим образом.

В *первом* разделе приводятся основные сведения об используемой модели для описания интонационной структуры речи.

Во *втором* разделе описывается акустическая база данных эталонных фраз, положенная в основу обучающей системы.

В *третьем* разделе приводится структурная схема разработанной компьютерной обучающей системы.

В *четвёртом* разделе описываются особенности программной реализации и интерфейс системы.

В *заключении* перечисляются проблемы, успешное решение которых легло в основу описанной системы, а также перечислены примеры возможных практических приложений.

## 1. Интонационная модель

В соответствии с предложенной в [Lobanov, 2007] ПАЕ-моделью, минимальной просодическим компонентом, из которого составляется интонация фразы, является Акцентная Единица (АЕ). АЕ может состоять из одного или более фонетических слов, причём одно из них должно иметь в своём составе полноударный (ядерный) слог. Каждая АЕ, в свою очередь, состоит из ядра (полноударная гласная фонема), предъядра (все фонемы, предшествующие полноударной гласной) и заядра (все фонемы за полноударной гласной). ПАЕ-модель предполагает, что для определенного типа интонации топологические свойства мелодического контура АЕ не зависят от количественного и качественного содержания предъядра, ядра и заядра.

В [Lobanov, 2014]) показано, что ПАЕ-модель обеспечивает наглядное представление интонационных конструкций русской речи (ИК1 — ИК7) — в виде набора их Универсальных (унифицированных) Мелодических Портретов (УМП) в нормированных координатах «*Частота — Время*». Нормализация по времени осуществляется путём приведения к стандартной длине элементов АЕ: предъядерных, ядерных и заядерных участков. Этот вид нормализации устраняет различия мелодической кривой, связанные с количественным составом предъядерных и заядерных участков АЕ.

Для нормализации по частоте определяются минимальное ( $F_{0\ min}$ ) и максимальное ( $F_{0\ max}$ ) значения частоты основного тона —  $F_0$  для всего ансамбля мелодических кривых  $\{IK_i\}$  в произнесении данного диктора. Нормализация осуществляется в соответствии с формулой:  $F_0^N = (F_0 - F_{0\ min}) / (F_{0\ max} - F_{0\ min})$ . Этот вид нормализации устраняет различия мелодической кривой, связанные с индивидуальными дикторскими различиями в высоте голоса. В то же время величина  $D = [(F_{0\ max} / F_{0\ min}) - 1]$ , выраженная в октавной шкале, может быть использована для оценки диапазона изменения ЧОТ, характеризующего данную ИК — (*широкий-средний-узкий*).

Таким образом, нормированное пространство для отображения УМП ИК<sub>i</sub> может быть представлено в виде прямоугольника с координатными осями ( $T_N, F_0^N$ ), как это представлено на схематическом рисунке 1. При этом интервалам на оси абсцисс соответствуют: [0–1/3] — предъядро, [1/3–2/3] — ядро, [2/3–1] — заядро. Интервалам на

оси ординат соответствуют: [0–1/3] — низкий уровень тона, [1/3–2/3] — средний, [2/3–1] — высокий. Такое представление обеспечивает основное внимание к особенностям формы кривой ЧОТ на ядре при меньшем внимании на количественные и качественные состав пред- и за-ядра. В рамках УМП предоставляется возможным достаточно строго описывать мелодическую кривую, пользуясь широко распространёнными терминами:

- «низкий-средний-высокий» для уровня тона и
- «восходящий-ровный-нисходящий» для направления движения на каждом из участков УМП,
- «широкий-средний-узкий» для диапазона изменения ЧОТ на всех участках в целом.

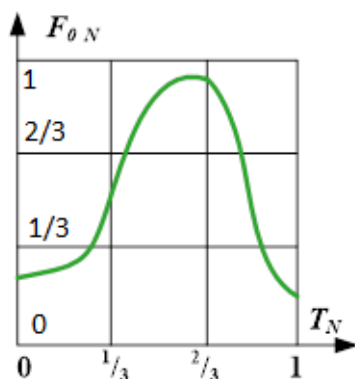


Рис. 1. Общий вид УМП ИК

В [Lobanov, 2015]] описан положительный опыт создания мелодических портретов сложных повествовательных предложений русской речи с использованием ПАЕ-модели и УМП, а в [Lobanov, 2016]] показано, что представление интонации в виде УМП позволяет выявить характерные различия при сравнении мелодических портретов английских и русских фраз диалоговой речи.

## 2. Акустическая база данных

В качестве акустической базы данных эталонных ИК используются образцы фраз, произносимых мужским и женским голосами, содержащихся в мультимедийном учебнике РКИ [Odintsova, 2011]. В используемой БД присутствуют по 4–5 эталонных образцов для каждой из интонационных конструкций (ИК1 — ИК7), а также несколько образцов диалоговой речи и отрывков прозы в дикторском прочтении. Кроме того, в дополнение к наиболее употребительным образцам в состав БД включены примеры различного вида употреблений каждой из ИК, начитанные профессиональным русскоязычным диктором. В качестве примера в таблицах 1–3 приведены тексты фраз, в которых реализуются различные способы употребления ИК1, ИК2, ИК3. Выбранные фразы — одноакцентные. Знаком (+) обозначены ядерные гласные фраз.

Таблица 1. Примеры употребления ИК 1

№	Употребление	Примеры
1	Конец простого предложения	<i>На горизонте появилось о+блако. Наконец пришёл сто+рож.</i>
2	Нейтральный ответ	<i>— Приеду завтра у+тром. — А печь уже прото+плена.</i>
3	Часть предложения	<i>Потом зашла санита+рка, чтобы погре+ться Они отъе+хали, и стрельба прекрати+лась</i>
4	Название — объявление	<i>Фабрика номер де+сять. Куплю персидский ковё+р.</i>
5	Желание — совет	<i>Хотелось бы поговори+ть с вами. Советую Вам не заде+рживаться.</i>

Таблица 2. Примеры употребления ИК 2

№	Употребление	Примеры
1	С вопросительным словом	<i>О ко+м ты плачешь? Когда+ отправляемся?</i>
2	Альтернативный вопрос	<i>Иванов или Пота+пов? Занятия будут или не бу+дут?</i>
3	Вопрос-уточнение	<i>Ты смо+жешь придти? В пя+тницу придёшь?</i>
4	Вопрос-удивление	<i>Неужели пра+вда? Что за шу+м такой?</i>
5	Выделение слова	<i>Маши+на вчера сломалась! Машина вчера+ сломалась!</i>
6	Обращение	<i>Дорогой Иван Петро+вич! Господа+ присяжные!</i>
7	Приветствие — прощание	<i>Здра+встуйте товарищи! Проща+й Одесса!</i>
8	Выражение вежливости	<i>Пожа+луйста не сердитесь! Прости+те меня!</i>
9	Требование — побуждение	<i>Проходи+те! не заде+рживайтесь! Молча+ть! шагом ма+ри!</i>
10	Положительная оценка	<i>Это здо+рово! Замеча+тельная идея!</i>
11	С частицей ВОТ	<i>Во+т вам за это! Во+т ваши книги!</i>
12	С усилительными частицами	<i>Даже Ва+ня пришёл! Хоть бы ве+тер стих!</i>
13	С местоименными словами	<i>Такой ве+чер! Ведь это здо+рово!</i>

Таблица 3. Примеры употребления ИК 3

№	Употребление	Примеры
1	Общий вопрос	— Вы вчера бы+ли в городе? — Вы вчера были в го+роде?
2	Однословный переспрос	— Куда+? На пра+ктику? — Отку+да? Из институ+та?
3	Повторение вопроса	— Где я гуля+ла? В парке. — Приехал ли Са+ша? Не знаю.
4	Просьба-вопрос	Дава+й я тебе помогу? Мо+жно я пойду в кино?
5	Альтернативный вопрос	Ивано+в или Потапов? Занятия бу+дут или не будут?
6	Вопрос-удивление	— Неуже+ли Наташа? — Разве это Ната+ша?
7	Вопрос-уточнение	Ты приедешь по+ездом? Голова+ заболела?
8	Волеизъявление-просьба	— Закро+йте дверь! — Прошу не опа+здывать!
9	Выражение оценки	Она такая у+мница! Как здесь краси+во!
10	Незавершенность 1-го вида	На полу лежала су+мка, забытая кем-то. Это была служа+нка, ставшая хозяйкой.
11	Незавершенность 2-го вида	Трещит ого+нь, стук посу+ды и ужин готов. Все вста+ли, помолча+ли и приступили к трапезе.

В соответствии с рекомендациями, приведенными в учебнике РКИ [Odintsova, 2011], аналогичные таблицы составлены для остальных ИК4 — ИК7.

Акустические сигналы, реализующие каждую из эталонных фраз, размечены путём указания границ каждой из входящих в неё гласных и сонорных с указанием их функционального значения во фразе: **предъядра, ядра, заядра**. На рис. 2 приведен пример такого рода разметки для двух различных фраз. Знаком (+) отмечены ядерные гласные.

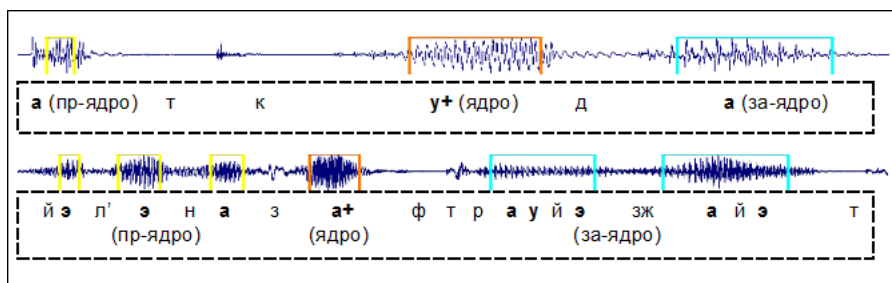


Рис. 2. Пример разметки сигнала фразы «Откуда?» (вверху) и фразы «Елена за+втра уезжает?» (внизу)

### 3. Структурная схема компьютерной обучающей системы

На рисунке 3 приведена структурная схема компьютерной обучающей системы. Назначение системы — предоставить учащемуся возможность компактного и наглядного отображения результатов анализа мелодического и энергетического контуров эталонных фраз с различными ИК, а также получить численную оценку качества обучения интонации речи путём сопоставления контура произносимой фразы с эталонным контуром.

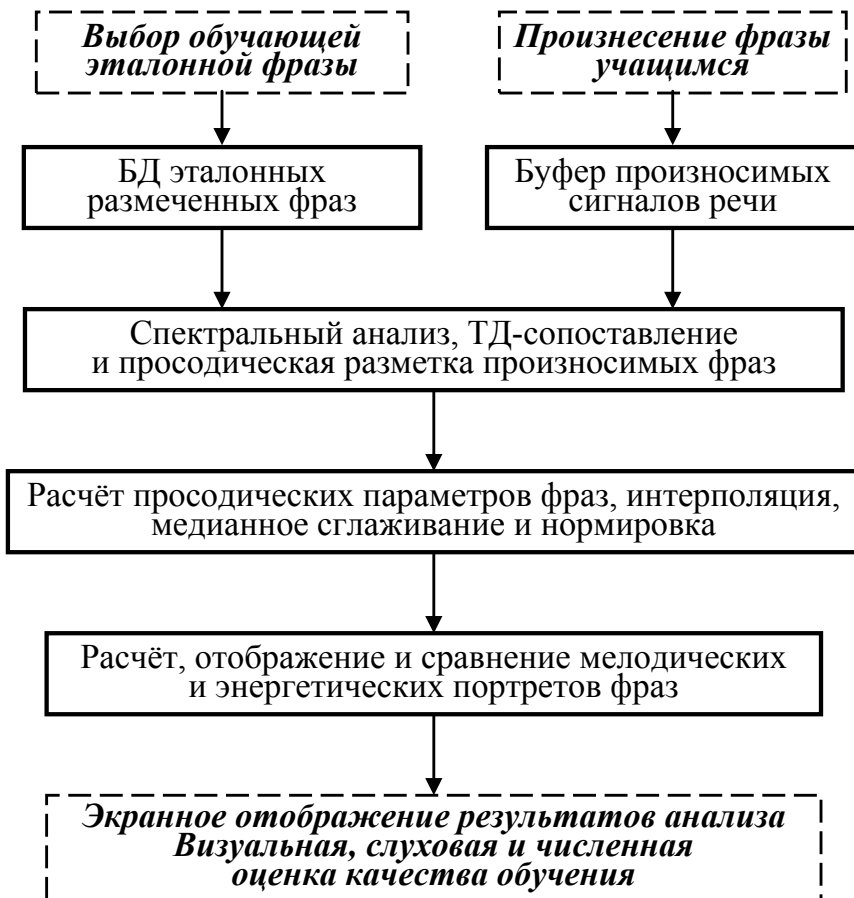


Рис. 3. Структурная схема компьютерной системы обучения интонации речи

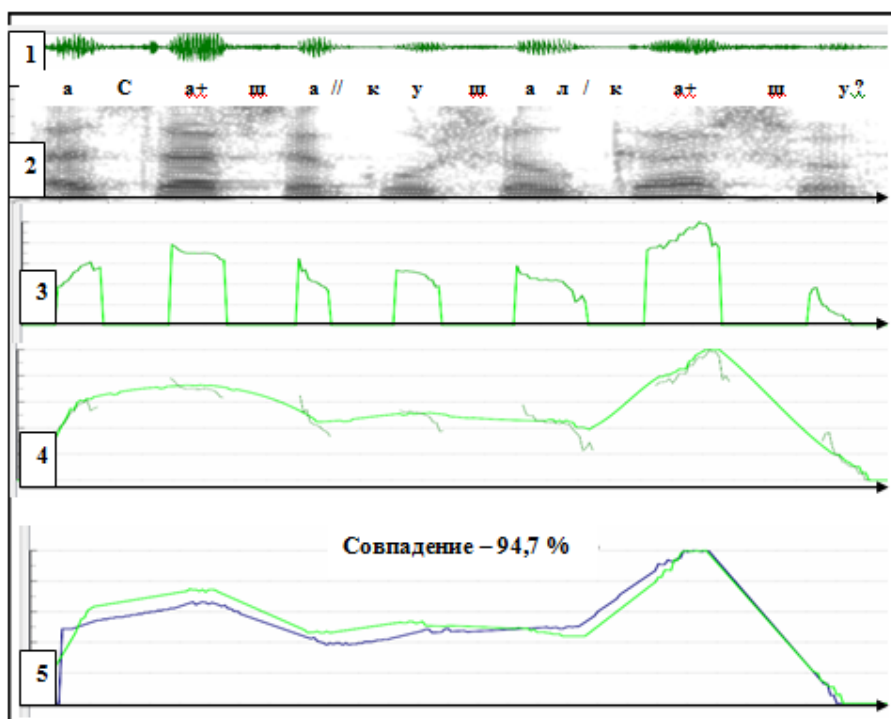
В БД эталонов хранятся образцы фраз различных ИК и примеры их употребления. Для каждой из эталонных фраз проводится предварительная просодическая разметка, включающая указание фразовых (синтагматических) границ, а также положение ядерных, предъядерных и заядерных гласных во фразе. Учащийся, в соответствии с предлагаемой методикой обучения интонационным навыкам, выбирает нужные фразы, прослушивает их и затем



произносит их. При этом осуществляется запись речевого сигнала в буфер, а затем производится описываемая ниже обработка эталонного и произнесённого сигналов.

Эталонный и произнесённый сигналы подвергаются спектральному анализу, затем с использованием модифицированного метода динамического программирования осуществляется их сопоставление, перенос просодических меток и разметка произносимых фраз. Далее производится расчёт просодических параметров фраз — частоты основного тона  $F_0$  и энергии сигнала  $A_0$ . Над этими параметрами осуществляется операции интерполяции на не голосовых участках, медианное сглаживание и их нормировка.

На рисунке 4 представлена иллюстрация произнесённой 2-х акцентной вопросительной фразы: «А Са+ша кушал ка+шу?»: её осциллограмма, спектр и пример обработки частоты основного тона —  $F_0$ , в результате которой получается искомая непрерывная мелодическая кривая фразы  $F_0(t)$ . Аналогичная обработка осуществляется для получения энергетической кривой  $A_0(t)$ . Полученные кривые  $F_0(t)$  и  $A_0(t)$  являются основой для построения мелодического и энергетического портретов фразы.



**Рис. 4.** Иллюстрация обработки сигналов речи:

- 1 — осциллограмма, 2 — спектр, 3 —  $F_0$  (до обработки),  
4 —  $F_0$  (после интерполяции и медианного сглаживания), 5 — сравнение мелодических кривых  $F_0(t)$  эталонной и произнесённой фразы

## 4. Программная реализация и интерфейс прототипа системы

Прототип системы (*IntonTrainer*) реализован в виде отдельного приложения под ОС Windows (XP, 7, 8, 10) и Linux (Debian, Ubuntu). В реализации используется библиотека Qt для реализации отладочного пользовательского интерфейса и ядра приложения, и библиотека MathGL для вывода графиков. Для реализации основного пользовательского интерфейса используется приложение, написанное на HTML/JS (используется ReactJS фреймворк). Алгоритмы расчёта и обработки данных написаны на языке C с использованием библиотек GNU Scientific Library, Speech Signal Processing Toolkit (SPTK) и OpenAL. Такая структура приложения позволяет решить следующие задачи:

- быструю и лёгкую адаптацию и стилизацию приложения под новый язык, за счёт изменения лишь HTML/JS составляющей;
- возможность интеграции ядра приложения (алгоритмов) в сторонние продукты за счёт отсутствия внешних зависимостей и использования C;
- запуск приложения под различными системами и окружениями за счёт использования кросс-платформенных решений

В приложении выделены следующие слабосвязанные модули:

- ядро приложения (реализация на C)
- взаимодействие с аудио устройствами (реализация на C и OpenAL)
- основной пользовательский интерфейс (реализация на HTML/JS запускаемое в V4 Javascript Engine)
- отладочный интерфейс и режим обучения (реализация на C++/Qt)

Начальное окно приложения представлено на рис. 5.

По нажатию кнопки «Начать» мы попадаем на структурированный перечень обучающих заданий (рис. 6)

После нажатия на выбранную эталонную фразу будет открыто окно тренировки произношения (рис. 7). Пользователь может прослушать эталонную фразу и просмотреть кривые изменения F0, A0 для выбранной фразы.

В режиме тренировки пользователь, нажимая на клавишу ввода, произносит фразу и видит результат сравнения эталонной и произнесённой фраз. На рисунке 8 представлены исходные кривые изменения ЧОТ вопросительной фразы «*Еле+на завтра уезжает?*» для эталона (светлая кривая) и произнесения (тёмная кривая).

Далее при нажатии клавиши «Show UMP» (см. рис. 9) высвечивается результирующее окно, в котором показаны сравнительный вид УМП фраз и диапазоны изменения ЧОТ, представленные столбиками в октавной шкале.

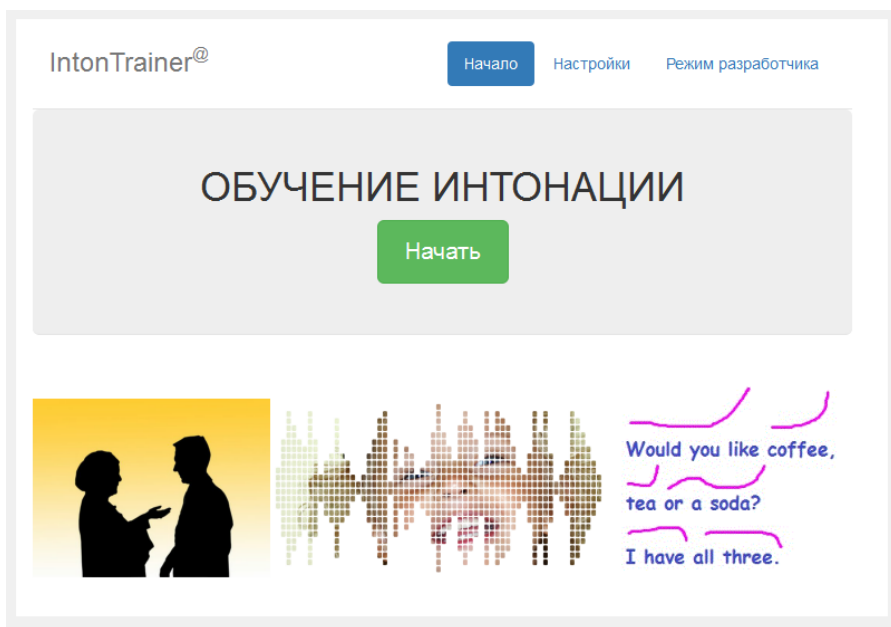


Рис. 5. Начальное окно приложения

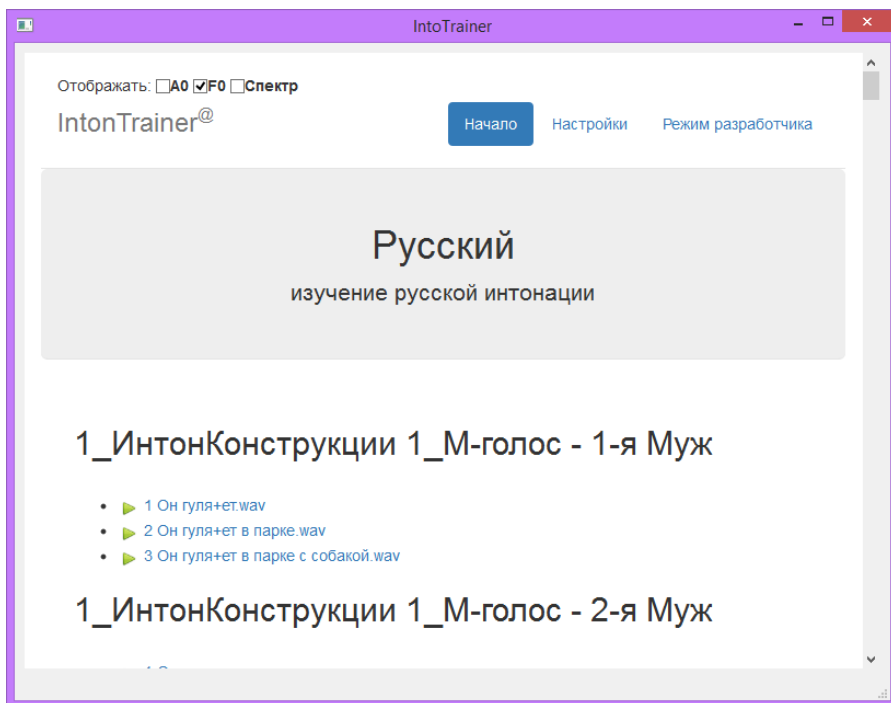
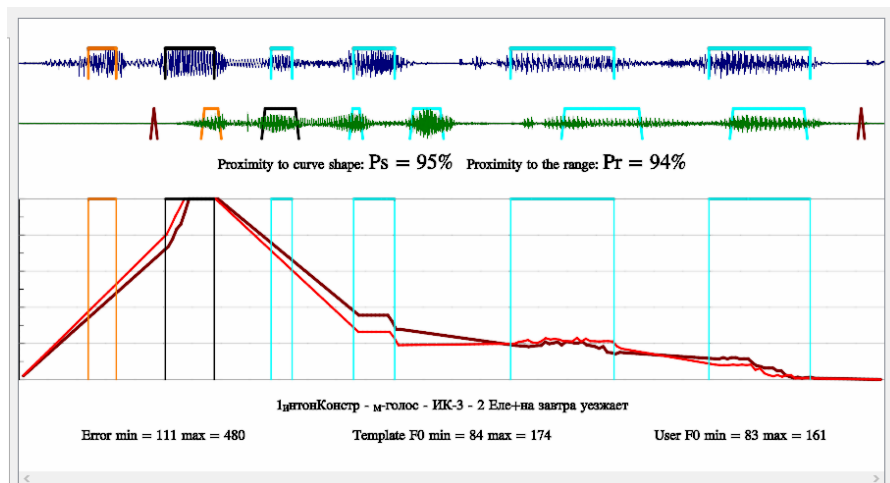


Рис. 6. Перечень обучающих заданий

**Рис. 7.** Окно тренировки. Прослушивание эталонной фразы и просмотр F0, A0



**Рис. 8.** Исходный результат сравнения F0 произнесённой фразы (тёмная линия) с эталонной (светлая линия)

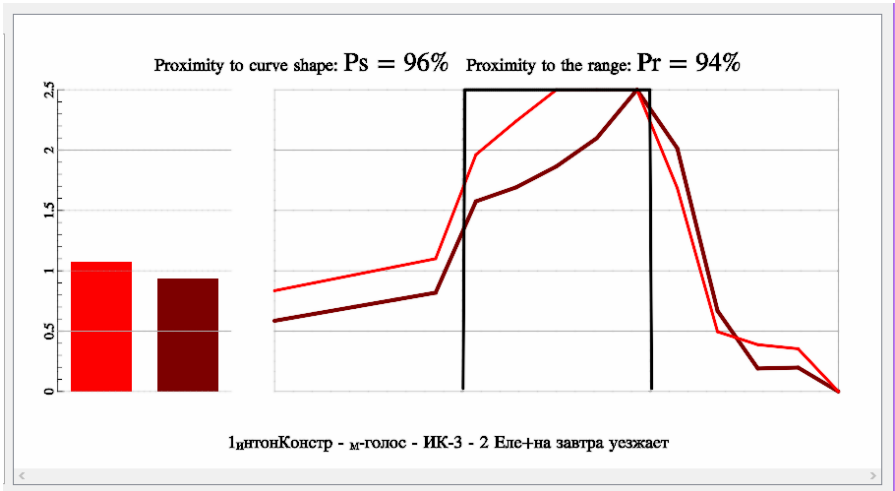
В верхней части окна представлены численные оценки (в %) близости эталонной и произнесённой форм УМР — Ps и диапазонов ЧОТ — Pr, рассчитанных по следующим формулам:

$$Ps = |r| * 100\%,$$

где  $r$  — выборочный коэффициент корреляции между кривыми ЧОТ, вычисляемый по стандартной формуле статистики;

$$Pr = \{(F0max / F0min)_{np} * 100\% \} / \{(F0max / F0min)_{эт}\},$$

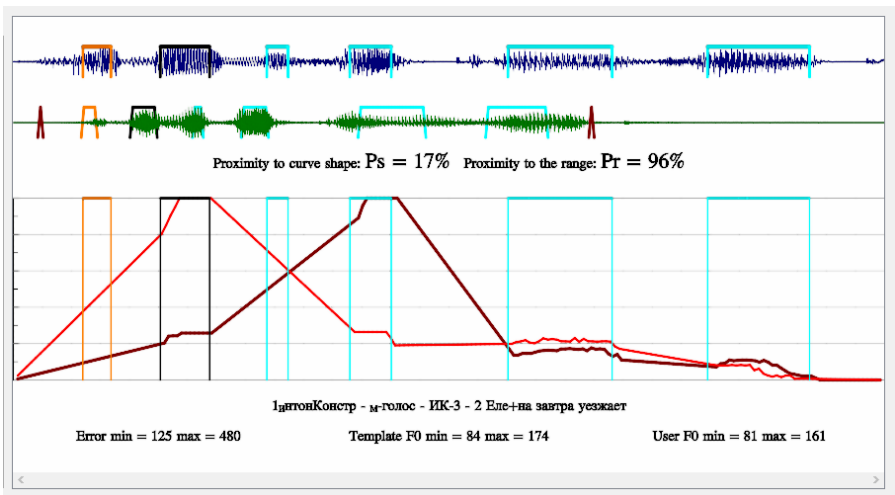
здесь (/) — знак деления.



**Рис. 9.** Рассчитанный результат сравнения УМП F0 произнесённой фразы (тёмная линия) с эталонной (светлая линия) и диапазонов их изменения

Для сравнения эффективности описанной процедуры на рисунках 10 и 11 представлены кривые изменения ЧОТ эталонной фразы «Елена завтра уезжает?» и неправильно произнесённой той же фразы со смещённым центром: «Елена за+втра уезжает?»

Как видно из сравнения результатов, представленных на рисунках 8 и 9, а также 10 и 11, использование предложенного УМП-представления даёт более строгую и более реалистичную оценку близости интонации эталонной и произносимой фраз.



**Рис. 10.** Исходный результат сравнения F0 произнесённой фразы (тёмная линия) с эталонной (светлая линия)

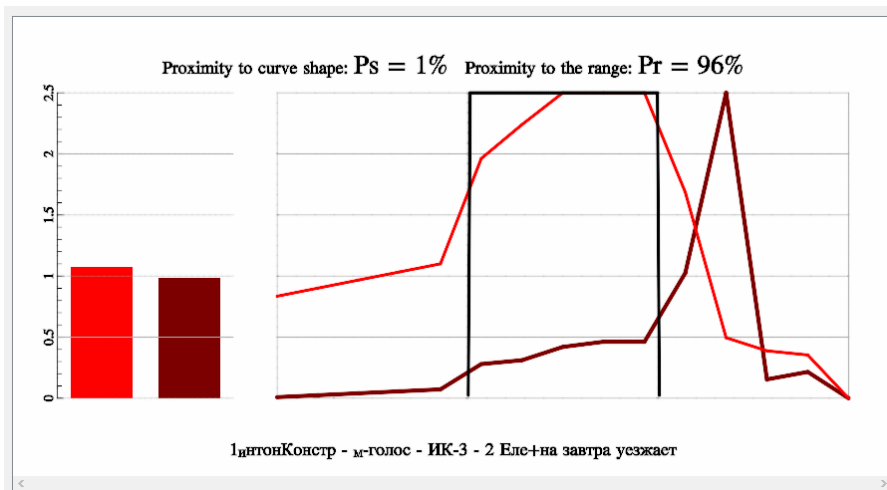


Рис. 11. Рассчитанный результат сравнения УМП F0 произнесённой фразы (тёмная линия) с эталонной (светлая линия) и диапазонов их изменения

## Заключение

В заключение укажем некоторые перспективы развития и возможных практических приложений разработанной системы:

- Описанная русскоязычная система интонационного тренинга — «*IntonTrainer*» — сможет найти применение не только при обучении РКИ, но, при соответствующей доработке, также в практике обучения нормативной интонации дикторов радио и ТВ, операторов колл-центров и людей других профессий, где требуется идеальное владение устной речью.
- Система «*IntonTrainer*» может стать инновационным продуктом, предназначенным для углублённого обучения интонации устной речи на различных других иностранных языках. Насколько нам известно, существующие аналоги такого рода систем ещё далеки от идеальных. К настоящему времени созданы демо-версии системы «*IntonTrainer*», ориентированные на обучение интонации **английского, немецкого и китайского** языков.
- Анализатор интонационных характеристик целесообразно использовать также в составе систем распознавания речи, что повысит точность распознавания благодаря снятию семантической неопределённости при интонационном членении непрерывной речи.
- Анализатор интонационных параметров, безусловно, будет полезен в подсистемах идентификации индивидуальных и эмоциональных характеристик речи говорящего, в сопоставительных исследованиях естественных языков, а также при создании адекватных интонационных моделей в системах синтеза речи по тексту для повышения интонационной выразительности синтезированной речи.

Авторы приносят свою благодарность анонимным рецензентам этой работы за их ценные замечания, которые мы постарались учесть при подготовке её окончательной редакции.

## Литература

1. *Catia Cucchiarini, J. van Doremalen, and Helmer Strik* (2008), DISCO: Development and Integration of Speech technology into Courseware for language learning. In: Proceedings of Interspeech, pp. 791–794.
2. *Sixuan Zhao* (2010), Computer Aided Evaluation of Intonation for Language Learning Based on Prosodic Unit Segmentation Biopolis, Singapore, Proceedings of the Second APSIPA Annual Summit and Conference, pp. 788–793.
3. *Lobanov B. M.* (1997) Continuous Speech Recognizer for Aircraft Application / Lobanov B. M., Levkovskaya T. V. // Proceedings of the 2<sup>nd</sup> International Workshop “Speech and Computer” — SPECOM’97 — Cluj-Napoca, pp. 97–102.
4. *Lobanov, B. M.* (2008) «IntoClonator» — Computer system of cloning prosodic characteristics of speech [Komp'yuternaya sistema klonirovaniya prosodicheskikh kharakteristik rechi] / B. Lobanov, L. Tsirulnik, O. Sizonov // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2008”, Bekasovo, pp. 330–338.
5. *Shimamura T.* (2001) Weighted Autocorrelation for Pitch Extraction of Noisy Speech // IEEE Transactions on Speech and AudioProcessing, Vol. 9, pp. 727–730.
6. *Bryzgunova E. A.* (1968), Sounds and Intonation of Russian Speech [Zvuki i Intonatsiya Russkoy Rechi] — Nauka, Moscow.
7. *Odintsova I. V.* (2011), Sounds. Rhythm. Intonation. [Zvuki. Ritm. Intonatsiya.] — Flinta-Nauka, Moscow.
8. *Lobanov, B. M.* (2006), Language- and speaker specific implementation of intonation contours in multilingual TTS synthesis / B. Lobanov, L. Tsirulnik, D. Zhadinets, E. Karnevskaia // Speech Prosody: Proceedings of the 3-rd International conference, Dresden, Germany, V. 2. — pp. 553–556.
9. *Lobanov B. M.* (2014) Universal Melodic Portraits of Intonation Patterns in Russian Speech [Universalnye melodicheskie portrety intonatsionnykh konstruktsiy russkoy rechi] / Lobanov B., Okrut T. // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2014”, Bekasovo, pp. 330–339.
10. *Lobanov B. M.* (2015) An Experience of Melodic Portraits Creation of Complex Declarative Sentences of Russian [Opyt sozdaniya melodicheskikh portretov povestvovatelnykh predlozheniy russkoy rechi] // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2015” Moscow, pp. 414–426.
11. *Lobanov, B. M.* (2016) Comparison of Melodic Portraits of English and Russian Dialogic Phrases / B. M. Lobanov // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2016” Moscow, pp. 382–392.

# СИНТАКСИС И СЕМАНТИКА ИМЕННОГО КОНТРАФАКТИВА<sup>1</sup>

**Лютикова Е. А.** (lyutikova2008@gmail.com),

**Татевосов С. Г.** (tatevosov@gmail.com)

МГУ имени М. В. Ломоносова; МПГУ, Москва, Россия

# SYNTAX AND SEMANTICS FOR NOMINAL COUNTERFACTUALS<sup>2</sup>

**Lyutikova E. A.** (lyutikova2008@gmail.com),

**Tatevosov S. G.** (tatevosov@gmail.com)

Lomonosov Moscow State University; Moscow State University  
of Education, Moscow, Russia

The paper explores syntax and semantics of nominal counterfactuals, in which the antecedent contains a DP instead of a clause (“if BY not DP”). We address a number of issues related to their internal structure and semantic interpretation, including obligatoriness of negation and its inability to license negative polarity items. We argue that despite the apparent nominal character of the antecedent, its interpretation is not different from that of regular counterfactuals whereby the antecedent provides the restrictor for the universal quantifier in the domain of possible worlds. We propose that the nominal constituent undergoes re-interpretation along the following path. A pronominal element *pro* adjoins to DP, which takes its reference in the domain of events. This element finds an antecedent event in the standards way regular pronominals do. It is related to the denotation of the DP by an unpronounced thematic relation. This derives a proposition of the form “that there is an event identical to the referent of *pro*, which || DP || is a participant of”. Since the referent of *pro* is an eventuality in the evaluation world, the counterfactual only get a coherent interpretation if the antecedent is negated. For the same reason, the antecedent is not an environment where NPIs can be licensed.

**Keywords:** counterfactual conditionals, event semantics, coercion

---

<sup>1</sup> Данное исследование поддержано грантом РФФ (проект 16-18-02003 «Структура значения и ее отображение в системе лексических и функциональных категорий русского языка», реализуемый в МПГУ).

<sup>2</sup> The research has been supported by Russian Science Foundation (project 16-18-0003 “Structure of meaning and its mapping into lexical and functional categories of Russian” at MSPU).



## 1. Введение

Цель этой статьи — обсудить синтаксис и семантику конструкции, иллюстрируемой в примерах (1a–b). В (Добрушина 2016, раздел 7.4.1.2) она охарактеризована как условная конструкция с протасисом жесткой структуры: «если + бы + не + существительное в именительном падеже».

- (1) а. *Если бы не Жанна, пропали бы мы обе: я и Лялька.* [НКРЯ]  
 б. *Так бы и закончилась вечеринка без происшествий, если бы не одна девушка, которая не знала, что много пить вредно.* [НКРЯ]

Данная конструкция на первый взгляд напоминает условное предложение с контрфактической семантикой (2)<sup>3</sup>, однако обладает рядом нетривиальных особенностей.

- (2) а. *И ничто не спасло бы Чикобаву, если бы Сталин не снял трубку и не оставил его жить.* [НКРЯ]  
 б. *И если бы один человек не снял копии с бумаг, концов в этой истории никогда бы не нашли.* [НКРЯ]

Во-первых, там, где в (2) находится синтаксическая составляющая, выражающая пропозицию, в (1) представлена именная группа, имеющая непропозициональное семантическое содержание<sup>4</sup>. Во-вторых, имеются семантические ограничения на референциальный статус этой именной группы: она должна быть референтной, хотя и необязательно определенной, ср. (1a–b). Появление в этой позиции выражения, не допускающего референтного прочтения, ведет к неграмматичности (3).

- (3) а. *\*Так бы и закончилась вечеринка без происшествий, если бы не какая-нибудь девушка.*  
 б. *\*Если бы не любой врач, пропали бы мы обе.*

В-третьих и в-четвертых, в этой конструкции отрицание играет примечательную роль. Оно обязательно: утвердительные аналоги предложений типа (1) неграмматичны.

<sup>3</sup> Сходство между условными контрфактическими предложениями и исследуемой конструкцией поддерживается также тем, что в обоих случаях наряду с союзом *если* могут использоваться союзы *когда* и *коли*:

<sup>(i)</sup> *Когда бы не чистописание, мальчик был бы самым счастливым учеником в классе.* [НКРЯ]

<sup>(ii)</sup> *Давно бы уж ему быть фейерверкером, коли бы не характер его.* [НКРЯ]

<sup>4</sup> Мы придерживаемся стандартного понимания пропозиции как характеристической функции множества миров, которая применительно к каждому конкретному миру дает истинностное значение. Именные группы, даже обозначающие свойства и характеристики, такие, как *способность* в (5), не могут быть истины или ложны в мире и тем самым не являются пропозициональными. Мы признательны рецензенту «Диалога» за комментарий, побудивший нас внести это уточнение.

- (4) а. \*Если бы Жанна, пропали бы мы обе: я и Лялька.  
б. \*Так бы и закончилась вечеринка без происшествий, если бы одна девушка, которая не знала, что много пить вредно.

Кроме того, отрицание неспособно лицензировать отрицательно-полярные элементы внутри именной группы, ср. (5). Это неожиданно, если учесть, что не только сентенциальное, но и термовое отрицание в русском языке может лицензировать NPI внутри той составляющей, к которой оно присоединяется, ср. пример (6) из (Рожнова 2009)<sup>5</sup>, а граница именной группы, как показано в (Корнакова и др. 2016), не является барьером для лицензирования NPI со стороны сентенциального отрицания, ср. (7а–б).

- (5) а. \*Если бы не способность ни на какую подлость...  
б. \*Если бы не способность на какую-либо подлость...  
(6) а. его неспособность ни на какую подлость (Рожнова 2009: 50)  
б. его неспособность на какую-либо подлость  
(7) а. Необходима была «прописка» в государственной филармонии, иначе мы не имели права давать **никаких** представлений. [НКРЯ]  
б. Комментируя это назначение, СМИ отметили, что генерал Бурутин не замечен в связях с **какими-либо** группировками ВПК. [НКРЯ]

Н. Р. Добрушина справедливо отмечает, что конструкции в (1) не могут рассматриваться как результат эллипсиса бытийного глагола (такую гипотезу выдвигает А. В. Гладкий (Гладкий 1982: 164)) либо другого предикативного материала. Следовательно, именной характер протасиса является конституирующей характеристикой конструкции. Далее мы будем называть эту конструкцию именным контрфактивом. В следующем разделе мы обсудим характеристики именных контрфактивов более подробно, а в разделе 3 наметим возможные пути объяснения некоторых их свойств.

## 2. Синтаксис и семантика именного контрфактива

В этом разделе мы сделаем два взаимосвязанных наблюдения. Во-первых, именная группа в протасисе именного контрфактива должна быть референтной. Во-вторых, протасис имеет имплицитную референцию к событию. Событие имеет место в актуальном мире и описывается определенной дескрипцией, содержание которой иногда задается явным образом, но чаще проясняется из контекста.

Рассмотрим эти обобщения по порядку. Составной частью протасиса именного контрфактива выступает именная группа, которая располагается справа

---

<sup>5</sup> М. А. Рожнова отмечает, что способностью лицензировать отрицательные местоимения в русском языке обладает отрицание при большинстве отлагольных и деадъективных существительных (Рожнова 2009: 50–55). Предметные существительные допускают термовое отрицание только в эксплицитном или подразумеваемом противительном контексте (ср. *Не чай он там пьет*).

от отрицательной частицы *не*. Важной характеристикой конструкции является ограничение на категориальный статус составляющей, входящей в протасис: она может представлять собой только именную группу. Составляющие других категорий, например, предложные группы, группы прилагательного, группы наречия, в данном контексте недопустимы:

- (8) а. *Если бы не одно обстоятельство...*  
 б. *\*Если бы не в одном случае...*  
 в. *\*Если бы не очень похожий...*  
 д. *\*Если бы не холодно...*

Все обычные типы зависимых, возможные в именной группе, допустимы и в именных контрфактивах. В примерах (9а-д) показаны различные модификаторы: указательное местоимение, притяжательное местоимение, причастный оборот, рестриктивное относительное предложение, а также аппозитивное относительное предложение, особый статус которого проявится чуть позже.

- (9) а. *Вода теплая какая, если бы не моя поясница, и я бы с вами поплескалась.* [НКРЯ]  
 б. *Если бы не этот пьяница пожарник, валявшийся в коридоре, Валиев наверняка не заметил бы окна под потолком.* [НКРЯ]  
 в. *Если бы не те средства, которые перешли от брата моей матери... то моей матери и ее сестре было бы очень трудно жить...* [НКРЯ]  
 д. *Если бы не спутники, которые оттащили его, всё кончилось бы скверно.* [НКРЯ]

Именная группа должна быть референтной. Примеры (1а-б) иллюстрируют две доступные возможности: определенные референтные и неопределенные референтные именные группы. Родовые именные группы, выступающие именами типов (kinds), представляют собой частный случай определенности и, соответственно, также допускаются в именных контрфактивах.

- (10) а. *Если бы не шалфей, то у меня было бы очень много проблем с лактацией.* [Я]<sup>6</sup>  
 б. *Если бы не муха-дрозофила, ученые еще долго бы не открыли законы наследования у людей.*

Нереферентные именные группы в составе именных контрфактивов чаще всего не допускаются. Невозможны, в частности, нереферентные отрицательно-полярные *-нибудь*, *-либо*, *бы то ни было* местоимения, как видно из примеров типа (11):

- (11) а. *\*Если бы не кто-нибудь из вас, пропали бы мы обе.*  
 б. *Если бы кто-нибудь из вас (не) пришел, пропали бы мы обе.*

Как показывает (11б), протасис условного предложения — это контекст, в котором отрицательно-полярные единицы допускаются, хотя одни из них,

<sup>6</sup> Примеры, обнаруженные с помощью поисковой системы «Яндекс», здесь и далее помечены как [Я].

возможно, выглядят естественнее, чем другие. Это означает, что неприемлемость (11а) — это именно свойство интересующей нас конструкции, а не более общее ограничение, присущее условным предложениям<sup>7</sup>.

Другие именные группы, имеющие логический тип обобщенного квантора ( $\langle\langle e,t\rangle,t\rangle$ ), в именных контрфактивах чрезвычайно редки. Так, кванторное выражение *минимум два* кажется несовместимым с исследуемой конструкцией, ср. (12а). При этом, однако, обнаруживаются примеры, в которых именной контрфактив включает *каждый*, как в (12b).

- (12) а. \*Если бы не минимум два мальчика, (которые...)  
б. Если бы не я, если бы не каждый из нас, твой отец заставил бы тебя вколоть сыворотку и сожрать Эрена. [Я]

Именные группы с *все* в этой конструкции также допускаются, как видно из (13). Известно, однако, что именные группы с кванторными словами типа *все* имеют индивидуную (тип *e*), а не кванторную (тип  $\langle\langle e,t\rangle,t\rangle$ ) интерпретацию, в частности, когда соединяются с определенной дескрипцией, см. (Brisson 1998). Именно такие интерпретации с необходимостью возникают у *все* в именных контрфактивах.

- (13) Я могу сказать одно, что если бы не все эти тяжелые обстоятельства, я бы рада была. [НКРЯ]

Кроме именных групп с индивидуной референцией, которые мы видели выше, в именных контрфактивах встречаются именные группы с событийной референцией, возглавляемые существительными, которые представляют собой дескрипции ситуаций.

- (14) Если бы не приход Даши, она сделала бы себе пару гренок с сыром и колбасой и запила их огромной чашкой крепкого кофе. [НКРЯ]

На первый взгляд, именные группы этого типа допускают тот же диапазон интерпретаций, что и индивидуальные именные группы. Примеры в (15) показывают, в частности, что они не могут иметь нереферентного прочтения.

- (15) а. \*Если бы не какой-нибудь из приездов Даши, ...  
б. \*Если бы не приезд кого-либо из нас, ...

Выясняется, однако, что событийные именные группы демонстрируют более сильное ограничение: помимо перечисленных, они не допускают также неопределенную референтную интерпретацию. Пара примеров (16а-б) показывает контраст в приемлемости неопределенной референтной интерпретации для индивидуальных и событийных именных групп.

- (16) а. Я, казалось, впал в настоящее бешенство, и если бы не какой-то беззаботный негр, который — это было написано на его лице — «никогда не волнуется и всегда счастлив», я просто оторвал бы ручку. [НКРЯ]

<sup>7</sup> Мы признательны рецензенту «Диалога» за комментарий, который склонил нас сделать это уточнение.

- b. *\*Если бы не какое-то ограбление банка, произошедшие в прошлом месяце, я бы не опасался за сохранность своего вклада*<sup>8</sup>.

Таким образом, в именных контрфактивах мы видим два типа именных групп: определенные референтные событийные и референтные индивидуальные<sup>9</sup>. С чем связана такая дистрибуция? Ответ на этот вопрос подсказывает следующее наблюдение: даже в тех именных контрфактивах, где именные группы вводят в рассмотрение индивиды, содержание высказывания включает в себя неперенное указание на событие с участием этого индивида. Вернемся к примеру (1a). Предложение (1a) не просто вводит в рассмотрение индивид, обозначенный именем Жанна, но и указывает на ситуацию с ее участием. Единственно возможное понимание (1a) состоит в том, что если бы эта ситуация не имела места в актуальном мире, произошло бы то, что описывается пропозицией в аподосисе. Ситуация может быть ясна из контекста или введена эксплицитно с помощью доступных языковых средств. Более подробно мы обсудим разные возможности ниже. Для целей этого раздела достаточно описательного обобщения: за antecedентами любых именных контрфактивов стоит указание на ситуацию. Если это так, перед нами открывается возможность увязать ограничения на референциальный статус событийных и индивидуальных именных групп в рамках единого обобщения:

- (17) Обобщение о семантике именного контрфактива

В протасисе именного контрфактива содержится указание на определенное событие из актуального мира.

Определенность событийных именных групп вытекает из (17) непосредственно. Референтность индивидуальных именных групп следует из (17) через промежуточное обобщение:

- (18) Обобщение о референтности участников референтного события

Участники события, имеющего место в мире  $w$ , существуют в мире  $w$ .

(18) соответствует очевидному элементу метафизики события: события с несуществующими участниками невозможны. Требование, которое налагает на индивиды (18), эквивалентно требованию референтности, что предсказывает диапазон возможных в именных контрфактивах именных групп без всяких дополнительных допущений.

<sup>8</sup> С точки зрения рецензента «Диалога», приемлемость местоимения *какой-то* повышается в следующем предложении:

(i) *Если бы не какие-то ограбления банков в прошлом месяце, Васю бы уже сделали генералом.*

Мы не разделяем это суждение. Можно, однако, заметить, что *какой-то* в некоторых специальных условиях, требующих дополнительного изучения, допускается и с определенными именными группами: *Тут какой-то Вася приходил*. Возможно, различие в суждениях связано с тем, что носители, считающие (i) приемлемым, способны реконструировать эти условия.

<sup>9</sup> Маргинальная возможность кванторных индивидуальных именных групп будет обсуждаться ниже.

Вопрос о том, откуда приходит ограничение, согласно которому в протасисе должно быть представлено (имплицитное) событие с определенным референциальным статусом, требует дополнительного прояснения. Мы осуществим это в следующем разделе. Это позволит придать обобщению (17) более точную форму.

Рецензент «Диалога» отмечает, что «протасис может отсылать и к будущему событию (например, в одной из интерпретаций примера (14))». Как это соотносится с обобщением (17)? Не имея возможности предложить здесь развернутый комментарий, сформулируем общую идею. Будущее можно понимать как модальный оператор, производящий квантификацию по метафизически доступным мирам (Sorley 2009), то есть таким мирам, которые имеют с актуальным общее прошлое. Если в (14) сделать временную референцию к будущему обязательной, например, введя в него соответствующее наречие (*Если бы не приход Даши завтра утром, она сделала бы себе пару гренков с сыром к завтрашнему обеду...*), мы получим предложение, в котором нет ни малейшей эпистемической неопределенности по поводу будущего события. Такое предложение предполагает, что событие имеет место во всех метафизических доступных вариантах развития актуального мира. Если, кроме того, принять в качестве исходного допущения, что индивиды и события не утрачивают идентичность при переходе от мира к миру, для (17) потребуется лишь самое незначительное уточнение: мы говорим об актуальном мире в прошлом и настоящем и обо всех метафизически доступных версиях актуального мира, если речь идет о будущем.

### 3. Анафора к событию

Интуитивное понимание именных контрфактивов, которые мы обсудили в разделе 2, сводится к тому, что они всегда вводят указание на событие (ср. Добрушина 2016). Это то общее, что происходит и с индивидуальными, и с событийными именными группами. В этом разделе нам предстоит придать этой интуиции более строгий смысл. Кроме всего прочего, следует увязать с этой интуицией стандартное понимание контрфактических предложений (как и других условных предложений) в терминах квантификации по возможным мирам (Kratzer 1986, von Stechow 1998, Iatridou 2000).

Контрфактические предложения типа (2a), повторяемого здесь как (19a) имеют интерпретацию типа (19b) (мы приводим неформальный перифраз; последние варианты формального анализа см., например, Romero 2014, Schultz 2014, Khoo 2015):

(19) а. *И ничто не спасло бы Чикобаву, если бы Сталин не снял трубку и не оставил его жить.* [НКРЯ]

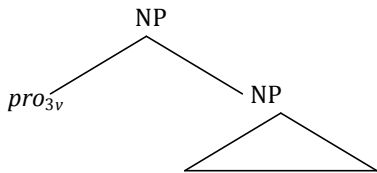
b. *оптимальные миры*<sup>10</sup>, в которых Сталин не снял трубку, одновременно являются мирами, в которых ничто не спасло Чикобава  
и  
наш мир не входит в множество миров, где Сталин не снял трубку.

Первая часть толкования присуща любым условным предложениям: протасис и аподосис понимаются как множество миров, связанных отношением подмножества. Вторая часть специфична для контрфактических условных предложений<sup>11</sup>. Таким образом, условное предложение (в том числе и с контрфактическим условием) — кванторная структура, в которой роль рестриктора и сферы действия играют, соответственно, протасис и аподосис, а квантор со значением всеобщности чаще всего не имеет поверхностного выражения. Как и всякая кванторная структура, она устанавливает отношение между двумя множествами и в этом смысле похожа на кванторные структуры типа *Все люди смертны* ( $\equiv \{x: x \text{ — человек}\} \subseteq \{x: x \text{ смертен}\}$ ). Единственная особенность условных предложений — тип множеств, связанных отношением подмножества, а именно, тип множества миров, или пропозиций.

Наблюдение, что именные контрфактивы вводят указание на событие, полностью укладывается в понимание условных предложений, представленное в (19b). Ключевой вопрос — как возникает указание на событие, причем событие, которое имеет место в актуальном мире? Наша гипотеза состоит в том, что именные контрфактивы содержат имплицитный прономинальный элемент с референцией к событию.

Этот элемент представляет собой событийный аналог обычных анафорических местоимений с индивидуальной референцией, типа *она*. С другой стороны, его можно понимать как фонологически не выраженный аналог местоимения *это* с событийной референцией (например, *Вася заснул, и это ее очень расстроило*).

(20)



где  $v$  — логический тип событий

В этом качестве *pro* может получать референцию так же, как это делают анафорические местоимения. В теоретико-модельной литературе, например,

<sup>10</sup> Понятие оптимального мира появилось в анализе модальности (в том числе в анализе условных предложений) благодаря А. Кратцер (см. Kratzer 1981 и последующие работы). Не имея возможности ввести и прокомментировать его здесь, мы отсылаем читателя к первоисточнику; в качестве введения в проблематику весьма удачно также учебное пособие (von Stechow, Heim 2015).

<sup>11</sup> В литературе ведется оживленная дискуссия о статусе этого семантического компонента. Традиционное понимание контрфактивов, которому следует, например, (Добрушина 2016), предполагает у него статус ассерции. Есть, однако, факты, которые склоняют целый ряд семантистов к мысли, что более правильно было бы анализировать его как импликатуру (см. Anderson 1951, Iatridou 2000, McCay 2015). Мы надеемся, однако, что сказанное ниже не зависит от окончательного решения этой проблемы.

это делается посредством функции, присваивающей значения несвязанным переменным, то есть соотносящей переменную с конкретной сущностью описываемого мироздания. Традиционная техническая реализация этой идеи состоит в том, что функция применяется к индексу, который несет местоимение (например, «3» в (20)), и возвращает конкретную сущность. Стандартная запись для референта местоимения имеет, например, вид  $g(3)$ , где  $g$  — присваивающая функция, а 3 — индекс местоимения.

Самый очевидный случай, известный как анафора, — это когда у местоимения есть антецедент, эксплицитно введенный в рассмотрение в предшествующем контексте. Другая возможность, которая обычно описывается как дейктическая, — когда местоимение непосредственно указывает на некоторую сущность в текущем контексте. Анафорические местоимения по необходимости определенно-референтны (если только не связаны каким-либо оператором; мы вернемся к этому ниже). Мы предполагаем, что именно таким способом у именных контрфактивов возникает указание на конкретное событие.

Следующий вопрос — как с помощью *pro* возникает пропозиция, необходимая для интерпретации условной конструкции. Для именных групп с событийной и индивидной референцией мы предполагаем близкие, но не идентичные механизмы.

Событийные именные группы — более простой случай, поскольку они с самого начала представляют собой выражение, описывающее событие. Мы предполагаем, что такие именные группы обозначают событийные предикаты, или функции логического типа  $\langle v, \langle s, t \rangle \rangle$ , которые принимают события в качестве аргумента и возвращают множество миров, в которых это событие имеет место. В частности, именную группу типа *приезд Даши* следует анализировать как функцию, которая характеризует множество событий  $e$ , таких, что  $e$  удовлетворяет предикату *приезжать*, а агенсом  $e$  выступает Даша. Соединяясь с *pro* и принимая *pro* как аргумент, эта функция создает множество миров, таких что в них представлено событие  $g(3)$ , которое удовлетворяет дескрипции *приезд Даши*.

Существенный для нашего сюжета компонент этой системы состоит в том, что областью значений присваивающей функции выступают сущности актуального мира. Соответственно, референт *pro*, такой, как  $g(3)$ , — это всегда событие актуального мира. Денотат клаузы протасиса — это пропозиция, которая показана в (21). Отрицание имеет в нем регулярную интерпретацию<sup>12</sup>.

(21) а. (Если бы) не *приезд Даши*...

б.  $\neg \exists e [e = g(3) \wedge \text{приезд}(\text{Даша})(e)]$

<sup>12</sup> В (21b) мы не даем интерпретацию союза *если* и частицы *бы*, отвечающей за «контрфактичность». Мы сохраняем нейтралитет в вопросе о том, есть ли в русском языке союз *если бы*, или *если* и *бы* следует понимать как независимые компоненты контрфактической конфигурации, каждый с собственной семантикой. Вслед за (Kratzer 1986) мы исходим из того, что пропозиция в протасисе используется для формирования рестриктора модального квантора, представленного в любом условном предложении. См., например, введение в семантическую теорию модальности и условного предложения в (von Stechow, Heim 2015).



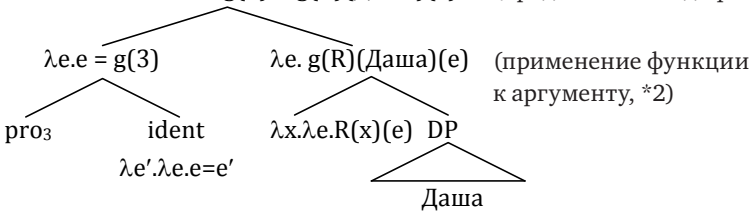
Для индивидуальных именных групп требуется дополнительное допущение. *pro* представляет собой событийную переменную, тогда как именные группы в именных контрфактивах типа (9a–d) обозначают индивиды. Чтобы они могли соединиться в осмысленное целое, необходимо ввести в структуру переменную, обозначающую семантическую роль R индивида в событии (которая также интерпретируется с помощью присваивающей функции). Конечный результат деривации протасиса в (22a) показан в (22b), а полная деривация до применения отрицания — в (22c)<sup>13</sup>:

(22) а. (Если бы) не Даша...

б.  $\neg \exists e [e = g(3) \wedge g(R)(\text{Даша})(e)]$

с.           ②  $\exists e [e = g(3) \wedge g(R)(\text{Даша})(e)]$  (экзистенциальное закрытие)

              ①  $\lambda e. e = g(3) \wedge g(R)(\text{Даша})(e)$  (предикатная модификация)



Это соответствует интуитивному пониманию именных контрфактивов. Они не просто описывают контекстно-релевантное событие. Роль, которую индивид играет в этом событии, также контекстно-зависима, ср. *Если бы не Даша, которая успела нажать на тормоз, старушка бы погибла* и *Если бы не Даша, переехав которую, автомобиль потерял скорость, старушка бы погибла*.

Теперь мы можем обратиться к анализу именных контрфактивов, содержащих в протасисе обобщенные кванторы. На именные группы, имеющие тип обобщенных кванторов, накладываются значительные ограничения. Например, предложения типа (23) с именными группами вида [каждый N] в протасисе находятся, по-видимому, по ту сторону грамматичности, ср. (23a–b):

<sup>13</sup> Рецензент «Диалога» делится с нами следующими недоумением: «В (21) было „приезд(Даша)(e)“, а здесь вместо описывающего событие предиката возникает только роль в событии — в то время как на самом деле должны быть и событие, и роль участника в нем. Наверное, это значит, что либо в (21) нужно писать „приезд(e) & Agent(Даша)(e)“, либо в (22) сказать, что R — это не только семантическая роль, а роль+событийная дескрипция». Спешим разрешить недоумение. В семантике событий есть два способа соотношения событий и их участников — дэвидсоновский и неодэвидсоновский. Во втором случае теория предполагает одноместные предикаты над событиями («приезд(e)»), к которым с помощью двухместных отношений типа *agent*, *theme* и т.д. привязываются индивиды («Agent(Даша)(e)»). Все это вместе означает: ‘имеется событие, которое удовлетворяет описанию «приезд», и оно состоит в отношении «агенса» с индивидом по имени Даша’. В первом, дэвидсоновском формате отношение между событиями и индивидами устанавливаются непосредственно самим глаголом. (21) следует читать как ‘имеется такое событие, что оно состоит в отношении «приезд» к индивиду по имени Даша’. При таком подходе глагол — это и есть семантическая роль, только не обобщенная (агенса, пациенса и т.д.), а максимально конкретная (приезд, поедание, написание и т.д.) Об эмпирических соображениях, позволяющих оценить достоинства и недостатки дэвидсоновской и неодэвидсоновской семантики событий, см. (Татевосов 2013).

(23) а. *Если бы не волонтеры, проведение олимпийских игр в Сочи или в любом другом месте было бы просто невозможно.* [НКРЯ]

б. *\*Если бы не каждый волонтер, проведение олимпийских игр в Сочи было бы просто невозможно.*

В этом месте мы сталкиваемся не с полным запретом на именные группы такого класса, а на очень специфические ограничения на их дистрибуцию, которые делают большинство из них неприемлемыми в произвольно взятом контексте. В корпусах находятся примеры типа (12b) или (24), которые, по интуиции носителей русского языка, полностью приемлемы:

(24) а. *Ведь если бы не каждый отдельно взятый участник, этот лагерь был бы уже не таким, каждый внес частичку своего сердца.*

б. *В общении стараюсь использовать идею взаимной благодарности за совместную работу: если бы не каждый из нас, работа была бы иной и не столь удачной.*

Есть ли разница между (23b) и (24a–b)? Мы предполагаем, что разница есть и что, более того, она полностью предсказывается нашим анализом. Неприемлемость (23b) и то, как она преодолевается в (24a–b), связана с характером взаимодействия кванторного слова *каждый* с событием, вводимым, согласно гипотезе, местоимением *pro*.

Местоимение *pro*, как и любое (несвязанное) местоимение, обозначает конкретную единичную сущность. Однако кванторное слово *каждый* является сильным дистрибутивным (Татевосов 2001), то есть наводит дистрибутивность не только по индивидам, но и по событиям. Предложение типа *Каждый мальчик видел двух девочек* можно понять так, что *каждый мальчик* имеет более узкую сферу действия, чем *двух девочек*; в этом случае речь идет об одних и тех же двух девочках. Однако тогда предложение с необходимостью сообщает, что на каждого мальчика приходится собственная ситуация встречи с двумя девочками и что встречи разных мальчиков происходят в разное время. Это же интерпретационное ограничение иллюстрируется примерами типа (25):

(25) *??Бомба разрушила каждый дом в поселке.*

Предложение (25) аномально потому, что единичность акта разрушения, производимого единственной бомбой, входит в противоречие с требованием множественности событий, которое наводится кванторным словом *каждый*.<sup>14</sup>

Это означает, что конфигурация вида (26), которая могла бы стоять за именными контрфактивами типа (23b), неприемлема по той же причине, что

<sup>14</sup> Отвечая на вопрос рецензента «Диалога», мы хотели бы подчеркнуть, что аномальность (25) не следует из того, что в единичном акте разрушения невозможно уничтожить все дома. Напротив: предложение *Бомба разрушила все дома в поселке* полностью приемлемо. Это показывает, что *каждый*, в отличие от *все*, нуждается в множественности событий, которые он включает в свою сферу действия. В английском языке так же различаются *all* и *every*, с одной стороны, и *each* с другой: *The bomb destroyed every building* || *all the buildings in the village* vs. *??The bomb destroyed each building in the village*.

и (25): *каждый* предполагает дистрибутивность по событиям, однако  $g(3)$  — это по необходимости единичное событие.

(26) [...  $\forall x$  [волонтер ( $x$ )  $\rightarrow$  ... ( $x$ )( $g(3)$ ) ...]]

(24a–b) подтверждают это: ни то, ни другое предложение нельзя понять так, что *каждый отдельно взятый участник* или *каждый из нас* выступает участником одного и того же события, гарантирующего истинность аподосиса. Напротив, для каждого индивида интуитивно предполагается собственное такое событие («каждый внес частицу собственного сердца»).

По всей видимости, в этом случае мы имеем дело с так называемой функциональной интерпретацией:

(27) [ ...  $\forall x$  [отдельно.взятый.участник ( $x$ )  $\rightarrow$   $\exists f_c$  ... ( $x$ )( $f_c(3)$ ) ... ]]

Можно предположить, что при функциональной интерпретации индекс у событийного местоимения *pro* не обрабатывается присваивающей функцией  $g$ , оценивающей все несвязанные переменные в пропозиции. Это происходит при помощи контекстно доступной функции (в (27) обозначена как  $f_c$ ) того же логического типа  $\langle n, e \rangle$ , которая связывается внутри пропозиции. Если связывание ее квантором существования происходит внутри сферы действия квантора всеобщности, вводимого кванторным словом *каждый*, как раз и возникает та интерпретация, которую мы наблюдаем в (24): каждого участника контекст обеспечивает (собственной) функцией  $f_c$ , которая сопоставляет индексу 3 на местоимении событие, в котором участвовал этот отдельно взятый участник. Или, проще: каждому отдельно взятому участнику контекст дает возможность сопоставить событие с его участием. От суммы этих событий контрфактически зависит поддержание лагеря в наилучшем возможном виде.<sup>15</sup>

В этом месте становится ясно, что возможность выполнить условие в (27) существенно зависит от того, что в текущем контексте известно о множестве-рестрикторе. Если оно конечно, компактно и четко очерчено, в любом контексте делается намного проще определить для каждого его элемента такую функцию, которая задает его вклад в сохранение лагеря в (24a) или его роль в поддержании работы в (24b). Для (23b) существенно труднее подобрать контекст, приписывающий каждому волонтеру событие, цепочка которых приводит к успешному проведению олимпийских игр. Максимальное сужение и конкретизация множества-рестриктора существенно улучшает это предложение:

(28) *Если бы не каждый из вас, присутствующих здесь волонтеров, проведение олимпийских игр было бы просто невозможно.*

Мы предполагаем, таким образом, что именно это ограничение отвечает за сравнительную предпочтительность именных контрфактивов с кванторными

<sup>15</sup> У нас нет возможности подробно рассматривать природу функциональных интерпретаций и их семантическую деривацию. Читатель, возможно, найдет полезным их обсуждение в связи с употреблением кванторных именных групп в вопросах (см. Chierchia 1993 и последующую литературу).

именными группами с *каждый*. По всей видимости, это ограничение вовсе нельзя соблюсти с множествами, у которых не задана четкая кардинальность, как это происходит в случае кванторных слов *минимум/максимум два* и т.п., см. (12а).

#### 4. Синтаксические корреляты пропозиции в именных контрфактивах

В этом разделе мы подробнее остановимся на возможных языковых средствах, позволяющих именной группе выступать представителем пропозиции.

Во-первых, именные контрфактивы могут содержать собственно анафорическую проформу для отсылки к событию — местоимение *это* (см. (Пекелис 2016)). В примере (29) antecedентом местоимения *это* выступает событие, которая эксплицитно описывается предыдущей клаузой.

(29) *Популярен он был необычайно, если бы не это, его наверняка давно бы опять арестовали.* [НКРЯ]

Следующую группу случаев образуют именные конструкции с существительными типа *факт, событие, обстоятельство*, при которых выражающее пропозицию изъяснительное придаточное выступает в качестве дополнения (30а). Существительное-вершина в таких конструкциях может отсутствовать, так что придаточное предложение оказывается вложено в местоименную оболочку (30b). В (Князев 2016) и (Лютикова 2016) для таких конструкций развивается анализ, в соответствии с которым CP заполняет позицию комплемента функциональной вершины D ([<sub>DP</sub> *то* [<sub>CP</sub> *что ...* ]]).

(30) а. *В позиции экономистов не было бы ничего странного, если бы не тот факт, что все они являются постоянными консультантами... российского правительства!* [НКРЯ]

б. *Если бы не то, что через час начинался рабочий день и в любое время могли прийти люди, я бы ни за что не расстался с ним.* [НКРЯ]

Как отмечалось выше, пропозицию может выражать именная группа, возглавляемая событийным существительным. В этом случае материала именной группы оказывается достаточно для создания дескрипции ситуации. Например, в (31) ситуация, к которой отсылает протасис, по-видимому, описывается как *Вере Михайловне разрешили приезжать ко мне*.

(31) *Я очень скоро протянул бы ноги, если бы не разрешение Вере Михайловне приезжать ко мне.* [НКРЯ]

Достаточно часто, однако, соотношение между именной группой, в том числе и событийной, в протасисе именных контрфактивов и подразумеваемой пропозицией не однозначно и основано на метонимии. Так, пример (32а) совместим с различными прочтениями протасиса: *если бы он не уехал, если бы Надя не уехала, если бы Наде не предстоял отъезд*, и т.п. Как отмечает Н. Р. Добрушина, «<i>i>менная группа с частицей *бы* обозначает целую ситуацию, и часто контекст содержит определенные фрагменты этой ситуации, например, в форме

причастного оборота, относительного придаточного или большого количества определений» (Добрушина 2016: 7.4.1.2). В (32b) именная группа строится на базе предметного существительного, однако модификатор именной группы — атрибутивное страдательное причастие — указывает на дополнительные дескриптивные характеристики события (например, *Ремни были пристегнуты*).

- (32) а. *Если бы не отъезд, вполне возможно, Надя вышла бы за него замуж.* [Я]  
 б. *Если бы не пристёгнутые ремни, я бы мог стукнуться головой о ветровое стекло.* [НКРЯ]

Один из примечательных случаев отсылки к ситуации при помощи именной группы — это использование аппозитивных относительных предложений. Рассмотрим (33).

- (33) *Если бы не спутники, которые оттащили его, всё кончилось бы совсем скверно.* [НКРЯ]

По-видимому, наиболее естественная интерпретация (33) состоит в том, что относительное предложение в составе именного контрфактива имеет аппозитивную интерпретацию: речь идет об определенных спутниках, и относительное предложение не производит рестрикции множества спутников до множества спутников, оттащивших героя. Функция относительного предложения в (33) иная: оно как раз и создает дескрипцию для той ситуации, которая выступает в протасисе контрфактива (*Если бы спутники его не оттащили, всё кончилось бы скверно*)<sup>16</sup>. Легко видеть, что наличие аппозитивного относительного предложения может лицензировать реинтерпретацию событийной именной группы в протасисе именного контрфактива: например, в (34) искомой пропозицией является не *Климат потеплел*, а скорее *Потепление климата нагрело антарктические воды до нужной крабам температуры*.

- (34) *Но переселенцы не смогли бы там выжить, если бы не потепление климата, которое нагрело антарктические воды до нужной крабам температуры.* [НКРЯ]

Однако не любые аппозитивные относительные предложения в именных контрфактивах создают дескрипцию для пропозиции. Например, в (35), как следует из более широкого контекста, пропозиция в протасисе в действительности имеет вид *Слава повторял реплики собеседников на все лады*; внимательность Славы к людям и их отношениям выступает лишь в качестве обоснования интереса Славы к релевантному обмену репликами.

- (35) <*Беседа Р. и Д. привлекла внимание С., который повторял реплики собеседников на все лады*> *Если бы не Слава, который всегда был внимателен к людям и их отношениям ... , я бы не запомнил обмена репликами с Диной Морисовной ...* [НКРЯ]

<sup>16</sup> Здесь напрашиваются параллели с французскими клефовыми структурами *c'est ... qui* и *(il) у а ... qui*, которые, помимо фокусной интерпретации — узкого фокуса в первом случае и репрезентационного фокуса во втором — могут использоваться для создания структуры вида *topic-comment* (Karssenberг, Lahousse 2015).

Можно высказать предположение, что интерпретация аппозитивного относительного предложения в именных контрфактивах коррелирует с независимо отмечаемой во многих работах дихотомией пресуппозитивных (*not-at-issue*) vs. нарративных (континуативных, *at-issue*) аппозитивных конструкций (см. Del Gobbo 2003; Nouwen 2007; AnderBois et al. 2010, 2015; Лютикова 2012; Koev 2013). Эта гипотеза, однако, требует дополнительных исследований.

Напрашивается вывод: рассматриваемая конструкция предъявляет к контексту лишь одно ограничение — он должен поддерживать ре-интерпретацию именной группы как пропозиции, предоставив информацию о ситуации актуального мира, в котором участвует соответствующий индивид. На языке предложенного нами анализа это равносильно тому, во-первых, что событийное местоимение *pro* должно найти себе антецедент, а во-вторых, что должна быть доступна минимальная информация о характере участия индивида в событии. Более никаких ограничений нет: подходит любой способ, которым всего этого можно добиться.

Чем может быть мотивирован именно такой сценарий превращения индивида в пропозицию? Полный ответ на этот вопрос возможен, по-видимому, лишь в более широком теоретическом контексте, а именно, в контексте теории, описывающей возможные реинтерпретации языковых выражений, происходящих, когда они оказываются в чужеродной семантико-синтаксической среде. В литературе явления, связанные с реинтерпретацией, чаще всего обсуждаются в связи с коэрсией<sup>17</sup>, преобразованием, которому подвергается значение выражения *a*, чтобы *a* могло соединиться с другим выражением *b* в осмысленное целое. Разные авторы обсуждают широкий диапазон коэрссионных явлений — онтологическую коэрссию, коэрссию типа, аспектуальную коэрссию, каузативную коэрссию и целый ряд других (Moens & Steedman 1988, Jackendoff 1997, Michaelis, Lambrecht 1996, de Swart 1998, Zucchi 1998, Coulson, Fauconnier 1999, Michaelis 2004, 2006, Koontz-Garboden 2007, Partee 2010, Lyutikova, Tatevosov 2015). Происходящее в именованном контрфактиве — это, по-видимому, явление того же ряда.

Условная конструкция связывает отношением подмножества два множества миров: миры, обозначаемые пропозицией в протасисе и миры, обозначаемые пропозицией в аподосисе. Наличие множества миров представляется необходимым условием того, чтобы условная конструкция могла иметь интерпретацию. Именные группы индивидуального типа, которые мы видим в именных контрфактивах, этому условию не удовлетворяют и вследствие этого должны подвергнуться реинтерпретации. Им приходится проделать путь от индивида к пропозиции. Возможностей здесь, вероятно, есть больше одной, но в любом случае их не слишком много. Та, которую обсуждаем мы, кажется в высшей степени естественной. От индивида мы переходим к утверждению о ситуации с его участием, то есть к пропозиции, которая истинна в мире *w* в том случае, если наш индивид принял участие в событии *e*. Как можно идентифицировать событие *e*, если от именованного именованного контрфактива не требуется, чтобы оно

---

<sup>17</sup> Словом «коэрсия» мы за неимением лучшего переводного эквивалента обозначаем явление, в англоязычной литературе известное как *coercion*. Английский термин представляется общеупотребительным. Мы не знаем, кто первым ввел его в научный оборот.

было эксплицитно выражено или описано внутри него? Способ, который, как мы подозреваем, задействован в описываемом случае, также представляется совершенно естественным. Это событие надо обнаружить в текущем контексте — например, как уже упоминавшееся или дейктически доступное. Как только это происходит, пропозиция, утверждающая участие референта именной группы в этом событии, готова, и вся конструкция получает осмысленную интерпретацию. Если это верная интуиция, то самый простой способ сделать ее явной — предположить имплицитное местоименное средство, выбирающее себе референт среди событий.

## 5. Загадка отрицания

Согласно предположению, сформулированному в предыдущих разделах, указание на событие, вводимое в протасисе именного контрфактива, имеет анафорический характер. Важное эмпирическое преимущество такого анализа состоит в том, что он позволяет нам объяснить асимметрию по полярности в примерах (1) и (3). (1a) повторяется как (36a), (3a) — как (36b):

- (36) а. *Если бы не Жанна, пропали бы мы обе: я и Лялька.*  
 б. *\*Если бы Жанна, пропали бы мы обе.*

Как показывает этот пример, утвердительный аналог протасиса невозможен. Мы утверждаем, что он невозможен благодаря анафорическому элементу *pro*, который мы полагаем частью механизма реинтерпретации именной группы, как показано в (22). После реинтерпретации создается пропозиция в (37).

- (37)  $\neg \exists e [e = g(3) \wedge g(R)(\text{Жанна})(e)]$

Как обсуждалось выше, *pro* получает референцию в актуальном мире; это местоимение обозначает происходящее в нем событие, в которое присваивающая функция отображает индекс на местоимении, например,  $g(3)$ .

Предложение с отрицательным именным протасисом, такое, как (36a), обозначает отношение подмножества между пропозицией в протасисе и пропозицией в аподосисе. Его семантика неформально показана в (38):

- (38) а. Те миры, в которых неверно, что имеется событие  $e$ , которое идентично контекстно-релевантному событию  $g(3)$  и в котором участвует Жанна,  
 — это подмножество миров, в которых верно, что мы обе пропали.  
 б. Наш мир не входит в множество миров, в которых неверно, что имеется событие  $e$ , которое идентично контекстно-релевантному событию  $g(3)$ , такому, что Жанна участвует в  $g(3)$ .

Обратим внимание на контрфактический элемент этой семантики в (38b). Нетрудно увидеть, что в отрицательном именном контрфактиве этот элемент имеет когерентную интерпретацию.  $g(3)$  — событие актуального мира. Пропозиция в (38b) 'Имеется такое событие, которое идентично  $g(3)$  и в котором

участвует Жанна', находится под отрицанием. Под отрицанием мы получаем множество миров, где нет такого события. Поскольку в нашем мире  $w_0$  событие  $g(3)$  есть, можно быть уверенным, что  $w_0$  —не один из них. Тем самым (38) удовлетворяется.

Рассмотрим, что происходит, если в аподосисе нет отрицания, (36b). Контрфактический компонент (38b) приобретает вид (39):

(39) Наш мир не входит в множество миров, где имеется событие  $e$ , которое идентично контекстно-релевантному событию  $g(3)$ , такому, что Жанна участвует в  $g(3)$ .

(39) — очевидное противоречие. Контрфактивность требует, чтобы наш мир не содержал в себе события, которое произошло в нашем мире. Нетрудно убедиться, что этот результат не зависит от лексического наполнения условной конструкции и с неизбежностью наступит у любого именной контрфактива, не имеющего отрицания над именной группой в аподосисе. Мы предполагаем, что именно в этом заключен источник неприемлемости предложения (36b) и аналогичных.

Это обсуждение показывает, что именной контрфактив не является, в отличие от обычного контрфактива (и прочих условных предложения) контекстом снятой утвердительности. Согласно нашему предположению, так происходит потому, что в реинтерпретации именной группы как выражения, обозначающего пропозицию, задействуется указание на событие, имеющее место в актуальном мире. В той степени, в которой снятая утвердительность — необходимое условие лицензирования отрицательно-полярных единиц, мы получаем объяснение их невозможности в именном контрфактиве, о которой говорилось в разделе 1. Иными словами, отрицательно-полярные единицы невозможны в именном контрфактиве по той же причине, по которой они невозможны в утвердительном матричном предложении в индикативе, в прошедшем времени совершенного вида типа *Кто-либо пришел*.

## Литература

1. *AnderBois S., Brasoveanu A., Henderson R.* (2010), Crossing the appositive-at-issue meaning boundary, *Semantics and Linguistic Theory*, Vol. 20, pp. 328–346.
2. *AnderBois S., Brasoveanu A., Henderson R.* (2015), At-issue proposals and appositive impositions in discourse, *Journal of Semantics*, Vol. 32(1), pp. 93–138.
3. *Anderson A. R.* (1951), A note on subjunctive and counterfactual conditionals, *Analysis*, Vol. 12(2), pp. 35–38.
4. *Copley B.* (2009), *The Semantics of the Future*, Routledge Outstanding Dissertations in Linguistics, New York.
5. *Coulson S., Fauconnier G.* (1999), Fake Guns and Stone Lions: Conceptual Blending and Privative Adjectives, in *Cognition and Function in Language*, edited by Fox B., Jurafsky D. and Michaelis L., CSLI Publications, Stanford, pp. 143–158.
6. *De Swart H.* (1998), Aspect shift and coercion, *Natural Language and Linguistic Theory*, Vol. 16, pp. 347–385.



7. *Del Gobbo F.* (2003), *Appositives at the interface*, Doctoral dissertation, University of California, Irvine.
8. *Dobrushina N. R.* (2016), *Semantics of the subjunctive: a corpus-based study of grammatical polysemy [Semantika kosvennogo naklonenije: korpusnoe issledovanie grmmaticheskoy polisemii]*. Dr. Sc. dissertation, HSE.
9. *Gladkij A. V.* (1982), *On the meaning of esli [O znachenii sojuza "jesli"]*, in *Semiotics and informatics [Semiotika i Informatika]*, Vol. 18, pp. 43–75.
10. *Iatridou S.* (2000), *The grammatical ingredients of counterfactuality*, *Linguistic Inquiry*, Vol. 31(2), pp. 231–270.
11. *Jackendoff R.* (1997), *The architecture of the language faculty*, MIT Press, Cambridge.
12. *Karssenbergh, L., Lahousse K.* (2015), *Two types of clefts in French: Information structure, existential presupposition & syntax*, Paper presented at IGG 41, Perugia, 27 February 2015, URL: <https://www.arts.kuleuven.be/ling/presentationalclefts>
13. *Khoo J.* (2015), *On indicative and subjunctive conditionals*, *Philosophers' Imprint*, Vol 15(32), pp. 1–40.
14. *Knyazev M.* (2016), *Licensing clausal complements: The case of Russian čto-clauses*. LOT Publications, Vol. 418, Utrecht.
15. *Koev T.* (2013), *Apposition and the Structure of Discourse*, Ph. D. dissertation, Rutgers University.
16. *Koontz-Garboden A.* (2007), *Aspectual coercion and the typology of change of state predicates*, Ph. D. dissertation, Stanford University.
17. *Kornakova E. V., Lyutikova E. A., Grashhenkov P. V.* (2016), *Licenzirujushhie svojstva otricanija v russkom jazyke*, *Rhema*, Vol. 4, pp. 58–82.
18. *Kratzer A.* (1981), *The notional category of modality, in Words, worlds, and contexts: New approaches in word semantics*, edited by Eikmeyer H.-J. and Riesen H., de Gruyter, Berlin, pp. 38–74.
19. *Kratzer A.* (1986), *Conditionals*, *Chicago Linguistics Society*, Vol. 22(2), pp. 1–15.
20. *Lyutikova E. A.* (2012), *Interpretation of the pronoun kotoryj in appositive relative clauses [Interpretacija mestoimenija kotoryj v appozitivnyh odnositel'nyh predlozhenijah]*, *Izvestija RAN. Language and Literature*, Vol. 71(1), pp. 3–23.
21. *Lyutikova E. A.* (2016), *Syntax of noun phrases in the articleless language [Sintaksis imennoj gruppy v bezartiklevom jazyke]*. Dr. Sc. dissertation, MGU.
22. *Lyutikova E., Tatevosov S.* (2015), *Recursive passivization: a causative coercion account*, in *Affix Ordering Across Languages and Frameworks*, edited by Manova S., OUP, Oxford, pp. 17–52.
23. *Mackay J.* (2015), *Actuality and fake tense in conditionals*, *Semantics and Pragmatics*, Vol. 8(12), pp. 1–12.
24. *Michaelis L.* (2004), *Type shifting in construction grammar: An integrated approach to aspectual coercion*, *Cognitive Linguistics*, Vol. 15, pp. 1–67.
25. *Michaelis L., Lambrecht K.* (1996), *Toward a Construction-Based Model of Language Function: The Case of Nominal Extraposition*, *Language*, Vol. 72, pp. 215–247.
26. *Moens M., Steedman M.* (1988), *Temporal ontology and temporal reference*, *Computational Linguistics*, Vol. 14, pp. 15–28.
27. *Nouwen R.* (2007), *On appositives and dynamic binding*, *Research on Language and Computation*, Vol. 5, pp. 87–102.

28. *Partee B. H.* (2010), Privative adjectives: subjective plus coercion, in *Presuppositions and Discourse: Essays offered to Hans Kamp*, edited by Bauerle R., Reyle U. and Zimmermann Th. E., Emerald Group Publishing, Bingley, pp. 273–285.
29. *Pekelis O. E.* (2016), The pronoun *eto* as a truth-conditional modifier (on the construction “if (when, even though) R, *eto* Q) [*Mestoimienie eto kak sredstvo modifikacii istinnostnogo statusa (o konstrukcii vida esli (kogda, hotja...) R, eto Q)*], *Voprosy jazykoznanija*, Vol. 5, pp. 49–78.
30. *Rozhnova M. A.* (2009), Syntactic characteristics of negative pronouns in Spanish and Russian [*Sintaksicheskie svojstva otricateľnyh mestoimenij v ispanskom i russkom jazykah*], Diploma thesis, RGGU.
31. *Tatevosov S. G.* (2001), Semantics of nominal constituents: universal quantifiers [*Semantika sostavljajushchix imennoj gruppy: kvantornuje slova*], *Nasledie*, Moscow.
32. *Tatevosov S. G.* (2013), Empirical issues in event semantics [*Semantika sobytija kak empiricheskaya problema*], *Philosophy of language and formal semantics [Filosofija jazyka i formalnaja semantika]* edited by Kusliy P. Alpha-M, Moscow, pp. 8–42.
33. *von Fintel K.* (1998), Quantifiers and ‘if’-clauses, *The Philosophical Quarterly*, Vol. 48(191), pp. 209–214.
34. *Von Fintel K., Heim I.* (2015), Intensional semantics. MIT lecture notes. <https://github.com/fintelkai/fintel-heim-intensional-notes/blob/master/Intensional-Semantics.pdf>
35. *Zucchi A.* (1998), Aspect shift, in *Events and grammar*, edited by Rothstein S., Kluwer, Dordrecht, pp. 349–370.

# ТОЛЕРАНТНОСТЬ ПО-РУССКИ И ПО-АНГЛИЙСКИ: КОРПУСНОЕ ИССЛЕДОВАНИЕ<sup>1</sup>

**Малафеев А. Ю.** (amalafeev@yandex.ru),  
**Орехова Е. А.** (kate.orekhova@gmail.com)

Национальный исследовательский университет «Высшая школа экономики», Нижний Новгород, Россия

Толерантность — сложное и отчасти противоречивое понятие, которое может быть неоднозначно оценено не только в разных культурах, но и в рамках одной культуры. В статье представлено сопоставительное исследование восприятия русско- и англоговорящими феномена толерантности на основе анализа корпусных данных. На начальном этапе исследования был создан пилотный корпус текстов о толерантности на основе текстов из сети Интернет. Корпус состоит из русскоязычного подкорпуса объемом 199607 словоупотреблений и из англоязычного подкорпуса объемом 210898 словоупотреблений. Результаты анализа мини-корпуса текстов о толерантности были верифицированы на материале корпусов общего профиля ruTenTen11 и enTenTen13 на платформе Sketch Engine. Авторы изучили и сопоставили «скетчи» лексем *толерантность*, *tolerance*, *толерантный* и *tolerant*. В частности, были проанализированы различные лексико-семантические поля и тематические группы коллокатов данных лексем, а также следующие конструкции: *X толерантности* и *X of tolerance*, *толерантность к X* и *tolerance towards X*, *толерантность и/или X* и *tolerance and/or X*. Кроме того, на основе корпусных данных было сформировано и проанализировано словообразовательное гнездо для слов *tolerance/толерантность*, в том числе большое количество окказионализмов. Корпусный анализ позволил выявить ряд характерных особенностей восприятия толерантности русско- и англоговорящими участниками Интернет-коммуникации.

**Ключевые слова:** толерантность, отношение к толерантности, составление специализированного корпуса, корпусный анализ, языковая картина мира

---

<sup>1</sup> Публикация подготовлена в рамках поддержанного РГНФ научного проекта № 15-04-00534 «Толерантность как культурная, политическая, лингвистическая проблема (анализ материалов СМИ и политического дискурса)».

# TOLERANCE IN RUSSIAN AND IN ENGLISH: A CORPUS STUDY

**Malafeev A. Y.** (amalafeev@yandex.ru),  
**Orekhova E. A.** (kate.orekhova@gmail.com)

National Research University Higher School  
of Economics, Nizhny Novgorod, Russia

Tolerance is a complex and partly contradictory concept that can be understood differently not only in different cultures, but also within the same culture. This paper presents a comparative study of the perception of tolerance by Russian and English speakers based on analysis of corpus data. At the initial stage of the study, the authors semi-automatically compiled a pilot web-based corpus of texts about tolerance. The corpus consists of a Russian-language subcorpus of 199,607 words and an English-language subcorpus of 210,898 words. After the mini-corpus was analyzed, the results were verified on the data from the general corpora ruTenTen11 and enTenTen13 using the Sketch Engine platform. The authors compared the word sketches for *толерантность* (*tolerantnost'*), *tolerance*, *толерантный* (*tolerantnyi*) and *tolerant*. In particular, this implied analyzing various lexical-semantic fields and thematic groups of collocates, as well as the following patterns: *X толерантности* (*tolerantnosti*) and *X of tolerance*, *толерантность к* (*tolerantnost' k*) *X* and *tolerance towards X*, *толерантность и/или* (*tolerantnost' i/ili*) *X* and *tolerance and/or X*. In addition, various derivatives of *толерантность* (*tolerantnost'*) / *tolerance* were discovered in the corpora and analyzed, including numerous nonce words. The corpus analysis enabled a deep insight into the way tolerance is perceived by Russian and English speakers.

**Key words:** tolerance, views on tolerance, specialized corpus building, corpus analysis, language worldview

## 1. Введение

Начиная с последнего десятилетия XX века, слово *толерантность* стало общеупотребительным, и содержание понятия «толерантность» активно обсуждается в философии, политологии, религиоведении и социологии. В зависимости от сферы *толерантность* определяется по-разному, например, как «социальная норма, определяющая устойчивость к конфликтам в полиэтничном межкультурном обществе» (Асмолов 2000: 7).

Существуют лингвистические исследования, посвящённые проблеме *толерантности*. Их можно условно разделить на два блока. Первый посвящён проблематике языковой *толерантности*. Согласно А. Г. Кудрявцеву (Кудрявцев 2011: 8), языковой *толерантностью* называется «совокупность вербальных и невербальных средств общения, которые индивид использует при взаимодействии с человеком или группой людей, так или иначе отличающихся от него».

К другим исследованиям в данном направлении можно отнести следующие: (Вахрушев 2010; Вахрушев, Нестерова 2011; Юрова 2015).

Второй блок исследований нацелен на выявление содержания понятия толерантности, определение её значимости для человека и культуры. Это, в основном, лингвокультурологические исследования, определяющие содержание концепта ТОЛЕРАНТНОСТЬ, а также его место в русской языковой картине мира (Стернин, Шилихина 2001; Михайлова 2005; Сумина 2014; Южакова 2014; Романова, Малафеев 2016). В русле такого подхода толерантность рассматривается как формирующийся концепт русской языковой картины мира, со свойственной ему национальной спецификой.

В данной работе предпринята попытка проанализировать и сопоставить различные аспекты понимания и восприятия русско- и англоговорящими толерантности на основе корпусных данных. Были поставлены следующие задачи: выявить отношение русско- и англоговорящих к феномену толерантности, узнать, в какой степени им ясно содержание данного понятия, с чем оно ассоциируется, а также определить, какое место занимает толерантность в языковой картине мира. Особенностью данного исследования является сочетание корпусного метода с элементами контент-анализа.

## 2. Материал исследования

### 2.1. Обоснование выбора материала

Существующие большие корпуса русского и английского языков, такие как ruTenTen11 и enTenTen13, состоящие из текстов из сети Интернет (14,55 млрд и 19,69 млрд словоупотреблений, соответственно) и доступные на платформе Sketch Engine (Kilgarriff и др. 2004), были рассмотрены нами в качестве возможного материала для исследования. Действительно, лексема *толерантность* употребляется в ruTenTen11 102 847 раз, а лексема *tolerance* в enTenTen13 — 248 647 раз. Однако приведенные частоты относятся к употреблению данных слов в нескольких значениях. Так, в Большом энциклопедическом словаре приводится три значения слова *толерантность*, из которых лишь одно актуально для проводимого исследования: «терпимость к чужим мнениям, верованиям, поведению»<sup>2</sup>. В Оксфордском словаре приведены четыре значения слова *tolerance*, в т. ч. “[t]he ability or willingness to tolerate the existence of opinions or behaviour that one dislikes or disagrees with”<sup>3</sup>.

В качестве небольшого эксперимента мы сформировали средствами Sketch Engine две случайные выборки по 100 контекстов употребления лексем *толерантность* и *tolerance*. Анализ показал, что в случае *толерантности* значение «терпимость к чужим мнениям, верованиям, поведению» актуализировалось в 84 контекстах, а в случае *tolerance* — лишь в 38.

<sup>2</sup> <http://dic.academic.ru/dic.nsf/enc3p/294398>

<sup>3</sup> <https://en.oxforddictionaries.com/definition/tolerance>

При автоматическом извлечении коллокатов или составлении скетчей (Kilgarriff and Rundell 2002) учитываются употребления заданных слов во всех значениях, что — в контексте нашего исследования — снижает полезность получаемых данных и затрудняет их интерпретацию, особенно в случае английского языка. Разумеется, это не означает невозможности использования корпусных данных для решения поставленных нами задач.

Следует отметить, что смещение извлекаемых коллокатов по значению центрального слова сильно зависит от конкретного морфолого-синтаксического шаблона. В качестве примера рассмотрим две группы коллокатов, входящих в скетч лексемы *tolerance*, сформированный в Sketch Engine на материале enTenTen13. Как видно из таблицы 1, в первом случае лишь шесть коллокатов относятся к изучаемому лексико-семантическому варианту (ЛСВ), зато во втором случае таких коллокатов — 41.

**Таблица 1.** Смещение коллокатов *tolerance* по значению центрального слова (на примере двух типов коллокатов)

Коллокаты	Количество	Примеры
<b>Прилагательные, являющиеся препозиционными определениями <i>tolerance</i></b>		
Релевантные	6 (12%)	<i>religious, racial, good-natured</i>
Нерелевантные	30 (50%)	<i>glucose, zero, fault</i>
Ошибочные <sup>4</sup>	24 (38%)	<i>patience, exercise, exacting</i>
<b>Существительные, связанные с <i>tolerance</i> через сочинительные союзы <i>and/or</i></b>		
Релевантные	41 (82%)	<i>acceptance, respect, patience</i>
Нерелевантные	9 (18%)	<i>dependence, resistance, withdrawal</i>

Приведённый пример показывает, что при анализе коллокатов слов *толерантность/tolerance* на базе корпуса общего профиля приходится в каждом случае обращаться к контекстам, чтобы удостовериться, что коллокат относится именно к изучаемому ЛСВ. Если же собрать коллекцию текстов, в которых обсуждается толерантность именно как «терпимость к чужим мнениям, верованиям, поведению», то, при условии репрезентативности этой коллекции, можно выявить языковые особенности и закономерности, свойственные дискурсу «вокруг толерантности». По этой причине было принято решение создать пилотный корпус текстов о толерантности. Результаты, полученные при анализе этого корпуса, были впоследствии верифицированы на материале больших корпусов общего профиля.

## 2.2. Создание корпуса текстов о толерантности

На первом этапе работы базовая часть русскоязычного и англоязычного подкорпусов была автоматически сформирована с помощью BootCaT (Baroni, Bernardini 2004). Этот инструмент позволяет создавать специализированные

<sup>4</sup> Выделенные системой, но на самом деле не относящиеся к данной группе коллокатов.

корпусы небольшого объёма по заданным пользователем ключевым словам, выгружая тексты из сети Интернет.

В интерфейсе BootCaT в поисковую машину Microsoft Bing были отправлены следующие запросы (порядок не имеет значения):

- (для русского языка) *толерантность, толерантное поведение, толерантность в social media, понятие толерантности, взгляд на толерантность, толерантность в СМИ, понимание толерантности, нетолерантный, толерантный, мнение о толерантности.*
- (для английского языка) *tolerant attitude, tolerant, tolerance, "mass media" tolerance, behavior tolerance, notion intolerant, intolerant, tolerance "social media", tolerance discussion, "mass media" intolerant, tolerant meaning, tolerance notion, tolerant behavior, meaning tolerance, meaning intolerant, intolerant attitude, intolerant behavior, tolerant "social media".*

На втором этапе мы вручную удалили тексты, не раскрывающие тематику толерантности, ошибочно попавшие в состав подкорпусов в результате автоматического сбора текстов. Затем подкорпусы были увеличены путём добавления вручную дополнительных текстов из сети Интернет (см. табл. 1). В корпусе представлены тексты следующих жанров: новостные, научные, словарные и энциклопедические статьи, записи в блогах, на форумах и др.

**Таблица 2.** Объём корпуса

Подкорпус	Русский		Английский	
	I	II	I	II
<b>Этап</b>				
<b>Типов</b>	16 718	33 467	11 661	13 426
<b>Токенов</b>	73 952	199 607	121 040	210 898

Несмотря на относительно небольшой объём собранного нами корпуса, релевантность входящих в него текстов для поставленных исследовательских задач была обеспечена ручной обработкой. Собранный корпус текстов (в лемматизированном и лемматизированном виде) о толерантности размещён авторами в сети Интернет для свободного использования в исследовательских целях<sup>5</sup>.

После получения первых результатов и их интерпретации мы воспользовались инструментарием платформы Sketch Engine для верификации сделанных выводов на более обширном материале (корпусы TenTen).

### 3. Анализ скетчей слов толерантность/ толерантный/ tolerance/ tolerant

С помощью Sketch Engine были сгенерированы скетчи лексем *толерантность/ толерантный/ tolerance/ tolerant* по данным собранного нами корпуса. При анализе каждой коллокации, а их было рассмотрено около 300, мы обращались к контекстам, чтобы более точно интерпретировать семантические связи между

<sup>5</sup> <https://sites.google.com/site/toleranceresearchproject/#corpus>

коллокатами и центральными словами. После этого мы также сформировали и проанализировали скетчи тех же лексем на основе корпусов ruTenTen11 и enTenTen13. В этом случае было рассмотрено около 900 коллокаций; при этом значительная часть коллокаций, приблизительно 35–40%, была отсеяна из-за нерелевантности (многозначность лексем *толерантность* и *толерантный* и ошибки Sketch Engine). Часть выводов, сделанных на основе анализа нашего корпуса, была подтверждена, часть — опровергнута. Кроме того, при анализе данных ruTenTen11 и enTenTen13 были получены новые результаты, которых нельзя было бы добиться при использовании только нашего небольшого корпуса. Учитывая ограничения на объём публикации, приведём лишь наиболее характерные наблюдения.

### 3.1. «Толкование» толерантности

Нами был проведён частотный анализ глаголов тематической группы с семантическим компонентом ‘означать’. В таблице 3 приведены частоты, полученные в Sketch Engine при обработке как нашего корпуса, так и корпусов TenTen. Здесь и далее относительные частоты коллокатов вычисляются следующим образом: абсолютная частота коллоката делится на суммарную частоту коллокатов данного типа. Например, в случае enTenTen13 частота глагола *mean* как коллоката, принадлежащего к группе “verbs with tolerance as subject”, делится на суммарную частоту коллокатов в данной группе, которая составляет 22 656.

Таблица 3. «Толкование» толерантности

Коллокат	Мини-корпус		TenTen	
<i>Означать</i>	19	0,093	131	0,031
<i>Предполагать</i>	11	0,054	83	0,020
<i>Подразумевать</i>	4	0,020	19	0,005
<b>Сумма</b>	<b>34</b>	<b>0,167</b>	<b>233</b>	<b>0,056</b>
<i>Mean</i>	12	0,065	395	0,017
<i>Presuppose</i>	0	0,000	9	0,000
<i>Imply</i>	2	0,011	47	0,002
<b>Сумма</b>	<b>14</b>	<b>0,076</b>	<b>451</b>	<b>0,020</b>

Анализ глаголов, часто выступающих сказуемым при подлежащем *толерантность/tolerance* и относящихся к рассматриваемой тематической группе, показал, что в речи русскоговорящих больше проявляется стремление объяснить, истолковать содержание понятия «толерантность», чем в речи англоговорящих. Возможно, это связано с относительной новизной этого понятия в русской культуре.

### 3.2. «Научный взгляд» на толерантность

Анализ словосочетаний *X толерантности* (где *X* — имя существительное, главное слово в словосочетании) для русского языка и *X of tolerance* для английского показал, что есть некоторые основания говорить о том, что «научный взгляд» на толерантность сильнее выражен в русскоязычных текстах. Однако



в данном случае корпусные данные довольно противоречивы и не могут интерпретироваться однозначно. Одна из возможных интерпретаций результатов, приведённых в таблице 4, может заключаться в том, что, по данным корпусов TenTen, русскоговорящие пишут о толерантности в среднем более научнообразно, чем англоговорящие: толерантность чаще становится объектом исследования, анализа. Что же касается данных мини-корпуса текстов о толерантности, они не подтверждают наблюдение, сделанное на основе TenTen. Скорее всего, это артефакт отбора текстов: так, словосочетание *notion(s) of tolerance* является неожиданно частотным в нашем корпусе, по сравнению с данными enTenTen13.

Таблица 4. «Научный взгляд» на толерантность

Коллокат	Мини-корпус		TenTen	
<i>проблема</i>	63	0,075	1711	0,037
<i>вопрос</i>	8	0,010	996	0,022
<i>понятие</i>	41	0,049	475	0,010
<i>определение</i>	32	0,038	331	0,007
<i>понимание</i>	12	0,014	201	0,004
<i>феномен</i>	11	0,013	156	0,003
<i>исследование</i>	5	0,006	155	0,003
<i>аспект</i>	12	0,014	149	0,003
<i>концепция</i>	4	0,005	134	0,003
<i>модель</i>	5	0,006	80	0,002
<i>тип</i>	8	0,010	52	0,001
<i>явление</i>	7	0,008	46	0,001
<i>теория</i>	4	0,005	46	0,001
<i>проблематика</i>	0	0,000	42	0,001
<i>анализ</i>	17	0,020	27	0,001
<b>Сумма</b>	<b>229</b>	<b>0,272</b>	<b>4601</b>	<b>0,099</b>
<i>issue</i>	6	0,034	206	0,008
<i>concept</i>	3	0,017	195	0,008
<i>notion</i>	28	0,161	134	0,005
<i>understanding</i>	6	0,034	131	0,005
<i>definition</i>	12	0,069	110	0,004
<i>view</i>	10	0,057	78	0,003
<i>meaning</i>	3	0,017	66	0,003
<i>type</i>	1	0,006	61	0,002
<i>problem</i>	0	0,000	49	0,002
<i>aspect</i>	1	0,006	45	0,002
<i>study</i>	0	0,000	45	0,002
<i>phenomenon</i>	0	0,000	40	0,002
<i>nature</i>	1	0,006	33	0,001
<i>analysis</i>	1	0,006	14	0,001
<i>theory</i>	0	0,000	9	0,000
<b>Сумма</b>	<b>72</b>	<b>0,414</b>	<b>1216</b>	<b>0,048</b>

### 3.3. Сопоставление различных групп коллокатов

Кратко представим основные результаты сравнения различных лексико-семантических полей и тематических групп коллокатов, относящихся к лексемам *толерантность/толерантный/tolerance/tolerant*. Некоторые тенденции наблюдаются уже на малом корпусе, но мы будем использовать более надёжные данные о частотности на основе корпусов TenTen.

В русском языке среди коллокатов *толерантности* более выражено лексико-семантическое поле с семантическим признаком «обучение» — *воспитание, урок, воспитывать, тренинг, педагогика* и др. (суммарная частота 4983) — по сравнению с полем «провозглашение»: *пропаганда, демонстрировать, проповедь* и др., (478). В английском ситуация обратна: поле «обучение» представлено значительно слабее (*foster, teach, instill, lesson*, 1034), чем поле «провозглашение» (*promote, preach, demonstrate* и др., 3242).

Также на русскоязычном материале более ярко выделяется коммуникативная сторона выражения толерантности: *коммуникативная толерантность* (173), *толерантное общение* (137), *толерантность и политкорректность* (569), тогда как в англоязычных текстах это гораздо менее заметно: *?communication/communicative tolerance* (0), *speech tolerance* (1), *tolerant communication* (2), *tolerance and political correctness* (64).

Интересно соотношение между количеством упоминаний религиозной и этнической толерантности:

**Таблица 5.** Религиозная и этническая толерантность

Коллокат (частота)	В сумме
<i>Религиозная</i> (790), <i>межконфессиональная</i> (182), <i>конфессиональная</i> (100), <i>межрелигиозная</i> (74)	1146
<i>Этническая</i> (1019), <i>межэтническая</i> (652), <i>межнациональная</i> (397), <i>этнонациональная</i> (22), <i>межрасовая</i> (11)	2101
<i>Religious</i> (4645), <i>inter-religious</i> (40)	4685
<i>Racial</i> (524), <i>ethnic</i> (122), <i>inter-ethnic</i> (20)	666

Эта разница в суммарных частотах (в 1,8 и 7 раз, соответственно) может означать, что для русскоговорящих этническая толерантность значительно более актуальна, чем религиозная, а для англоговорящих — наоборот. То же самое наблюдается и при анализе конструкций *толерантность к X* и *tolerance towards X*:

- *Конфессиям* (7) — *religions* (48), *faiths* (20)
- *Национальностям* (16) — *nationality* (2)

Характерно употребление в этих же конструкциях слов, обозначающих гомосексуализм и гомосексуалистов:

- (*Толерантность к*) *гомосексуализму* (16), *секс-меньшинствам* (8), *извращенцам* (8), *гомосексуалам* (7), *гей-сообществу* (4), *содомитам* (4).
- (*Tolerance towards*) *minorities* (25), *gays* (11), *homosexuality* (8), *homosexuals* (7).

Здесь обращает на себя внимание отсутствие среди англоязычных примеров пейоративов, аналогичных словам *извращенец* и *содомит*.

### 3.4. Аксиологический аспект толерантности

Также мы сопоставили конструкции *толерантность и/или X* и *tolerance and/or X*. В большинстве случаев *X* — это аксиологема, контекстуально синонимичная — в той или иной степени — слову *толерантность* или *tolerance*. Анализ данных конструкций может помочь понять, какие ценности, по мнению русско- и англоговорящих, наиболее близки к толерантности, а значит, и сделать выводы о различиях в понимании содержания слов *толерантности/tolerance*.

Мы рассмотрели по десять первых коллокатов, входящих в данные конструкции и ранжированных по убыванию статистической меры, для русского и английского языков (на материале корпусов TenTen). Для каждого из коллокатов были подобраны максимально близкие переводные эквиваленты. После этого мы сопоставили относительные частоты коллокатов и их эквивалентов, которые рассчитывались по тому же принципу, что и в разделе 3.1. Для ранжирования пар слов применялось деление меньшей из двух относительных частот на большую (см. табл. 6, в которой приведены некоторые примеры). В верхней части таблицы находятся пары близких по значению слов русского и английского языка, одно из которых гораздо чаще встречается рядом с *толерантностью* или *tolerance*, чем другое. В нижней части таблицы, напротив, приведены пары близких по значению коллокатов, которые не столь сильно различаются по частоте употребления в сочинительной связи со словом *толерантность* или *tolerance*.

**Таблица 6.** Примеры аксиологем, контекстуально синонимичных словам *толерантность/tolerance*

Коллокат	Абс. ч.	Относ. ч.	Эквивалент	Абс. ч.	Относ. ч.	Соотношение
<i>патриотизм</i>	610	0,0253	<i>patriotism</i>	21	0,0003	0,010
<i>acceptance</i>	2188	0,0267	<i>приятие (приятие — 0)</i>	10	0,0004	0,016
<i>гуманизм</i>	320	0,0133	<i>humanism</i>	36	0,0004	0,033
<i>политкорректность</i>	560	0,0233	<i>political correctness</i>	64	0,0008	0,034
<i>forgiveness</i>	597	0,0073	<i>прощение</i>	21	0,0003	0,074
<...>						
<i>equality</i>	729	0,0089	<i>равенство</i>	107	0,0044	0,499
<i>Плюрализм<sup>6</sup></i>	305	0,0127	<i>pluralism</i>	608	0,0074	0,586
<i>openness</i>	506	0,0062	<i>открытость</i>	251	0,0104	0,592
<i>respect</i>	3258	0,0398	<i>уважение</i>	938	0,0390	0,980

<sup>6</sup> И *плюрализм*, и *pluralism* входили в первые десять коллокатов в рассматриваемых конструкциях.

Разница в частоте употребления пар *патриотизм/patriotism, acceptance/приятие* и др. со словами *толерантность/tolerance* может свидетельствовать о лингвокультурных особенностях восприятия толерантности русско- и англоговорящими. Близость же относительных частот может быть интерпретирована как некоторое сходство между соответствующими фрагментами двух языковых картин мира.

### 3.5. Выявление словообразовательного гнезда

Приведём также результаты исследования словообразовательных гнёзд слов *толерантность* и *tolerance* по данным ruTenTen11 и enTenTen13. Для поиска родственных слов были использованы следующие запросы: *\*толер\** и *\*toler\**, соответственно.

Среди окказионализмов, выявленных в корпусе ruTenTen, наиболее заметным является окказионализм *толерастия* (731), образованный методом контаминации. В enTenTen13 также были обнаружены примеры окказионализмов, образованных методом контаминации: *tolerage* (4), *Britolerant* (1).

Примечательно, что вокруг вершины *толерастия* образуется собственное словообразовательное гнездо: *толераст* (1133), *толерастический* (101), *толерастный* (57), *дотолераститься* (14), *толерастить* (11) и др. Высокая частотность употребления, а также появление дериватов может свидетельствовать о переходе слов *толерастия* и *толераст* в разряд неологизмов.

В качестве других русскоязычных примеров могут служить слова *толерантнущий* (409), *толерант* (201), *толеразм* (19), *толеростия* (2), *толерантство* (5). Характерно, что в обоих подкорпусах были обнаружены слова, обозначающие «носителя» толерантности: *толераст*, *толерант*, *толерантщик* (18), *tolerator* (32), *tolerist* (4) и др.

## 4. Заключение

Таким образом, нами было проведено исследование различных аспектов восприятия толерантности русско- и англоговорящими на основе корпусных данных. На ранних этапах исследования мы использовали собственный мини-корпус текстов о толерантности; впоследствии мы верифицировали первоначальные наблюдения на данных больших корпусов общего профиля на платформе Sketch Engine.

Как показал анализ корпусного материала, восприятие толерантности русско- и англоговорящими довольно сильно различается. В этих различиях проявляются культурно-исторические особенности стран и народов, что вполне ожидаемо. Так, англоговорящие воспринимают толерантность как нечто самоочевидное и привычное, в то время как для многих русскоговорящих это относительно новый феномен, который необходимо осмыслить и правильно истолковать. Этническая толерантность обсуждается на русском языке значительно чаще, чем религиозная, тогда как англоговорящие, по-видимому, уделяют веротерпимости ещё большее внимание, чем расовой и этнической толерантности.

Мы также продемонстрировали некоторые различия в ценностных установках, касающихся толерантности: например, русскоговорящие чаще рассуждают о патриотизме, гуманизме и политкорректности в тесной связи с толерантностью, по сравнению с англоговорящими, хотя и те, и другие подчёркивают необходимость уважения к окружающим. Что же касается обилия в русскоязычном корпусе ироничных производных окказионализмов, образованных от слова *толерантность*, одна из возможных интерпретаций — это неприятие некоторыми носителями русской культуры идеи толерантности, сопротивление ей как чему-то навязанному извне. Впрочем, причины и особенности проявления этого неприятия, на наш взгляд, заслуживают специального изучения.

## Литература

1. *Asmolv A. G.* (2000), *Tolerance: from utopia to reality* [Tolerantnost': ot utopii — k real'nosti], *On the way to tolerant mind* [Na puti k tolerantnomu soznaniyu], Smysl, Moscow, pp. 4–7.
2. *Baroni M., Bernardini P.* (2004), *BootCaT: Bootstrapping corpora and terms from the Web*, *Proceedings of LREC*, Lisbon, pp. 1313–1316.
3. *Kilgarriff A. and Rundell M.* (2002), *Lexical Profiling Software and its lexicographic applications: a case study*, *Proceedings of EURALEX*, Copenhagen, pp. 807–818.
4. *Kilgarriff A., Rychly P., Smrz P., Tugwell D.* (2004), *The Sketch Engine*, *Proceedings of Euralex*, Lorient, France, pp. 105–116.
5. *Kudryavtsev A. G.* (2011), *Realization of tolerance in language as an element of multicultural competence* [Yazykovaya tolerantnost' kak element kul'turnoy kompetentnosti], *Bulletin of Nizhnevartovsk State University* [Vestnik Nizhnevartovskogo gosudarstvennogo universiteta], Vol. 4, pp. 8–12.
6. *Mikhaylova O. A.* (2005), *Tolerance and forbearance: a linguist's view* [Tolerantnost' i terpimost': vzglyad lingvista], *Philosophical and linguistic problems of tolerance* [Filosofskie i lingvokul'turologicheskie problem tolerantnosti], OLMA\*PRESS, Moscow, pp. 99–111.
7. *Romanova T. V., Malafeev A. Y.* (2016), *Tolerance from the national mentality viewpoint* [Tolerantnost' s tochki zreniya natsional'noy mental'nosti], *Issues of Psycholinguistics* [Voprosy Psikholingvistiki], Vol. 2 (28), pp. 226–243.
8. *Sharoff S.* (2006), *Creating General-Purpose Corpora Using Automated Search Engine Queries*, *Wacky! Working papers on the Web as Corpus*, Bologna, pp. 63–98.
9. *Sternin I. A., Shilikhina K. M.* (2001), *Communicative aspects of tolerance* [Kommunikativnye aspekty tolerantnosti], *Istoki*, Voronezh.
10. *Sumina E. S.* (2014), *Concept of tolerance as a component of the Russian philosophical worldview* [Kontsept tolerantnosti kak komponent russkoy filosofskoy kartiny mira], *Concept*, Vol. S3, pp. 36–40, available at: <http://e-koncept.ru/2014/14530.htm>
11. *Vakhrushev A. A.* (2010), *Verbal aggression and language tolerance in intercultural business communication* [Verbal'naya agressiya i yazykovaya tolerantnost' v mezhkul'turnom delovom obschenii], *Chelyabinsk State University Bulletin* [Vestnik Chelyabinskogo gosudarstvennogo universiteta], Vol. 4, pp. 53–58.

12. *Vakhrushev A. A., Nesterova N. M.* (2011), Contrastive analysis of language actualization, language aggression, and language tolerance means (based on English business communication) [Sopostavitel'nyy analiz sredstv yazykovoy aktualizatsii, yazykovoy agressii i yazykovoy tolerantnosti (na material angloyazychnoy delovoy koomunikatsii)], Chelyabinsk State University Bulletin [Vestnik Chelyabinskogo gosudarstvennogo universiteta], Vol. 37, pp. 20–25.
13. *Yurova T. A.* (2015), Prosodic features of language tolerance in the overall process of globalization [Prosodicheskiye cherty yazykovoy tolerantnosti v obschem protsesse globalizatsii], Philological Sciences. Issues of Theory and Practice [Filologicheskiye nauki. Voprosy teorii i praktiki], Vol. 11(53), pp. 212–215.
14. *Yuzhakova Y. V.* (2014), Tolerance from linguoculturological viewpoint [Tolerantnost' s pozitsiy lingvokul'turologii], Language and Culture (Novosibirsk) [Yazyk i kul'tura (Novosibirsk)], Vol. 10, pp. 64–68.

# ЭГОЦЕНТРИЧНОСТЬ В СЕМАНТИКЕ СЛОВА НИКАК<sup>1</sup>

**Падучева Е. В.** (elena.paducheva@yandex.ru)

ФИЦ ИУ РАН, Москва, Россия

Как пишет Э. Бенвенист [Бенвенист 1974: 295], «язык настолько глубоко отмечен выражением субъективности, что возникает вопрос, мог ли бы он, будучи устроенным иначе, вообще функционировать и называться языком». То, что Бенвенист называет **субъективностью** в языке, было впоследствии названо **эгоцентричностью**. Бенвенист предложил новый подход к изучению эгоцентричности, разграничив два плана ( $\approx$  контекста) употребления языка — план дискурса и план повествования. Впоследствии возникло понятие **режима интерпретации**: различается речевой (диалогический) и нарративный режим, см. Апресян 2003, Падучева 1986, 1996. Это разграничение принципиально для **эгоцентрических единиц языка** (эгоцентриков), семантика которых включает обращение к говорящему. Важную роль играет противопоставление **первичных** и **вторичных** эгоцентриков: вторичные допускают проекцию (когда роль говорящего выполняет какой-то иной субъект), тогда как первичные проекции не допускают. В работе показано, что вводное *никак* выражает предположение, сделанное на базе непосредственно воспринимаемой ситуации. Отмечаются следующие свойства *никак*. *Никак* часто употребляется в вопросительном предложении, однако в таком, которое не обязательно требует ответа; в контексте *никак* естественно встретить частицы и местоимение 2-го лица; *никак* часто выражает удивление; может быть синонимом *кажется*. Употребляется только в независимом предложении (и этим отличается от *неужели*, которое вполне допустимо в подчинительном контексте). Употребляется только в речевом режиме и является первичным эгоцентриком: всегда имеет подразумеваемым субъектом говорящего, не допуская ни гипотаксической, ни нарративной проекции.

## EGOCENTRICITY IN THE SEMANTICS OF THE RUSSIAN WORD *НИКАК*

**Paducheva E. V.** (elena.paducheva@yandex.ru)

Federal Research Center “Computer Science and Control”  
of the Russian Academy of Sciences, Moscow, Russia

---

<sup>1</sup> Работа выполнена при поддержке гранта 17-04-00554 «Семантические процессы в сфере местоименной и предикатной лексики: стратегии референции и концептуализации».

E. Benvenist (Benvenist 1974) proposed a new approach to the study of egocentricity in a language. He delimitates two contexts of language use: the **context of discourse** and the **context of narration** (*plan de discourse* and *plan de récit*). Thereafter the notion **register of interpretation** came into being. Two registers are distinguished: **dialogical register** and **narrative register**, see Apresjan 2003, Paducheva 1986, 1996. This distinction is of utmost importance for the egocentric entities of language (egocentrals) whose semantics appeals to the implied speaker. The opposition of **primary** (rigid) and **secondary** (shiftable) egocentrals is essential: secondary egocentrals undergo projections (when the role of the speaker is fulfilled by another subject), while primary egocentrals don't allow projections.

In the paper it is demonstrated that the parenthetical *nikak* expresses an uncertain statement made on the basis of directly observed situation. The following properties of *nikak* are distinguished. *Nikak* is often used in an interrogative sentence, but in an utterance that doesn't require an answer; different particles are often used in a *nikak*-sentence and the second person pronoun; *nikak*-sentence can express an explanation or interpretation of the situation before the eyes of the observer; *nikak*-sentence can express surprise; *nikak* may have *kazhetsja* 'it seems' as a synonym. *Nikak* is used only in an independent clause (thus differing from *neuzheli*, which is appropriate in a subordinate context). It is used only in the dialogical register and is a primary egocentral: it can only have a speaker as an implied subject and doesn't allow hypotactical or narrative projection.

**Key words:** egocentric linguistic entity register of interpretation primary egocentral secondary egocentral projection

Предметом анализа в работе является частица *никак* (как в контексте *Ты никак кататься задумал?*). К рассмотрению привлекается частица *неужели* как более основательно изученная, см. Булыгина, Шмелев 1997: 270–281. Предложение вида «*никак* Р» может быть сходно по смыслу с предложением вида «*неужели* Р». Так, в примере (1) можно заменить *никак* на *неужели* почти без изменения смысла:

- (1) Ты что, Мурзик, *никак*, на волю не хочешь? [Елена Хаецкая. Синие стрекозы Вавилона/ Обретение Энкиду (1997)]

Однако между *никак* и *неужели* имеется целый ряд различий.

### Различия между *никак* и *неужели*

1. Самый общий контекст для *неужели* — это когда «обсуждаемая ситуация», непосредственно соответствующая пропозиции Р, представляется говорящему невероятной, маловероятной. Например: — *Иванов ошибается.* — *Неужели?* В такой позиции *никак* исключено.
2. *Неужели* может выражать сомнение, решительное отрицание: *Неужели я тебя брошу?* Здесь тоже невозможно *никак*.
3. *Неужели* может выражать предположение. (В роли оснований, позволяющих сделать то или иное предположение, могут выступать опасения, надежды и прочее.) Пример из Булыгина, Шмелев 1997: *Неужели*



*он сейчас будет прощаться? — думала Елена (И. Тургенев) Никак* в таком контексте не исключено, но, в принципе, в сочетании с будущим временем затруднено, см. ниже.

4. *Неужели* может выражать простую заинтересованность — *Неужели Р?* Или просто *Неужели*? А *никак* так не употребляется.

### **Никак — реакция**

Слово *неужели* — это, как правило, реакция на предшествующее высказывание. Обычно вопросительное предложение играет инициативную, а не реактивную роль. Между тем, при анализе вопроса с *неужели* необходимо принимать во внимание предшествующую реплику, поскольку *неужели* реактивно. Предложение с *никак* тоже является реакцией. Но скорее не на фразу, а на ситуацию, которая представилась взору говорящего.

- (2) Рывком открыла дверцу машины. — Ты что, *никак*, кататься задумал? — Не кататься, а уехать, — сказал я. [Людмила Уварова. Одинокий с собакой снимет комнату // Библиотека «Огонек», 1990]
- (3) А ты, *никак*, и полы вымыл, смотрю? [Алексей Грачев. Ярый-3. Ордер на смерть (2000)]

Удивление — одна из естественных реакций на представившуюся взору ситуацию, и *никак* часто выражает удивление.

### **Никак в не вопросительном предложении**

*Неужели* употребляется только в вопросе. В отличие от *неужели*, *никак* может употребляться не только в вопросительном предложении. Примеры, где *никак* употребляется в утвердительном контексте.

- (4) Кривцов еле устоял, схватившись за фонарь. Ощупал голову, отнял ладонь — она была в крови. — Вася, я, *никак*, его убила... — охнула женщина, разъяренность исчезла из ее глаз, появилась бабья жалостливая деловитость. [Евгений Евтушенко. Ягодные места (1982)]
- (5) Что это братец нейдёт, вон, *никак*, гроза заходит. [А. Н. Островский. Гроза (1860)]
- (6) — Да ты, *никак*, уже поработал, и на славу... [Марина Дяченко, Сергей Дяченко. Привратник (1994)]
- (7) Батюшки мои, да, *никак*, это она радуется! Это она так веселится, что вот и ей с карусели платочком машут. [Федор Кнорре. Каменный венок (1973)]
- (8) — Оно <солнце>, *никак*, тоже на воскресник вышло, — поддержал его Кузьмин, также прикрываясь пятерней от солнца. [Борис Можаев. Саня (1957)]

## **Никак в словарях**

Словарь Даля определяет значение вводного *никак* как ‘кажется, видится, должно быть, по-видимому, я думаю’: *Никак приходил кто-то*. Отмечается склонность *никак* к употреблению в вопросительном предложении.

МАС трактует вводное *никак* как частицу (просторечную), которая указывает на неуверенность, предположительность высказывания (наречие *никак*, как в контекстах *никак не мог заснуть, без кровли никак нельзя* трактуется как омоним; ср. также эллиптическое *без него мне никак <нельзя>*).

Эти определения не отражают уникальной семантики и сочетаемостных особенностей частицы *никак*.

## **Никак и речевой режим интерпретации**

Предложения с вводным *никак*, вопросительным и невопросительным, употребляются только в прямой речи. *Никак* — это исключительная принадлежность речевого режима: *никак* невозможно в нарративе (см. Падучева 2011). В примере (9) это не очевидно, но более широкий контекст показывает, что это прямая речь:

- (9) Он, *никак*, узнав наше соединение, позавидовал нашему благополучию и отнял её опять у меня; [М. Д. Чулков. Пересмешник, или Славенские сказки (1766–1768)]

## **Никак и другие частицы**

В предложении с *никак* очень часто имеются разного рода частицы: *что* (*Ты что, никак, кататься задумал?*), *да* (*Да ты, никак, тонешь?!*); местоимение и частица *это* (*Батюшки мои, да, никак, это она радуется!*)

*Никак* можно сопоставить с еще более просторечным *нешто*, только вопросительным (— *Нешто и в самом деле выкупаться?..* [И. Грекова. На испытаниях (1967)]) и *небось*. В Жукова 2003 *никак* сравнивается с оборотом *не иначе, как*.

## **Никак и местоимение 2-го лица**

Чаще всего предложение с *никак* — это высказывание, обращенное ко второму лицу: подлежащее очень часто *ты* или *вы* (*Ты, никак, меня сватать задумал?*).

## **Никак и удивление**

Реакция на представившуюся взору ситуацию — это, как правило, удивление. Очень часто идею удивления в семантике *никак* передает текст, вводящий прямую речь:

- (10) — Ты, *никак*, меня сватать задумал? — от неожиданности растягивая слова, спросила она. — Задумал. Сватаю уж... [Валентин Распутин. Изба (1999)]
- (11) — Да это, *никак*, вояки короля Смоля?! — удивился Пешкин. [Ефим Чеповецкий. Приключения шахматного солдата Пешкина (1986)]
- (12) Он даже вилкой повел в воздухе, перечеркивая этот день. — Да ты, *никак*, крестишься? — удивилась Анна Николаевна. [Лазарь Карелин. Головокружение // «Юность», 1971]
- (13) Вдруг она всплеснула руками: — Коля, да, *никак*, это ты? Я тебя совсем не признала! [Алексей Грачев. Ярый-3. Ордер на смерть (2000)]

### ***Никак* и непосредственно наблюдаемая ситуация**

Высказывание с *никак* обычно является реакцией на нечто непосредственно наблюдаемое:

- (14) — Эй-эй, — закричал чиж. — Да ты, *никак*, тонешь?! — Буль, — кивнул головой Пешкин, и пузыри снова заплясали по луже. [Ефим Чеповецкий. Приключения шахматного солдата Пешкина (1986)]
- (15) На пороге стояли: васьлевская баба, Царьков-Коломенский и давешний молодой человек. — Да вы мне, *никак*, чудитесь? — сказал Васильев. [Евгений Попов. Темный лес (1970–2000)]

Примеры (16), (17) — редкий случай, когда высказывание не является реакцией на непосредственно наблюдаемые события:

- (16) — Не-ет, тому фамилья, *никак*, Фомин была... да, Фомин. [Л. М. Леонов. Вор. Части 1–2 (1927)]
- (17) Что это, думаю, *никак*, закупить этих спичек и этой соли так, чтобы и ходить не надо было. [Андрей Битов. Записки из-за угла (1964)] = разве что

Связь высказывания с непосредственно наблюдаемой ситуацией может быть опосредованная:

- (18) — А вы с жизнью, *никак*, простились, товарищ лейтенант? — залился булькающим смешком Нечаев. [Юрий Бондарев. Горячий снег (1969)]

### ***Никак* и время**

Обычно в предложении с *никак* время настоящее (поскольку это реакция на ситуацию перед глазами). Реже — прошедшее СВ в значении перфектного состояния.

- (19) — Что это, И.Д., *никак* ты на старости лет за изучение латыни принялся? [Р. Л. Антропов. Гений русского сыска И. Д. Путилин (1908)]

С будущим временем *никак* сочетается плохо. Разве что если будущее в значении 'собирается':

(20) *Никак* он сейчас будет прощаться (сконструированный пример) = 'собирается прощаться'.

## ***Никак* и вопрос**

Чаще всего *никак* употребляется в вопросительном предложении:

(21) А ты, *никак*, сегодня под эропланом ночевал? [Владимир Войнович. Жизнь и необычайные приключения солдата Ивана Чонкина (1969–1975)]

Но вопрос может быть такой, который не требует ответа.

(22) — Эй, да ты, *никак*, в моей одежде, Сфинкс? [Мариам Петросян. Дом, в котором... (2009)]

(23) — Да ты, *никак*, не узнаешь меня? [Людмила Гурченко. Аплодисменты (1994–2003)]

(24) Да ты, друг мой, *никак*, давно спишь?! [Василий Голованов. Остров, или оправдание бессмысленных путешествий (2002)]

Знак вопроса может быть условностью:

(25) А ты, *никак*, и полы вымыл, смотрю? [Алексей Грачев. Ярый-3. Ордер на смерть (2000)]

В (26), (27), напротив, можно было бы поставить знак вопроса:

(26) — Да что ты, *никак*, Прохорыч, с ума спятил! [М. Е. Салтыков-Щедрин. Пропала совесть (1878)]

(27) МОЦАРТ (пытаясь помешать). Ты, *никак*, с ума сошел! Ты что! [Людмила Петрушевская. Сцена отравления Моцарта (1985–1990)]

О том, что вопросительное предложение может не быть вопросом, свидетельствуют диалогические реакции. На вопрос должен следовать ответ. А Угадал в (28) — это реакция на утверждение:

(28) — Да это, *никак*, Пыжик?! — Угадал, — ответил Василий. [Ефим Чеповецкий. Приключения шахматного солдата Пешкина (1986)]

(29) «Эх, мать-перемать! — догадался шофер. — Да это ж, *никак*, лошади! [Нина Воронель. Без прикрас. Воспоминания (1975–2003)]

Согласно Жукова 2003, высказывание с *никак* — это неуверенная констатация наличия некоторой (наблюдаемой) ситуации (*Ты никак постригся?*).

Впрочем, вводящий текст может трактовать высказывание с *никак* как вопрос:

(30) — Да ты, *никак*, на нее обиделся? — спросила Поля. — На кого? — будто бы удивился я. [Альберт Лиханов. Кикимора (1983)]

(31) — Да ты, *никак*, выпивши? — спросила жена, встречая Сергея Митрофановича на крыльце. [Виктор Астафьев. Ясным ли днем (1967)]

В (32)–(34) это вопрос, на который дается ответ:

(32) — Да вы, батенька, *никак*, трус? — Все мы, знаете ли, трусоваты в известных обстоятельствах». [Василий Голованов. Остров, или оправдание бессмысленных путешествий (2002)]

(33) — Ты, *никак*, боишься меня? — Я вообще не боюсь мужчин... [Татьяна Моспан. Подиум (2000)]

(34) — Ты что, *никак*, кататься задумал? — Не кататься, а уехать, — сказал я. [Людмила Уварова. Одинокий с собакой снимет комнату // Библиотека «Огонек», 1990]

Однако чаще вопросительное предложение не трактуется как вопрос; скорее это предположение, основанное на непосредственном восприятии; здесь, по Далю, *никак* = 'кажется':

(35) — Ты, *никак*, заговариваешься? — ужаснулась Татьяна. [Ирина Муравьева. Филемон и Бавкида (1995)]

В Жукова 2003 отмечается, что *никак* может быть гипотетическим объяснением наблюдаемой ситуации (*А ты похорошела. Никак зубы вставила?*).

Если нет ни удивления, ни вопроса, то *никак* сближается с *небось*, которое выражает предположение:

(36) — Их уж, *никак*, лет двадцать не продают? [Борис Можаяев. Живой (1964–1965)]

### ***Никак* и вводность**

При начальной позиции во фразе у *никак* пропадает вводность; *никак* становится равно как *будто*:

(37) Отпирай же двери!.. — Ладно!.. *никак* с ума спятила... Пусти да пусти, а куды я пушу?.. [В. В. Крестовский. Петербургские трущобы. Книга о сытых и голодных. Роман в шести частях. Ч. 4 (1864)]

### ***Никак* и подчинительный контекст**

*Никак* может употребляться только в независимом предложении. Большая часть примеров с *никак* вообще не может быть поставлена в подчинительный контекст из-за местоимений и частиц. Пример (38б) показывает, что *никак* неподчинимо.

(38) а. Их уж, *никак*, лет двадцать не продают [Борис Можаяев. Живой (1964–1965)];  
б. \*Он сказал, что их уж, *никак*, лет двадцать не продают (надо сказать — *наверно*).

Между тем *неужели* может употребляться в подчиненной позиции — в значении косвенного вопроса или пропозициональной установки.

(39) Спрашивать его о том, что *неужели* он, с его опытом жизни, не знал о возможных последствиях походов в Sozialamt, мне не хотелось. [Михаил Гиголашвили. Адский рай // «Зарубежные записки», 2007] — косвенный вопрос: спрашивать, верно ли, что он не знал; маловероятно, что не знал.

Думал о людях и обезьянах: в том смысле, что — *неужели* люди произошли от обезьян? [Василий Шукшин. Штрихи к портрету (1973)] = что не может быть, чтобы произошли

— В разговоре вы сказали, что *неужели* советская власть не заступится за ваше ремесло? [В. А. Солоухин. Владимирские просёлки (1956–1957)] = что не может быть, чтобы не заступилась;

Профессор, горько усмехаясь, начинал выговаривать ей, что *неужели* она за двадцать пять лет жизни с ним не научилась разбираться в солнечной системе, что Сатурна в это время и в этой четверти неба не бывает. [М. М. Зощенко. Возвращенная молодость (1933)] = что удивительно, что она не научилась.

У нее, против воли, промелькнула мысль, что *неужели* Ирина сама не замечает, какая она стала? [П. С. Романов. Человеческая душа (1926)] = что не может быть, чтобы не замечала

— Я вот часто думаю, — продолжала она, — что *неужели* же я должна была после такой ужасной семейной жизни умереть для всего и не позволить себе полюбить другого... [А. Ф. Писемский. Мещане (1877)] = что не может быть, чтобы должна была;

— Я все-таки, любезный друг, желаю знать определительно, что *неужели* же между тобой и Домной Осиповной совершенно и навсегда все кончено? [А. Ф. Писемский. Мещане (1877)] = желаю знать, верно ли, что все кончено; я не могу в это поверить.

И меня теперь пугает одно, что *неужели* это когда-нибудь изменится и мы должны будем расстаться!.. [А. Ф. Писемский. Просвещенное время (1875)] = что может быть, что это изменится.

Я спросил г-на Пейтэе, что *неужели* сии гульбища никогда более не посещаются? [Ф. Ф. Матюшкин. Журнал кругосветного плавания на шлюпе «Камчатка» под командою капитана Головина (1817)] = верно ли, что не посещаются; удивительно, если это так.

Так что *никак* в этом отношении отлично от *неужели*. В этом *никак* отлично и от *кажется*, которое хоть и не без труда, но все-таки подчинимо:

(40) Говорит, что, *кажется*, и Леонов не подписал. [А. Т. Твардовский. Рабочие тетради 60-х годов (1968) // «Знамя», 2003]

Итак, частица *никак* — типичный первичный эгоцентрик: не употребляется нигде, кроме как в диалоге — в прямой речи.

## Литература

1. *Apresjan Ju. D.* (2003), Linguistic terminology of the dictionary [Lingvističeskaja terminologija slovarja] Apresjan Ju. D. e.a. New explanatory dictionary of synonyms of the Russian language. The third issue/ [Novyj ob"jasnitel'nyj slovar' sinonimov russkogo jazyka. Tretij vypusk] Moscow, Jazyki slavjanskoj kultury, pp. XVIII–XLIX.
2. *Benvenist E.* (1974), General linguistics [Obščhaja lingvistika] Moscow, Progress.
3. *Bulygina T. V., Shmelev A. D.* (1997), Linguistic conceptualization of the world (on the basis of Russian grammar) Jazykovaja konceptualizacija mira (na materiale russoj grammatiki) Moscow, Jazyki russoj kultury.
4. *Paduceva E. V.* (1986), The semantics of aspect and the point of reference [Semantika vida i točka oščeta] Izvestija AN SSSR. Series literature and language. Vol. 45, № 5, pp. 413–424.
5. *Paduceva E. V.* (1996), Semantic investigations [Semantičeskie issledovanija] Moscow, Jazyki russoj kultury.
6. *Zhukova O. V.* (2003), *Nikak/ne inache kak*: a comparative analysis of the two discursive words [Nikak/ne inache kak: sravnitel'nyj analiz dvux diskursivnyx slov]. D. Palliard (editor) Discursive words of Russian: contextual variation and semantic unity [Diskursivnye slova russkogo jazyka: kontekstnoe var'irovanie i semantičeskoe edinstvo] Moscow, pp. 131–143.

# SEMANTIC HALO OF A METER: A KEYWORD-BASED APPROACH

**Piperski A. Ch.** (apiperski@gmail.com)

Russian State University for the Humanities; National Research University Higher School of Economics, Moscow, Russia

The semantic halo of a meter (*semantičeskij oreol metra*) is a notion that was introduced by Mikhail Gasparov to describe semantic invariants of poetic texts composed using the same metrical scheme. Most studies addressing this phenomenon have been based on expert knowledge of the text corpus. In this study, I propose an automated approach to analyzing the semantic halo of various meters based on keyword extraction, using a simple measure of keyness as developed by Adam Kilgarrieff. The method is applied to texts from the Poetic subcorpus of the Russian National Corpus. It allows us to discern basic motifs that are very close to those identified by literary scholars, which proves that it is a promising way to analyze the semantic halo of a meter. Some novel associations can also be inferred from the keyword lists. Clearly, keyword extraction cannot replace profound expert knowledge, but it can serve as a useful first step of the analysis.

**Keywords:** semantic halo of a meter, verse, keyword extraction, Russian National Corpus

# СЕМАНТИЧЕСКИЙ ОРЕОЛ МЕТРА И ИЗВЛЕЧЕНИЕ КЛЮЧЕВЫХ СЛОВ

**Пиперски А. Ч.** (apiperski@gmail.com)

РГГУ; НИУ ВШЭ, Москва, Россия

Понятие семантического ореола метра было введено М. Л. Гаспаровым для описания семантических инвариантов поэтических текстов, написанных по одной и той же метрической схеме. Большинство исследований в этой области базируются на экспертном знании текстов, однако в данной работе предлагается автоматизированный подход к анализу семантических ореолов различных метров. Он опирается на извлечение ключевых слов методом, предложенным А. Килгарриффом. Этот подход иллюстрируется на материале текстов из Поэтического корпуса Национального корпуса русского языка. Он позволяет выявить базовые мотивы, которые близки к мотивам, выделяемым для соответствующих метрических схем литературоведами, что демонстрирует перспективность описываемого метода. Кроме того, он позволяет обнаружить некоторые ранее не отмечавшиеся ассоциации. Разумеется, извлечение ключевых слов не может заменить собой глубокого экспертного знания, однако может служить в качестве отправной точки для последующего анализа.

**Ключевые слова:** семантический ореол метра, стих, извлечение ключевых слов, Национальный корпус русского языка



## 1. Introduction

The semantic halo of a meter (*semantičeskij oreol metra* in Russian) is a notion that was introduced by Gasparov (1979) in order to describe semantic invariants of poetic texts composed using the same meter. Most studies addressing this phenomenon have been based on expert knowledge of the text corpus, the most prominent among them being the studies by Gasparov himself (1999). This book contains a collection of essays about different metrical schemes, including trochaic tetrameter with alternating dactylic and masculine endings (X4дмдм = T4dmdm), iambic trimeter (Я3 = I3), dactylic tetrameter with dactylic endings (Д4дддд = D4dddd), trochaic pentameter (X5 = T5), etc. All of these metrical schemes are shown to have some culturally rooted semantic associations, and the evolution of these associations can also be traced historically. For instance, Gasparov establishes five principal motifs present in the Russian trochaic pentameter, namely, Way, Night, Love, Landscape, and Life & Death (Gasparov 1999: 331–370). They are easily discernible in Lermontov’s “Vyxožu odin ja na dorogu” (“Alone I step out onto the road”, 1841), which is probably the best-known and the most influential poem in this meter, and can be encountered in various combinations in other poems employing trochaic pentameter. Yet another collection of analyses made along similar lines was published by Wachtel (1998). However, it is clear that the availability of large amounts of digitized poetical texts calls for an automated approach to this problem (Birnbaum & Thorsen 2015).

## 2. Methodology and data

The main hypothesis underlying the present study is that the semantic halo of a meter can be at least hinted at by the keywords used in the poems written in this meter. For instance, Lermontov’s “Vyxožu odin ja na dorogu” directly contains the words *doroga* ‘way, path’, *noč* ‘night’, *ljubov* ‘love’, *žizn* ‘life’, which correspond to four out of five motifs identified by Gasparov, as well as many words pertaining to Landscape, including *tuman* ‘fog’, *zvezda* ‘star’, *nebesa* ‘heaven’, *zemlja* ‘earth’, and *dub* ‘oak’. If some of these motifs are repeatedly occurring in poems in trochaic pentameter and are more frequent here than in poems using other metrical schemes, we might expect that the corresponding words will occur more frequently in trochaic pentameter—in other words, they will be the keywords of the iambic pentameter corpus as compared to the corpus consisting of poems written in other meters. These keywords can be identified using some keyness measure. Some scholars have expressed a skeptical attitude towards equating the semantic halo of a meter with its keywords (Brogan 2012: 1264), but it must be noted that keywords can be useful as a starting point for a further detailed analysis, which I aim to show.

In order to study semantic invariants of texts with different metrical schemes, I use the Poetic subcorpus of the Russian National Corpus (RNC), which contains a large collection of Russian poetry from the 18<sup>th</sup> century up to the 2<sup>nd</sup> half of the 20<sup>th</sup> century. It includes 78868 metrically annotated poetic texts, their total size amounting to 11 million words. For the purposes of this study, only texts containing no more than 500 words were included in the sample, since longer texts, such as Pushkin’s “Eugene Onegin”

or Nekrasov's "Komu na Rusi žit' xorošo" ("Who is Happy in Russia?"), are hardly ever homogeneous and can barely serve as good representatives of the semantic halo of a meter. Moreover, some recurring words and names from longer texts, e.g., *Tat'jana* 'Tatyana (proper name)' in "Eugene Onegin", make frequency lists skewed and are likely to be identified as keywords typical of the metrical scheme rather than of individual texts. This restriction led to the exclusion of some 3% of the texts, leaving a corpus of 76,559 poems, containing 8 million tokens (including punctuation marks which are treated as separate tokens). These texts were lemmatized and automatically disambiguated with *Mystem* (*pymystem3* library for Python). Clearly, this leads to some inaccuracies, but these are unavoidable and cannot be corrected manually because of the large corpus size.

There are many possible ways of identifying keywords in corpus linguistics (see, e.g., Scott 1997; Bondi & Scott (eds.) 2010; Culperer & Demmen 2015). For the purpose of this study, I use the Add- $N$  keyness measure proposed by Kilgarrieff (2009). Once the frequencies in the focus corpus ( $f_c$ ) and the frequencies in the reference corpus ( $r_c$ ) are expressed in ipm, the keyness measure  $K$  for a word  $w$  is calculated as follows based on the frequencies ( $f$ ) of this word in the two corpora:

$$K = (f_{f_c}(w) + N) / (f_{r_c}(w) + N).$$

Depending on the choice of  $N$ , this measure favors different types of keywords. If  $N$  is low (e.g.,  $N = 1$ ), it selects low-frequency words, their keyness being large in terms of effect size but not necessarily statistically significant; if  $N$  is high (e.g.,  $N = 1,000$ ), it selects high-frequency words, their keyness being not so large in terms of effect size but much more statistically significant. The value of  $N$  best suited for the purpose of a study has to be identified by the researcher. In the present case,  $N = 3,000$  was chosen in order to extract only high-frequent keywords.

### 3. Results and discussion

There are 20 metrical schemes represented by more than 50,000 tokens. For these schemes, keywords with  $K > 1.1$  were selected. The results are listed in the Appendix.

It is obvious that there are some metrical schemes that do not possess an identifiable semantic halo. One of them is (1) Iambic tetrameter with alternating feminine and masculine endings (I4fm = Я4ЖМ), which also happens to be the most popular metrical scheme overall. It is characterized by the iambic words *kogdá* 'when', *dušá* 'soul', and *mečtá* 'dream', whose high frequency can be explained by meter: a subordinating conjunction like *kogdá* 'when' is expected to appear at the beginning of a verse since syntactic boundaries are likely to coincide with verse boundaries (enjambements do of course occur, but they are rare), and iambic conjunctions are favored here because they fit the meter. The same is true of *dušá* 'soul' and *mečtá* 'dream', which are favored not only at the beginning of a verse, but also verse-internally after possessive pronouns denoting a singular possessor (*mojá* / *tvojá* / *egó* / *ejó dušá* 'my / your / his / her soul'). A similar preference for words for which prosodic template is identical to the verse foot is visible in the lists for all metrical schemes. For example, the lists for trochaic schemes in (2) and (9) contain trochaic words *mílyj* 'dear', *sérdce* 'heart', *pésnja* 'song', *étoť* 'this', *ésli* 'if', *tól'ko* 'only', etc. close to the top. The amphibrachic

scheme (10) favors the words *Rossija* ‘Russia’, *xolódnyj* ‘cold’, and *dalěkij* ‘faraway’, whereas the list for anapaest in (11) includes anapaestic words *molodój* ‘young’, *golubój* ‘light blue’, and *zlotój* ‘golden’.

Metrical scheme (2) Trochaic tetrameter with feminine and masculine endings (T4fm = X4жм) is linked to Love and Song, which is exemplified by the keywords *milyj* ‘dear’, *serdce* ‘heart’, and *pesnja* ‘song’. It also features emotional language and dialogs, which is reflected in the high frequency of quotation marks («) and exclamation marks (!).

Metrical scheme (7) Iambic hexameter with alternating feminine and masculine endings (I6fm = Ябжм) has obvious romantic associations. Its keywords include such nouns as *duša* ‘soul’, *ljubov* ‘love’, *luč* ‘ray’, *krasota* ‘beauty’, *nadežda* ‘hope’. A similar conclusion holds true for (3) Iambic tetrameter with various patterns of feminine and masculine endings (I4f,m = Я4ж,м), which is further linked to the topic of creativity through *pevec* ‘singer’, *muza* ‘Muse’, and *vдохновение* ‘inspiration’.

Metrical scheme (9) Trochaic pentameter with alternating feminine and masculine endings (T5fm = X5жм) is an interesting test case for the method presented here, since it is was thoroughly studied by Gasparov (see Section 1 above). Four out of five motifs listed by Gasparov are also observable in the keyword list, namely Way (*uxodit* ‘go away’, *put* ‘way’), Life & Death (*žit* ‘to live’), Night (*noč* ‘night’, *zvezda* ‘star’), and Landscape (*zemlja* ‘earth’, *zvezda* ‘star’). This shows that the keyword approach yields results similar to an in-depth analysis by a literary scholar.

Another case where keywords conform to Gasparov’s analysis is in metrical scheme (10) Amphibrachic trimeter with alternating feminine and masculine endings (Ab3fm = Aф3жм). One of its possible semantic associations is “elevated intonation” (Gasparov 1999: 194–202), which is closely linked to patriotism. In the keyword list, this is expressed by the high frequency of the words *Rossija* ‘Russia’ and *russkij* ‘Russian’. A closer look at the texts shows that these words are especially characteristic of the poems in this meter composed by Andrei Bely, Georgy Ivanov, and Yaroslav Smelyakov.

Metrical scheme (11) Anapaestic trimeter with alternating feminine and masculine endings (An3fm = АН3жм) seems to be closely connected to the motif of Landscape (*zemlja* ‘ground, earth’, *sneg* ‘snow’, *prostor* ‘open space’, *vpered* ‘in front’).

Metrical scheme (12) *Dolnik* with alternating feminine and masculine endings (Dk3fm = Дк3жм) is a form that has entered Russian poetry relatively recently, and most of the poems that represent it come from the 20<sup>th</sup> century. It shows a turn to more personal topics, which is exemplified by the prevalence of body part names (*glaz* ‘eye’, *ruka* ‘hand / arm’). It also

Metrical scheme (15) Iambic trimeter with alternating feminine and masculine endings (I3fm = Я3жм) is a metrical scheme associated with the motif of Love, which is confirmed by the fact that its keywords include words like *milyj* ‘dear’, *ljubit* ‘to love’, *ljubov* ‘love’, *ax* ‘o!’, *vesna* ‘spring’.

Metrical scheme (18) Trochaic tetrameter with feminine endings (T4f = X4ж) was quite popular in the middle of the 19<sup>th</sup> century and is often used for describing nature: *solnce* ‘sun’, *more* ‘sea’, *tuča* ‘cloud’, *nebo* ‘sky’, *les* ‘forest’.

These illustrative analyses of the keyword lists show clearly that keywords can be linked to the notion of the semantic halo of a meter. Of course, one must proceed with caution, since general keyword lists for a metrical scheme do not take into account

any historical development and may include keywords that were relevant in different time periods. Some of the items may be invited by the meter formally (cf. discussion of *kogda*) rather by a cultural semantic association, which also must be differentiated from the semantic halo.<sup>1</sup> One should also propose an operationalizable procedure for identifying semantic fields based on keywords, since the interpretations given above are often fairly impressionistic. However, keyword lists can definitely serve as a starting point for more detailed investigations.

## 4. Conclusion

The automated extraction of keywords from large poetic corpora has been shown to be a useful way for identifying the semantic halo of various meters. It cannot replace a detailed analysis of texts performed by a literary scholar, but this approach can facilitate the primary identification of relevant motifs and serve as a starting point for further analysis.

## References

1. *Birnbaum D., Thorsen E.* (2015), Enabling the automated identification and analysis of meter and rhyme in Russian verse, Paper presented at Computational approaches to poetry: a day of workshops, University of Helsinki, Finland, June 2015.
2. *Bondi M., Scott M.* (eds., 2010), *Keyness in texts*, John Benjamins, Amsterdam; Philadelphia.
3. *Brogan T. V. F.* (2012), Trochaic, In *Greene R.* (ed.), *The Princeton encyclopedia of poetry and poetics*, 4<sup>th</sup> ed., Princeton University Press, Princeton, pp. 1462–1463.
4. *Culperer J., Demmen J.* (2015), Keywords, In *Biber D., Reppen R.* (eds.), *The Cambridge handbook of English corpus linguistics*, Cambridge University Press, Cambridge.
5. *Gasparov M. L.* (1979), Semantic halo of a meter: Towards the semantics of the Russian iambic trimeter [Semantičeskij oreol metra: k semantike ruskogo trëhstopnogo jamba], In *Linguistics and poetics [Lingvistika i poëtika]*, Nauka, Moscow, pp. 282–308.
6. *Gasparov M. L.* (1999), Meter and semantics: On one of the mechanisms of cultural memory [Metr i smysl: ob odnom iz mehanizmov kul'turnoj pamjati], RGGU, Moscow.
7. *Kilgarriff A.* (2009), Simple maths for keywords, In *Proceedings of Corpus Linguistics Conference CL2009*, University of Liverpool, UK, July 2009.
8. RNC. Russian National Corpus. [www.ruscorpora.ru](http://www.ruscorpora.ru)
9. *Scott M.* (1997), PC analysis of key words — and key key words, *System* 25(2), pp. 233–245.
10. *Wachtel M.* (1998), *The development of Russian verse: Meter and its meanings*, Cambridge University Press, Cambridge.

---

<sup>1</sup> This was pointed out by an anonymous reviewer to whom I am very grateful.

## Appendix

Word	Focus corpus freq.	Reference corpus freq.	K
<b>1. Iambic tetrameter with alternating feminine and masculine endings (I4fm = Я4жм; 695,663 tokens)</b>			
<i>твой</i>	4,914.7	3,433.7	1.230
<i>когда</i>	3,017.3	1,975.3	1.209
<i>но</i>	6,858.2	5,507.9	1.159
<i>свой</i>	4,216.1	3,241.2	1.156
<i>мир</i>	2,396.3	1,751.8	1.136
<i>мечта</i>	1,085.3	671.3	1.113
<i>душа</i>	3,018.7	2,427.9	1.109
<b>2. Trochaic tetrameter with feminine and masculine endings (T4fm = X4жм; 331,449 tokens)</b>			
<i>милый</i>	1,777.0	940.6	1.212
<i>!</i>	4,634.2	3,312.7	1.209
<i>сердце</i>	3,330.8	2,446.0	1.162
<i>«</i>	1,834.4	1,196.8	1.152
<i>песня</i>	1,653.3	1,071.5	1.143
<i>ж</i>	1,412.0	873.1	1.139
<i>наш</i>	2,380.5	1,726.5	1.138
<i>ходить</i>	775.4	358.7	1.124
<i>ах</i>	956.4	522.7	1.123
<i>уж</i>	1,390.9	942.2	1.114
<i>свет</i>	2,139.1	1,616.2	1.113
<i>словно</i>	902.1	516.0	1.110
<i>ль</i>	1,173.6	772.8	1.106
<i>око</i>	992.6	612.1	1.105
<i>волна</i>	1,249.1	857.1	1.102
<i>тихо</i>	859.9	507.5	1.100
<b>3. Iambic tetrameter with various patterns of feminine and masculine endings (I4f,m = Я4ж,м; 292,724 tokens)</b>			
<i>когда</i>	3,279.5	2,019.7	1.251
<i>вы</i>	4,195.1	2,860.6	1.228
<i>мой</i>	8,007.5	5,962.2	1.228
<i>друг</i>	2,551.9	1,540.9	1.223
<i>поэт</i>	1,298.2	610.7	1.190
<i>мечта</i>	1,369.9	682.1	1.187
<i>душа</i>	3,388.9	2,444.7	1.173
<i>она</i>	5,298.5	4,072.3	1.173

Word	Focus corpus freq.	Reference corpus freq.	K
<i>ваш</i>	1,164.9	590.6	1.160
<i>любовь</i>	2,473.3	1,727.4	1.158
<i>стих</i>	1,311.8	733.5	1.155
<i>твой</i>	4,516.2	3,526.1	1.152
<i>свой</i>	4,249.7	3,290.8	1.152
<i>сей</i>	1,028.3	514.3	1.146
<i>муза</i>	683.2	244.3	1.135
<i>взор</i>	1,246.9	743.6	1.134
<i>милый</i>	1,472.4	956.3	1.130
<i>певец</i>	591.0	198.5	1.123
<i>но</i>	6,603.5	5,588.0	1.118
<i>жизнь</i>	2,695.4	2,108.0	1.115
<i>око</i>	1,021.4	612.9	1.113
<i>надежда</i>	744.7	371.3	1.111
<i>сердце</i>	3,064.3	2,460.5	1.111
<i>природа</i>	761.8	391.3	1.109
<i>слава</i>	949.7	563.2	1.108
<i>вдохновение</i>	440.7	108.7	1.107
<i>пора</i>	1,195.7	789.5	1.107
<i>живой</i>	1,168.3	764.3	1.107
<i>ж</i>	1,287.9	880.5	1.105
<i>ее</i>	1,117.1	729.1	1.104
<i>ум</i>	806.2	448.0	1.104
<i>его</i>	1,612.4	1,186.7	1.102
<b>4. Iambic pentameter with alternating feminine and masculine endings (I5fm = Я5жм; 291,540 tokens)</b>			
<i>но</i>	7,011.0	5,572.8	1.168
<i>сонет</i>	548.8	64.2	1.158
<i>мечта</i>	1,238.3	687.2	1.149
<i>мир</i>	2,483.4	1,782.2	1.147
<i>сон</i>	2,112.9	1,494.2	1.138
<i>когда</i>	2,733.8	2,040.5	1.138
<i>душа</i>	3,203.7	2,451.8	1.138
<i>день</i>	3,430.1	2,672.8	1.133
<i>и</i>	50,970.7	44,841.9	1.128
<i>как</i>	11,970.9	10,410.7	1.116
<i>о</i>	5,049.0	4,218.5	1.115
<i>любовь</i>	2,233.0	1,736.6	1.105

Word	Focus corpus freq.	Reference corpus freq.	K
<b>5. Iambic tetrameter with alternating masculine and feminine endings (I4mf = Я4мж; 219,314 tokens)</b>			
<i>твой</i>	4,582.5	3,533.5	1.161
<i>но</i>	6,917.0	5,588.7	1.155
<i>когда</i>	2,799.6	2,045.1	1.150
<i>.</i>	7,477.9	6,142.5	1.146
<i>день</i>	3,469.9	2,678.7	1.139
<i>душа</i>	3,082.3	2,462.2	1.114
<i>тоска</i>	1,108.0	707.1	1.108
<i>иной</i>	797.9	440.6	1.104
<b>6. Iambic pentameter with alternating masculine and feminine endings (I5mf = Я5мж; 173,737 tokens)</b>			
<i>но</i>	7,897.0	5,574.7	1.271
<i>сонет</i>	627.4	69.7	1.182
<i>как</i>	12,329.0	10,426.2	1.142
<i>с</i>	460.5	38.6	1.139
<i>.</i>	7,315.7	6,153.9	1.127
<i>чей</i>	719.5	305.0	1.125
<i>вновь</i>	949.7	548.0	1.113
<i>мир</i>	2,336.9	1,796.0	1.113
<i>когда</i>	2,618.9	2,053.5	1.112
<i>тень</i>	1,421.7	988.0	1.109
<i>сон</i>	1,991.5	1,506.2	1.108
<i>миг</i>	794.3	441.2	1.103
<b>7. Iambic hexameter with alternating feminine and masculine endings (I6fm = Я6жм; 144,418 tokens)</b>			
<i>душа</i>	3,967.6	2,451.9	1.278
<i>когда</i>	3,365.2	2,041.9	1.262
<i>я</i>	26,679.5	20,894.4	1.242
<i>твой</i>	5,061.7	3,534.7	1.234
<i>сонет</i>	782.5	68.9	1.232
<i>любовь</i>	2,783.6	1,735.7	1.221
<i>ты</i>	13,114.7	10,510.3	1.193
<i>себя</i>	2,264.3	1,461.1	1.180
<i>мечта</i>	1,357.2	695.3	1.179
<i>свой</i>	4,417.7	3,305.8	1.176
<i>толпа</i>	948.6	403.1	1.160
<i>их</i>	1,620.3	995.4	1.156

Word	Focus corpus freq.	Reference corpus freq.	K
<i>о</i>	5,352.5	4,228.5	1.155
<i>луч</i>	1,225.6	657.5	1.155
<i>вновь</i>	1,073.3	547.3	1.148
<i>но</i>	6,827.4	5,603.0	1.142
<i>любить</i>	1,973.4	1,362.3	1.140
<i>мой</i>	7,277.5	6,014.1	1.140
<i>нет</i>	2,797.4	2,110.2	1.134
<i>красота</i>	865.5	410.1	1.134
<i>то</i>	4,016.1	3,189.8	1.133
<i>миг</i>	893.2	440.7	1.132
<i>сомнение</i>	512.4	125.7	1.124
<i>и</i>	50,845.5	44,958.7	1.123
<i>лишь</i>	1,744.9	1,224.4	1.123
<i>сон</i>	2,049.6	1,506.9	1.120
<i>надежда</i>	768.6	377.9	1.116
<i>мир</i>	2,354.3	1,797.7	1.116
<i>ни</i>	2,735.1	2,156.7	1.112
<i>тишина</i>	920.9	526.2	1.112
<i>жизнь</i>	2,686.6	2,119.2	1.111
<i>здесь</i>	1,620.3	1,166.4	1.109
<i>перед</i>	983.3	591.7	1.109
<i>который</i>	1,211.8	802.6	1.108
<i>как</i>	11,861.4	10,441.9	1.106
<i>тьма</i>	865.5	498.4	1.105
<i>природа</i>	754.8	398.4	1.105
<i>путь</i>	1,405.6	998.3	1.102
<b>8. Iambic pentameter with various patterns of feminine and masculine endings (I4f,m = Я5ж,м; 134,355 tokens)</b>			
<i>и</i>	55,070.5	44,894.2	1.212
<i>но</i>	7,256.9	5,597.2	1.193
<i>я</i>	24,561.8	20,937.8	1.151
<i>он</i>	9,214.4	7,925.5	1.118
<i>его</i>	1,667.2	1,194.3	1.113
<i>ни</i>	2,701.8	2,158.0	1.105
<i>когда</i>	2,575.3	2,057.0	1.102
<b>9. Trochaic pentameter with alternating feminine and masculine endings (T5fm = X5жм; 125,119 tokens)</b>			
<i>этот</i>	4,459.8	2,694.8	1.310
<i>если</i>	1,726.4	1,017.1	1.177

Word	Focus corpus freq.	Reference corpus freq.	K
только	2,509.6	1,679.6	1.177
земля	2,693.4	1,872.6	1.168
словно	1,095.0	523.0	1.162
год	2,134.0	1,447.9	1.154
ночь	2,837.3	2,058.9	1.154
мочь	2,349.8	1,650.9	1.150
такой	1,598.5	1,009.1	1.147
чтобы	959.1	472.6	1.140
что	9,295.2	7,807.7	1.138
сколько	743.3	288.5	1.138
над	3,316.8	2,564.3	1.135
звезда	1,582.5	1,049.5	1.132
город	1,095.0	623.9	1.130
уходить	1,158.9	682.5	1.129
синий	967.1	513.1	1.129
ветер	1,726.4	1,195.7	1.126
каждый	1,422.6	926.7	1.126
путь	1,486.6	998.0	1.122
золотой	1,031.0	608.4	1.117
голубой	839.2	450.5	1.113
по	4,419.8	3,679.1	1.111
темный	1,134.9	731.1	1.108
жить	1,830.3	1,378.0	1.103
разве	511.5	191.6	1.100
<b>10. Amphibrachic trimeter with alternating feminine and masculine endings (Ab3fm = Аф3жм; 100,252 tokens)</b>			
туман	1,216.9	522.0	1.197
и	53,594.9	44,956.7	1.180
над	3,551.1	2,563.7	1.177
идти	2,194.5	1,541.7	1.144
Россия	728.2	260.4	1.143
холодный	887.8	423.1	1.136
далекий	857.8	396.0	1.136
под	3,032.4	2,319.9	1.134
за	5,157.0	4,208.6	1.132
она	4,897.7	4,107.2	1.111
на	15,371.3	13,563.8	1.109
а	6,214.3	5,350.2	1.103
шуметь	638.4	300.1	1.102
русский	698.2	360.5	1.101

Word	Focus corpus freq.	Reference corpus freq.	K
<b>11. Anapaestic trimeter with alternating feminine and masculine endings (An3fm = Ан3жм; 90,658 tokens)</b>			
этот	5,085.0	2,695.3	1.420
только	3,011.3	1,677.4	1.285
словно	1,456.0	521.4	1.265
над	3,915.8	2,560.7	1.244
это	2,603.2	1,689.0	1.195
молодой	1,070.0	492.0	1.166
голубой	1,014.8	450.2	1.164
земля	2,669.4	1,876.4	1.163
ты	12,729.2	10,532.4	1.162
ни	2,945.1	2,158.2	1.153
золотой	1,158.2	608.7	1.152
чтобы	1,003.8	474.2	1.152
по	4,665.9	3,679.5	1.148
дорога	1,290.6	748.4	1.145
за	5,140.2	4,209.9	1.129
я	23,991.3	20,964.3	1.126
сколько	705.9	290.9	1.126
.	7,291.1	6,166.3	1.123
снег	1,125.1	725.8	1.107
простор	606.7	259.7	1.106
до	2,117.8	1,628.0	1.106
на	15,310.3	13,566.7	1.105
не	19,468.8	17,330.0	1.105
от	4,081.3	3,407.3	1.105
впереди	430.2	105.9	1.104
кто-то	606.7	272.9	1.102
<b>12. Dolnik with alternating feminine and masculine endings (Dk3fm = Дк3жм; 86,034 tokens)</b>			
это	2,812.8	1,687.3	1.240
только	2,615.2	1,682.5	1.199
.	7,787.6	6,161.6	1.177
ветер	1,917.8	1,196.2	1.172
до	2,417.6	1,625.0	1.171
этот	3,661.3	2,712.1	1.166
словно	1,034.5	526.5	1.144
глаз	2,068.9	1,433.2	1.143
наш	2,406.0	1,746.4	1.139

Word	Focus corpus freq.	Reference corpus freq.	K
на	15,831.0	13,562.0	1.137
надо	1,022.9	602.6	1.117
дверь	790.4	410.0	1.112
последний	1,115.8	700.7	1.112
вечер	941.5	552.9	1.109
рука	2,510.6	1,975.2	1.108
по	4,382.0	3,683.1	1.105
море	1,383.2	971.9	1.104
<b>13. Iambic hexameter with paired feminine and masculine endings (16ffmm = ЯБЖЖММ; 77,957 tokens)</b>			
сей	3,412.1	504.8	1.830
себя	3,194.1	1,458.7	1.389
то	5,541.5	3,181.8	1.382
ты	15,329.0	10,510.4	1.357
который	2,129.4	797.0	1.351
свой	5,413.2	3,305.3	1.334
коль	1,090.3	164.9	1.292
зреть	1,103.2	178.4	1.291
когда	3,412.1	2,052.5	1.269
иль	1,770.2	783.3	1.261
твой	5,079.7	3,547.3	1.234
стих	1,603.4	746.3	1.229
чтоб	2,321.8	1,330.6	1.229
сие	795.3	87.9	1.229
иметь	897.9	206.2	1.216
о	5,772.4	4,233.8	1.213
тот	3,976.6	2,766.2	1.210
он	10,159.4	7,925.4	1.204
во	2,424.4	1,508.8	1.203
дух	1,423.9	680.4	1.202
мой	7,824.8	6,019.3	1.200
сам	1,860.0	1,061.3	1.197
!	4,566.6	3,355.5	1.191
уж	1,693.2	953.5	1.187
всякий	872.3	268.3	1.185
ль	1,475.2	782.7	1.183
ныне	756.8	178.4	1.182
любезный	679.9	116.8	1.181
кто	3,117.1	2,203.6	1.176
мочь	2,462.9	1,654.0	1.174

Word	Focus corpus freq.	Reference corpus freq.	K
кой	615.7	85.9	1.172
она	5,310.6	4,105.4	1.170
муза	795.3	255.1	1.166
уже	1,359.7	748.8	1.163
являть	564.4	67.1	1.162
я	24,821.4	20,961.0	1.161
великий	949.2	411.6	1.158
ум	1,000.6	455.8	1.158
наука	590.1	118.1	1.151
слава	1,103.2	572.1	1.149
к	5,644.1	4,557.9	1.144
труд	936.4	449.4	1.141
око	1,128.8	622.9	1.140
взор	1,257.1	757.1	1.133
желать	590.1	174.0	1.131
они	4,605.1	3,731.8	1.130
утеха	461.8	72.3	1.127
другой	1,372.6	887.8	1.125
хотя	602.9	206.0	1.124
всегда	885.1	457.3	1.124
богиня	487.4	102.5	1.124
тогда	962.1	527.7	1.123
от	4,194.6	3,407.3	1.123
мысль	949.2	521.7	1.121
муж	564.4	180.7	1.121
дар	628.6	242.3	1.119
писать	679.9	288.0	1.119
быть	8,722.8	7,493.8	1.117
днесь	397.7	41.6	1.117
сатира	384.8	30.3	1.117
пред	1,013.4	594.7	1.116
б	1,180.1	745.3	1.116
однако	423.3	71.5	1.115
знать	2,309.0	1,766.5	1.114
приятный	436.1	84.9	1.114
глас	487.4	138.9	1.111
достойный	423.3	85.0	1.110
вид	602.9	246.9	1.110
красота	782.5	414.8	1.108
котор	346.3	22.7	1.107
оставлять	667.0	313.0	1.107



Word	Focus corpus freq.	Reference corpus freq.	K
славный	474.6	137.5	1.107
какой	1,257.1	850.2	1.106
день	3,296.7	2,694.5	1.106
прекрасный	821.0	455.4	1.106
российский	359.2	47.6	1.102
царь	744.0	396.5	1.102
честь	551.6	226.0	1.101
ко	756.8	413.7	1.100
<b>14. Trochaic tetrameter with various patterns of feminine and masculine endings (T4f,m = X4ж,м; 72,217 tokens)</b>			
дева	1,135.5	294.9	1.255
конь	1,218.5	470.9	1.215
милый	1,758.6	968.1	1.199
это	2,534.0	1,691.8	1.180
если	1,661.7	1,022.4	1.159
или	1,343.2	762.6	1.154
очень	706.2	216.2	1.152
марья	429.3	9.6	1.139
вода	1,620.1	1,057.4	1.139
тихо	969.3	518.0	1.128
ну	803.1	387.4	1.123
из	3,988.0	3,241.4	1.120
ходить	775.4	372.3	1.120
нет	2,714.0	2,117.2	1.117
волна	1,315.5	869.3	1.115
бить	706.2	324.7	1.115
река	1,066.2	661.9	1.110
дед	443.1	106.8	1.108
пить	775.4	411.1	1.107
речка	415.4	85.8	1.107
грудь	1,273.9	865.1	1.106
у	3,586.4	2,963.7	1.104
мчатся	498.5	168.7	1.104
месяц	637.0	298.3	1.103
плих	304.6	0	1.102
чаша	512.3	187.6	1.102
муха	387.7	76.5	1.101

Word	Focus corpus freq.	Reference corpus freq.	K
<b>15. Iambic trimeter with alternating feminine and masculine endings (I3fm = ЯЗжм; 69,813 tokens)</b>			
милый	1,776.2	968.2	1.204
она	5,471.8	4,105.2	1.192
.	7,878.2	6,164.1	1.187
жить	2,177.2	1,378.1	1.183
лишь	1,962.4	1,227.4	1.174
я	25,081.3	20,962.7	1.172
любить	2,105.6	1,366.9	1.169
любовь	2,506.7	1,748.0	1.160
хоть	1,188.9	638.2	1.151
ах	1,045.7	536.2	1.144
петь	1,733.2	1,139.3	1.143
сам	1,618.6	1,064.2	1.136
пять	558.6	138.6	1.134
а	6,460.1	5,351.3	1.133
о	5,171.0	4,240.6	1.128
мой	7,162.0	6,027.0	1.126
нет	2,735.9	2,117.2	1.121
весна	1,203.2	761.0	1.118
окно	1,303.5	855.3	1.116
но	6,617.7	5,616.3	1.116
ода	415.4	68.7	1.113
хотеть	1,317.8	887.8	1.111
семьдесят	343.8	11.8	1.110
друг	2,034.0	1,573.8	1.101
<b>16. Iambic hexameter with alternating masculine and feminine endings (I6mf = Ябмж; 65,649 tokens)</b>			
твой	4,737.3	3,552.5	1.181
давно	1,127.2	502.6	1.178
...	2,223.9	1,437.7	1.177
взор	1,416.6	756.6	1.176
я	25,042.3	20,965.2	1.170
их	1,675.6	1,001.1	1.169
и	52,795.9	45,000.9	1.162
мой	7,433.5	6,025.4	1.156
день	3,564.4	2,693.2	1.153
ты	12,445.0	10,541.6	1.141
средь	868.3	399.0	1.138

Word	Focus corpus freq.	Reference corpus freq.	K
душа	3,214.1	2,473.1	1.135
где	3,579.6	2,800.3	1.134
сонет	487.4	78.4	1.133
весь	5,224.8	4,269.0	1.131
пора	1,279.5	800.4	1.126
но	6,687.1	5,616.3	1.124
толпа	822.6	409.6	1.121
свой	4,051.9	3,319.8	1.116
любить	1,873.6	1,369.2	1.115
таинствен- ный	517.9	159.5	1.113
волна	1,294.8	869.8	1.110
прекрасный	822.6	456.0	1.106
себя	1,934.5	1,471.8	1.103
здесь	1,599.4	1,171.1	1.103
мечта	1,081.5	704.1	1.102
лишь	1,660.3	1,230.3	1.102
<b>17. Iambic hexameter with various patterns of feminine and masculine endings (16f,m = Ябж,м; 64,109 tokens)</b>			
свой	4,944.7	3,312.8	1.259
их	1,996.6	998.6	1.250
они	5,178.7	3,728.7	1.215
который	1,606.6	803.6	1.211
...	2,339.8	1,436.9	1.203
ты	13,196.3	10,535.9	1.197
твой	4,835.5	3,551.9	1.196
сей	1,216.7	527.5	1.195
себя	2,324.2	1,468.7	1.191
и	53,814.6	44,994.2	1.184
когда	2,979.3	2,058.3	1.182
бог	1,887.4	1,152.2	1.177
душа	3,400.5	2,471.7	1.170
толпа	982.7	408.4	1.169
!	4,398.8	3,359.0	1.164
где	3,712.4	2,799.4	1.157
взор	1,325.9	757.4	1.151
вдруг	1,341.5	779.4	1.149
перед	1,123.1	594.5	1.147
он	9,515.0	7,934.5	1.145
иль	1,310.3	788.7	1.138

Word	Focus corpus freq.	Reference corpus freq.	K
природа	857.9	401.2	1.134
слеза	1,263.5	773.1	1.130
певец	608.3	209.7	1.124
среди	811.1	399.6	1.121
уж	1,435.1	956.9	1.121
его	1,700.2	1,198.2	1.120
раб	577.1	206.0	1.116
здесь	1,653.4	1,170.7	1.116
внимать	670.7	289.2	1.116
красота	811.1	415.2	1.116
ум	857.9	457.9	1.116
улыбка	748.7	375.6	1.111
с	11,121.7	9,716.8	1.110
тогда	889.1	529.0	1.102
<b>18. Trochaic tetrameter with feminine endings (T4f = X4ж; 60,204 tokens)</b>			
-	5,298.7	3,583.7	1.260
у	4,285.4	2,959.4	1.223
этот	3,969.8	2,712.9	1.220
«	2,126.1	1,216.4	1.216
солнце	2,043.1	1,160.2	1.212
море	1,710.8	970.8	1.186
туча	1,046.4	444.7	1.175
ти	531.5	10.4	1.173
ка	631.2	107.2	1.169
песня	1,777.3	1,090.4	1.168
белый	1,760.7	1,090.9	1.164
...	2,126.1	1,439.0	1.155
»	1,694.2	1,117.5	1.140
ах	1,029.8	536.9	1.139
роза	1,029.8	544.2	1.137
к	5,547.8	4,561.0	1.131
небо	2,690.9	2,041.3	1.129
словно	980.0	528.5	1.128
слышать	1,029.8	579.5	1.126
птица	946.8	505.8	1.126
рыцарь	448.5	79.0	1.120
полный	1,013.2	585.4	1.119
девица	465.1	99.5	1.118
это	2,242.4	1,695.3	1.117
лес	1,295.6	859.3	1.113

Word	Focus corpus freq.	Reference corpus freq.	K
если	1,478.3	1,024.8	1.113
с	11,128.8	9,717.4	1.111
народ	897.0	507.7	1.111
сердце	3,072.9	2,478.1	1.109
посмотреть	465.1	128.0	1.108
тройка	365.4	36.6	1.108
тихо	897.0	519.2	1.107
только	2,192.5	1,688.7	1.107
птичка	415.3	88.3	1.106
отвечать	614.6	270.0	1.105
выходить	847.1	483.9	1.104
дама	431.9	116.0	1.101
<b>19. Iambic tetrameter with masculine endings (14m = Я4м; 55,349 tokens)</b>			
мой	8,726.4	6,018.2	1.300
он	11,093.2	7,925.2	1.290
когда	3,450.8	2,056.1	1.276
один	3,432.8	2,166.4	1.245
день	3,866.4	2,692.3	1.206
я	25,872.2	20,964.7	1.205
звезда	1,879.0	1,052.1	1.204
дом	1,644.1	884.4	1.196
тень	1,770.6	992.0	1.195
кто	3,179.8	2,205.7	1.187
тот	3,794.1	2,770.9	1.177
сон	2,294.5	1,511.3	1.174
свет	2,439.1	1,632.2	1.174
жизнь	3,017.2	2,123.2	1.174
свой	4,372.3	3,318.5	1.167
огонь	1,806.7	1,117.9	1.167
дно	831.1	291.0	1.164
жить	2,095.8	1,380.1	1.163
луна	1,156.3	578.4	1.161
путь	1,626.0	1,001.3	1.156
как	12,484.4	10,453.4	1.151
снег	1,282.8	726.5	1.149
идти	2,222.3	1,545.2	1.149
там	2,493.3	1,779.0	1.149
петь	1,752.5	1,140.3	1.148
тишина	1,047.9	529.7	1.147
твой	4,498.7	3,555.7	1.144

Word	Focus corpus freq.	Reference corpus freq.	K
пора	1,337.0	800.7	1.141
земной	1,029.8	563.2	1.131
так	3,830.2	3,039.0	1.131
чтоб	1,897.1	1,336.4	1.129
лес	1,355.0	859.2	1.128
звон	704.6	287.0	1.127
ночь	2,692.0	2,066.8	1.123
стих	1,210.5	751.4	1.122
час	1,463.4	980.3	1.121
Ру	361.3	2.8	1.119
ночной	867.2	457.6	1.118
пустой	795.0	397.6	1.117
грудь	1,318.9	865.7	1.117
конь	867.2	475.0	1.113
последний	1,120.2	702.2	1.113
труд	831.1	451.5	1.110
ты	12,014.7	10,547.1	1.108
душа	3,053.4	2,475.2	1.106
процать	722.7	373.8	1.103
пока	668.5	329.2	1.102
и	49,901.5	45,031.1	1.101
нет	2,637.8	2,119.0	1.101
уж	1,355.0	958.0	1.100
<b>20. Iambic hexameter with paired masculine and feminine endings (16mmff = Ябммжж; 53,793 tokens)</b>			
ты	19,296.2	10,498.2	1.652
то	6,934.0	3,179.5	1.608
сей	2,379.5	520.6	1.528
твой	6,878.2	3,539.8	1.510
себя	3,346.2	1,462.9	1.422
когда	3,848.1	2,053.7	1.355
свой	5,186.5	3,313.2	1.297
во	2,825.6	1,508.9	1.292
который	1,877.6	802.8	1.283
мой	8,402.6	6,020.9	1.264
свет	2,844.2	1,629.6	1.262
тот	4,182.7	2,768.4	1.245
мочь	2,751.3	1,654.5	1.236
коль	910.9	169.0	1.234
зреть	892.3	182.6	1.223

Word	Focus corpus freq.	Reference corpus freq.	K
всякий	985.3	269.4	1.219
чтоб	2,268.0	1,334.0	1.215
быть	9,703.9	7,490.9	1.211
иметь	873.7	208.5	1.207
честь	892.3	224.7	1.207
стих	1,524.4	749.4	1.207
труд	1,134.0	449.5	1.198
сам	1,859.0	1,063.7	1.196
она	5,484.0	4,107.8	1.194
желать	780.8	174.0	1.191
любезный	706.4	118.4	1.189
что	9,852.6	7,817.2	1.188
через	687.8	104.7	1.188
уж	1,691.7	955.8	1.186
к	5,948.7	4,559.1	1.184
слава	1,226.9	572.9	1.183
они	4,926.3	3,732.3	1.177
ум	1,059.6	457.1	1.174
день	3,680.8	2,693.7	1.173
я	25,077.6	20,971.0	1.171
имя	966.7	409.1	1.164
кто	3,048.7	2,206.8	1.162
приятный	576.3	85.0	1.159
сие	576.3	91.5	1.157
тогда	1,078.2	528.2	1.156
отрада	632.1	143.9	1.155
век	1,394.2	813.3	1.152
утеха	539.1	72.9	1.152
лишь	1,859.0	1,229.6	1.149
другой	1,468.6	888.6	1.149
хотя	669.2	206.7	1.144
судья	501.9	61.9	1.144
муж	632.1	181.4	1.142
счастливы	836.5	361.3	1.141
час	1,543.0	979.9	1.141
достойный	520.5	85.4	1.141
велеть	539.1	105.4	1.140
любовь	2,416.7	1,750.1	1.140
его	1,784.6	1,198.3	1.140
весь	5,279.5	4,270.0	1.139
уже	1,264.1	751.3	1.137
ее	1,245.5	739.9	1.135
писать	725.0	288.8	1.133
муза	687.8	257.5	1.132
внимать	725.0	289.3	1.132
всегда	910.9	458.4	1.131

Word	Focus corpus freq.	Reference corpus freq.	K
хоть	1,115.4	639.8	1.131
российский	446.2	48.0	1.131
<	520.5	113.5	1.131
несчастный	520.5	117.2	1.129
монархиня	390.4	4.9	1.128
б	1,226.9	746.3	1.128
ль	1,264.1	786.2	1.126
величество	390.4	17.2	1.124
представ- лять	427.6	55.5	1.122
глас	520.5	139.7	1.121
иль	1,245.5	789.8	1.120
днесь	409.0	42.6	1.120
видеть	1,747.4	1,242.7	1.119
ко	818.0	414.4	1.118
сатира	390.4	31.4	1.118
должный	594.9	215.9	1.118
любить	1,877.6	1,369.9	1.116
надпись	409.0	56.7	1.115
счастье	1,226.9	789.8	1.115
прекрасный	855.1	456.3	1.115
покой	818.0	426.4	1.114
езде	520.5	161.9	1.113
рок	501.9	147.2	1.113
один	2,751.3	2,171.3	1.112
ныне	539.1	181.6	1.112
являть	409.0	69.7	1.111
толь	353.2	19.6	1.110
он	9,146.2	7,939.0	1.110
взор	1,171.2	759.2	1.110
кой	427.6	88.7	1.110
плод	501.9	154.3	1.110
закон	557.7	208.1	1.109
красота	780.8	415.9	1.107
стараться	390.4	66.7	1.106
мучение	409.0	82.2	1.106
взирать	390.4	68.0	1.105
но	6,525.0	5,619.0	1.105
наука	446.2	120.5	1.104
пит	334.6	22.8	1.103
град	446.2	128.5	1.102
забава	409.0	93.4	1.102
краса	539.1	214.8	1.101
польза	353.2	45.9	1.101
можно	539.1	216.3	1.100
догадка	316.0	14.1	1.100

# «Я СКАЖУ ТЕБЕ С ПОСЛЕДНЕЙ ПРЯМОТОЙ»: ПРЯМАЯ И КОСВЕННАЯ РЕЧЬ ПО ДАННЫМ КОРПУСА С ПРОСОДИЧЕСКОЙ РАЗМЕТКОЙ

**Подлесская В. И.** (vi\_podlesskaya@il-rgggu.ru)

Российский государственный гуманитарный университет; Российская академия народного хозяйства и государственной службы; Москва, Россия

**Ключевые слова:** прямая речь, косвенная речь, русский язык, корпус, устная речь, просодия

# “JA SKAZHU TEBE S POSLEDNEJ PRJAMOTOJ”: DIRECT AND INDIRECT SPEECH VIEWED THROUGH THE PRISM OF PROSODICALLY ANNOTATED CORPUS DATA

**Podlesskaya V. I.** (vi\_podlesskaya@il-rgggu.ru)

Russian State University for the Humanities;  
Russian Academy of National Economy and Public Administration;  
Moscow, Russia

The paper focuses on how prosody complements grammar in differentiating between two main strategies of quoting—direct and indirect speech (*Peter said: “I am a linguist”* vs. *Peter said that he was a linguist*). Basing on corpus data from spontaneous spoken Russian, the paper addresses two research questions: (1) how do grammatical, lexical and prosodic parameters correlate in canonical cases of direct and indirect speech in Russian, and (2) what are typical prosodic deviations from canonical patterns. Answering the first question, we indicated prosodic features (pitch movements, pitch reset, localization of phrasal accents, pauses etc.) that help to identify a prosodic break between the reporting frame (*Peter said*) and the following reported utterance in the case of direct deixis (when personal and spatiotemporal indexes are oriented towards the reported situation, cf. *I am a linguist*) and the lack of such a break in the case indirect deixis (when personal and spatiotemporal indexes are shifted towards the reporting situation, cf. *that he was a linguist*). Answering the second question, we indicated particular

prosodic patterns used (i) to integrate otherwise prosodically disintegrated constructions with direct deixis, and (ii) to disintegrate otherwise prosodically integrated constructions with indirect deixis.

**Key words:** direct speech, indirect speech, Russian, corpus, natural discourse, prosody

## 1. Постановка вопроса

Данная работа посвящена феномену цитации в живой речи<sup>1</sup>. На корпусном материале исследуются просодические средства, позволяющие говорящему включить так называемую «чужую» речь в состав текущего повествования. Основная цель — уточнить, как в русском неподготовленном дискурсе просодия вносит вклад в реализацию известного в грамматической традиции противопоставления прямой и косвенной речи.

Я рассматриваю противопоставление «прямая vs. косвенная речь» не как бинарное, считая, что прямая и косвенная цитация представляют собой концы шкалы, промежуточные значения на которой связаны с действием разнородных факторов — в том числе, прагматических, грамматических, просодических (в устной речи), графических (в письменной речи). На русском материале плодотворность такой точки зрения убедительно продемонстрирована в работах Е. В. Падучевой (1996 а,б, *inter alia*), показавшей, что скалярность этого противопоставления связана с разной степенью редукции свойств диалогического режима использования языка: дейктичности, экспрессивности и ориентации на синхронного адресата. В типологической перспективе такая позиция была обоснована в работах А. Ю. Айхенвальд (Aikhenvald 2008, 2011a,b), систематизировавшей возможные параметры внутриязыкового и межъязыкового варьирования цитационных конструкций. Разнообразие средств, грамматикализирующих противопоставление прямой и косвенной речи, и в первую очередь разнообразие форм индексального сдвига (грамматически детерминированных правил выбора дейктических элементов) было исследовано Н. Эвансом (Evanse 2013). Многофакторный анализ цитационных конструкций в устной русской речи на корпусном материале был впервые предпринят в рамках коллективного исследования Литвиненко и др. 2009. Я в значительной степени опираюсь на сформулированные в этой работе подходы.

Я использую материал корпуса «Веселые истории из жизни» электронной коллекции «Рассказы о сновидениях и другие корпуса звучащей речи» (SpokenCorpora 2013). Корпус содержит 40 устных монологов (аудиофайлы с синхронизированными просодически размеченными транскриптами, респонденты от 18 до 60 лет, около 10 000 словоупотреблений). В корпусе было обнаружено 162 эпизода включения чужой речи.

Дальнейшее изложение будет строиться следующим образом. Сначала, по необходимости кратко, будут рассмотрены прагматические параметры

---

<sup>1</sup> Исследование выполнено при поддержке РФФ, грант 17-18-01184

противопоставления прямой и косвенной речи и задокументированные в корпусе прототипические случаи прямой и косвенной речи, в которых прагматические параметры полностью согласуются с лексико-грамматическими и просодическими (раздел 2). Далее я продемонстрирую обнаруженные в устном подкорпусе три основных типа рассогласования прагматических и просодических параметров (раздел 3). В разделе 4 будут приведены количественные данные и подведены итоги.

## 2. Прагматика и просодия конструкций с «чужой речью». Канонические случаи прямой и косвенной речи

Вслед за Литвиненко и др. 2009 я буду использовать термин «цитационная конструкция», считая, что последняя состоит из собственно цитации — отрезка дискурса, передающего чужую речь, и авторской ремарки, центральным элементом которой в прототипическом случае является глагол речи (конструкции с глаголами мысли, чувства и им подобные остаются за рамками данного исследования).

Главное прагматическое свойство прямой речи состоит в том, что цитация формирует самостоятельную иллокуцию. Вследствие этого все дейктические элементы цитации ориентированы на внутреннюю коммуникативную ситуацию, а не на внешнюю ситуацию нарратива, в который интегрирована чужая речь (я буду использовать рабочий термин «прямой» дейксис). Другим важным следствием иллокутивной самостоятельности цитации является отсутствие ограничений на тип иллокуции: в прототипических случаях прямой речи возможны типы иллокуций, отличные от сообщения (вопросы, обращения, восклицания, директивы).

В прототипических случаях прямой речи лексико-грамматические свойства цитационной конструкции согласуются с прагматическими, что приводит, в частности, к тому, что в составе цитации становятся возможными так называемые *main-clause phenomena* «явления главного предложения», или, согласно важному уточнению, сделанному Е. В. Падучевой (1996а:299), «явления предложения, которому соответствует отдельный речевой акт». Это средства выражения субъективной модальности, экспрессивно диалогические элементы, реактивные реплики и проч. Ср. обнаруженные в исследованном корпусе конструкции, неперебиваемые в режим косвенной речи: вопросы, вводимые дискурсивным маркером *a* (*А где здесь сервис-центры?; Ну а вы не можете сказать?*); восклицания, вводимые междометиями (*Ой какие попугайчики!*); вопросы с клитизированной частицей *-то* (*Чего вообще случилось-то?*); звуковые имитации (*Бээ*); интродуктивные формулы (*Так вот: ...*), дискурсивные повторы (*Очень много значит. Очень.*) и др.

В устном дискурсе в прототипических случаях прямой речи с прагматическими свойствами цитационной конструкции согласуются и просодические. Во-первых, для выражения иллокутивных значений в цитации используются те же стандартные коммуникативно-просодические паттерны, что и в независимых иллокуциях в диалогическом режиме: в частности, быстрый подъем

в высокий регистр по типу ИК3 в терминологии интонационных конструкций (Брызгунова 1982) для выражения общего вопроса, подъем с последующим падением по типу ИК5 для восклицания и т.п. Во-вторых, авторская ремарка и цитация реализуются как две отдельных коммуникативно-просодических составляющих, каждая — со своим фразовым акцентом. Препозитивная авторская ремарка вдобавок обычно реализуется с падением тона в главном фразовом акценте, типичным маркером просодической завершенности, что создает уникальную дискурсивную коллизию: просодическая реализация не проецирует продолжения, но, в то же время, наличие предиката речи с незаполненной валентностью, напротив, предполагает дальнейшее развертывание дискурса.

Следующие два примера иллюстрируют согласование всех параметров в прототипическом случае прямой речи. В примере (1) в составе цитации имеется обращение и специальный и общий вопросы, выражающие также удивление. Маркирующая вопрос-удивление частица *же* — типичный пример *main-clause phenomena*, делающего невозможным перевод вопроса в режим косвенного цитирования. Авторская ремарка формирует отдельную коммуникативно-просодическую составляющую и реализуется с падением тона в главном фразовом акценте, локализованном на глаголе речи. В строке 12 представлен типичный для специального вопроса паттерн — подъем на вопросительном слове и падение тона в конце:

(1) FS32-sp<sup>2</sup>

10. — ... она \спросила:
11. «/↑Мама@
12. А /как же ... за́мок может \танцевать?
13. .. У него же ножек /нету?»

Сходным образом, с падением тона на глаголе речи реализована авторская ремарка в примере (2) в строке 54. Кроме того, в цитации в строке 55 использован паттерн, типичный для эмпазы — удлинение и тональное искривление слога — носителя главного фразового акцента (Янко 2008: 83–97), см. слог *ве-е* в составе слова *Медве-едева*, рис. 1. Очевидно, что подобное просодически выраженное значение эмпазы делает невозможным перевод этой иллюстрации в режим косвенного цитирования. Замечу также, что если коннектор *к тому же* допускает перевод в режим косвенной речи (*говорит, что к тому же она землячка...*), то комплекс *и к тому же* — не допускает, т.е. является *main-clause phenomenon* (\*?*говорит, что и к тому же она землячка...*):

<sup>2</sup> Индекс примера содержит ссылку к ярлыку текста в составе корпуса. Номера строк соответствуют нумерации строк в рассказе. Об используемой системе дискурсивной транскрипции см. Кибрик, Подлеская (ред.) 2009, SpokenCorpora 2013. Существенно, что в отличие от письменной речи, где решение об использовании кавычек принимает говорящий (пишущий), при документировании устной речи решение об использовании кавычек принадлежит транскриптеру. Согласно конвенции, принятой для SpokenCorpora 2013, кавычками обрамляются те цитации, в которых использован прямой дейксис и/или имеются *main-clause phenomena*.



(2) FS13-sp

52. ... А /потом .. ну такая /пауза,

53. /тишина,

54. и он \говорит:

55. ... «И /ещё .. к \тому же .. она /землячка /\Медве-едева!»

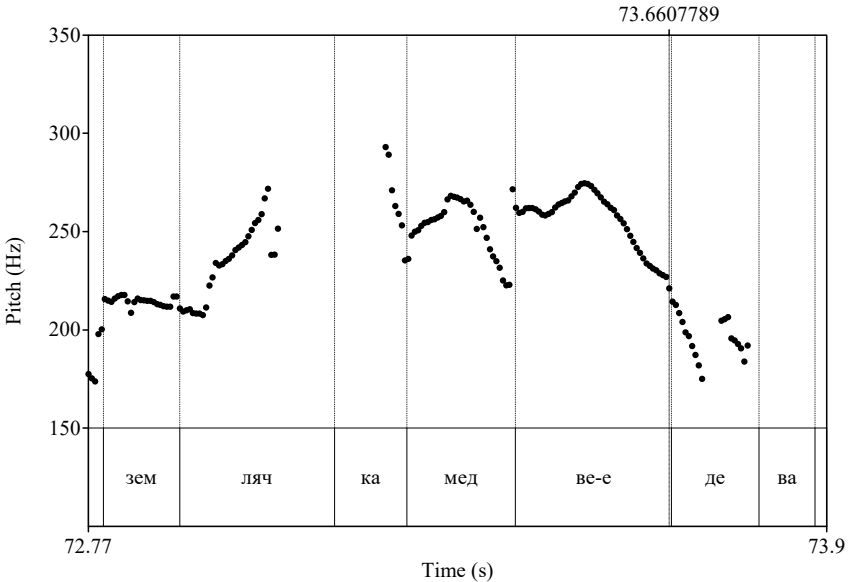


Рис. 1. FS13-sp: «землячка Медведева»

Обратимся теперь к цитационным конструкциям в формате косвенной речи. Их главное прагматическое свойство состоит в том, что авторская ремарка и собственно цитация образуют единую иллокуцию. Вследствие этого все дейктические элементы цитации (личные, временные, пространственные) ориентированы не на внутреннюю коммуникативную ситуацию, а на внешнюю ситуацию нарратива, в который интегрирована чужая речь (для таких случаев я буду использовать рабочий термин «косвенный» дейксис).

Другое важное следствие иллокутивной несамостоятельности цитации состоит в том, что в составе цитации затруднены типы иллокуции, отличные от сообщения: мы имеем дело не с речевыми актами вопроса, директива, обращения и т.д., а с сообщениями о них; такими сообщениями являются в частности, косвенные вопросы и косвенные директивы. В прототипическом случае лексико-грамматические свойства таких конструкций согласуются с прагматическими. Цитация в формате косвенной речи грамматически строится как сентенциальный актанта вводящего предиката. Такой сентенциальный актанта оформляется по тем же моделям, что и другие сентенциальные актанта в данном языке: может вводиться комплементаризатором (в русском — это союзы *что* и *чтобы*), может требовать ограничений на грамматические формы предиката

(например, требовать так называемого согласования времен, как в английском и других европейских языках), может требовать ограничений на порядок слов (например, –заключительной позиции глагола, как в немецком языке) и т.д.

Далее, в косвенной речи в состав цитации могут быть внедрены дискурсивные маркеры, связанные с развертыванием внешней речевой ситуации — показатели речевых затруднений (повествователя, а не «внутреннего» говорящего!), аппроксиматоры (*мол, типа, как бы* и др.), которые нужны, чтобы показать, что чужая речь приводится не *verbatim*: ее форма, а часто и смысл, преподносятся в редакции повествователя. Явления такого рода отражают текущий процесс речепроизводства и не могут интерпретироваться как исходно принадлежавшие цитируемому фрагменту.

В прототипических случаях косвенной речи с прагматическими свойствами цитационной конструкции согласуются и просодические. Авторская ремарка и собственно цитация демонстрируют существенно более высокую степень просодической интеграции, чем в режиме прямого цитирования. Во-первых, препозитивная авторская ремарка реализуется обычно как просодически незавершенная, т.е. проецирующая дальнейшее развертывание дискурса. Просодическая незавершенность авторской ремарки маркируется чаще всего «прототипическим русским подъемом» с падением на заударных, если они есть, т.е. по типу ИК-3 (Янко 2008:31).

Так, в следующем примере хорошо согласованы (а) косвенный дейксис («мы спросили [у него]<sub>i</sub>, [о<sub>n</sub>]<sub>i</sub> ли сажает»); (б) косвенный вопрос; (в) наличие аппроксиматора-репорта *мол* (строка 32), (г) речевое затруднение, которое проявилось в необходимости оборвать слово в строке 32 (*e* = — возможно, первоначально предполагалось *его цветы*); (д) речевое затруднение, которое проявилось в наложении двух синтаксических структур (*ему кто-то сажает и у него кто-то приходит [сажать]*); (е) речевое затруднение, которое проявилось в необходимости переформулировать строку 34 в строке 35 с использованием дискурсивного маркера уточнения *в смысле*; и, наконец, (ж) подъем тона в главном фразовом акценте авторской ремарки (в строке 31 на глаголе *спросили*), который маркирует незавершенность и, тем самым интегрирует авторскую ремарку и собственно цитацию (строки 32–35) в единую иллокуцию:

(3) FS04-sp

30. а /там у него \цветы всякие в горшках,
31. .. ну и мы /спроси<sub>д</sub>ли,
32. –мол это е= || он /сам сажает,
33. или это ему-у там-м у него кто-то-о \приходит ему /сажа<sub>т</sub>ет,
34. или может быть это .. /домохозя<sub>й</sub>ка,
35. ..” ээ в смысле-е .. которая \квартиру-у ... {СМEX} мм –сда<sub>е</sub>т ему,

Во-вторых, авторская ремарка может произноситься вовсе без акцента, и тогда она еще более тесно интегрируется с цитируемым фрагментом. Так, в примере (4), в строке 84, глагол речи реализован атонически, авторская ремарка и цитация объединены в одну коммуникативно просодическую составляющую с единым фразовым акцентом:

(4) FS05-sp

- 83. позвонил /-маме,,,
- 84. сказал что он /-приехал,,,
- 85. /мама очень \обрадовалась.

Заметим также, что в этом примере и строка 83, и строка 84 реализованы с акцентом типа ИК-6 по Е. А. Брызгуновой — с подъемом на ударном слоге, за которым не следует падения на заударных. Эта конфигурация, которую условно можно назвать интонацией многоточия, широко используется для выражения незавершенности при описании череды событий, открытого списка (Янко 2008: 109–113, 166–167). Таким образом, цитация интегрируется в нарратив, описывающий последовательность событий «позвонил», «сказал (что приехал)», «мама обрадовалась».

Итак, прагматический фактор является ведущим при разграничении прямой и косвенной речи — он задает функциональный статус цитационной конструкции. Остальные факторы проецируют функциональное задание на конкретный языковой ландшафт, эти факторы менее универсальны и более лингвоспецифичны (ср. например, японский язык, где комплементаризер одинаково употребителен и с прямым, и с косвенным дейксисом).

Наши корпусные данные показывают, что в русской устной речи говорящий, выбрав между прямым и косвенным дейксисом, последовательно придерживается этого выбора в пределах цитационной конструкции. Крайне редки случаи, когда говорящий по ходу построения конструкции меняет прямой дейксис на косвенный или наоборот. Таким редким исключением является следующий пример, где в строках 71 и 73 имеется косвенный дейксис, подкрепленный соответствующим синтаксисом и просодией — сентенциальным актантом, просодически интегрированным в главное предложение, однако в строке 75 происходит сбой (при попытке построить директив с *пусть*, который требует специальных ухищрений в режиме косвенного цитирования), и в строке 76 возникает прямой директив (хотя и «приправленный» аппроксиматором-репортивом типа) *типа*):

(5) FS02-sp

- 70. В /итоге,
- 71. он сказал что он нам не /верит,
- 72. и он уже столько очень много \сказок /наслушался,
- 73. и говорит что если мы хотим .. не /платить штраф,
- 74. и просто /уехать,
- 75. значит пусть ==
- 76. ... ну типа «\Уезжайте!»

Рассмотрим теперь более подробно, как просодически реализуются цитационные конструкции в живой неподготовленной речи.

### 3. Неканонические случаи просодической реализации цитационных конструкций

В корпусе «Веселых историй из жизни» представлены три основных типа просодических отклонений от канонов прямой и косвенной речи в цитационных конструкциях: первый можно условно назвать «гиподиалогическим» вариантом прямой речи, второй — «дезинтегрирующим» вариантом косвенной речи, третий — «гипердиалогическим» вариантом прямой речи.

#### 3.1. Гиподиалогический тип

При гиподиалогических отклонениях цитация с прямым дейксисом утрачивает просодические симптомы отдельной иллокуции, интегрируется (подобно цитации с косвенным дейксисом) с авторской ремаркой в единую иллокуцию, которая, в свою очередь, встраивается далее в структуру рассказа. Проиллюстрирую это на примере (6) ниже. В строках 22–23 описывается речевое событие — высказывание героини. Ее речь строится с прямым дейксисом: вначале следует речевой акт обращения, а затем следует сообщение, в котором непосредственный адресат индексируется с помощью притяжательного местоимения второго лица. В диалоге и при каноническом цитировании в формате прямой речи это сообщение должно было бы просодически реализоваться с падением на реме (главном фразовом акценте). Однако происходит иное: речевое событие связывается в нарративе причинно-следственным отношением со следующим событием — имитированием опьянения. В этой цепочке речевое событие просодически маркировано как незаключительное: в главном фразовом акценте строки 23, на слове *решены-ы*, имеется подъем по типу ИК-6 с ровным участком на растянутом ударном слоге, см. рис. 2. Кроме того, авторская ремарка лишена собственного акцента и просодически интегрирована с обращением. Таким образом, строки 22–23 выступают как цельный узел нарративной структуры, и просодия маркирует не иллокутивное намерение «внутреннего» говорящего, а внешнюю связь этого узла с другими в нарративной структуре:

(6) FS25-sp

22. она говорит «/\Алексей@  
 23. Твои проблемы все ... /-решены-ы.»,,,  
 24. ... \во-от,  
 25. .. и-и .. собственно говоря тогда мне пришлось понимаешь  
 /эмулировать бессознательное \опьянение.

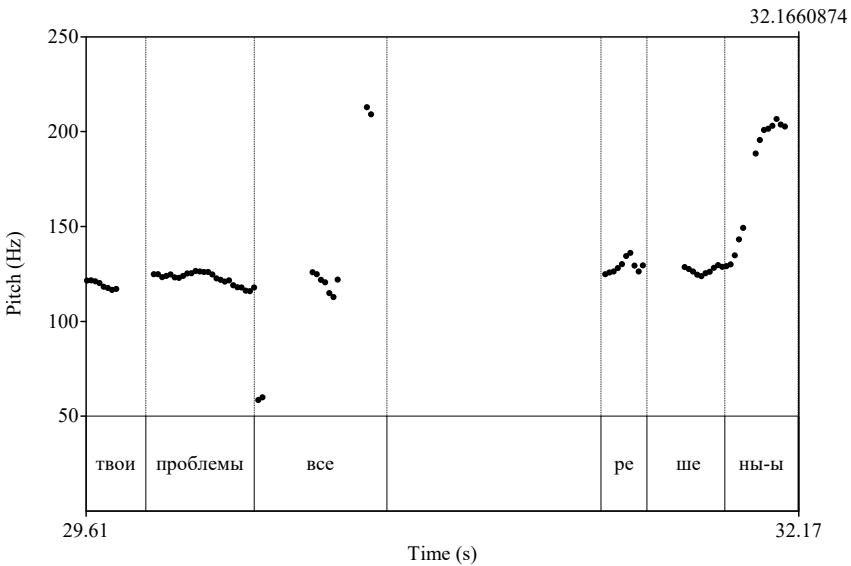


Рис. 2. FS25-sp: «проблемы решены»

Вариантом гиподиалогического паттерна является такой, при котором и авторская ремарка, и сама цитация являются незаключительными узлами нарративной цепочки. Так, в примере (7) строка 28, строка 29 и строка 30 являются незаключительными узлами и реализуются с одинаковыми подъемами в реме по типу ИКЗ — в высокий регистр с резким падением на заударных, см. рис. 3. Цепочка завершается лишь в строке 31, которая реализуется с падением тона в реме по типу ИК1 — прототипическое финальное падение в максимально низкий для данного говорящего уровень. Опять же сама цитация лишена иллокутивной просодии, хотя и исполнена с прямым дейксисом (второе лицо адресата) — подъем тона лишь отражает ее незаключительный статус в цепочке:

(7) FS08-sp

- 28. \постоянно /подкалывал,
- 29. «Ты как /парень»,
- 30. я ещё /стесняюсь этого,
- 31. а он меня ещё и \подкалывает.

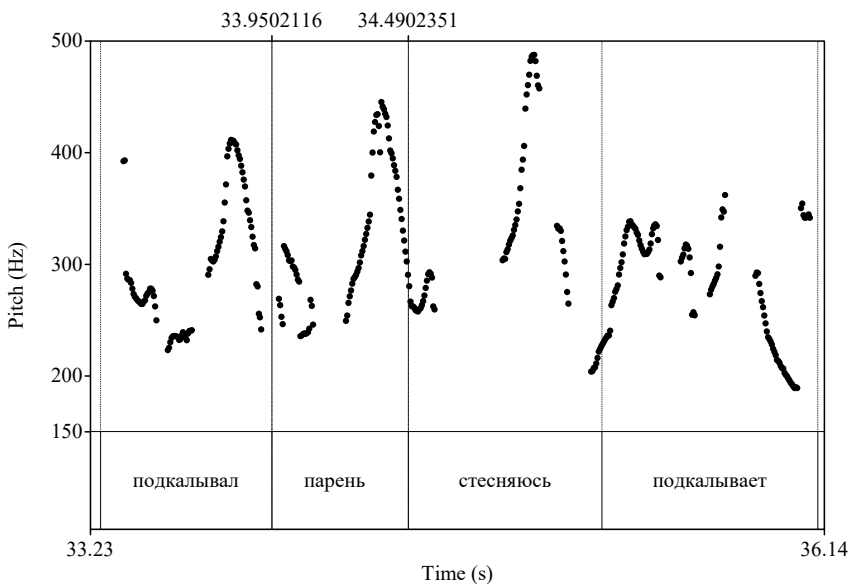


Рис 3. FS08-sp: «ты как парень»

Гиподиалогический тип хорошо сочетается с лексико-грамматическими отклонениями от прототипа при прямом дейксисе. См. пример (8), где все строки, включая 24 (авторскую ремарку) и 25–27 (цитируемый фрагмент), реализованы просодически как незаключительные узлы нарративной цепочки: 24 и 26 с подъемом по типу ИК3, а 25 и 27 — с подъемом по типу ИК6, см. рис. 4. При этом в строке 25 имеется «запрещенный» в прямой речи комплементаризер *что*, в той же строке — аппроксиматор-имитатив *вот*, а в строке 26 — аппроксиматор-репрезентатив *тыры-пыры* (аппроксиматоры не могли принадлежать исходному цитируемому фрагменту):

(8) FS05-sp

- 23. ...пошёл –бегать .. ээ по всяким /авиакассам,
- 24. .../объяснять,
- 25. что «/–Вот,,,
- 26. тыры-/пыры,
- 27. надо /–уехать.»,,,

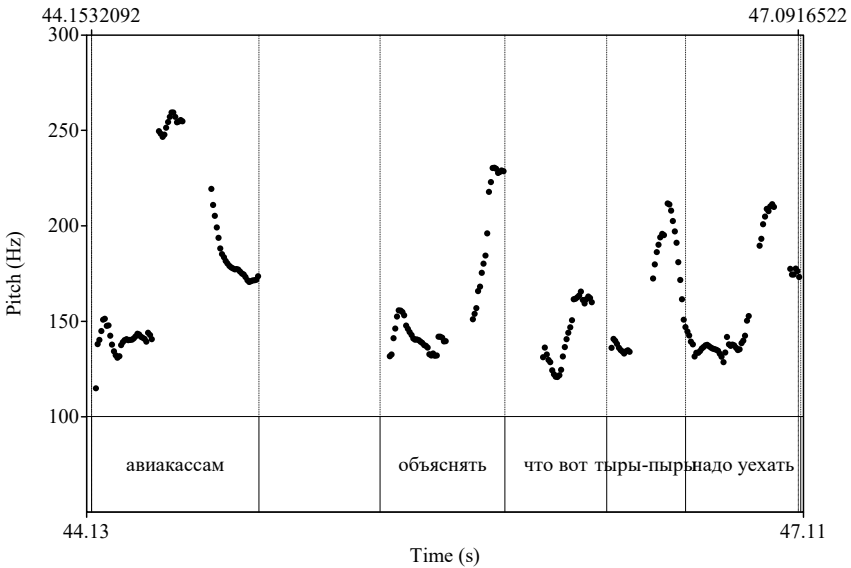


Рис. 4. FS05-sp «надо уехать»

Таким образом, гиподиалогический тип реализации цитационных конструкций с прямым дейксисом приводит к большей интегрированности конструкции, по сравнению с каноническим вариантом.

### 3.2. Дезинтегрированный тип

В конструкциях с косвенным дейксисом встречается просодический паттерн, обратный гиподиалогическому: это реализация авторской ремарки как независимой иллокуции — с падением тона в главном фразовом акценте. Если гиподиалогический паттерн сдвигает конструкцию с прямым дейксисом в сторону косвенной речи, то падение тона в авторской ремарке, наоборот, сдвигает конструкцию с косвенным дейксисом в сторону прямой. Проиллюстрирую это примером (9). Здесь в строке 193 авторская ремарка не интегрируется просодически с цитацией (что было бы типично для косвенного дейксиса): глагол речи произносится с понижением тона, свидетельствующий о том, что первоначально планировалась прямая цитация. Слушающий в этом месте ожидает начала цитации в формате прямой речи, однако этого не происходит — далее

следует маркер хезитации *значит*, знаменующий, по всей видимости, речевое затруднение, перестройку, и переход к косвенной цитации, вводимой комплементаризером *что*:

(9) FS18-sp

193. .... И тут Brenton ... мне \говорит:

194. .... \значит,

195. что он этот знак берёт с \собой!

Еще один возможный контекст для конструкций дезинтегрированного типа — изложение последовательности событий с использованием особой, «перечислительной» стратегии: говорящий монотонно «бубнит», произнося все клаузы, будь то главные или придаточные, с одинаковой интонацией открытого списка. Обычно в рематическом акценте каждой клаузы списка — ровный тон или небольшой пологий подъем тона с последующим ровным участием, т.е. ИК6 в терминологии интонационных конструкций Е. А. Брызгуновой. В строгом смысле при перечислительной стратегии нельзя говорить о том, что авторская ремарка артикулируется с интонацией завершенности, здесь дезинтеграция связана фактически с отказом противопоставления авторской ремарки и цитируемого фрагмента по завершенности: оба становятся равноправными элементами списка. Так, в следующем примере одинаково — с акцентом по типу ИК6 в реме — реализуются и строка 71 (авторская ремарка) и строки 72 и 74 (цитируемый фрагмент):

(10) FS18-sp

71. .. /—спрашивали у нас,,,

72. /—откуда мы,,,

73. .... \вот,

74. чего мы тут /—делаем,,,

Таким образом, данный тип реализации цитационных конструкций с косвенным дейксисом приводит к меньшей интегрированности конструкции, по сравнению с каноническим вариантом.

### 3.3. Гипердиалогический тип

Гипердиалогический тип конструкций с прямым дейксисом говорящие используют для намеренной театрализации описываемого речевого события. Утрирование исходного иллокутивного намерения, с одной стороны, сохраняет прагматическую перспективу «внутреннего» говорящего, но, с другой стороны, привносит и оценку ситуации внешним повествователем. Оценочные значения могут быть любого «знака» от восхищения до недоумения. Часто такая педалированная просодия, передразнивание, понимается как ироническая, сопровождается улыбкой или смехом (особенно, учитывая, что материал — «веселые истории»). При этом просодическая реализация включает выраженные черты эмфазы, в том числе, уже упоминавшиеся выше удлинение и тональное искривление акцентированного слога, а также сдвиг в более высокий регистр, вплоть



до фальцета. Так, в примере (11) гипердиалогически реализуется строка 15 — с усиленной громкостью, в очень высоком регистре, с восходящее-нисходящим тоном на ударном слове, см. рис. 5:

(11) FS12-sp

12. ·и /потом он ·· просто на всю /маши-ину,
13. я бы даже сказала /\бо-ольше,
14. больше чем на всю /машину,
15. ·· «/\**Кто-о** это?!»
16. {СМЕХ} Начинает /кричать,

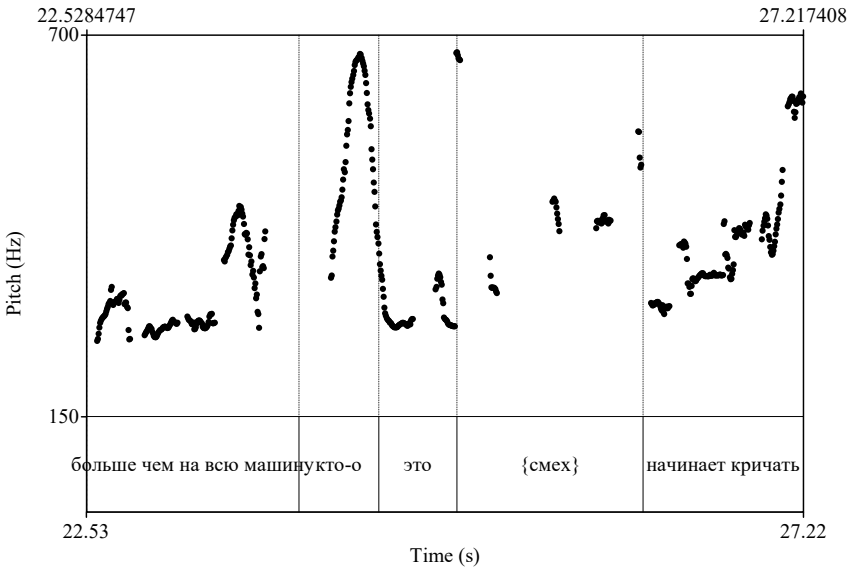


Рис. 5. FS12-sp: «кто это»

Аналогичным образом, в примере (12) с явной иронией гипердиалогически реализуется цитируемое обращение, строка 19: фальцетом, растянуты и ударный и заударный слоги, оба — с восходящее-нисходящим тоном, см. рис. 6:

(12) FS21-sp

- 15. … {ЧМОКАНЬЕ} Одна бабуля —
- 16. … /поняв,
- 17. что она не успевает выйти на /остановке,
- 18. — … неожиданно \кричит:
- 19. … {ЦОКАНЬЕ} … «/\↑↓Води-ите-ель@!

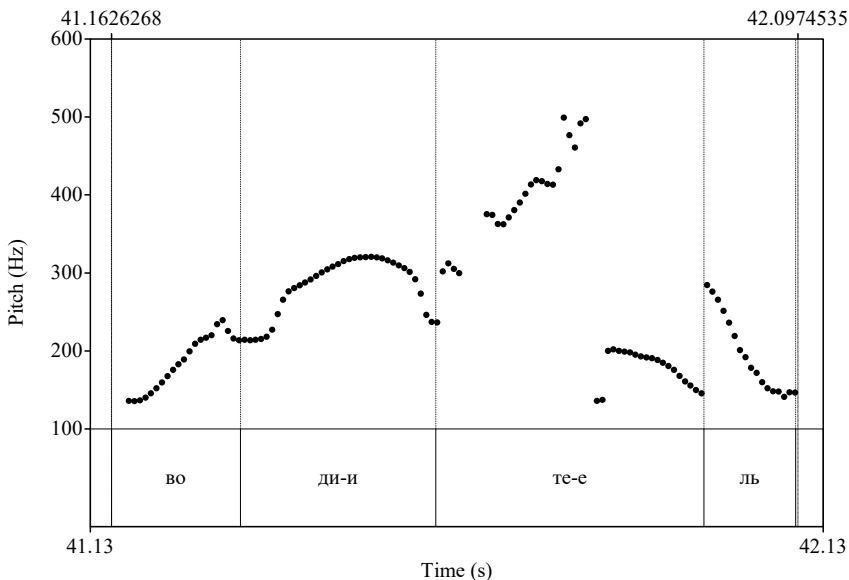
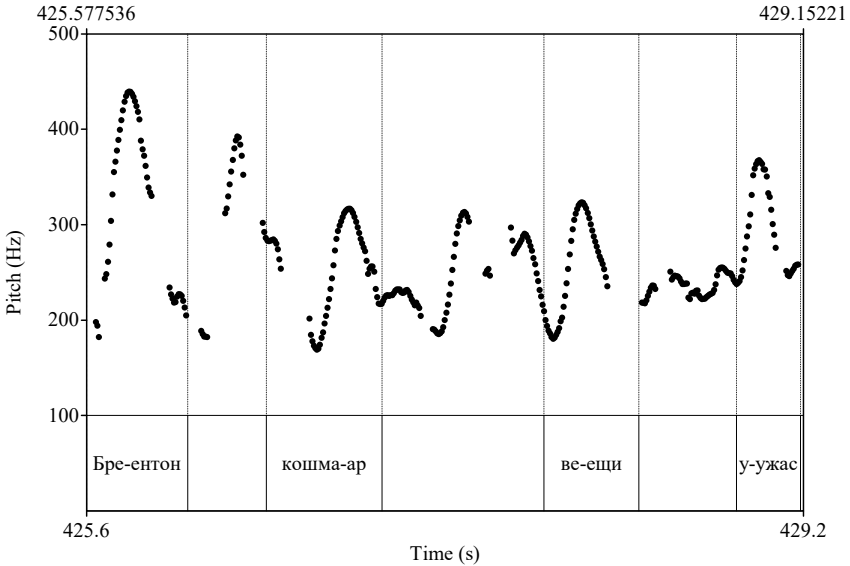


Рис. 6. FS21-sp: «Водитель!»

Гиподиалогические и гипердиалогические отклонения от канона прямой речи в редких случаях могут наблюдаться совместно в цитационной конструкции с прямым дейксисом. Так, в следующем примере цитация реализуется иронически утрированно, с намеренной театрализацией (высокий регистр, растянутые ударные слоги, «блеющий» голос, Янко 2008: 85), ср. рис. 7. Но при этом строки 238, 239 и 240 произносятся с явным просодическим маркированием незавершенности, т.е. как незаключительные в нарративной последовательности:

(13) FS18-sp

237. я говорю «/\Бре-ентон@!  
 238. Какой /\кошма-ар!,,,  
 239. нельзя такие /\ве-ещи!,,,  
 240. это же /\у-ужас просто!,,,  
 241. .. нас посадят в /\тюрьму непременно!».



**Рис. 7.** FS18-sp: «кошмар, ужас»

Итак, реализация конструкции с прямым дейксисом в гипердиалогическом варианте способствует ее большей интеграции в текущий нарратив за счет «внедрения» точки зрения «внешнего» говорящего непосредственно в цитируемый фрагмент.

#### 4. Количественные данные и итоги

В предыдущих разделах были обсуждены основные типы просодических отклонений от канонических случаев прямой и косвенной речи, которые были обнаружены в обследованном корпусе. Приведем теперь некоторые количественные данные, которые позволят составить представление о распространенности задокументированных нами явлений.

Из общего числа эпизодов включения чужой речи (162) в корпусе «Веселых историй из жизни» лишь 19 либо имеют признаки косвенного дейксиса, либо вообще не имеют однозначно интерпретируемых прагматических маркеров. Абсолютное большинство (143) имеют признаки прямого дейксиса. В соответствии со сказанным выше, к этому классу были отнесены цитационные

конструкции, где в составе цитации личные местоимения и личные формы глагола употреблены «прямо» (без индексального сдвига), а в отсутствие свидетельств личного дейксиса (или в дополнение к нему) — где в составе цитации имеются *main-clause phenomena*. Количественные сведения об обнаруженных неканонических явлениях них суммированы в таблице 1:

**Таблица 1.** Отклонения от прототипов прямой и косвенной речи в цитационных конструкциях в корпусе «Веселых историй из жизни»

	Всего	С просодическими отклонениями от канона		
Эпизоды включения чужой речи	162	40		
Имеют признаки прямого дейксиса	143	37	гиподialogический тип	24
			гиперdialogический тип	13
Не имеют признаков прямого дейксиса	19	3	дезинтегрированный тип	3

Как видим, количественное распределение обнаруженных цитационных конструкций позволяет сделать несколько важных выводов. Во-первых, становится очевидным, что доля конструкций с прямым дейксисом в общем числе отраженных в нарративе речевых событий очень высока (143 из 162). Во-вторых, высока доля конструкций с просодическими отклонениями от канона — четверть от общего числа (40 из 162), при этом отклонения от канона в конструкциях с прямым дейксисом встречаются чаще (37 из 143), чем в конструкциях с косвенным дейксисом (3 из 19). При этом гиподialogический тип представлен гораздо шире (24 из 37), чем гиперdialogический (13 из 37, имевшиеся в выборке два случая, где были представлены оба типа, считались в таблице как гиперdialogические). Разумеется, проведенное исследование приоткрывает лишь малую часть всего комплекса проблем, связанных с передачей чужой речи в устном неподготовленном дискурсе. Дальнейшая работа с этим и другими корпусами звучащей речи позволит установить, в частности, насколько просодические отклонения от канона коррелируют с лексико-грамматическими отклонениями, и, во-вторых, насколько наличие просодического ресурса влияет на номенклатуру и частотность лексико-грамматических отклонений от канона в устной речи по сравнению с письменной.

## Литература

1. Aikhenvald A. (2008) Semi-direct speech: Manambu and beyond // *Language Sciences* 30: 383–422
2. Aikhenvald A. (2011a) *Speech Reports: A Cross-Linguistic Perspective* // Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.) *Language at Large: Essays on Syntax and Semantics*. Brill: Leiden, Boston. 2011, 290–326
3. Aikhenvald A. (2011b) *Semi-Direct Speech in Typological Perspective* // Alexandra Y. Aikhenvald and R. M. W. Dixon (eds.) *Language at Large: Essays on Syntax and Semantics*. Brill: Leiden, Boston. 2011, 327–368

4. *Bryzgunova E. A.* (1982) *Intonatsiya* [Intonation], *Russkaya grammatika* [Russian Grammar]. Vol. 1, Nauka, Moscow, pp. 98–118.
5. *Evans, Nicholas.* (2013). Some problems in the typology of quotation: A canonical approach. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*. Oxford: Oxford University Press, 66–98
6. *Janko T. E.* (2008) *Intonacionnye strategii russkoj rechi v tipologicheskom aspekte* [Intonational strategies in spoken Russian from a comparative perspective]. Moskva: Jazyki Slavjanskix Kul'tur.
7. *Kibrik A. A., Podlesskaya V. I.* [Eds.] (2009) *Rasskazy o snovidenijax: korpusnoe issledovanie usntogo russkogo diskursa* [Night Dream Stories: A corpus study of spoken Russian discourse]. Moscow: Jazyki Slavjanskix Kul'tur.
8. *Paducheva E. V.* (1996a) *Subjektivnaja modal'nost': illokutivnye pokazateli i vvodnye slova* [Subjective modality: illocutive markers and parentheticals] // *Semantičeskie issledovanija. Semantika vremeni i vida v Russkon jazyke. Semantika narrativa* [Semantic investigations Semantics of tense and aspect in Russian. Semantics of narrative.] *Jazyki russkoj kul'tury*, Moscow, 297–320
9. *Paducheva E. V.* (1996b) *Nesobstvennaja prjamaja rech i svobodnyj kosvennyj diskurs* [Semi-direct speech and free indirect discourse] // *Semantičeskie issledovanija. Semantika vremeni i vida v Russkon jazyke. Semantika narrativa* [Semantic investigations Semantics of tense and aspect in Russian. Semantics of narrative.] *Jazyki russkoj kul'tury*, Moscow,
10. *Spokencorpora* (2013) *Prosodically Annotated Corpus of Spoken Russian (PrACS-Russ)*. Pilot version. Online: <http://spokencorpora.ru>

# ОРФОЭПИЧЕСКИЙ ВЗГЛЯД НА НЕКОТОРЫЕ ВАРИАНТНЫЕ ЯВЛЕНИЯ РУССКОГО ЛИТЕРАТУРНОГО ЯЗЫКА В ЭПОХУ КОРПУСНОЙ ЛИНГВИСТИКИ<sup>1</sup>

**Пожарицкая С. К.** (sofkonst@yandex.ru)

МГУ им. М. В. Ломоносова, Москва, Россия

**Добрушина Е. Р.** (edobrush@gmail.com)

Православный Свято-Тихоновский Государственный университет, Москва, Россия

В статье обсуждается статус орфоэпии как научной дисциплины: предмет и границы орфоэпии; характер орфоэпических рекомендаций и эволюция словарных помет в связи с динамикой акцентных и грамматических вариантов литературного произношения. Рассматриваются возможности использования корпусных данных для формулирования некоторых прогнозов относительно динамики орфоэпической нормы.

1. Утверждается, что предметом орфоэпических рекомендаций не могут быть позиционные модификации фонем типа [см']ех/[с'м']ех, которые не фиксируются слухом «наивного говорящего» и не могут быть предметом его сознательного выбора.

2. Предметом рекомендаций и оценки по шкале «старое»/«новое» могут быть варианты слов, отличающиеся звукотипами, как ж[а]лётъ/ж[ы]лётъ, [сейф]/[с'ейф], [ч]то/[ш]то.

3. Акцентные варианты слова являются самым подвижным и востребованным участком орфоэпии. Учитывая динамику конкурирующих вариантов, орфоэпические словари должны избегать запретительных помет, отдавая предпочтение рекомендательным типа «старое-новое».

4. Рассматривая варианты некоторых грамматических форм (*поезжай/езжай, одеть/надеть, тычу/тыкаю, мучу/мучаю, их/ихний поезжай/езжай, одеть/надеть, тычу/тыкаю, мучу/мучаю, их/ихний*), мы приходим к выводу, что данные НКРЯ и сопоставление Основного и Газетного корпусов может дать объективные сведения о современном состоянии орфоэпической нормы и прогнозировать ее динамику.

В заключение следует отметить разницу помет в орфографических и орфоэпических словарях: если орфографическим пометам естественно быть регламентирующими и запретительными, то орфоэпические пометы должны быть по преимуществу рекомендательными, не претендующими на блокирование изменения языка.

**Ключевые слова:** система, узус, нормативные пометы, динамика, корпус

---

<sup>1</sup> Исследование выполнено при поддержке гранта РГНФ 17-04-00517

## ORTHOEPY IN AN ERA OF CORPUS LINGUISTICS: SOME CASE STUDIES IN VARIATION IN STANDARD RUSSIAN

**Pozharickaja S. K.** (sofkonst@yandex.ru)

Moscow State University, Moscow, Russia

**Dobrushina E. R.** (edobrush@gmail.com)

St. Tikhon's Orthodox University (PSTGU), Moscow, Russia

We discuss the status of orthoepy as a linguistic discipline; its scope and limitations; the nature of orthoepic prescriptions and the evolution of dictionary marks in its connection to the evolution of the stress and grammar of standard Russian. We consider the use of Russian National Corpus as an instrument of prediction of further orthoepic changes and envisaging prescriptions resulting from this change.

1. Prescription can not apply to contextual modifications of phonemes ([см'] ех/[с'м']ех), because they are not perceived by a "lay speaker", are not subject to his or her conscious choice and can not lead to communicative failures.

2. On the other hand, prescription may apply to variants that are different in 'sound types' (звукотип, e.g. ж[а]лётъ/ж[ы]лётъ, [сейф]/[с'ейф], тй[хъй]/тй[х'ий], [ч]то/[ш]то, ё[жж]у/ё[ж'ж']у); these variants may be suggested to represent newer vs. older norms.

3. Stress variants are the most dynamic domain of orthoepy. Special focus is on the change in verbal stress as reflected in the difference between the prescribed vs. real usage and in the chronology of dictionary marks (*дружит→дружит, родился→родился* etc.).

4. The use of the data of Russian National Corpus and, more specifically, comparative analysis of the main corpus and the corpus of newspapers is an important method of analysis of grammatical variation. We consider the competition between such forms as *поезжай/езжай, одеть/надеть, тычу/тыкаю, мучу/мучаю, их/ихний* and provide charts showing positive dynamics of the more recent variants.

In conclusion, we stress the difference in the nature of orthographic and orthoepic marks in the dictionaries. Orthographic marks may be prescriptive and prohibitive. Orthoepic marks may only, in most cases, have a status of recommendations; they can not (be regarded as means to) block language change.

**Key words:** language usage, prescription, linguistic change, corpus linguistics

Орфоэпия не является самостоятельным разделом науки о языке. У нее нет своего отдельного предмета исследования; она вторична по отношению к фонетике, акцентологии, грамматике, возможно — к интонологии, синтаксису и лексике. Это дисциплина прикладного характера, которую называют «правилами пользования устной речью». Вопрос о границах ее применения является предметом дискуссии и определяет различия в содержании орфоэпических

словарей. В их названиях обычно присутствует слово «произношение», которое имеет неопределенный смысл (от индивидуальных особенностей артикуляции до выбора лексем), а потому не может считаться термином и не проливает свет на содержание понятия «орфоэпия».

1. Существует устойчивая традиция связывать орфоэпию с фонетикой, понимаемой как изучение и позиционных реализаций фонем, и вариантов фонемного состава слов, хотя второе не является фонетикой в настоящем смысле слова.

Орфоэпический словарь 2012 года (далее ККК 2012) добавляет к этому проблемы ударения. Словарь 1983 г. (далее ОС-83) включает также варианты грамматических форм слов, и Р. И. Аванесов пишет по этому поводу: «Объем понятия орфоэпии не является вполне устоявшимся. Одни понимают орфоэпию суженно, как совокупность норм произношения в собственном смысле слова, исключая из нее вопросы ударения, практически очень важные. Другие, напротив, понимают орфоэпию очень широко — как совокупность не только специфических норм устной речи (т.е. норм произношения и ударения), но также и правила образования грамматических форм (ср., например, в русском языке *свечей* или *свеч*, *колыхáется* или *колышется*, *тяжелее* или *тяжéле*), которые в равной мере относятся также и к письменному языку. Такое понимание представляется наиболее целесообразным» (Аванесов 1984: 13).

Поле деятельности орфоэпических рекомендаций или границы орфоэпии следует определять в связи с функцией орфоэпии как практической дисциплины, цель которой — в ситуации вариативности и возможности (потребности) выбора рекомендовать говорящему оптимальный («правильный») звуковой состав слова либо место ударения в нем. «Правильный» звуковой облик слова выявляется в процессе коммуникации, который должен быть бесконфликтным, — этим определяется диапазон вариативности звучания. Но звуковой состав слова усваивается в основном путем подражания; в процессе речи он воспроизводится автоматически и может контролироваться сознанием только тогда, когда различие достигает уровня звукотипов. Так, различие безударных гласных в системе эающего и икающего вариантов литературной нормы [н'и<sup>е</sup>сú] / [н'и]сú не фиксируется слухом «наивного говорящего», не приводит к сбою коммуникации, и орфоэпическая рекомендация в этом случае не имеет смысла. Так же обстоит дело и с разницей в степени позиционно обусловленной мягкости [с] / [с'] / [с'] в слове *смех* в отличие от самостоятельной (фонематической) мягкости согласного в [се]ссия / [с'е]ссия, а также с составом гласных звуков в многосложном слове, например, *многоквартирный*: [мнóгъ] *квартíрный* / [мнъгъ] *квартíрный*. Изобилие предупредных гласных в разговорной речи может компенсироваться эллипсисом слога — [мнъг] *квартíрный* — без ущерба для восприятия смысла.

Коммуникативный конфликт, который проявляется в негативной реакции пассивного участника акта коммуникации, может быть спровоцирован динамикой узуса, как *e[ж']y* / *e[ж]y*, *ти[x'эй]* / *ти[x'эй]*, либо привнесением элемента другой фонетической системы, как *стри*[r'ót], *но*[y]á или [н'е]сú, [н'о]



сú, [n'a]сú. Первое заслуживает оценочной словарной пометы по шкале «старое» — «новое», второе — как элемент иной системы — не является предметом орфоэпии.

Для описания позиционно обусловленных (фонетических) вариантов существует термин **орфофония**. Так, С. С. Высотский пишет: «У славистов есть понятие не только орфоэпии, но и орфофонии. Лингвисты время от времени об этом напоминают. Вот Дмитрий Николаевич Ушаков давно уже об этом хлопотал, чтобы лингвисты не путали эти понятия. И Надия Александровна Янко-Триницкая писала об этом, чтобы не путали по плохой традиции разные вещи, когда к фонетике относят фонемный состав слова, что на деле принадлежит области лексики и морфологии, а не собственно фонетике, не живому произношению» (Высотский 1984: 24).

Идея существования двух разных объектов описания и, соответственно, двух разных научных дисциплин лежит в основе описания ленинградского произношения (называемого Петербургской нормой) Л. А. Вербицкой: «Исходя из понимания фонемы представителями ленинградской фонологической школы, можно сказать, что орфоэпия занимается выявлением фонемного состава слов (например, в слове поэт есть фонема /a/, а не /o/: /рает/, а не /роет/, орфофония занимается вопросами качества произнесения отдельных звуков, а именно тем, как должен произноситься тот или иной звук (например, в слове поэт /a/ — гласный открытый, после губного согласного /р/ несколько огубленный, начальная фаза этого гласного более задняя, чем стационарный участок и т. п.). В самом общем виде можно сказать, что орфоэпия — это правила, определяющие нормативный фонемный состав слова, а орфофония — это правила произношения вариантов фонем» и далее: «Сознательной кодификации подвергается именно орфоэпия, так как фонемный состав слов хорошо осознается носителями языка» (Вербицкая 1976: 27).

Транспонируя этот текст на идеологию Московской фонологической школы и заменяя в ряде случаев слово «фонема» словом «звукотип», можно согласиться с тем, что орфоэпия и орфофония имеют разные объекты исследования, при этом орфофония допускает оценку только по шкале «старое-новое», а орфоэпия может оценивать варианты по шкале «правильно-неправильно», если имеет к тому объективные основания.

2. Мена звукотипов обычно осознается говорящими и может быть предметом выбора и оценки по параметру «правильно-неправильно», которые обычно совмещаются с оценкой по шкале «старое-новое», например: *ж[а]леть / ж[ы]леть, е[жж]у / е[ж'ж']у, [ч']то / [шт]то, ти[x'ь]й / ти[xъ]й, [т]error / [т']error*.

Трудность получения сведений о звуковых и акцентных вариантах для формулирования орфоэпических рекомендаций состоит в том, что здесь не применим прямой опрос: у говорящих часто отсутствует адекватное представление о звуковом составе произносимых ими слов, поскольку на это накладывается зрительный, орфографический облик слова. Об этом свидетельствуют известные случаи несовпадения между реально практикуемым произношением и тем, как его декларирует сам говорящий (это называют одним из парадоксов

социолингвистики). Так, при составлении «Вопросника по современному русскому произношению» всего 18,5% опрошенных заявили, что они произносят мягкое [з'] в слове *зверей*, а в магнитофонной записи произношения тех же лиц оказалось 72% случаев мягкого [з'] в этом слове (Барина, Ильина, Кузьмина 1971: 325). Близкие к этому результаты показывает и сопоставление акцентных вариантов, реально употребляемых (*actual usage*), с теми, которые говорящий считает правильными и приписывает себе (*reported usage*).

С этим связан главный вопрос орфоэпии — о критериях оценки существующих произносительных вариантов. В дискуссиях на эту тему выявляются две позиции, одну из которых называют субъективной, другую — объективной.

**2.1. Субъективной** называют позицию априорного представления о нормативности, основанного на традиции, следовании образцам и авторитетам и личном вкусе автора рекомендаций. К этому отрицательно относились в свое время еще А. М. Пешковский и А. А. Шахматов, говоря о «литературно-языковом идеале»: «Первой и самой замечательной чертой его является поразительный консерватизм, равного которому мы не встречаем ни в какой другой области духа. Из всех идеалов это единственный, который лежит целиком позади. «Правильной» всегда представляется речь старших поколений, предшествовавших литературных школ. Нормой признается то, что было, и отчасти то, что есть, но отнюдь не то, что будет» (Пешковский 1959 : 55), и «Главный и единственный авторитет в языке — это обычай, употребление» и далее: «Странно было бы вообще, если бы учебное заведение вместо того, чтобы указывать, как говорят, решалось указывать, как надо говорить» (Шахматов 1899 : 30). Тем не менее, такой способ оценки вариантов продолжает существовать.

**2.2. Объективная** оценка орфоэпических вариантов может быть только результатом социолингвистических исследований, которые на материале русского языка были начаты в 60-е годы XX-го столетия под руководством М. В. Панаева. Они выявили сдвиги в произносительном узусе середины XX-го века по сравнению с концом XIX-го, («младшая» и «старшая» орфоэпические нормы) и определили пути дальнейших исследований, которые могут быть продолжены возможностями корпусной лингвистики; в частности, Национального корпуса русского языка (НКРЯ) и других (например, ГИКРЯ), а также путем специально создаваемых для конкретных задач баз данных (см., например, Пéroва, Бондаренко, Добрушина 2016).

**3.** Акцентные варианты слова являются самым «востребованным» и самым сложным участком орфоэпического нормирования, поскольку тенденции изменения акцентных норм реализуются на конкретных лексемах и постепенно распространяются на другие слова той же грамматической категории, При этом рекомендации словарей обычно отстают от узуса, а их предпочтения часто не имеют объективных оснований. «Отсутствие фиксированного места ударения, разноместность его делает ударение в русском языке индивидуальным признаком каждого слова» (Аванесов 1984: 13).

Происходящие здесь процессы неоднородны: наблюдается усиление подвижности ударения в одних классах слов (именах) и стабилизация — в других (глаголах). Так, в акцентной парадигме существительных мужского рода происходит смещение ударения на окончание в субпарадигме мн. числа и мена <и> на <а> в им. мн. у многосложных имен: (*ли́фты / лифты́* и *трéнеры / тренерá*); в глаголе наблюдается сдвиг ударения с приставки на основу в субпарадигмах прошедшего времени глаголов типа *нали́ть, переня́ть* и с окончания на основу — в парадигме глаголов типа *дру́жить, звони́ть, вклю́чить*; ударение переходит с возвратного постфикса на основу: *роди́лся / роди́лся*. При этом данные опросов и наблюдений показывают существенные расхождения между нормативными пометами словарей и узусом. Так, по данным Э. Марклунд Шараповой, форму *нали́л*, имеющую в разных словарях пометы *доп.* и *разг.* при нормативном *на́лил*, употребляют и признают правильной 92,5% говорящих; *отда́л* (*доп., разг.*) — 64,5% (N *отдал*); *роди́лся* (N *роди́лся* и *родился́*) — 93,5%; *нача́лся* (*неправ.*) — 58,5% (N *начался́*); *включи́шь* (*неправ., разг.*) — 85,8% (N *включи́шь*); *позво́ним* (*не рек., простореч.*) — 32% (N *позвони́м*) (Marklund Sharapova 2000).

Динамику акцентных вариантов форм презенса глагола *дру́жить* показывают словари:

**Табл. 1.** Динамика акцентных вариантов презенса глагола *дру́жить*<sup>2</sup>

	дру́жит	дру́жит
Словарь 1847	N	-
Огиенко 1914	N	-
Ушаков 1935–1940	N	-
БАС 1950–1965	N	просторечие
Аванесов-Ожегов 1960	N	и допустимо
МАС 1957–1961	N	N
Ожегов-Шведова 1972	N	N
Горбачевич 1973	устаревающее	N
ОС 1983	допустимо устаревающее	N
РПП 1996	устаревающее	N
БАС 2006	перех. устар. <sup>2</sup>	N
ККК 2012	—	N

Акцентная динамика этого глагола просматривается и по акцентологическому корпусу НКРЯ: из 23-х форм со старым акцентом типа *дру́жит* 8 зафиксированы в XIX в., 12 в XX в. и 3 в XXI в., при этом количество форм с новым акцентом преобладает в 11 раз (252). Они появляются в 20-е годы XX в., составляя 20% от общего их количества и оставляя 80% на долю XXI века. форм с новым акцентом типа *дру́жит*.

Орфоэпические рекомендации относительно места ударения должны учитывать направление и темпы происходящих изменений и избегать запретов в тех случаях, когда изменение системно; так, «запретительная» помета

<sup>2</sup> В словарной статье *дру́жить* примеры из А. С. Пушкина и Н. А. Некрасова.

ОС 83 *неправ.* для формы *включит* заменена на «разрешительную» и *допуст. младш.* в ККК 2012, и дальнейшее продвижение «новых» форм таких глаголов в направлении нормы не подлежит сомнению.

Орфоэпические прогнозы относительно конкурирующих вариантов часто переоценивают значение системы и недооценивают власть узуса, и нельзя согласиться, например, с А. Б. Шапиро, когда он говорит: «Даже если девяносто процентов будут говорить *документ*, это не может стать литературной нормой» (цит. по Крысин 2004: 294–310). Почему? Никаких системных запретов здесь нет (правда, большая часть имен на *-мент* имеет ударную финаль), а 90% говорящих — это уже норма, и остается только понять причину ее становления.

#### 4. Варианты грамматических форм, обсуждаемые с позиций «правильно — неправильно»

4.1. Императив глаголов *ехать*, *ездить* (*езжай?* *поезжай?* *ехай?* *едь?* *езди?*). В Грамматическом словаре А. А. Зализняка относительно *ехать* говорится: «... *повел. нет* (в знач. *повел. употр. поезжай* [*также простореч. езжай*], с отрицанием не *езди*)» (Зализняк 2016 : 684). Но рекомендуемая нормой форма *поезжай* находится вне системы, поскольку в СРЛЯ отсутствуют соотносительные с ней формы *\*поезжать*, *\*поезжаю*<sup>3</sup>, в этом смысле у нее нет преимущества перед отвергаемой нормой формой *езжай*, которая тоже не имеет опоры в виде *\*езжать*, *\*езжаю*. Обе формы аномальны по отношению к системе, однако частотность *езжай(те)* имеет положительную динамику, а *поезжай(те)* — отрицательную. При этом в Основном корпусе НКРЯ нормативная форма *поезжай(те)* имеет некоторое преимущество (60% : 40%), а в Газетном, который в большей степени ориентирован на язык СМИ XXI века — наоборот (20% : 80%)



Рис. 1. Частота форм *поезжай(те)* / *езжай(те)* в Основном корпусе НКРЯ

Таким образом, учитывая динамику варьирующихся форм, словарные рекомендации относительно их употребления несомненно заслуживают смягчения.

<sup>3</sup> Заметим, что в севернорусских Архангельских говорах отмечаются редкие императивы *поезжай* и перех. *новезай*, *понесай*, полная парадигма которых неизвестна (и есть ли она?).

4.2. Динамику изменения формообразовательного класса с устранением чередования демонстрирует вариативное деепричастие глагола *тыкать* «указывать»: формы *тыкая(сь)* / *тыча(сь)*.

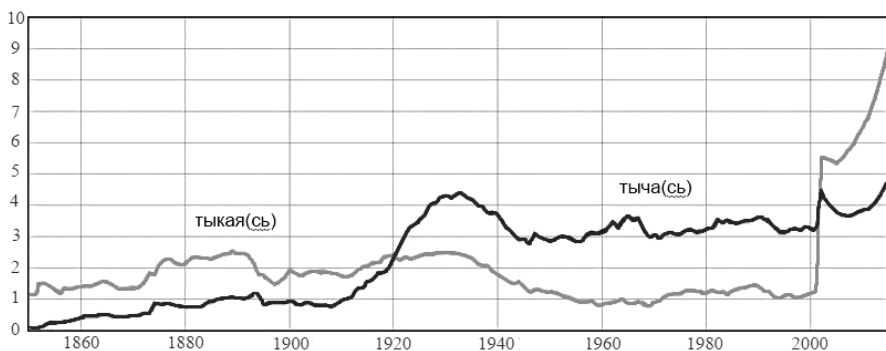


Рис. 2. Частота деепричастий *тыкая(сь)* и *тыча(сь)* в Основном корпусе НКРЯ

Как видим, «новая» форма *тыкаясь*, уже в середине XIX века имевшая преимущество перед «старой» формой *тычась*, уступила ей в 20-е годы XX века, а в начале XXI «взяла реванш» в виде резкого скачка опережения. Однако, судя по данным ГИКРЯ, можно думать, что здесь кроется не выявленная нами разница между спонтанной речью: (*тыка(сь):тыча(сь)*) — (60:40 из 11700) и профессиональной: *тыка(сь):тыча(сь)* — (26:74 из 800).

4.3. Варианты ударения в формах презенса глаголов *мучусь ... мучатся* и *мучаюсь ... мучаются* ОС-83 дает как равноправные, однако судя по корпусным данным, их узуальный статус различен: «новые» формы типа *мучаю* сменяют «старые» типа *мучу*, и процесс вытеснения форм *мучусь*, *мучатся* новыми *мучаюсь*, *мучаются* происходит очень интенсивно и практически уже закончен: в Газетном корпусе «старых» форм ничтожно мало: для инфинитива *мучать*: *мучить* 97 : 3 (из 1861), для презенса *мучаются* : *мучатся* 96 : 4 (из 372) в КЖЗ и 86 : 14 (из 14237) в КЖЖ.



Рис. 3. Частота форм *мучатся* и *мучаются* в Основном корпусе НКРЯ

Тем самым, орфоэпическая рекомендация, уравнивающая статус старой и новой форм, должна быть скорректирована в пользу новой.

**4.4.** Однокоренные глаголы *одеть* и *надеть* конкурируют в аналогичных контекстах: *надеть шапку / одеть шапку*. Реально существующее семантическое различие между этими глаголами, узаконенное нормой как необходимость говорить *одеть ребенка*, но *надеть шапку*, не ощущается говорящими как значительное; игнорирование его не провоцирует коммуникативного конфликта, и это различие стирается в современной речи, как устной, так и письменной. При этом «агрессивное поведение» наблюдается со стороны *одеть* и возможности сказать *надеть / одеть шапку*, но только *одеть ребенка*. Вытеснение *надеть* его «парой» *одеть* наблюдается по крайней мере с 40-х годов XIX века, ср: (1) *А какое платье вы оденете в четверг?* (И. И. Панаев 1840), (2), *одев очки* (Н. А. Некрасов), *одеть пальто* (Ф. М. Достоевский), *одевши на себя две шубы* (Н. И. Костомаров, 1862–1875), и возрастает не ослабевающими темпами. Исследованная А. А. Бонч-Осмоловской представительная база интернет-текстов XXI века (комментарии, живые журналы и т.д.) показывает равноправие глаголов при прямых дополнениях со значением артефактов, и это один из примеров победы узуса (имплицитной нормы) над декларируемой эксплицитной нормой (Бонч-Осмоловская 2008).

**4.5.** Стойко отрицательная помета *неправ.* при местоимении *ихний* как варианте личного и притяжательного *их* завоевала в речевой практике некоторое право на существование в «отпочковавшемся» значении «чуждый, принадлежащий другой культуре»: в исследованном (Добрушина 2014) Газетном корпусе из 151 обнаруженных форм *ихний* 30% зафиксированы в контекстах на тему правильности речи или в качестве речевой характеристики некультурных персонажей, а остальные 70% имеют относительное значение: (1) *все больше «ихние» Анджелины да Памелы...* (о западных актрисах), (2) *у ихних «франклинов» с «рузвельтами»* (о долларах) (3) *интересна мотивация ихнего Первого* (о немецком Первом канале телевидения), (4) *не устраиваем бузу на нашей Болотной или ихней Уолл-стрит* и т.д. При этом если в художественных текстах второй половины XIX в. *ихний* в устах персонажа является знаком разговорной речи, возможно, в рамках литературного языка, то в XXI веке *ихний* как вариант притяжательного *их* находится за рамками литературной нормы. Но нельзя исключить возможности того, что в будущем *ихний* войдет в нормативные словари с положительной пометой в этом специальном значении.

Краткий обзор актуальной орфоэпической практики и соотнесение ее с корпусными данными показывает следующее:

- содержание орфоэпических словарей должно определяться практическими вопросами, возникающими в связи с реально существующей вариативностью фонемного состава слов, места ударения и структуры грамматических форм;
- учитывая динамику смены вариантов, пометы орфоэпических словарей должны иметь по преимуществу рекомендательный, а не запретительный характер;
- нормативные рекомендации словарей подлежат контролю корпусными эмпирическими данными, и нет сомнений в целесообразности их использования в практических целях.

## Литература

1. *Аванесов Р. И.* (1984). Русское литературное произношение. Москва.
2. *Аванесов Р. И., Ожегов С. И.* (1959). Русское литературное произношение и ударение. Москва.
3. *Баранова Г. А., Ильина Н. Е., Кузьмина С. М.* (1971). О том, как проверялся вопросник по произношению // Развитие фонетики современного русского языка. М.
4. БАС (1948–1960). Словарь современного русского литературного языка: В 17 томах. Москва-Ленинград.
5. БАС (2006) — Большой академический словарь русского языка. т.5 (ред. К. С. Горбачевич) СПб.
6. *Бонч-Осмоловская А. А.* (2006). Практика против грамматики: «одеть» и «надеть» в современном русском языке // Собрание сочинений: к шестидесятилетию Льва Иосифовича Соболева. Москва.
7. *Вербицкая Л. А.* (1976). Русская орфоэпия. (К проблеме экспериментально-фонетического исследования особенностей современной произносительной нормы), Ленинград.
8. *Высотский С. С.* (1984) О московском народном говоре // Городское просторечие. Проблемы изучения. Москва.
9. *Горбачевич К. С.* (1973). Трудности словоупотребления и варианты норм русского литературного языка. Словарь-справочник, Л.
10. *Добрушина Е. Р.* (2013) Между нормой-интуицией и нормой-кодификацией или Двести лет вместе с ихней. // Русский язык в научном освещении. 2013. № 2(26). :181–204.
11. *Зализняк А. А.* (1977). Грамматический словарь русского языка. Москва.
12. *ККК* (2012)— *Каленчук М. Л., Касаткин Л. Л., Касаткина Р. Ф.* Большой орфоэпический словарь русского языка. Москва.
13. *Крысин Л. П.* (2004) Языковая норма в проекции на современную речевую практику, Русский язык сегодня, № 4, Москва.
14. МАС (1957–1961)— *Евгеньева А. П.* (ред.). Словарь русского языка: В 4-х томах. Москва.
15. *Огиенко И. И.* (1914) Русское литературное ударение (справочник и словарь русского ударения), Киев.Русский язык по данным массового обследования. *Крысин Л. П.* (ред) Москва.
16. *Ожегов С. И., Шведова Н. Ю.* (1972) Словарь русского языка, Москва.
17. ОС (1983) — *Аванесов Р. А.* (ред.) Орфоэпический словарь русского языка. Произношение. Ударение. Грамматические формы. Москва.
18. *Перова Д. М., Бондаренко К. Е., Добрушина Н. Р.* (2016) База данных для исследования вариативности твердых/мягких согласных перед е в заимствованных словах // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной международной конференции «Диалог» 2016. Вып. 15, Москва.
19. *Пешковский А. М.* (1959) Объективная и нормативная точка зрения на язык. *Пешковский А. М.* Избранные труды. М.

20. РПП — Введенская Л. А., Червинский П. П. (1996). Русское произношение и правописание. Словарь-справочник. Феникс, Ростов-на-Дону.
21. Словарь 1847 — Востоков (ред.) (1847) Словарь церковнославянского и русского языка, составленный Вторым отделением Императорской академии наук, Типография Императорской академии наук, СПб.
22. Ушаков — Толковый словарь русского языка. Под ред. Д. Н. Ушакова. Москва, 1935–1940. (4 т.).
23. Шахматов А. А. (1899) Несколько слов по поводу записки И. Х. Пахмана, План нового Академического словаря с точки зрения иноязычного, Записка И. Х. Пахмана, Сб. ОРЯС ИАН LXVII (1).
24. *Marklund Sharapova Elisabeth* (2000) Implicit and Explicit Norm in Contemporary Russian Verbal Stress, *Studia Slavica Upsaliensia*. 40, Uppsala.

## References

1. *Avanesov R. I.* (1984), Standard Russian pronunciation [Russkoe literaturnoe proiznoshenie]. Moscow.
2. *Avanesov R. I.* (ed.) (1983) Orthoepic dictionary of Russian. Pronunciation. Stress. Inflection [Orfoepicheskii slovar' russkogo iazyka. Proiznoshenie. Udarenie. Grammaticheskie formy], Moscow.
3. *Avanesov R. I., Ozhegov S. I.* (1959) Standard Russian pronunciation and stress [Russkoe literaturnoe proiznoshenie i udarenie].
4. *Barinova G. A., Il'ina N. E., Kuz'mina S. M.* (1971) How the pronunciation questionnaire was checked [O tom, kak proverialsia voprosnik po proiznosheniju] Evolution of modern Russian phonetics [Razvitie fonetiki sovremennogo russkogo jazyka], Moscow.
5. BAS (1948–1960) — Dictionary of modern standard Russian: in 17 volumes [Slovar' sovremennogo russkogo literaturnogo jazyka: V 17 tomakh]. Moscow—Leningrad.
6. BAS (2006) — The Great Academic Dictionary of Russian. V.5. SPb [Bolshoj akademicheskij slovar' russkogo jazyka. T. 5 (Gorbachevich K. S. ed.) SPb].
7. *Bonch-Osmolovskaja A. A.* (2006). Usage vs. grammar: «odet'» and «nadet'» in modern Russian [Praktika protiv grammatiki: «odet'» i «nadet'» v sovremennom russkom jazyke], Collected Works: In honor of Lev Iosifovich Sobolev [Sobranie sochinenij: k shestidesiatiletiju L'va Iosifovicha Soboleva], Moscow.
8. *Gorbachevich K. S.* (1973) Difficult cases of word usage and competing norms in standard Russian. A reference dictionary [Trudnosti slovoupotreblenija i varianty norm russkogo literaturnogo iazyka. Slovar'-spravochnik], Leningrad.
9. *Dobrushina E. R.* (2013) Intuitive norm vs. codifying norm, or Two hundred years together with ихний [Mezhdju normoj-intuiciej i normoj-kodifikacij ili Dvesti let vmeste s ihnij]. *Russkij jazyk v nauchnom osveshhenii* — [Russian language and linguistic theory], 2013. no 2 (26), pp. 181–204.
10. *Kalenchuk M. L., Kasatkin L. L., Kasatkina R. F.* (2012). A comprehensive dictionary of Russian orthoepics [Bol'shoi orfoepicheskii slovar' russkogo jazyka], AST — Press, Moscow.



11. *Krysin L. P.* (2004). Language norm as projected on modern language usage [Jazykovaja norma v proektsii na sovremennuiu recheviju praktiku], Russian language today [Russkii jazyk segodnia], no 4.
12. *Marklund Sharapova Elisabeth* (2000) Implicit and Explicit Norm in Contemporary Russian Verbal Stress, *Studia Slavica Upsaliensia*. 40, Uppsala.
13. *MAS* (1985–1988), Dictionary of Russian, in 4 volumes Evgen'eva A.P. (ed.) [Slovar' russkogo jazyka v 4 tomakh], Russkii jazyk, Moscow.
14. *Ogienko I. I.* (1914) Stress in standard Russian (reference dictionary of Russian stress) [Russkoe literaturnoe udarenie (spravochnik i slovar' russkogo udarenija)]. Kiev.
15. *Ozhegov S. I., Shvedova N. Yu.* (1972), Dictionary of Russian. Moscow. [Словарь русского языка] 1972.
16. *Perova D. M., Bondarenko K. E., Dobrushina N. R.* (2016). A database for variational studies of palatalized/velarized consonants preceding [e] in loanword [Baza dannykh dlia issledovanija variativnosti tverdykh/miagkikh soglasnykhpered e v zaimstvovannykh slovakh], Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2016” [Komp'yuternaja Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoi Konferentsii “Dialog 2016”], Vol. 15, Moscow. pp. 528–539.
17. *Peshkovski A. M.* (1959) Objective and prescriptive view on language [Ob'ektivnaja i normativnaja tochka zrenija na jazyk], Peshkovski A. M. Selected works [Peshkovskii A. M. Izbrannye trudy], Moscow.
18. *Shakhmatov A. A.* (1899) A few words on I.Kh. Pahman's notes. Plan of a new academic vocabulary from the point of view of a foreign [Neskol'ko slov po povodu zapiski I.Kh. Pakhmana, Plan novogo Akademicheskogo slovaria s točki zrenija inojazychnogo], Note by I. H. Pahman [Zapiska I. Kh. Pakhmana], Coll. ORYAS IAN LXVII (1).
19. *Ushakov D. N.* (ed.) (1935–1940) Dictionary of Russian (Tolkovy slovar' russkogo yazyka).
20. *Verbitskaia L. A.* (1976) Russian orthoepy (Towards experimental phonetic studies of the modern pronunciation standard) [Russkaia orfoepia (K probleme eksperimental'no-foneticheskogo issledovaniia osobennostei sovremennoi proiznositel'noi normy)]. Leningrad.
21. *Vostokov A. Kh.* (ed.) (1847) Dictionary of Church Slavonic and Russian, compiled by the second department of the Imperial Academy of Sciences [Slovar' tserkovnoslavianskogo i russkogo iazyka, sostavlennyi Vtorym otdeleniem Imperatorskoi akademii nauk]. St. Petersburg.
22. *Vvedenskaja L. A., Chervinsky P. P.* (1996) Russian pronunciation and spelling. A reference dictionary [Russkoe proiznoshenie i pravopisanie. Slovar'-spravochnik]. Feniks, Rostov-na-Donu.
23. *Vysotsky S. S.* (1984) On Moscow vernacular [O moskovskom narodnom govore], Urban vernacular. Issues in empirical research [Gorodskoe prostorechie. Problemy izucheniia], Moscow.
24. *Zaliznyak A. A.* (2016) Russian grammatical dictionary [Grammaticheskij slovar' russkogo jazyka], Russkij jazyk, Moscow.

# ГОВОРЮ Я, КАРЛ...<sup>1</sup>

**Рахилина Е. В.** (rakhilina@gmail.com)

НИУ Высшая Школа Экономики, Институт РЯ РАН,  
Москва, Россия

В статье описывается история употребления русской конструкции с глаголом *говорить*, которая сегодня служит обычным маркером прямой речи говорящего. Исследование проводится с привлечением текстов XIX века, и опирается среди прочего на особым образом размеченный корпус текстов этого периода.

**Ключевые слова:** корпусные исследования, история русского языка, языковые контакты, дискурсивные маркеры, лексическая семантика, грамматикализация

**Rakhilina E. V.** (rakhilina@gmail.com)

National Research University Higher School of Economics,  
Institute for Russian Language / Russian Academy of Sciences,  
Moscow, Russia

The paper describes the diachronic development of Russian construction with the verb *govorit'* 'to say', the main marker introducing the direct speech. The analysis relies heavily on the data from XIX century Russian as presented in a special annotated corpus of Early Modern Russian.

**Key words:** corpus studies, history of Russian, language contacts, discourse markers, lexical semantics, grammaticalization

## 0. Введение

Настоящая работа — одна из первых, построенных на материале пилотного варианта нового размеченного корпуса XIX века, который разрабатывается по проекту «Пограничный русский». Корпус предназначен для исследования микродиахронических явлений — небольших семантических сдвигов, изменений в лексике и словообразовании, управлении, порядке слов и проч., а также в слабо-композициональных и идиоматичных конструкциях разного рода.

---

<sup>1</sup> Работа выполнена при поддержке гранта РНФ 16-18-02071 «Пограничный русский: оценка сложности восприятия русского текста в теоретическом, экспериментальном и статистическом аспектах» (рук. Н. А. Слюсарь).

В заголовке статьи дается отсылка к широко распространенному сейчас интернет-мему, о котором подробнее см. Кронгауз и др. 2016, с. 51.

Тексты в корпусе автоматически размечаются морфологически и затем подвергаются детальной ручной разметке двумя-тремя независимыми разметчиками в отношении отклонений нормы XIX века от современной. В процессе разметки была выработана система помет, которая используется в корпусе в его настоящем варианте. Она дает общее представление о том, какого рода изменения в системе русского языка за его двухсотлетний период превалируют и могли бы считаться особенно значимыми. Всего помет около 30, они отражают параметры, значения которых отклоняются от современной нормы. Приведем здесь несколько иллюстраций:

- **asp (изменения в видовой форме):** *Когда я ему заметил, что он мог бы побеспокоиться в пользу хотя моего чемодана, за которым я вовсе не желал лазить* (→ *лезть*) в эту бездну
- **compag (изменения в форме или конструкции сравнительной степени):** *жизнь моя становится пустее* (→ *все более пустой*) день ото дня
- **deriv (мена словообразовательной модели или морфемы):** *чтоб докончить* (→ *закончить*) портрет, я скажу, что у него был немного вздернутый нос, зубы ослепительной белизны и карие глаза
- **num (изменение числовой формы):** *оно или служит объяснением цели сочинения, или оправданием и ответом на критики* (→ *критику*)
- **wo (изменения в порядке слов):** *Он скептик и матерьялист, как все почти медики* (→ *почти все*) до нравственной цели и до журнальных нападок

Заметим, что язык XIX века не так часто описывался системно: хорошо известны два основополагающих труда, но оба относятся к середине прошлого века (Виноградов, Шведова 1964 и Булаховский 1948) и соответственно, отражают уже отошедший на целый век узус. Кроме того, у авторов того времени (а было и немало более частных исследований, в частности, лексики, и тоже в 60–70-е годы) не было возможности проверить свои гипотезы по такому широкому материалу, который теперь предоставляет исследователю поиск в корпусе, в том числе исторический.

В наших исследованиях соответствие той или иной формы или конструкции XIX-го века сегодняшнему дню отслеживается по данным Национального корпуса русского языка (НКРЯ). С его помощью определяется и степень уникальности замеченного лингвистического эффекта: по корпусу мы проверяем, встречаются ли такого же типа отклонения в других текстах того же или более раннего времени. Таким образом, НКРЯ оказывается для исследований этого рода необходимой точкой отсчета. Его возможности должны быть использованы: данные XX века заново проверены и получены новые — пока мы слишком мало знаем о том, насколько отделилась от нас в отношении языка, например, русская классическая литература. Именно для этого предназначен создаваемый нами дополнительный инструмент, о котором мы говорим: размеченный корпус XIX-го века.

В качестве пилотного текста мы взяли хрестоматийный роман М. Ю. Лермонтова «Герой нашего времени». По общему представлению, он до сих пор хорошо читается, в том числе школьниками — как если бы был написан в наше время. Ожидаемую трудность (которая должна быть в современном издании скомпенсирована историческим комментарием) вызывают характерные для

этого текста военные и специальные термины и историзмы — *кафтан, штабс-капитан, десятина, пыж, разводная ванна, лорнет, слободка* и проч. Они — утраченная нами часть реальности, о которой ведется повествование, их нельзя исключить из романа или «перевести». Другое дело — примеры типа (1):

(1) *Сами посудите, что ж я мог отвечать против этого?*

В них встречается по несколько форм, отличных от современных, но таких, которые можно было бы привести к более привычному нам виду (1'), «исправив»: вопросительную конструкцию *что ж*, устранив из нее частицу → *что*, аспектуальную форму *отвечать* → *ответить* и управление: *ответить против этого* → *ответить на это*. Ср.:

(1') *Сами посудите, что я мог ответить на это?*

Лингвистически содержательной задача такого «перевода» на современный язык становится только тогда, когда мы можем исследовать причины замеченных нами языковых изменений.

Первая гипотеза, которую мы каждый раз проверяем — это «след» калькирования с французского. Действительно, языковая ситуация того времени располагает именно к этой логике: как известно, в начале века интерференция русского с французским была слишком сильна, а в последующем ослабла. Радикальное изменение социолингвистической ситуации могло послужить естественной причиной расхождения между русским того времени и современным русским.

В этой небольшой работе мы рассмотрим некоторые сюжеты, связанные именно с идеей французского влияния на русский первой трети XIX века.

## 1. Кальки

В естественном и очень живом для современного русского читателя дискурсе романа кальки, тем не менее, встречаются. Прежде всего, они затрагивают синтаксические конструкции и порядок слов<sup>2</sup>.

В этом отношении показателен пример (2):

(2) *Я слез и подкрался к окну: неплотно притворенный ставень позволил мне видеть пирующих...*

Такого рода конструкции (с неодушевленным субъектом при модальном глаголе *позволить* и дательным падежом субъекта подчиненной ситуации) с точки зрения современной русской нормы выглядят странно. В обычном частотном случае субъектом при *позволить* оказывается абстрактное имя, а не неодушевленный физический предмет (*помощь, прибыль, проект, развитие* и под.) или лицо / группа лиц (*его превосходительство, общество* и проч.).

---

<sup>2</sup> Но также и собственно лексику, где, по словам В. М. Живова, «семантическое калькирование сочетается с переработкой семантики, унаследованной от предшествующих эпох» (Живов 2009, с.19), о которой мы здесь специально говорить не будем.

Для неодушевленного неактивного субъекта в этом случае в русском лучше безличная конструкция, как в (3):

- (3) *Сквозь неплотно притворенный ставень я мог видеть VS. можно было увидеть VS. было видно пирующих...*

Напротив, во французском, и в том числе французском того времени, в аналогичных контекстах встречается ровно такое же распределение синтаксических ролей, как у Лермонтова. Хорошим свидетельством этому служит пример (4):

- (4) *Fr. Une autre porte entrebaillée me laissa voir un vieillard une Bible à la main ...*  
 [« La France en Afrique et l’Orient à Paris » Benjamin Gastineau Paris, G. Barba 1864 p. 9]

Что касается порядка слов, то это тоже тот фрагмент синтаксического поля, в котором кальки вполне ожидаемы. С этой точки зрения любопытны атрибутивные и адвербиальные конструкции, ср., например, (5) с постпозицией одного из прилагательных в цепи однородных членов:

- (5) *Со всех сторон горы неприступные, красноватые скалы, обвешанные зеленым плющом и увенчанные купами чинар, желтые обрывы <...>*

В современном дискурсе такая постпозиция является признаком сказового устаревшего высокого стиля, ср. (6) — однако тогда этот стиль выдерживается во всем фрагменте текста.

- (6) *И взошел я на гору высокую, где ни солнце не сияет, ни древа нет, ни трава не растет, только гады и змеи, свищущие и скрежещущие зубами.* [Михаил Шишкин. Венерин волос (2004) // «Знамя», 2005]

В контексте (5) прямой порядок в атрибутивной конструкции выглядит так же естественно, как и инверсный — совершенно нейтрально. Для Лермонтова такой порядок не служит никаким стилистическим маркером, он индуцирован французским и может считаться синтаксической калькой (Булаховский 1948, Успенский 1985 и др.).

Заметим, что постпозиция прилагательных смотрится более системно, если принять во внимание многочисленные (и неестественные в современном русском) случаи постпозиции наречия в адъективных конструкциях, которые широко представлены в нашем корпусе, ср., например, (7–10):

- (7) *Тут толпилось шумно десятка два грузин и горцев; поблизости караван верблюдов остановился для ночлега.* [→ *шумно толпились*]
- (8) *...и скоро чайник зашипел приветливо* [→ *приветливо зашипел*]
- (9) *...и я боялся в ней исчезнуть совершенно* [→ *совершенно исчезнуть*]
- (10) *Вы ошибаетесь опять: я вовсе не гастроном* [→ *опять ошибаетесь*]

## 2. Не-кальки

Не все расхождения в синтаксической структуре текстов XIX и XXI века можно объяснить калькированием. В частности, как хорошо известно, существенную роль в архаизации текста играет постпозиция местоимений, выражающих субъект глагола или посессора, как в (11–13):

(11) *Присел я у забора и стал прислушиваться, стараясь не пропустить ни одного слова* [→ *Я присел ... и стал прислушиваться*]

(12) *пополз я по густой траве вдоль по оврагу* [→ *Я пополз ... вдоль по оврагу*]

(13) *Исполнив долг свой, сел я к нему на кровать* [→ *Исполнив свой долг, я сел к нему на кровать*]

Этот эффект объясняется постепенным превращением этих форм в русском языке из безударных клитик в полноударные лексические единицы [Зализняк 2008]. В XXI веке этот процесс давно завершился, и такие примеры возможны только в стилизациях, но в первой трети XIX-го века он еще полностью не закончился, так что в «Герое нашего времени» такие примеры все еще встречаются во множестве.

Именно с этой точки зрения нужно рассматривать встречающиеся в тексте Лермонтова случаи так называемого Pro-drop.

Pro-drop'ом принято называть явление опущения приглагольных субъектных местоимений, ср. *пришел, увидел, победил VS. он пришел, он увидел, он победил* (ср., например, [Тестелец 2001: 287]). Языки, которые допускают такое опущение, называются Pro-drop языками. К ним, в чисел прочих, относится русский — в отличие от французского и многих других европейских, ср.:

(14) Fr. *J'arrive* ~ русск. *Иду (сюда)*.

Понятно, что на самом деле в каждом конкретном языке есть специальные условия и ограничения, регулирующие Pro-drop, так что даже в Pro-drop-языках опущение возможно далеко не в любом контексте (подробнее см., например, Фужерон, Брейар 2004). Тем не менее, если говорить в самом общем виде, то современный русский и французский в этом отношении хорошо противопоставлены. Но тогда мы вполне могли бы ожидать здесь какие-то рефлексии интерференции русского и французского в прошлом, два века назад. Между тем анализ Лермонтовской прозы показывает, что современный русский, как ни странно, много ближе к французскому, чем язык Лермонтова, в котором опущения встречаются гораздо чаще, чем сегодня, ср. (15–16):

(15) *<...> и все грустила, напевала свои песни вполголоса, так что, бывало, и мне становилось грустно, когда слушал ее из соседней комнаты.*  
[→ *когда я слушал*]

(16) *Поверь мне, аллах для всех племен один и тот же, и если он мне позволяет любить тебя, отчего же запретит тебе платить мне взаимностью?»*  
[→ *отчего же он запретит*]

Объяснение этому довольно простое, оно связано с тем же эффектом исчезновения класса клитических местоимений, который мы упоминали выше.

### 3. Особый случай: все-таки кальки?

Убедившись на только что разобранных примерах в сложности задачи, связанной с определением калькирования, обратимся к следующему абзацу текста романа:

- (17) *Наконец — уж бог знает откуда он явился, только не из окна, потому что оно не отворялось, а должно быть, он вышел в стеклянную дверь, что за колонной, — наконец, говорю я, видим мы, сходит кто-то с балкона...*

В этом абзаце много несоответствий современной норме и интересных с лингвистической точки зрения явлений, но следуя тематике статьи, мы выделим последний фрагмент:

*наконец, говорю я, видим мы, сходит кто-то с балкона...*

и в нем обратим внимание на любопытную дискурсивную вставку *говорю я* с инверсным порядком слов.

В русском языке похожие маркеры очень частотны и имеют разную форму и семантику<sup>3</sup>. Во-первых, *я говорю* — прежде всего, с прямым порядком слов, которое вводит прямую или косвенную речь. С местоимением первого лица как субъекта — это речь говорящего, «авто-пересказ» (или, скорее, попытка точного воспроизведения) сказанного им ранее. Такой маркер необходим, когда происходит пересказ диалога, ср.:

- (18) *«Что, я говорю, нравится ли он тебе?.. да говори правду...» Сначала замялась, ну, а потом призналась, что ты ей очень нравишься. Что же, ты бы уж кончал это дело... Зачем в долгий ящик откладывать...*  
[И. И. Панаев. Опыт о хлыщах (1854–1857)]

В НКРЯ это значение зафиксировано для прямого порядка слов прямо с начала этой конструкции, с XVIII века, и живо до сих пор.

Однако в современном русском для этого маркера есть еще одно характерное значение, которое можно признать производным. Можно считать, что тот же дискурсивный показатель претерпевает небольшой семантический сдвиг — и вводит фрагмент монолога, а не диалога. В таком случае, как правило, он маркирует не буквальное воспроизведение уже произнесенного говорящим текста, а так сказать, его «краткий реферат». Этот «реферат» служит напоминанием со стороны говорящего другим участникам речевого акта о том, что уже было им сказано — когда основной рассказ был почему-то прерван, и нить его может быть упущена. Употребив этот дискурсивный маркер, говорящий возвращает нас к исходной теме, ср. (19).

<sup>3</sup> Во всем разнообразии фразеологизованные конструкции современного русского языка с глаголами говорения представлены в статье Безяева 1997.

(19) *Сам, грешным делом, люблю поспать в рабочее время. Особенно на учёных советах. Золотой сон! — Виноват, товарищ генерал! — А вы не стесняйтесь. Так вот, я говорю, спать на учёном совете — самое милое дело! Только не надо распускаться: носом клевать, изо рта пузыри пускать и так далее. [И. Грекова. На испытаниях (1967)]*

Приглагольная позиция местоимения, как мы уже обсудили в предыдущем разделе, может меняться в зависимости от временного периода, который представляет данный пример. В XIX веке местоимение часто находится в постпозиции к глаголу, поэтому мы могли бы ожидать, что найденный нами пример (17) попадал в один из этих классов.

Между тем, ни та, ни другая интерпретация *я говорю* для примера (17) не подходит. (17) описывает не диалог, а монолог Грушницкого — но этот монолог никто не прерывал. *Говорю я* стоит в самом центре рассказа Грушницкого, в самой его кульминационной, захватывающей точке. По своей роли в дискурсе он может быть отнесен к маркерам поддержания контакта со слушающими, предназначенного выделить, усилить данный момент повествования. В современном дискурсе вставка *говорю я* смотрелась бы в этом контексте неестественно и должна была бы быть заменена на какой-то другой маркер, что-то вроде: *представляете?! —* или, более лаконично и современно: *Карл!*

Примеры употребления *говорю я* именно этой семантики встречаются в начале XIX века, ср., например, (20), в котором говорящий действительно хочет специально акцентировать на внимание на том, что *теперь* все действительно *иначе* [да, Карл]:

(20) *Куда ж прежде девалось? И сам не знаю, а чувствую, что его нет более, и уверен, что, увидя предмет, при воспоминании о котором прежде душа моя пылала сладостным огнем, и все бытие перерождалось, — теперь, говорю я, не иначе взгляну на предмет тот, как на листы бумаги, на которых красным карандашом пачкал я харицы в первые месяцы ученичества <...> [В. Т. Нарезный. Российский Жилблз, или Похождения князя Гаврилы Симоновича Чистякова (1814)]*

Наша гипотеза состоит в том, что такое употребление является (неполной) калькой с французского того времени. Некоторые примеры об этом свидетельствуют, ср. (21):

(21) *M. Jules vit entrer dans son cabinet Fouquereau tout radieux. « Monsieur, il vient de venir une vieille femme, mais soignée, je dis, une fine mouche. Elle a demandé Monsieur, a paru contrariée de ne point le trouver, et m'a donné pour Madame une lettre que voici ».*

‘<...> к Жюлю в кабинет вошёл сияющий Фукеро. — Сударь, только что приходила старуха, гладкая такая. Продувная бестия, доложу я вам. Спрашивала вас и словно бы огорчилась, что вас дома нет, а барыне велела вот это письмо передать.’ [Бальзак, «Феррагус»; перевод М. Казас]

Особенность этого случая в том, что в современном французском так сказать нельзя, но и в современном русском, как мы знаем, это значение тоже



утрачено. Требуется дополнительное исследование, чтобы доказать это обстоятельно. Пока мы можем рассматривать это как правдоподобное предположение, которое, если бы не оправдалось, следовало бы выдумать: французское влияние на русский язык после первой трети девятнадцатого века заметно идет на убыль (Булаховский 1948, Успенский 1985 и др.). Судьба устаревших калек (а такие непременно должны быть) чрезвычайно интересна с лингвистической точки зрения.

#### 4. О дальнейшем развитии

Любопытно, что та же семантика встречается в первой трети XIX века у другого, модифицированного маркера, а именно: **я вам говорю** — с препозицией местоимения и выраженным адресатом, ср. (22). Как и в монологе Грушницкого (17), с которого мы начали свое обсуждение, здесь речь идет о дополнительном дискурсивном знаке усиления контакта с собеседником через привлечение его особенного внимания к тому, что и как было сказано в данный момент. Ср. также (23–24).

(22) Явление XIV. Прежние и Алинской.

[Александр Иванович Алинской, муж] (в некотором замешательстве).

*Я никак не мог уехать без того, чтобы не засвидетельствовать вам лично моей благодарности... и не извиниться перед вами.*

[Константин Прохорович Яззов, муж] (весело и дружески) *Руку, сосед!*

*Что тут за благодарность и извинения. Мы давеча вас насмешили, я вам говорю, не правда ли?* [Д. Т. Ленский. Хороша и дурна, и глупа и умна (1833)]

(23) — *Удивительное, я тебе говорю, стечение обстоятельств!..*

[А. Ф. Писемский. Масоны (1880)] **ЗНАЕШЬ ЛИ**

(24) *Это делается для того только, я тебе говорю, чтобы избежать огласки.*

[Д. Т. Ленский. Хороша и дурна, и глупа и умна (1833)]

В современном русском *я тебе говорю* существует как дискурсивный маркер, но он несколько изменил и свое дискурсивное значение, и область употребления: *я (тебе / вам) говорю* хорошо сохранилось в императиве, как своего рода маркер «строгости» приказа, ср. (25–27):

(25) *Горкин ему строго говорит: «Вася, я тебе говорю усни!»* [И. С. Шмелев. Лето Господне (1927–1944)]

(26) — *Нет, Ася, — уже строго сказала мать, я тебе говорю, совершенно одинаковые.* [М. И. Цветаева. Сказка матери (1934)]

(27) — *Вставай, я тебе говорю, ишак кавказский, идиот осетинский!* [А. И. Куприн. Яма (1909–1915)]

Как правило, в императиве строгость приказания достигается за счет настойчивости говорящего, апеллирующего к тому, что он повторяет то же требование, которое уже было выдвинуто им некоторое время назад, но до сих пор не выполнено. Идея настойчивости и связанного с ним императивного повтора прослеживается в семантике этого маркера и в некоторых косвенных приказа-ниях / просьбах, не собственно императивных, как в (28) или (29):

(28) — *Врешь, я тебе говорю, мерзавец! Ты Ксенофонт!* [А. А. Шкляревский. Русский Тичборн (Из уголовной хроники) (1903)]

(29) *Это Саша, я вам говорю, это Саша!* ≈ ‘<обязательно> поверьте, я повторяю, что это Саша (а не кто-то другой)’ [Вера Белоусова. Второй выстрел (2000)]

Для этого вида императивов следовало бы ввести особую категорию «стратива», который, впрочем, в широком смысле, входит в тот же класс дискурсивных маркеров, что и *Карл!* — но только с поправкой на условия императивного контекста, в котором настойчивость часто сопряжена с повтором. Другими словами, мы можем считать, что здесь действует, в терминологии Construction Grammar, известный эффект *coercion*: воздействия структуры и семантики конструкции (в нашем случае — императива) на встраивающуюся в нее языковую единицу (дискурсивный маркер).

## 5. Заключение

Мы рассматриваем новый пилотный корпус XIX века как лингвистический инструмент, с помощью которого можно исследовать новую область — микродиакронию. Она даст нам возможность увидеть новые задачи через сопоставление двух языковых норм, разделенных двумя веками. В том сюжете, который мы здесь рассказали, есть две важные для нас перспективы. Первая — сложнейшее взаимодействие русского и французского того времени, которое требует тщательного и профессионального исследования. Вторая — конструкции с непростой и некомпозиционной семантикой, семантическая сложность которых становится очевидна только тогда, когда нам как читателям или как лингвистам бросается в глаза неполное их соответствие нашей современной норме. Действительно, фразы *я (тебе) говорю, говорю я*, давно и хорошо освоенные лингвистами глаголы речи, смотрятся настолько обыденно, что требуется специальный инструмент, чтобы вскрыть их нетривиальность.

## Литература

1. *Безяева М. Г.* «Как вам сказать?» Об участии конструкций с глаголами речи в формировании коммуникативной целеустановки высказывания // Вестник Московского университета. Серия 9: Филология. 1997, № 2. С. 62–87.
2. *Булаховский Л. А.* Русский литературный язык первой половины XIX века. Киев, 1948.

3. *Виноградов В. В., Шведова Н. Ю.* Очерки по исторической грамматике русского литературного языка XIX века. М.: Наука, 1964.
4. *Живов В. М.* История понятий, история культуры, история общества // В. М. Живов (ред.) Очерки исторической семантики русского языка раннего нового времени. М.: Языки славянских культур, 2009.
5. *Зализняк А. А.* Древнерусские энклитики. М.: Языки славянских культур, 2008.
6. *Кронгауз М. А., Литвин Е. А., Мерзлякова В. Н., Пиперски А. Ч., Сомин А. А., Черненко Ю. А.* Словарь языка интернета. М., АСТ-ПРЕСС, 2016.
7. *Лермонтов М. Ю.* Герой нашего времени. М.: АСТ, 2015.
8. *Рахилина Е. В., Бородин М. А., Резникова Т. И.* «Тамань сегодня»: корпусное исследование русского языка XIX века // Труды Института русского языка им. В. В. Виноградова, Вып. 11. М., 2016.
9. *Тестелец Я. Г.* Введение в общий синтаксис. М.: РГГУ, 2001.
10. *Успенский Б. А.* Краткий очерк истории русского литературного языка (XI–XIX вв.). М. Гнозис. Серия: Язык. Семиотика. Культура. 1994.
11. *Фужерон, И. И.; Брейар, Ж.* 2004. Местоимение «я» и построение дискурсивных связей в современном русском языке // Т. М. Николаева (ред.). Вербальная и невербальная опоры пространства межфразовых связей. М.: Языки славянской культуры, 147–165.

# РЕВЕРСИВНЫЙ ПЕРЕВОД КАК ИНСТРУМЕНТ ЛИНГВИСТИЧЕСКОГО АНАЛИЗА ДИСКУРСИВНЫХ СЛОВ<sup>1</sup>

**Шмелев А. Д.** (shmelev.alexei@gmail.com)

Московский педагогический государственный университет;  
Институт русского языка им. В. В. Виноградова РАН;  
Православный Свято-Тихоновский гуманитарный  
университет, Москва, Россия

**Зализняк Анна А.** (anna.zalizniak@gmail.com)

Институт языкознания РАН; Институт проблем  
информатики ФИЦ ИУ РАН, Москва, Россия

## REVERSE TRANSLATION AS A TOOL FOR ANALYSIS OF DISCOURSE WORDS

**Shmelev A. D.** (shmelev.alexei@gmail.com)

Moscow Pedagogical State University; Vinogradov Institute  
of Russian Language; Russian Academy of Sciences;  
St Tikhon's Orthodox University, Moscow, Russia

**Zalizniak Anna A.** (anna.zalizniak@gmail.com)

Institute of Linguistics, RAS; Institute of Informatics Problems  
FRC CSC RAS, Moscow, Russia

The paper demonstrates the possibilities of “one-direction analysis” in contrastive studies based on the parallel corpora. We assume that one may regard translation equivalents and paraphrases of a linguistic unit extracted from real translated texts as a source of information about its semantics; translations into Russian may be even more revealing in this respect. Our methods are based on the above hypothesis: we correlate the explications of the discourse words given in earlier studies with the “stimuli to translation”, that is, fragments of the original text that may cause the appearance of these discourse words in a Russian translation as a “reaction” to those “stimuli”. Using this methodology, we seek to validate, disprove or improve the semantic analysis of these words made without recourse to electronic corpora. The results of the analysis of the Russian discourse words *eshche*, *vidimo*, *po-vidimomu* and *vidno* are set forth.

**Key words:** Russian language, semantics, discourse markers, corpus linguistics, parallel corpora, translation, reverse translation

---

<sup>1</sup> Работа выполнена при финансовой поддержке РФФИ, грант №16-06-00339.

## 1. Унидирекционный подход в контрастивных исследованиях

Унидирекционный подход к контрастивным корпусным исследованиям был разработан авторами в рамках проекта «Контрастивное корпусное исследование специфических черт семантической системы русского языка» (см. Зализняк 2015, Шмелев 2015<sup>2</sup>). Он состоит в том, что множество переводных эквивалентов высказываний, содержащих анализируемую языковую единицу, рассматривается как источник информации о ее семантике и прагматике (ср. подход, развиваемый в Anderman, Rogers 2008). Соответственно, перевод используется как инструмент лингвистического анализа, целью которого является уточнение семантического анализа языковой единицы одного (в данном случае — русского) языка и, в частности, выявление степени ее лингвоспецифичности. «Унидирекционным» данный подход назван потому, что в обоих направлениях перевода предметом нашего анализа является только один из двух языков (в нашем случае — русский). Целью контрастивного анализа текстов оригинала и перевода является, таким образом, не установление типологических или каких-то иных сходств и различий, т. е. не собственно сравнение, а семантический анализ тех или иных единиц одного из двух языков<sup>3</sup>.

Наиболее информативным оказывается «реверсивный»<sup>4</sup> перевод, т. е. перевод с иностранного языка на русский, и в особенности тот случай, когда в оригинальном иностранном тексте отсутствует конвенциональный эквивалент некоторой единицы русского языка, появляющейся в переводе. Поскольку дискурсивные слова по большей части в высокой степени лингвоспецифичны (и, в частности, как правило, не имеют конвенционального эквивалента в других языках), подобная ситуация довольно типична<sup>5</sup>. Методология унидирекционного контрастивного анализа состоит в выявлении тех семантических и прагматических признаков высказывания на иностранном языке, которые послужили «стимулом»<sup>6</sup> появления анализируемого дискурсивного слова в русском переводе и, далее, в сопоставлении толкований дискурсивных слов, предложенных ранее, с полученными результатами, т. е. фрагментами текста иноязычного оригинала, «реакцией» на которые является появление в русском

<sup>2</sup> В указанных работах описаны также возможности применения данного подхода к анализу русских лингвоспецифичных единиц.

<sup>3</sup> Ср. принцип ‘translation as evidence for semantics’, предложенный в статье Noël 2003, и его применение к исследованию английского дискурсивного слова *Oh* в Aijmer 2008.

<sup>4</sup> Термин «реверсивный перевод» предложен Н. В. Бунтман в ходе работы над проектом «Контрастивное корпусное исследование дискурсивных слов русского языка».

<sup>5</sup> Так, в Базе данных дискурсивных единиц русского языка и их французских эквивалентов, создаваемой на основе французского параллельного подкорпуса НКРЯ под руководством одного из авторов (см. Зализняк 2016), в реверсивных парах отсутствие какого-либо «стимула перевода» отмечено в 13,3% случаев.

<sup>6</sup> Термин «стимул перевода» в том значении, в котором он используется в контрастивных корпусных исследованиях, был введен в Loiseau et al. 2013.

переводном тексте данных дискурсивных слов, — что позволяет верифицировать и уточнить эти толкования.

При этом различаются случаи, когда при переводе с иностранного языка на русский появление дискурсивного слова обязательно и когда употребление в русском переводе того или иного дискурсивного слова представляет собою индивидуальное решение переводчика. В первом случае необходимо выявить те признаки русского предложения, которые обуславливают обязательность употребления в нем данного дискурсивного слова для выражения данного смысла, а втором — те признаки иноязычного текста, которые могли побудить переводчика употребить данное дискурсивное слово.

В результате проведенного исследования удалось уточнить семантический анализ ряда русских дискурсивных слов, предложенный в «докорпусную» эпоху: подтвердить или опровергнуть наличие тех или иных семантических компонентов и условий употребления, а также обнаружить новые. В данной статье результаты исследования демонстрируются на материале русских дискурсивных слов *еще* (в одном из значений), *видимо*, *по-видимому* и *видно*.

## 2. Прагматическая обязательность дискурсивного слова: *еще*

В статье Levontina, Shmelev 2005 было введено понятие «прагматической обязательности» частицы. Имеется в виду ситуация, когда отсутствие некоторой дискурсивной единицы, не нарушая грамматической правильности предложения, порождает ложную импликацию или нарушает связность текста. Так, описанная в указанной статье частица *еще* является прагматически обязательной в одном из ее значений, реализуемом в контексте будущего времени: *Я еще вернусь*. В зависимости от контекста это значение может реализоваться в качестве угрозы (ср. *Ты у меня еще попляшешь!*; *Ты еще об этом пожалеешь*) или, наоборот, утешения (ср. *Мы еще повоюем*; *Ты еще всех нас переживешь*; *Все еще наладится*).

Значение, которое вносится частицей *еще*, включает следующие компоненты:

1. Имеется противоположная гипотеза о том, как будут развиваться события; во многих случаях фраза, содержащая частицу *еще*, начинается с союза *но*: *Но вы меня еще узнаете*; *Но мы еще дойдем до Ганга*.
2. Событие, о котором идет речь, произойдет не непосредственно после момента речи: *Мы еще повоюем* — значит, сейчас дела плохи. *Мы еще поговорим об этом* — часто используется для прекращения разговора на данную тему.
3. В момент речи ситуация развивается таким образом, что прогноз никак из нее не вытекает, ср.: *Заканчивалась поэма в том смысле, что, дескать, вы победили, но мы еще отомстим* (В. Ходасевич, Неудачники)

Использование частицы *еще* обязательно или почти обязательно, если человек вынужден отказаться на данный момент от совершения какого-то действия (или ему не удалось осуществить желаемое) и утешается тем, что это

не окончательно и когда-нибудь в неопределенном, пусть не близком, будущем ситуация изменится. Например, если сорвалась поездка, можно сказать: *Да ты еще 10 раз съездишь на Кипр*. Утешая человека, пережившего любовное разочарование, говорят: *Ты еще встретишь настоящую любовь* или *Да она еще сама к тебе прибежит*, и т. п.

С другой стороны, еще невозможно, если в данном высказывании формулируется условие, при котором данная ситуация будет иметь место, так как это входит в противоречие с вводимым еще смыслом 'в данный момент ничто не предвещает наступление данной ситуации': *Стучите, и (\*еще) отворят вам; Позвоните, и я (\*еще) приеду; Если что, я (\*еще) помогу; Жди меня, и я (\*еще) вернусь*.

## 2.1. Корпусные данные

Анализ данных реверсивного перевода на материале английского параллельного подкорпуса НКРЯ позволяет выделить следующие «стимулы» для появления *еще* в переводе:

- *but* (указывающее на наличие противоположной гипотезы о том, как будут развиваться события); ср. примеры (1), (6);
- *later, wait, tomorrow, next time* (событие, о котором идет речь, произойдет не непосредственно после момента речи); ср. примеры (1)–(5), (7), (8);
- признаки, свидетельствующие о том, что ситуация развивается таким образом, что прогноз никак из нее не вытекает; ср. примеры (4), (5), (7)–(9);
- *warn*; угроза; ср. пример (6);
- утешение; ср. примеры (3), (4), (9).

В приводимых ниже примерах фрагмент английского предложения, представляющий собой «стимул» появления слова *еще* в русском переводе, а также его русский эквивалент в русском переводе, выделены подчеркиванием.

- (1) Very well. But *we're going to talk* a lot more about this, Leah. Tomorrow... [Miranda Lee. *Fugitive Bride* (1998)] — Хорошо. Но мы еще поговорим обо всем подробнее, Лия. Завтра... [Миранда Ли. В любви все средства хороши (М. Авдокушина, 2000)]
- (2) But did he send it to me for its pictures of women naked except for black garterbelts, or for a remarkable science fiction story in there, *The Protocols of the Elders of Tralfamadore*? But more about that later. [Kurt Vonnegut. *Hocus Pocus* (1990)] — Но вот чего ради он прислал этот номер — чтобы я разлакомился, глядя на девочек, одетых только в черные пояски для подвязок, или чтобы я прочел потрясный научно-фантастический рассказ, который назывался Протоколы Тральфамадорских Мудрецов? Но к этому я еще вернусь. [Курт Воннегут. Фокус-покус (М. Ковалева, 1993)]
- (3) ...you wait, you'll be right famous at Hogwarts. [Joanne Kathleen Rowling. *Harry Potter and the Sorcerer's Stone* (1997)] — ...погоди, еще будешь гордостью «Хогварца». [Дж. К. Роулинг. Гарри Поттер и Волшебный камень (М. Спивак, 2001)]

- (4) “Didn’t you enjoy it?” “A little. In between terror.” “You’ll have less terror next time, and still less the time after, and eventually you’ll just experience the enjoyment and *I’ll make a racer out of you.*” [Isaac Asimov. The Gods Themselves (1972)] — Разве вы не испытывали удовольствия? — Иногда. В промежутках между припадками ужаса. — Ну, так в следующий раз припадков ужаса будет меньше, а потом еще меньше, и в конце концов останется одно удовольствие. *Я еще сделаю из вас чемпиона.* [Айзек Азимов. Сами боги (Р. Рыбакова, 1975)]
- (5) “Why did you go up there at six o’clock in the morning?” “Because I wanted to get into the office. I found out that the cleaning woman started on Mr. Mason’s office at six o’clock in the morning.” “We’ll go into some of this later and in private,” Hamilton Burger said grimly. “What I want to know now is what you did with the gun.” [Erle Stanley Gardner. The Case of the Daring Divorcee (1964)] — Почему вы решили сделать это в шесть часов утра? — Потому что мне нужно было войти в кабинет мистера Мейсона. Я выяснил, что уборку в офисе Мейсона начинают в шесть часов утра. — Хорошо, мы *еще* вернемся к этому вопросу, — угрюмо сказал Бюргер. — Я хочу знать, что вы сделали с револьвером [Эрл Стэнли Гарднер. Дело смелой разведки (М. Кудрявцева, 1990)]
- (6) All right, but I warned you, *you just remember* what I said when you’re on the train home tomorrow, you’re so. [Joanne Kathleen Rowling. Harry Potter and the Sorcerer’s Stone (1997)] Хорошо, но я вас предупредила, вы *еще* вспомните мои слова — в поезде, когда вас отправят домой, какие же вы все-таки. [Дж. К. Роулинг. Гарри Поттер и Волшебный камень (М. Спивак, 2001)]
- (7) “Good-by, Mr. Mason,” she said. “Bye now,” Mason said. “*I’ll be seeing you.*” [Erle Stanley Gardner. The Case of the Daring Divorcee (1964)] — До свидания, мистер Мейсон. — До свидания, — ответил Мейсон. — Мы *еще* встретимся. [Эрл Стэнли Гарднер. Дело смелой разведки (М. Кудрявцева, 1990)]
- (8) Wait a moment! *We shall meet again* soon. [J. R. R. Tolkien. The Lord of the Rings: The Two Towers (1954)] — Подожди, мы *еще* встретимся. [Дж. Р. Р. Толкин. Властелин колец: Две башни (М. Каменкович, В. Каррик, 1994)]
- (9) He tore his eyes away from his mother’s face, whispered, “*I’ll come back,*” and hurried from the room. [Joanne Kathleen Rowling. Harry Potter and the Sorcerer’s Stone (1997)] Он оторвал взгляд от лица матери, прошептал: «я *еще* приду» и поспешил прочь из комнаты. [Дж. К. Роулинг. Гарри Поттер и Волшебный камень (М. Спивак, 2001)]



### 3. Решение переводчика: *видимо, по-видимому, видно*

Перечисленные слова давно обсуждаются в литературе (см. в частности Виноградов 1975, Яковлева 1099, 1994, Иоанесян 1993, Разлогова 1996, 2004, Шмелев 1996 и др.); специально этим словам посвящена статья Киселева, Пайяр 2003. Мы предложим свою версию анализа данных слов, основанную на материале параллельных корпусов и двуязычной надкорпусной Базы данных дискурсивных слов.

Е. С. Яковлева (1988, 1994) обнаружила противопоставление двух классов модальных слов в зависимости от типа информации, которую они вводят: слова *кажется, как будто, вроде, явно, определено* (и некоторые другие) вводят «характерную» информацию, т. е. полученную говорящим в результате чувственного восприятия, а слова *наверно, видимо, по-видимому, вероятно, должно быть, может быть, бесспорно, безусловно, несомненно* — «нехарактерную», т. е. обозначают логический вывод, в основе которого лежат косвенные данные (Яковлева 1994: 219). В статье Иоанесян 1993 было предложено выделить дополнительный, как бы промежуточный, класс модальных слов — показателей «контекстно-связанной» информации. К этой категории были отнесены слова *видимо* и *по-видимому* — на том основании, что данные единицы употребляются лишь в тех случаях, когда говорящий формирует свое суждение, опираясь на некоторые перцептивные данные. Так, фраза *Он, видимо, спит* уместна лишь в той ситуации, когда у говорящего есть какие-то опытные данные, позволяющие сделать такой вывод — напр., говорящий видит, что субъект неподвижно лежит с закрытыми глазами, или не отвечает на телефонный звонок.

В Булыгина, Шмелев 1997: 303 противопоставление этих трех типов установок уточняется следующим образом: «[М]ожно “ни с того ни с сего”, лишь на основе общих знаний о климатических условиях Крыма сказать *А в Крыму сейчас, должно быть, жарко*, но фраза *В Крыму сейчас, видимо, жарко* может быть уместна в ситуации, когда говорящий видит на экране людей, гуляющих по Крыму в майках с короткими рукавами и в шортах. Иными словами, гипотеза, вводимая посредством *видимо* и *по-видимому* служит для объяснения непосредственных данных (evidence), эти слова оказываются близки по значению обороту *судя по всему*. Пролежав из-за болезни месяц в постели, можно сказать *Наверно, я очень изменился за это время* — только на основании общих представлений о том, что внешность человека, который так долго пролежал в постели, может сильно измениться; высказывание *Видимо, я очень изменился за это время* может быть естественным образом продолжено: *...меня никто не узнаёт*; а высказывание *Кажется, я очень изменился за это время* могло бы быть уместно в устах человека, впервые после долгого перерыва посмотревшего в зеркало.

Таким образом, установка, выражаемая словами *видимо* и *по-видимому*, — это установка на объяснение наблюдаемых фактов. Естественно ожидать, что данная установка характерна для тех, кто занимается исследовательской работой. И действительно, слова *видимо* и *по-видимому* чрезвычайно часто встречаются научных текстах, вводя, как и в рассмотренных ранее случаях “гипотезы-объяснения”. Кроме того, нередко в научной или квазинаучной речи эти слова указывают на ограниченность имеющихся в распоряжении автора данных

и означают что-то вроде ‘насколько мне известно’, ср.: *По-видимому, это единственный случай одушевленного *singulare tantum* в русском языке*».

О различии между *видимо* и *по-видимому* см. ниже.

### 3.1. Корпусные данные

Данные параллельного англо-русского подкорпуса «Национального корпуса русского языка» подтверждают гипотезу о том, что все три слова — *видимо*, *по-видимому* и *видно* в современном русском языке выражают установку на объяснение наблюдаемых фактов посредством предположения. Особенно показательны случаи отсутствия в тексте английского оригинала непосредственного (или даже какого бы то ни было) «стимула перевода». Приведем примеры.

(10) ‘I am doing something!’ said Mr. Gloop, who was now taking off his jacket and getting ready to dive into the chocolate. [Roald Dahl. *Charlie and the Chocolate Factory* (1964)] — Я и делаю, — сказал мистер Глуп и снял пиджак, собираясь, **видимо**, нырнуть в шоколадные глубины. [Роальд Даль. Чарли и шоколадная фабрика (М. Барон, Е. Барон, 1991)]

Слово *видимо* добавлено в переводе, чтобы обозначить более явно отношение логического вывода между наблюдаемой ситуацией (снял пиджак) и ее естественным объяснением (собираясь нырнуть).

(11) Whoever had sent this fax was still on the line... waiting to talk. [Dan Brown. *Angels and Demons* (2000)] — Тот, кто отправил ему факс, все еще оставался на линии... хотел, **видимо**, с ним поговорить. [Дэн Браун. Ангелы и демоны (Г. Косов, 2004)]

В русском переводе добавлено слово *видимо*, отсутствующее в оригинале: оно указывает на то, что последующий вывод основан на предшествующем перцептивном впечатлении.

(12) “He set out to do something no scientist had ever thought to do. Something that no one has ever had the TECHNOLOGY to do.” She paused, *as though* uncertain how to speak the next words. [Dan Brown. *Angels and Demons* (2000)] — Отец затеял то, до чего пока не додумался ни один ученый. Он решил сделать нечто такое, для чего до настоящего времени даже не существовало технических решений. — Виттория замолчала, **видимо**, не зная, как произнести следующие слова. [Дэн Браун. Ангелы и демоны (Г. Косов, 2004)]

В этом примере *видимо* использовано как эквивалент для англ. *as though* ‘как если бы’, что указывает именно на перцептивную основу данного вывода. Но таких примеров немного.

(13) “Well,” said Rabbit, after a long silence in which nobody thanked him for the nice walk they were having, “we’d better get on, I suppose.” [A. A. Milne. *The House at Pooh Corner* (1928)] — Ну что ж, — сказал Кролик, **по-видимому**, все это время напрасно ожидавший, что его поблагодарят за приятную

прогулку. — Пожалуй, надо идти. [Алан Александр Милн. Дом на Пуховой Опушке (Б. Заходер, 1960)]

В переводе при помощи добавленного слова *по-видимому* эксплицирована логическая связь между предложениями (Кролик сказал: «Пожалуй, надо идти», *потому что* ожидал, что его поблагодарят и не дождался).

Вводное *видно* также может эксплицировать логическую связь между описываемыми явлениями, в том числе, имеющую перцептивную основу (ср. англ. стимул *as if* в примере (15)):

- (14) She gets very excited when she tells you that stuff. [J. D. Salinger. The Catcher in the Rye (1951)] Она вся раскраснелась, **видно**, ей было очень интересно рассказывать. [Дж. Д. Сэлинджер. Над пропастью во ржи (Р. Райт-Ковалёва, 1965)]
- (15) He sat watching him closely, *as if* memory were beginning to work. [Graham Greene. The Power and the Glory (1940)] Он пристально смотрел на священника; память его, **видно**, начала работать. [Грэм Грин. Сила и слава (Н. Волжина, 1970–1980)]

### 3.2. Видимо, по-видимому, видно: есть ли различия?

Относительно соотношения данных дискурсивных слов выдвигались следующие гипотезы:

1. *По-видимому* в большей степени основано на логическом выводе, а вводное *видно* — на зрительном восприятии.
2. Для *по-видимому* «характерна большая временная дистанция между стимулом и реакцией, т. е. между положением вещей, с которым столкнулся субъект, и высказыванием, основанным на субъективном восприятии этого положения вещей» (Киселева, Пайар 2003).
3. Для вводного *видно* характерно использование в составе формул «примирения с действительностью» (Шмелев 1996): *видно, не судьба*.

Гипотезы 1 и 2 не подтверждаются ни данными параллельного корпуса (точнее, недостаточно данных для того чтобы считать их верными), ни данными «Национального корпуса русского языка» в целом. Гипотеза 3, на первый взгляд, тоже не подтверждается данными параллельного подкорпуса НКРЯ: в англо-русской части параллельного подкорпуса почти не встретилось таких примеров, где появление вводного *видно* было бы обусловлено этим мотивом. Но причина, очевидно, состоит в том, что сам мотив не очень характерен для англоязычного дискурса. Показательно, что выражающий эту идею оборот *не судьба* ни разу не встретился в НКРЯ в переводах с английского языка. Любопытно, впрочем, что когда переводчик усматривает в оригинале данный мотив, это оказывается достаточным основанием для появления в переводе слова *видно*, иногда в сопровождении других формул «примирения с действительностью»:

- (16) “You are a very strange creature by way of a friend! —always wanting me to play and sing before anybody and everybody! If my vanity had taken a musical turn,

you would have been invaluable; but as it is, I would really rather not sit down before those who must be in the habit of hearing the very best performers.” On Miss Lucas’s persevering, however, she added, “Very well, *if it must be so, it must.*” And gravely glancing at Mr. Darcy, “There is a fine old saying, which everybody here is of course familiar with: ‘Keep your breath to cool your porridge’; and I shall keep mine to swell my song.” [Jane Austen. *Pride and Prejudice* (1813)] [...] Однако когда мисс Лукас начала настаивать, она сказала: — Ну что ж, чему быть — того не миновать! — И, хмуро глядя на мистера Дарси, добавила: — В наших краях помнят еще старое правило: чем сказать, лучше смолчать. **Видно, уж надо мне смириться**, да тем и кончить дело. Джейн Остин. Гордость и предубеждение (И. С. Маршак, 1967)]

(17) There, *’twas to be, I suppose*. Your lot is your lot, and Scripture is nothing; for if you do good you don’t get rewarded according to your works, but be cheated in some mean way out of your recompense. [Thomas Hardy. *Far from the Madding Crowd* (1874)] — Ну что ж, **видно, такова судьба**. Кому что на роду написано [...] [Томас Гарди. Вдали от обезумевшей толпы (М. Богословская, Н. Высоцкая, 1970)]

(18) However, *what is must be...* [Thomas Hardy. *Far from the Madding Crowd* (1874)] — Но, **видно, такая уж мне выпала судьба**. [Томас Гарди. Вдали от обезумевшей толпы (М. Богословская, Н. Высоцкая, 1970)]

(19) The village of St. Petersburg still mourned. The lost children had not been found. Public prayers had been offered up for them, and many and many a private prayer that had the petitioner’s whole heart in it; but still no good news came from the cave. The majority of the searchers had given up the quest and gone back to their daily avocations, saying that it was plain the children could never be found. [Mark Twain. *The Adventures of Tom Sawyer* (1876)] — Городок Сент-Питерсберг все еще оплакивал пропавших детей. Они так и не нашлись. [...] Многие из горожан бросили поиски и вернулись к своим обычным делам, говоря, что детей, **видно**, уж не найти. [Марк Твен. Приключения Тома Сойера (Н. Дарузес, 1950)]

Заметим, что в основном корпусе НКРЯ из 300 (с небольшим) примеров на выражение *не судьба* в 36 (более чем в 10% случаев) это выражение встречается в контексте *видно*, и это служит некоторым свидетельством в пользу обсуждаемой гипотезы. Приведем некоторые примеры:

(20) Я написал к Вам наудачу в Карлсбад — *poste restante* уже давно, но **видно не судьба** попасть этому письму в Ваши руки. [П. В. Анненков. Письма к И. С. Тургеневу (1852–1874)]

(21) **Видно, не судьба** нам на этом свете быть счастливыми. [Василий Гроссман. Все течет (1955–1963) // «Октябрь», 1989]

(22) **Видно уж, не судьба**, — подумал он, снова карабкаясь в гору. — Ну и чёрт с ним. [Ю. О. Домбровский. Факультет ненужных вещей, часть 3 (1978)]

- (23) Ему бы полгода просидеть в горах, и он, может, до сих пор был бы с нами. Да **видно не судьба**. [Фазиль Искандер. Школьный вальс, или Энергия стыда // «Знамя», 1987]

Можно добавить, что в русско-английском подкорпусе НКРЯ оборот *не судьба* встретился шесть раз, и регулярно вызывал затруднения у переводчиков. При этом в примерах, когда он появился в сопровождении дискурсивного *видно*, переводчики опускали дискурсивное слово — по-видимому, считая, что его значение продублировано в выражении *не судьба*:

- (24) Прочитав, она возвратила мне письмо дрожащею рукою и сказала дрожащим голосом: «**Видно мне не судьба**... Родные ваши не хотят меня в свою семью. Буди во всем воля Господня! Бог лучше нашего знает, что нам надобно.» [А. С. Пушкин. Капитанская дочка (1836)] Having read the letter she returned it, and said in a trembling voice: “It was not my destiny. Your parents do not wish me in their family. May the will of God be done! He knows better than we what is best for us.” [Alexander Pushkin. Marie: a Story of Russian Love (Marie H. de Zielinska, 1877)]

- (25) ...с разочарованием грустно думал: «Верно, я еще не совсем большой, если не могу курить, как другие, и что, **видно, мне не судьба**, как другим, держать чубук между средним и безымянным пальцем, затягиваться и пускать дым через русые усы». [Л. Н. Толстой. Юность (1856)] — “Truly,” I thought to myself in my dejection and disillusionment, “I cannot be quite grown-up if I cannot smoke as other fellows do, and should be fated never to hold a chibouk between my first and second fingers, or to inhale and puff smoke through a flaxen moustache!” [Leo Tolstoy. Youth (C. J. Hogarth, 1910–1935)]

Кроме того, данные параллельного корпуса дают основания для следующей гипотезы, отчасти коррелирующей с двумя первыми:

4. *Видимо* характерно для устной речи, а *по-видимому* — в основном для письменной<sup>7</sup>.

Данные устного подкорпуса НКРЯ (объем 11 349 008 слов), рассмотренные на фоне основного корпуса подтверждают эту гипотезу: вводное *видимо* встречается в нем 102 раза, а *по-видимому* — 8; из них 2 раза — в лекции А. В. Исаченко, 2 раза — в беседе А. Максимова с Т. Себенцовой в программе «Времечко», и один раз — в чтении вслух Виктором Некрасовым своей повести: все это — не собственно спонтанная устная речь. Иными словами, *по-видимому* встречается в спонтанной устной всего в трех примерах устного корпуса, т.е. соотношение с *видимо* 1: 34. В основном корпусе соотношение иное: вводное *видимо* встречается несколько чаще, чем *по-видимому*, примерно в полтора раза, но никак не в 30 с лишним раз.

<sup>7</sup> Примеры, приводимые в подтверждение Гипотезы 2 в Киселева, Пайар 2003, указывают скорее именно на это противопоставление: *по-видимому* встречается в письменной речи, а *видимо* — в устной: в прямой речи персонажа (— *Это твоя бабушка, **видимо**, звонит*) и в комментарии от 1-го лица (*Он постучал в дверь и крикнул: — Можно попросить стакан холодной воды?! **Видимо**, рассчитывал на мою сентиментальность*).

Итак, мы продемонстрировали возможности применения унидирекционального подхода в контрастивных корпусных исследованиях в отношении использования реверсивного перевода для уточнения семантического анализа (несмотря на то, что объем корпусного материала — как моно-, так и двуязычного — пока мал) — на примере анализа нескольких дискурсивных слов русского языка. Анализ значения частицы *еще*, рассмотренного в Levontina, Shmelev 2005, подтвердился на материале англо-русского подкорпуса НКРЯ; из трех гипотез, высказанных в отношении дискурсивных слов *видимо*, *по-видимому*, *видно*, частично подтвердилась лишь одна, однако данные корпуса позволили сформулировать еще одну гипотезу, которая подтвердилась на более широком материале (устном и основном подкорпусе НКРЯ). Эффективность предложенного подхода будет, очевидно, возрастать по мере увеличения объема параллельных текстовых корпусов.

## Литература

1. Булыгина Т. В., Шмелев А. Д. (1997), Языковая концептуализация мира (на материале русской грамматики). М., 1997.
2. Виноградов 1975 — Виноградов В. В. О категории модальности и модальных словах в русском языке. // В. В. Виноградов. Исследования по русской грамматике: избранные труды. М., 1975.
3. Зализняк Анна А. (2015), Лингвоспецифичные единицы русского языка в свете контрастивного корпусного анализа // Компьютерная лингвистика и интеллектуальные технологии. По материалам международной конференции Диалог'2015. М., 2015. С. 651–662.
4. Зализняк Анна А. (2016), База данных межъязыковых эквиваленций как инструмент лингвистического анализа. // Компьютерная лингвистика и интеллектуальные технологии. По материалам международной конференции Диалог'2016. М., 2016. С. 763–775.
5. Иоанесян Е. Р. (1993), Классификация ментальных предикатов по типу вводимых ими суждений // Логический анализ языка: Ментальные действия. Москва, 1993. С. 89–95.
6. Киселева, Пайар 2003 — К. Киселева, Д. Пайар. Механизмы семантического варьирования на примере группы единиц с корнем *вид-*: ВИДИМО, ПО-ВИДИМОМУ, ВИДНО // Д. Пайар (ред.) Дискурсивные слова русского языка: контекстное варьирование и семантическое единство. М., 2003. С. 50–79.
7. Разлогова Е. Э. (1996), Модальные слова и оценка степени достоверности высказывания // Русистка сегодня, 1996/3. С. 21–47.
8. Разлогова Е. Э. (2004), Логико-когнитивные и стилистические аспекты семантики модальных слов.
9. Шмелев А. Д. (1996), Жизненные установки и дискурсивные слова. — Aspekta [ = Slavica Tampereusia, V]. Tampere, 311–322.
10. Шмелев А. Д. (2015), Русские лингвоспецифичные лексические единицы в параллельных корпусах: возможности исследования и «подводные камни» // Компьютерная лингвистика и интеллектуальные технологии. По материалам международной конференции Диалог'2015. М., 2015.

11. Яковлева 1994 — Яковлева Е. С. Фрагменты русской языковой картины мира (модели времени, пространства и восприятия). М., 1994.
12. Яковлева 1988 — Яковлева Е. С. Согласование модусных характеристик в высказывании. // Прагматика и проблемы интенциональности. М., 1988.
13. Aijmer K. (2008), *Translating Discourse Particles: A Case of Complex Translation* // Anderman G., Rogers M. (2008), *Incorporating Corpora: The Linguist and the Translator* / Ed. by G. Anderman and M. Rogers. Clevedon/Buffalo/Toronto: Multilingual Matters LTD.
14. Levontina I., Shmelev A. (2005), *The particles one cannot do without* // East—West Encounter: Second International Conference on Meaning ↔ Text Theory. Moscow, pp. 258–267.
15. Loiseau S., Sitchinava D. V., Zalizniak Anna A., Zatsman I. M. (2013), *Information technologies for creating the database of equivalent verbal forms in the Russian-French multivariant parallel corpus* // Информатика и ее применения, 2013. Т. 7. Вып. 2. С. 100–109.
16. Noël D. (2003), *Translations as evidence for semantics: An illustration*. *Linguistics* 41(4):757–785.

## References

1. Aijmer K. (2008), *Translating Discourse Particles: A Case of Complex Translation* // Anderman G., Rogers M. (2008), *Incorporating Corpora: The Linguist and the Translator* / Ed. by G. Anderman and M. Rogers. Clevedon/Buffalo/Toronto: Multilingual Matters LTD.
2. Bulygina T. V., Shmelev A. D. (1997), *Linguistic conceptualization of the world (based on the Russian grammar)* [Iazykovaia kontseptualizatsiia mira (na materiale russkoi grammatiki)], Shkola “Iazyki russkoi kul'tury”, Moscow.
3. Iakovleva E. S. (1988) *An agreement of modus characteristics in a proposition* [Soglasovanie modusnykh kharakteristik v vyskazyvanii], *Pragmatics and problems of Intentionality* [Pragmatika i problemy intensional'nosti], INION, Moscow.
4. Iakovleva E. S. (1994), *Fragments of the Russian Linguistic Worldview (Models of Time, Space, and Perception)* [Fragmenty russkoi iazykovoï kartiny mira (modeli vremeni, prostranstva i vospriatiia)], Gnozis, Moscow.
5. Ioanesian E. R. (1993), *Mental predicates as the types of their subordinate propositions* [Klassifikatsiia mental'nykh predikatov po tipu vvodimyykh imi suzhdenii], *Logical analysis of language: Mental acts* [Logicheskii analiz iazyka: Mental'nye deistviia], Nauka, Moscow, pp. 89–95.
6. Kiseleva K., Paillard D. (2003), *Mechanisms of semantic variation by example of items VIDIMO, PO-VIDIMOMU, VIDNO* [Mekhanizmy semanticheskogo var'irovaniia na primere gruppy edinit s kornem vid-: VIDIMO, PO-VIDIMOMU, VIDNO], D. Paillard (ed.) *Russian discourse words: contextual variation and semantic unity* [Diskursivnye slova russkogo iazyka: kontekstnoe var'irovanie i semanticheskoe edinstvo], Azbukovnik, Moscow, pp. 50–79.

7. *Levontina I., Shmelev A.* (2005), The particles one cannot do without // East—West Encounter: Second International Conference on Meaning ↔ Text Theory. Moscow, pp. 258–267.
8. *Loiseau S., Sitchinava D. V., Zalizniak Anna A., Zatsman I. M.* (2013), Information technologies for creating the database of equivalent verbal forms in the Russian-French multivariant parallel corpus // Информатика и ее применения, 2013. Т. 7. Вып. 2. С. 100–109.
9. *Noël D.* (2003), Translations as evidence for semantics: An illustration. *Linguistics* 41(4):757–785.
10. *Razlogova E. E.* (1996), — Modal words and the assessment of the reliability of a proposition [Modal'nye slova i otsenka stepeni dostovernosti vyskazyvaniia], *Russian Linguistics Today [Rusistika segodnia]*, 3, pp. 21–47.
11. *Razlogova E. E.* (2004), — Logiko-kognitivnye i stilisticheskie aspekty semantiki modal'nykh slov.
12. *Shmelev A. D.* (1996) — Life attitudes and discourse words [ZHiznennye ustanovki i diskursnye slova], *Aspekteja [= Slavica Tampereusia, V]*, Tampere, pp. 311–322.
13. *Shmelev A. D.* (2015), Russian language-specific lexical units in parallel corpora: Prospects of investigation and “pitfalls” [Russkie lingvospetsifichnye edinitsy v parallel'nykh korpusax: vozmozhnosti issledovaniia i “podvodnye kamni”], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference Dialogue' 2015 [Komp'iuternaja lingvistika i intellektual'nye tekhnologii]*, Moscow, pp. 584–594.
14. *Vinogradov V. V.* (1975), On the category of modality and modal words in Russian [O kategorii modal'nosti i modal'nykh slovakh v russkom iazyke], *V. V. Vinogradov, Issledovaniia po russkoi grammatike: izbrannye trudy*, Nauka, Moscow.
15. *Zalizniak Anna A.* (2015), Russian language-specific words as an object of contrastive corpus analysis [Lingvospetsifichnye edinitsy russkogo iazyka v svete kontrastivnogo korpusnogo analiza], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference Dialogue' 2015 [Komp'iuternaia lingvistika i intellektual'nye tekhnologii. Po materialam mezhdunarodnoi konferentsii Dialog'2015]*, Moscow, pp. 651–662.
16. *Zalizniak Anna A.* (2016), A database of cross-linguistic equivalences as an instrument of linguistic analysis [Baza dannykh mezh'iazykovykh ekvivalentsii kak instrument lingvisticheskogo analiza], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference Dialogue' 2016 [Komp'iuternaia lingvistika i intellektual'nye tekhnologii. Po materialam mezhdunarodnoi konferentsii Dialog'2016]*, Moscow, pp. 763–775.



# ОРФОГРАФИЧЕСКИЕ СОСЕДИ С ЗАМЕНОЙ БУКВЫ ПРИ ИЗУЧЕНИИ МЕХАНИЗМОВ ЛЕКСИЧЕСКОГО ДОСТУПА

**Слюсарь Н. А.** (slioussar@gmail.com)

НИУ ВШЭ, Москва, и СПбГУ, Санкт-Петербург, Россия

**Алексеева С. В.** (mail@s-alexeeva.ru)

СПбГУ, Санкт-Петербург, Россия

## SUBSTITUTION NEIGHBORS IN THE STUDY OF LEXICAL ACCESS

**Slioussar N. A.** (slioussar@gmail.com)

HSE, Moscow, and St. Petersburg State University,  
St. Petersburg, Russia

**Alexeeva S. V.** (mail@s-alexeeva.ru)

St. Petersburg State University, St. Petersburg, Russia

The role of orthographic neighbors (e.g. bank—tank) in word processing has been discussed in many experimental studies. However, these studies have been conducted on a limited pool of languages, and many important questions are still unresolved. After creating a lexical database StimulStat that contains various neighborhood parameters for Russian, we conducted the first experiment with substitution neighbors in Russian. We used lexical decision task with priming, and manipulated the following factors: whether the prime is more or less frequent than the target, whether the prime is a nominative singular (primary) form or an oblique form, and whether the substituted letter is word-final or in the middle of the word. The results suggest that noun forms undergo morphological decomposition at a very early stage and shed new light on the process of activating candidates during lexical. The results also have practical significance because it is well known that spelling errors are influenced by neighborhood effects.

**Keywords:** mental lexicon, orthographic neighbors, substitution neighbors, inflectional morphology, Russian

## 1. Введение

Важнейшей задачей при чтении является распознавание слов, или лексический доступ. Это сложный процесс, ключевым элементом которого является анализ буквенного состава слова. Общеизвестно, что если навык чтения не находится на самом начальном уровне, человек не читает слова буква за буквой. На интуитивном уровне об этом свидетельствует то, что мы можем перепутать похожие по написанию слова и часто не замечаем мелких опечаток. Значительная часть научных исследований в этой области опирается на феномен так называемого орфографического соседства (например, [Acha, Carreiras 2014; Davis et al. 2009; Grainger 1992, 2008]).

Соседями называются цепочки букв (слова или квазислова), отличающиеся друг от друга на одну букву (например, *крот* — *кот*, *крот* — *кров*) или за счет перестановки букв (например, *крот* — *корт*). Например, было показано, что на доступ к слову в ментальном лексиконе (в частности, на скорость его прочтения) влияет число слов-соседей и их частотность [Andrews 1996, Segui, Grainger 1990]. Исходя из этого можно сделать вывод, что в процессе распознавания графического облика слова мы рассматриваем несколько возможных кандидатов и перебираем их в зависимости от их частотности. Кроме того, в связи с этим различные факторы, связанные с орфографическими соседями, контролируются при подборе стимулов во многих экспериментах, которые напрямую не связаны с проблемой орфографического сходства [White et al. 2008; Whitney, Lavidor 2004 и др.].

Эффекты соседства наблюдаются и при написании слов: например, известно, что мы чаще допускаем опечатки в тех словах, у которых много соседей [Roux, Bonin 2009]. Автоматические системы проверки орфографии в качестве вариантов замены в первую очередь предлагают реальные слова, которые являются соседями несуществующего слова. Таким образом, изучение орфографических соседств (в частности, то, какие типы соседей чаще путаются при прочтении и написании) имеет практическое значение. Этому и посвящено представленное в данной работе исследование.

Подобрать слова, у которых есть соседи определенного типа, используя обычные словари, практически невозможно. Поэтому для ряда языков созданы лексические базы данных, в которых эти параметры учитываются наравне с другими психолингвистически релевантными характеристиками [Boudelaa, Marslen-Wilson 2010; Davis 2005; Davis, Perea 2005; Duchon et al. 2013; New et al. 2004; и др.]. Мы создали первую такую базу данных на материале русского языка — *StimulStat* (<http://stimul.cognitivestudies.ru/>) [Alexeeva et al. 2015, 2016]. В ней представлено более 50 тысяч наиболее частотных слов русского языка, охарактеризованных по более чем 70 различным параметрам, связанным с частотностью, буквенным и фонемным составом, просодическими особенностями, семантикой, грамматическими характеристиками лемм и словоформ, с наличием близких по написанию слов и др. В данной работе представлено экспериментальное исследование, материал для которого был подобран при помощи этого ресурса. Оно направлено на прояснение ряда нерешенных вопросов, связанных с распознаванием графического облика слова и с другими аспектами лексического доступа.

В базе *StimulStat* представлены все основные типы орфографических соседей, которые рассматриваются в психолингвистической литературе: соседи с заменой одной буквы (*крот* — *кров*) [например, Coltheart et al. 1977]; соседи с удалением одной буквы (*крот* — *рот*) [например, Davis et al. 2009]; соседи со вставкой одной буквы (*кара* — *карат*) [например, De Moor, Brysbaert 2000]; соседи с перестановкой двух соседних (*баня* — *баян*) или несоседних (*кот* — *ток*) букв [например, Andrews 1996]. В проведенном нами эксперименте в качестве стимулов были использованы пары соседей с заменой одной буквы. При этом, в отличие от всех прочих аналогичных ресурсов, все расчеты в базе сделаны как для лемм (начальных форм), так и для словоформ<sup>1</sup>, и мы воспользовались этим преимуществом в нашем эксперименте.

Так как исследований, в которых задействованы слова-соседи, очень много, ниже мы остановимся только на нескольких вопросах, имеющих непосредственное отношение к нашей работе. Большая часть экспериментов в этой области использует методику принятия лексического решения с праймингом. Испытуемый должен как можно быстрее определить, является ли предъявленная ему на экране компьютера цепочка букв реальным словом или нет. При этом перед стимульной последовательностью на небольшой промежуток времени показывается другая последовательность (прайм). Например, если прайм совпадает с целевым словом, последнее распознается значительно быстрее, чем в контрольном условии (где стимул и прайм никак не связаны) [Acha, Carreiras 2014]. Это называется эффектом прайминга. Он связан с тем, что ключевое слово уже было активировано в ментальном лексиконе, и активировать его повторно занимает меньше времени.

Если прайм является соседом целевого слова, результат зависит от соотношения их частотности. В ряде работ было показано, что более частотные праймы замедляют обработку целевых слов по сравнению с контрольным условием, то есть наблюдается отрицательный прайминг-эффект [Davis, Lupker 2006; Duñabeitia et al. 2009; Segui, Grainger 1990]. Это связывают с тем, что в процессе восприятия прайма активируются близкие по написанию слова, а затем все варианты, кроме избранного, подавляются, что и приводит к замедлению. При этом для праймов, которые менее частотны, чем ключевые слова, значимых эффектов выявлено не было. Это может объясняться тем, что для более частотных слов менее заметен эффект подавления. Наконец, если праймом является квазислово, обработка целевого слова ускоряется. Предполагается, что в этом случае близкие по написанию слова активируются, но не подавляются. Учитывая некоторую неопределенность в этом вопросе, мы включили в наш эксперимент пары праймов и целевых слов с разным соотношением частотности.

Второй вопрос, который не был разрешен в предыдущих работах, касается места перестановки. Многочисленные исследования свидетельствуют об особой роли первой и последней буквы в слове: например, они распознаются

<sup>1</sup> Для каждой группы слов-соседей подсчитано количество входящих в нее слов, их суммарная частотность и некоторые другие характеристики. Кроме того, база дает возможность одновременно учитывать целый ряд других параметров (длину, слоговую структуру, грамматические характеристики и многое другое).

значимо быстрее, чем буквы, находящиеся в середине [Boouma 1973; Gomez et al. 2008; Johnson et al. 2007; Tydgate, Grainger 2009]. Основываясь на этом, можно предположить, что место замены будет оказывать влияние на эффекты, связанные с соседством. Однако пока эта гипотеза не проверялась. В работе [Grainger 1988] сравнивались слова-соседи с заменой первой и последней буквы, и разница между ними оказалась незначимой. Однако Дж. Грейнджер и Х. Сеги [Grainger, Segui 1990] обращают внимание на то, что в этом эксперименте были использованы достаточно короткие слова, и предполагают, что при использовании более длинных слов могут обнаружиться различия как между этими двумя группами, так и между ними и соседями с заменой букв в середине слова. В проведенном нами эксперименте мы обращаемся к этому вопросу.

Наконец, эксперименты с использованием слов-соседей затрагивают целый ряд важных вопросов, не связанных непосредственно с распознаванием графического облика слова. Один из таких вопросов — проблема морфологического анализа при восприятии. Существующие модели рассматривают как словоформы [например, McClelland, Patterson 2002; Marslen-Wilson, Tyler 1997; Pinker 1991, 1999; Rumelhart, McClelland 1986; Ullman 2004], так и производные слова [Вааен et al. 1997; Pollatsek et al. 2000; Seidenberg, McClelland 1989; Taft, Forster 1975; Taft 2004 и др.] и ставят вопрос о том, хранится ли в ментальном лексиконе сложное слово или форма целиком или же содержащиеся в ней морфемы (в некоторых моделях реализованы обе опции).

Исследования с использованием слов-соседей проводятся прежде всего на материале соседей с перестановкой и основываются на предположении, что, если морфологический анализ начинается на самом раннем этапе, параллельно с распознаванием графического облика слова, связанные с соседством прайминг-эффекты будут значительно меньше или исчезнут наблюдаться в тех случаях, когда перестановка происходит на стыке морфем. Почти все существующие исследования сфокусированы на словообразовании [Beyersmann et al. 2013; Christianson et al. 2005; Duñabeitia et al. 2007; Masserang, Pollatsek 2012; Perea, Carreiras 2006; Rueckl, Rimzhim 2011; Sánchez-Gutiérrez, Rastle 2013]. Полученные результаты противоречивы: в одних исследованиях не было обнаружено прайминг-эффекта в условии с перестановкой букв на стыке морфем, в других прайминг-эффект наблюдался независимо от условия. Такое же противоречие наблюдается в работах, в которых рассматривается словоизменение (перестановки на стыке основы и окончания). Пока их всего две [Luke, Christianson 2013; Zargar, Witzel, 2016], поэтому очевидно, что необходимы дальнейшие исследования.

Мы обращаемся к этой задаче в нашем эксперименте, используя в качестве стимулов пары словоформ-соседей, где либо обе формы являются начальными, либо прайм стоит в форме косвенного падежа. Если процесс морфологического анализа при восприятии имеет место и начинается очень рано, мы можем ожидать связанные с соседством прайминг-эффекты в первом случае, но не во втором. Более того, в нашем эксперименте этот вопрос впервые рассматривается на материале соседей с заменой, а не с перестановкой.

## 2. Экспериментальное исследование

В эксперименте была использована методика принятия лексического решения с праймингом, описанная во введении. В качестве стимулов выступали пары словоформ-соседей с заменой одной буквы, подобранные с манипуляцией трех упомянутых выше параметров:

- место перестановки: конец слова (например, *казак* — *казан*) или середина слова (например, *пикет* — *пакет*);
- соотношение частотностей прайма и целевого слова: либо прайм частотнее, чем целевое слово (*ключ*, 32,3<sup>2</sup> — *клев*, 6,6), либо наоборот (*грот*, 1,2 — *гром*, 17,5);
- морфологический статус прайма: прайм — существительное в начальной форме (*голод* — *голос*) или в форме косвенного падежа (*ритма* — *рифма*).

В проведении эксперимента участвовала студентка СПбГУ Анастасия Петрова.

### 2.1. Метод

В эксперименте приняли участие 30 носителей русского языка в возрасте от 19 до 24 лет.

Используя базу данных *StimulStat*, мы отобрали стимульный материал: 64 существительных, которые имеют соседа с заменой одной буквы в начальной позиции. Средняя длина целевого слова — 4,7 буквы (диапазон: 4–7 букв), средняя частотность леммы — 55,7 на миллион (диапазон: 0,6–484,8 на миллион), а средняя частотность словоформы — 19,6 на миллион (диапазон: 0,17–155,3 на миллион). Отобранные пары целевых слов и праймов-соседей распределялись по восьми группам, представленным в Таблице 1 (по восемь пар на группу).

Таблица 1. Группы стимулов

Группа	Позиция перестановки	Форма прайма	Частотность прайма и цели	Пример
1a	середина слова	начальная	прайм частотней	<i>народ</i> — <i>наряд</i>
1b	середина слова	начальная	цель частотней	<i>пикет</i> — <i>пакет</i>
2a	середина слова	косвенная	прайм частотней	<i>тюрем</i> — <i>терем</i>
2b	середина слова	косвенная	цель частотней	<i>призов</i> — <i>призыв</i>
3a	конец слова	начальная	прайм частотней	<i>ключ</i> — <i>клев</i>
3b	конец слова	начальная	цель частотней	<i>грот</i> — <i>гром</i>
4a	конец слова	косвенная	прайм частотней	<i>зону</i> — <i>зонт</i>
4b	конец слова	косвенная	цель частотней	<i>меток</i> — <i>метод</i>

<sup>2</sup> Частотность указывается в количестве употреблений на миллион согласно словарю [Lyashevskaya O. N., Sharov 2009]. Здесь приведены частотности лемм, но при подборе стимулов учитывалась также частотность форм (согласно проекту «Частотная грамматика русского языка» [Lyashevskaya 2013]).

Чтобы подобрать праймы для контрольного условия, а также квазислова и праймы для них, для каждого уже отобранного нами прайма мы нашли по две словоформы, совпадающих с ним по длине и слоговой структуре и близких по частотности, но не обладающих орфографическим сходством с ним. На основании одной из этих словоформ создавалось квазислово<sup>3</sup>. Пример приведен в Таблице 2.

**Таблица 2.** Пример набора целевых стимулов и праймов

Целевое слово	Прайм 1 (эксп. условие для целевого слова)	Прайм 2 (контр. условие для целевого слова и квазислова)	Прайм 3 (эксп. условие для квазислова)	Квазислово
метла (лемма: 6,7, форма: 1,7)	метра (лемма: 115,2, форма: 12,9)	чёрта (лемма: 110,2, форма: 15,81)	листа (лемма: 127,6, форма: 13,78)	чёрса

Стимулы были распределены по двум экспериментальным протоколам. В первом протоколе целевое слово из первого набора предъявлялось после прайма 1 (т. е. в условии соседства), а квазислово — после прайма 2 (т. е. в контрольном условии). Целевое слово из второго набора предъявлялось после прайма 2 (в контрольном условии), а квазислово — после прайма 3 (в условии соседства) и т. д. Во втором протоколе условия были распределены зеркально. Каждый испытуемый проходил только один протокол.

Эксперимент проводился на персональном компьютере с использованием программы *PsyScope* [Cohen et al. 1993]. Сперва на экране предъявлялась инструкция:

«Уважаемый участник! Положите указательный палец правой руки на клавишу „J“ и указательный палец левой руки на клавишу „F“. Перед Вами на экране будет появляться знак „\*“, цепочка букв и затем другая цепочка букв. Ваша задача: как можно быстрее определить, является ли вторая цепочка букв словом русского языка. Если это слово, нажмите правую кнопку. Если это не слово, нажмите левую кнопку»<sup>4</sup>.

Затем испытуемым предъявлялись восемь тренировочных пар стимулов. После этого у них была возможность задать уточняющие вопросы. Далее следовал основной блок стимулов: 108 пар «прайм — целевое слово или квазислово».

Сперва испытуемый видел в центре экрана звездочку (точку фиксации). Временной интервал случайным образом выбирался из 1000, 1100, 1200, 1300, 1400 и 1500 мс. Это не позволяло испытуемым сформировать ожидания относительно того, когда появится прайм. Затем в центре экрана на 200 мс предъявлялся прайм. Он на 200 мс сменялся пустым экраном, после чего появлялся целевой стимул. Если испытуемый не нажимал на клавишу в течение 500 мс,

<sup>3</sup> Иначе испытуемые заметили бы, что сходство наблюдается только в парах с реальными словами.

<sup>4</sup> Половина испытуемых, проходивших каждый протокол, видела такую инструкцию, другая половина должна была нажимать на левую кнопку, если видит слово.

целевой стимул сменялся пустым экраном. Это было сделано для того, чтобы испытуемые отвечали быстрее. Если испытуемый не отвечал в течение последующих 1500 мс, на экране снова проявлялась звездочка, а затем — следующие прайм и стимул.

## 2.2. Результаты

В ходе эксперимента было сделано 3840 измерений времени реакции. Из них мы исключили 189 случаев (5%), где испытуемые дали неверный ответ. Ни один из участников не ошибался более чем в 15% случаев. Затем, как принято в такого рода исследованиях, мы удалили слишком быстрые (менее 100 мс) и слишком медленные (более 1500 мс) реакции, всего 109 случаев (3% от общего количества). Средние значения времени реакции для экспериментальных условий в различных группах, а также размер эффекта прайминга (разница между условием соседства и контрольным условием) даны в Таблице 3.

**Таблица 3.** Среднее время реакции и размер эффекта прайминга (в мс)

Группа	Целевой стимул	Конт-рольное условие	Условие соседства	Размер эффекта прайминга
1a «середина, нач. форма, прайм > цели»	слово	760,6	799,2	-38,6
1b «середина, нач. форма, прайм < цели»	слово	681,9	673,6	8,3
2a «середина, косв. форма, прайм > цели»	слово	758,1	769,7	-11,6
2b «середина, косв. форма, прайм < цели»	слово	729,6	703,5	26,1
3a «конец, нач. форма, прайм > цели»	слово	779,6	831,9	-52,3
3b «конец, нач. форма, прайм < цели»	слово	703,8	634,6	69,2
4a «конец, косв. форма, прайм > цели»	слово	718,0	722,3	-4,3
4b «конец, косв. форма, прайм < цели»	слово	703,2	691,0	12,3
1a	квазислово	765,6	779,2	-13,6
1b	квазислово	769,4	784,0	-14,7
2a	квазислово	745,6	748,4	-2,8
2b	квазислово	769,1	732,6	36,5
3a	квазислово	732,2	747,2	-15,1
3b	квазислово	772,0	763,5	8,5
4a	квазислово	755,7	757,2	-1,5
4b	квазислово	805,5	802,5	3,0

Чтобы установить, какие различия между условиями являются статистически значимыми, в программном пакете SPSS был проведен дисперсионный анализ с повторными измерениями (RM ANOVA). Сперва мы сравнили ответы в контрольном условии и в условии соседства внутри каждой группы. Значимая разница была обнаружена только для группы 3b, включающей стимулы-слова, где праймы являлись начальными формами и были менее частотны, чем целевое слово, а замена затрагивала последнюю букву ( $F(1, 230) = 10,02, p = 0,002$ ). Различия в остальных группах оказались незначимыми (для всех сравнений здесь и далее,  $p > 0,05$ ). Таким образом, можно предварительно сказать, что все исследуемые факторы играют роль для эффекта прайминга, связанного с наличием орфографических соседей с заменой одной буквы.

Затем мы взяли все реакции на стимулы-слова и протестировали значимость четырех факторов: позиции перестановки, формы прайма, соотношения частотности прайма и целевого слова, а также условия предъявления (условие соседства или контрольное). Прежде всего, значимым оказался фактор частотности ( $F(1, 1743) = 60,79, p < 0,001$ ). Это ожидаемый результат, так как частотность целевых слов в группах 1a, 2a, 3a и 4a была существенно выше, чем в группах 1b, 2b, 3b и 4b (средняя частотность лемм: 6,3 и 105,2 соответственно, средняя частотность форм: 2,2 и 37,0, средняя частотность лемм праймов: 95,5 и 11,9, средняя частотность форм праймов: 26,8 и 3,2).

Другие факторы не были значимыми, поэтому перейдем к их взаимодействиям. В первую очередь нас интересуют взаимодействия с фактором условия: они будут свидетельствовать о том, как другие факторы влияют на эффект прайминга. Значимости достигло взаимодействие факторов условия и частотности ( $F(1, 1743) = 7,88, p = 0,005$ ) и фактора условия и места перестановки ( $F(1, 1743) = 4,12, p = 0,043$ ). Это отражает тот факт, что в условиях, где прайм частотней целевого слова, время реакции в условии соседства в среднем больше, чем в контрольном, а в условиях, где целевое слово частотней прайма, наблюдается обратная картина, а также тот факт, что эффект прайминга более ярко выражен в условиях, где заменяется последняя буква, а не буква в середине.

Кроме того, значимым оказалось взаимодействие частотности и формы прайма ( $F(1, 1743) = 18,03, p < 0,001$ ). Можно заметить, что в условиях, где прайм был неначальной формой, разница между группами с более частотными и менее частотными стимулами меньше, чем в тех условиях, где прайм был начальной формой. Другие взаимодействия значимости не достигли.

### 3. Заключение

Результаты проведенного нами эксперимента проливают свет на целый ряд проблем, связанных с лексическим доступом. В предыдущих исследованиях, посвященных праймингу со словами-соседами, был обнаружен отрицательный прайминг-эффект, если прайм частотней целевого слова, и никаких значимых эффектов в обратной ситуации. В нашем исследовании также наблюдается тенденция читать целевое слово дольше, чем в контрольном условии, если прайм-сосед частотней, однако она не достигает статистической значимости. При этом



в одном из условий, где прайм-сосед менее частотен, чем целевое слово, впервые был обнаружен значимый положительный прайминг-эффект. И в целом взаимодействие фактора условия и фактора соотношения частотности оказалось статистически значимым. Это указывает на то, что подавление кандидатов, активированных и затем отвергнутых в процессе лексического доступа, может зависеть от соотношения частотностей этих кандидатов и целевого слова.

Роль места замены практически не исследована в существующей литературе. При этом ряд экспериментов, в которых перед участниками ставились другие задачи, прежде всего, поиск буквы в слове, указывает на особую роль первой и последней позиции. Однако полученные нами данные свидетельствуют о том, что делать на этом основании выводы об отборе слов-кандидатов в процессе лексического доступа, возможно, преждевременно. Замена буквы в конце слова оказалась в нашем эксперименте менее заметной, чем замена в середине (иначе говоря, праймы с заменой в конце воспринимались как более похожие на целевые слова, за счет чего прайминг-эффект был больше). Наконец, то, что в условии со значимым эффектом прайминга праймы предъявлялись в начальной форме, а не в форме косвенного падежа, свидетельствует о том, что морфологическая декомпозиция форм существительных не только имеет место, но и происходит на очень раннем этапе, параллельно с распознаванием графического облика слова.

В заключение заметим, что, эксперименты, в которых используются слова-соседи, многочисленны, но проводятся на небольшом числе языков. Это во многом связано с тем, что для подбора стимулов необходимы специализированные лексические базы данных. Теперь, когда подобная база создана для русского языка, такие эксперименты возможно проводить и на русском, и, как показывает проведенное нами исследование, они могут существенно дополнить общую картину. Отметим также, что эффекты соседства проявляются не только при чтении, но и при написании слов, влияя на количество и характер опечаток [Roux, Bonin 2009], поэтому полученные результаты имеют и практическое значение.

Исследование выполнено при поддержке гранта РФФИ №16-18-02071.

## Литература

1. *Acha J., Carreiras M.* (2014), Exploring the mental lexicon: A methodological approach to understanding how printed words are represented in our minds, *Mental Lexicon*, Vol. 9, pp. 196–231.
2. *Alexeeva S. V., Slioussar N. A., Chernova D. A.* (2015), StimulStat: A Lexical Database for Linguistic and Psychological Research on Russian Language [StimulStat: baza dannyih, ohvatyvayuschaya razlichnyie harakteristiki slov russkogo yazyika, vazhnyie dlya lingvisticheskikh i psihologicheskikh issledovaniy], available at: [http://www.dialog-21.ru/digests/dialog2015/materials/pdf/AlexeevaSV\\_SlioussarNACHernovaDA.pdf](http://www.dialog-21.ru/digests/dialog2015/materials/pdf/AlexeevaSV_SlioussarNACHernovaDA.pdf).
3. *Alexeeva, S. V., Slioussar, N. A., Chernova, D. A.* (2016), Stimulstat: a database for linguistic and psychological studies on russian language, The Seventh International Conference on Cognitive Science. Abstracts [Sed'maja mezhdunarodnaja konferencija po kognitivnoj nauke: Tezisy dokladov], Svetlogorsk, pp. 23–24.

4. *Andrews S.* (1996), Lexical retrieval and selection processes: Effects of transposed-letter confusability, *Journal of Memory and Language*, Vol. 35, pp. 775–800.
5. *Baayen R. H., Dijkstra T., Schreuder R.* (1997), Singulars and plurals in Dutch: Evidence for a parallel dual route model, *Journal of Memory and Language*, Vol. 37, pp. 94–117.
6. *Beyersmann E., Duñabeitia J. A., Carreiras M., Coltheart M., Castles A.* (2013), Early morphological decomposition of suffixed words: Masked priming evidence with transposed-letter nonword primes. *Applied Psycholinguistics*, Vol. 34, pp. 869–892.
7. *Boudelaa S., Marslen-Wilson W. D.* (2010), Aralex: A lexical database for Modern Standard Arabic, *Behaviour Research Methods*, Vol. 42, pp. 481–487.
8. *Bouma, H.* (1973), Visual interference in the parafoveal recognition of initial and final letters of words, *Vision Research*, Vol. 13, No. 4, pp. 767–782.
9. *Christianson K., Johnson R. L., Rayner K.* (2005), Letter transpositions within and across morphemes, *Journal of Experimental Psychology*, Vol. 31, pp. 1327–1339.
10. *Cohen J. D., MacWhinney B., Flatt M. R., Provost J.* (1993), Psyscope: A new graphic interactive environment for designing psychology experiments, *Behavioral Research Methods*, Vol. 25, pp. 257–271.
11. *Coltheart M., Davelaar E., Jonasson T., Besner D.* (1977), Access to the internal lexicon. *Attention and Performance VI*, New York, Academic Press. pp. 535–555.
12. *Davis C. J.* (2005), N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics, *Behavior Research Methods*, Vol. 37, pp. 65–70.
13. *Davis C. J., Lupker S. J.* (2006), Masked inhibitory priming in English: Evidence for lexical inhibition, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 32, No. 3, pp. 668–687.
14. *Davis C. J., Perea M.* (2005), BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish, *Behavior Research Methods*, Vol. 37, pp. 665–671.
15. *Davis C. J., Perea M., Acha J.* (2009), Re(de)fining the orthographic neighborhood: The role of addition and deletion neighbors in lexical decision and reading, *Journal of Experimental Psychology*, Vol. 35, pp. 1550–1570.
16. *De Moor W., Brysbaert M.* (2000), Neighborhood-frequency effects when primes and targets are of different lengths, *Psychological Research*, Vol. 63, pp.159–162.
17. *Duchon A., Perea M., Sebastián-Gallés N., Martí A., Carreiras M.* (2013), EsPal: one-stop shopping for Spanish word properties, *Behavior Research Methods*, Vol. 45, pp. 1246–1258.
18. *Duñabeitia J. A., Perea M., Carreiras M.* (2007), Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition, *Cognition*, Vol.105, pp. 691–703.
19. *Duñabeitia J. A., Perea M., Carreiras M.* (2009), There is no clam with coats in the calm coast: Delimiting the transposed-letter priming effect, *The Quarterly Journal of Experimental Psychology*, Vol. 62, pp. 1930–1947.
20. *Gomez P., Ratcliff R., Perea, M.* (2008), The overlap model: A model of letter position coding, *Psychological Review*, Vol. 115, No. 3, pp. 577–600.
21. *Grainger J.* (1988), Neighborhood frequency effects in visual word recognition and naming. *IPO Annual Progress Report*, Vol. 23, pp. 92–101.

22. *Grainger J.* (1992), Orthographic Neighborhoods and Visual Word Recognition, *Advances in Psychology*, Amsterdam, North-Holland. pp. 131–146.
23. *Grainger J.* (2008), Cracking the orthographic code: An introduction, *Language and Cognitive Processes*, Vol. 23, pp. 1–35.
24. *Grainger J., Segui J.* (1990), Neighborhood frequency effects in visual word recognition: A comparison of lexical decision and masked identification latencies, *Perception & Psychophysics*, Vol. 47, No. 2, pp. 191–198.
25. *Johnson R. L., Perea M., Rayner K.* (2007), Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 33, pp. 209–229.
26. *Luke S. G., Christianson K.* (2013), The influence of frequency across the time course of morphological processing: Evidence from the transposed-letter effect, *Journal of Cognitive Psychology*, Vol. 25, pp. 781–799.
27. *Lyashevskaya O. N.* (2013), Lexico-grammatical frequency dictionary: A preliminary design [Chastotnyj leksiko-grammaticheskij slovar': prospect proekta], *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2013" [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii "Dialog 2013"]*, Bekasovo, pp. 478–489.
28. *Lyashevskaya O. N., Sharov S. A.* (2009), Frequency dictionary of modern Russian [Chastotnyj slovar sovremennogo russkogo jazyka], *Azbukovnik*, Moscow.
29. *Marslen-Wilson W. D., Tyler L. K.* (1997), Dissociating types of mental computation, *Nature*, Vol. 387, pp. 592–594.
30. *Masserang K. M., Pollatsek A.* (2012), Transposed letter effects in prefixed words: Implications for morphological decomposition, *Journal of Cognitive Psychology*, Vol. 24, pp. 476–495.
31. *McClelland J.L., Patterson K.* (2002), Rules or connections in past-tense inflections: what does the evidence rule out?, *Trends in Cognitive Sciences*, Vol. 6, pp. 465–472.
32. *New B., Pallier C., Brysbaert M., Ferrand L.* (2004), Lexique 2: A new French lexical database, *Behavior Research Methods*, Vol. 36, pp. 516–524.
33. *Perea M., Carreiras M.* (2006), Do transposed-letter effects occur across lexeme boundaries?, *Psychonomic Bulletin and Review*, Vol. 13, pp. 418–422.
34. *Pinker S.* (1991), Rules of language, *Science*, Vol. 253, pp. 530–535.
35. *Pinker S.* (1999), *Words and rules: the ingredients of language*, New York, Harper Collins, 352 p.
36. *Pollatsek A., Hyönä J., Bertram R.* (2000), The role of morphological constituents in reading Finnish compound words, *Journal of Experimental Psychology*, Vol. 26, pp. 820–833.
37. *Roux, S., Bonin, P.* (2009), Neighborhood effects in spelling in adults. *Psychonomic Bulletin and Review*, Vol. 16, pp. 369–373.
38. *Rueckl J. G., Rimzhim A.* (2011), On the interaction of letter transpositions and morphemic boundaries, *Language and Cognitive Processes*, Vol. 26, pp. 482–508.

39. Rumelhart D., McClelland J. (1986), On learning the past tenses of English verbs, *Parallel distributed processing: explorations in the microstructure of cognition*. Vol. 2. Cambridge, MA, MIT Press. pp. 216–271.
40. Sánchez-Gutiérrez C., Rastle K. (2013), Letter transpositions within and across morphemic boundaries: Is there a cross-language difference?, *Psychonomic Bulletin and Review*, Vol. 20, pp. 988–996.
41. Seidenberg M. S., McClelland J. L. (1989), A distributed, developmental model of word recognition and naming, *Psychological Review*, Vol. 96, pp. 523–568.
42. Segui J., Grainger J. (1990), Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 16 No. 1, pp. 65–76.
43. Taft M. (2004), Morphological decomposition and the reverse base frequency effect, *Quarterly Journal of Experimental Psychology*, Vol. 57A, pp. 745–765.
44. Taft M., Forster K. I. (1975), Lexical storage and retrieval of prefixed words, *Journal of Verbal Learning and Verbal Behavior*, Vol. 14, pp. 638–647.
45. Tydgat I., Grainger J. (2009), Serial position effects in the identification of letters, digits, and symbols, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 35, No. 2, pp. 480–498.
46. Ullman M. T. (2004), Contributions of memory circuits to language: the declarative/procedural model, *Cognition*, Vol. 92, pp. 231–270.
47. White S. J., Johnson R. L., Liversedge S. P., Rayner K. (2008), Eye movements when reading transposed text: The importance of word-beginning letters, *Journal of Experimental Psychology*, Vol. 34, pp. 1261–1276.
48. Whitney C., Lavidor M. (2004), Why word length only matters in the left visual field, *Neuropsychologia*, Vol. 42, pp. 1680–1688.
49. Zargar, E. S., Witzel, N. (2016), Transposed-Letter Priming Across Inflectional Morpheme Boundaries, *Journal of psycholinguistic research*, pp. 1–16.

# ОТРИЦАНИЕ В РУССКИХ И ФИНСКИХ ВЕРСИЯХ ДВУСТОРОННИХ ДОГОВОРОВ МЕЖДУ РОССИЕЙ И ФИНЛЯНДИЕЙ: ОПЫТ КОРПУСНОГО ИССЛЕДОВАНИЯ

**Соума Ю. В.** (julia.souma@uta.fi),  
**Кудашев И. С.** (igor.kudashev@uta.fi),  
**Михайлов М. Н.** (mikhail.mikhailov@uta.fi)

Университет Тампере, Тампере, Финляндия

## NEGATION IN RUSSIAN AND FINNISH VERSIONS OF THE STATE TREATIES BETWEEN RUSSIA AND FINLAND: A CORPUS-BASED RESEARCH

**Souma Ju. V.** (iuliia.souma@uta.fi),  
**Kudashev I. S.** (igor.kudashev@uta.fi),  
**Mikhailov M. N.** (mikhail.mikhailov@uta.fi)

University of Tampere, Tampere, Finland

Different language versions of bilateral state treaties are supposed to be equally authentic and have identical meaning. This belief looks problematic from the linguistic point of view, at least on the sentence level and below. In this article, we discuss the use of negation markers in the Russian and Finnish versions of the state treaties between Russia and Finland. The research material is the Russian-Finnish subcorpus of the PEST corpus (Parallel Electronic Corpus of State Treaties) that includes all treaties starting from 1917, the year of the Russian revolution and Finland's independence. It was found out that in the Russian versions of the documents explicit negation markers (particles, prepositions, etc.) are used much more frequently than in the Finnish ones, while implicit negation markers (verbs and nouns expressing prohibition, refusal etc.) are more typical for the Finnish versions. This phenomenon was discovered by means of comparing frequencies of translation equivalents in the Russian and Finnish versions. The results were confirmed by studying parallel concordances: numerous examples of replacing explicit negative markers by implicit ones were found. This phenomenon can be explained by the differences between the two languages, the differences in diplomatic discourse conventions, and possible prevalence of translating the treaties from Russian into Finnish over composing texts directly in Finnish. At the current stage of the research, it is not yet clear which of these factors have an impact on the choice of negation markers and to what degree.

**Key words:** negation, negation markers, Russian language, Finnish language, parallel corpora, language for special purposes, legal language, language of state treaties

— Урхо, а вы всегда говорите «да»?

— Нет, Леонид Ильич.

(Из советского анекдота о телефонном разговоре Л. И. Брежнева с президентом Финляндии У. К. Кекконеном)

## 1. Введение

Процесс составления текста договора между государствами — это многократное редактирование текста, перевод его с одного языка на другой в обе стороны и сверка версий текста на разных языках. Иногда нельзя понять, какая из версий была исходной, а какая — переводом [см. напр. Probirskaja 2009: 47]. Таким образом, все версии международных договоров в некоторой степени являются переводами. В конце договора, как правило, сообщается:

- а) какие документы являются аутентичными;
- б) какую из языковых версий следует использовать в случае возникновения разногласий.

С точки зрения лингвиста пункт (б) несколько противоречит фразе о равной юридической силе: видимо, тот текст, к которому надлежит обращаться при возникновении разногласий, все же имеет «большую» юридическую силу. Кроме того, эта оговорка является косвенным признанием факта, что в текстах соглашений, несмотря на все сверки, могут быть расхождения.

Версии официальных документов на разных языках после подписания представителями стран получают статус оригинальных текстов, независимо от того, является ли документ переводом с другого языка или был составлен непосредственно на этом языке.

Однако даже визуальное сравнение текстов договоров демонстрирует отсутствие полной симметрии между языковыми версиями. В примере (1) русская версия текста явно длиннее финской, и в обоих текстах имеется некоторая непоследовательность и несогласованность в записи числительных цифрами или словами. Имеется и небольшое содержательное расхождение: в русском тексте говорится о неограниченном числе *шестилетних периодов*, в финской же версии слово *kuusivuotiskausi* 'шестилетний период' употреблено в единственном числе. В русском тексте употребляются формы будущего времени, финский же текст составлен в настоящем времени, поскольку в финском языке и настоящее, и будущее обычно выражаются формами настоящего времени. Последняя грамматическая нестыковка может показаться неизбежной, однако это не совсем так. Например, русский текст можно было бы составить в настоящем времени.

- (1) *Настоящее Соглашение будет оставаться в силе в течение десятилетнего периода. В случае, если действие Соглашения не будет прекращено какой-либо Договаривающейся Стороной путем письменного уведомления не менее чем за 12 месяцев до истечения этого периода, оно будет оставаться в силе в течение дальнейших шестилетних периодов при условии, что письменное уведомление о прекращении его действия*

*не будет направлено за 12 месяцев до истечения любого такого шестилетнего периода.*

*Tämä sopimus on voimassa 10 vuotta. Ellei kumpikaan sopimuspuoli kirjallisesti irtisano sopimusta vähintään kaksitoista kuukautta ennen tämän ajanjakson päättymistä, se jää edelleen voimaan uudeksi kuusivuotiskaudeksi, ellei sitä kirjallisesti irtisanoita vähintään kaksitoista kuukautta ennen sellaisen kuusivuotiskauden päättymistä.*

(= “Этот договор действует 10 лет. Если ни одна сторона договора не расторгнет письменно договора не менее чем за двенадцать месяцев до окончания этого временного периода, он остается по-прежнему в силе на новый шестилетний период, если его не расторгнут письменно не менее чем за двенадцать месяцев до окончания такого шестилетнего периода”)<sup>1</sup>

(Соглашение о взаимных отношениях в области рыболовства, 1981)

В данной статье рассматривается выражение отрицания в русских и финских версиях межгосударственных договоров между Россией и Финляндией. Мы собираемся выяснить, насколько симметрично выражаются в них сообщения об отсутствии, несогласии, отказе и запрете. Нашей отправной точкой является бытующее со времен Сепира и Уорфа представление о том, что тексты на разных языках не могут быть идентичными по содержанию, даже если в них закодировано одно и то же сообщение.

Отрицание выбрано в качестве предмета исследования в частности потому, что именно оно может оказаться проблематичным местом в договоре. Для решения проблемы одна из сторон может пойти на компромисс и сделать отрицание менее категоричным. В примере (2) в русском варианте договора выбран «мягкий способ» отрицания с помощью *чтобы не*, в финском же тексте оно выражено более категорически с помощью глагола *kieltää* ‘запрещать’.

- (2) *В частности же должно ставить рыболовную снасть так, чтобы она не пересекала другую отмеченную надлежащим образом снасть, или не была столько близко к ней, что она может быть повреждена или что рыбной ловле может быть причинен иногда ущерб.*

*Erytisesti on kielletty asettamasta pyydyksiä asianmukaisella tavalla merkittyjen pyydysten yli tai niin lähelle niitä, että ne voisivat vahingoittaa tai että niillä pyytäminen muulla tavoin tulisi kärsimään.*

(= “Отдельно **запрещено** устанавливать рыболовные снасти через надлежащим образом отмеченные снасти или так близко от них, что они могли бы быть поврежденными или ловля с помощью них могла бы пострадать”)

(Соглашение о порядке производства рыбного промысла, 1922)

<sup>1</sup> Финские примеры снабжаются подстрочными переводами, в которых по возможности сохранены лексико-грамматические особенности текста, однако для облегчения понимания может использоваться более подходящая по контексту лексика, а отсутствующие в русском языке грамматические формы заменяются на ближайшие русские аналоги. Возможны небольшие изменения порядка слов. Точное глоссирование с указанием грамматических форм представляется для таких длинных примеров неудобным.

Выражение отрицания в текстах договоров связано с их регламентирующим и предписывающим характером. Категоричность или дипломатичность выражения запрещения, недопустимости или невозможности чего-либо оказывает влияние на общую тональность текста договора.

## 2. Материал и методы

Настоящее исследование выполнено на материале параллельного корпуса межгосударственных договоров PEST (Parallel Electronic corpus of State Treaties), работа над которым ведется в Университете Тампере с 2015 г. В данной статье используется русско-финский раздел корпуса PEST.

В этом разделе 222 пары документов, объем русского субкорпуса около 300 000 словоупотреблений, финского — около 200 000 словоупотреблений. Тексты выровнены на уровне предложений, лемматизированы и содержат морфологическую и синтаксическую аннотацию. Обработка русских текстов выполнялась с помощью пакета TreeTagger (<http://corpus.leeds.ac.uk/mocky/>, см. тж. Sharoff & Nivre 2011), финских текстов — с помощью Finnish-dep-parser (<http://turkunlp.github.io/Finnish-dep-parser/>). По каждому тексту доступны метаданные (название, год подписания, тематика и т.п.).

Для работы с корпусом используется корпус-менеджер TextHammer (<https://mustikka.uta.fi/texthammer>), предназначенный для работы с параллельными корпусами текстов. Программа позволяет работать с выровненными аннотированными текстами и получать из них конкордансы, частотные списки, списки коллокаций и т.п.

Исследование выполняется количественными методами. Вначале анализируются лемматизированные частотные списки, полученные из корпуса текстов. Это позволяет выявить общую тенденцию и наиболее характерные языковые средства. Однако, оперируя только данными частотных списков, нельзя делать однозначные утверждения из-за многозначности слов и ошибок лемматизатора. Поэтому кроме частотных списков привлекаются конкордансы и коллокации.

## 3. Анализ частотных списков

Вслед за Е. В. Падучевой следует признать, что «отрицание в языке принадлежит к числу исходных, не толкуемых понятий. Оно входит в список семантических примитивов» (RKG 2011). Отрицание противопоставляется утверждению как маркированный член оппозиции и означает отсутствие или противопоставление, выраженное при помощи лексических, словообразовательных и синтаксических языковых средств. В таблице 1 представлен репертуар наиболее типичных средств выражения отрицания в русском и финском языках (за основу взяты RKG 2011 и Miestamo et al. 2015).



**Таблица 1.** Основные средства выражения отрицания в финском и русском языках

	Русский язык	Финский язык
Частицы	не, ни	—
Служебные глаголы	—	Отрицательный глагол ei (en, et, ei, ...)
Глагольные формы	—	3-й инфинитив глагола в абессиве (mA-ttA) Отрицательные причастия (mA-tOn)
Предикативы и модальные слова	нет, нельзя	-
Отрицательные местоимения и наречия	никто, ничто, никогда, незачем ...	(ei) kukaan, mitään, kumpikaan ... ensinkään, lainkaan, ollenkaan, yhtään
Предлоги	без	ilman, vailla/vaille
Падежные формы	—	абессив существительных (ttA)
Аффиксы	префиксы не-, без-, бес-, а-	суффикс -tOn, префиксы epä- и ei-

Отрицание может выражаться эксплицитно, т.е. с помощью специализированных языковых средств, и имплицитно, т.е. с помощью слов, в семантику которых входит элемент отрицания (например, *отказаться, воздержаться, запретить, запрет* и др., *kieltää* ‘запрещать, отрицать’, *evätä* ‘отклонять’ и др.). Коммуникативные акты с отрицанием могут быть выражены с помощью как эксплицитных, так и имплицитных средств выражения отрицания. Например, *Я не разрешаю вам приближаться* и *Я запрещаю вам приближаться* (см. т.ж. пример (6) в данной статье).

Для того, чтобы получить общее представление о частотности отрицания в русских и финских версиях документов, мы обработали лемматизированные словники, полученные из корпуса. Для анализа мы взяли первые 2000 слов лемматизированных словников для обоих языков. Были отмечены все леммы с эксплицитно выраженным отрицанием, например, в русском списке *без, безопасность, не, никакой* и др., а в финском *ei* ‘не’, *ilman* ‘без’, *irti* ‘отдельно’, *laiton* ‘незаконный’, *suolaton* ‘несоленый’ и др. Кроме того, по финскому корпусу была отдельно получена частота употребления формы абессива (253 вхождения). Результаты представлены в табл. 2.

**Таблица 2.** Эксплицитные и имплицитные средства выражения отрицания в текстах корпуса PEST

	Русский язык	Финский язык
Эксплицитные средства	3107	2059
Имплицитные средства	683	796
<b>Всего:</b>	<b>3790</b>	<b>2855</b>

Поскольку тексты являются параллельными, мы пользуемся абсолютными, а не относительными частотами. Числа показывают количество найденных маркеров, а не предложений, поскольку в одном предложении может встретиться несколько маркеров отрицания. Хотя далеко не все найденные слова действительно являются маркерами отрицания, в табл. 1 налицо довольно существенная разница в частоте эксплицитных маркеров отрицания. При этом, хотя в финских текстах немного больше случаев употребления имплицитных средств, это далеко не компенсирует разрыв в частоте употребления эксплицитных средств. Ожидаемого равенства количества отрицаний в русских и финских версиях не наблюдается, что говорит как о вероятных неточностях подсчетов случаев имплицитного отрицания, так и о дублировании выражения отрицания, и даже о возможной неполной идентичности версий.

Сравним теперь выборочно частоты лексических соответствий<sup>2</sup>.

**Таблица 3.** Сравнение частот русско-финских соответствий по корпусу

Русские слова	Частоты	Финские слова	Частоты
без	132	ilman	68
не	1584	ei, ettei, jollei	1352
запретить	17	kieltää	46
отказать, отказаться, отказываться	57	kieltäytyä	27
прекратить, прекращать, прекращаться, прекращение	147	keskeyttää, lopettaa, peruuttaa, lakata	122

Табл. 3 наглядно демонстрирует «сглаживание углов» в финских текстах. По всем группам слов таблицы, кроме одной (*запретить* — *kieltää*) частоты в русских текстах значительно выше. С чем могут быть связаны эти отличия? Могут ли они быть целиком обусловлены языковыми различиями? Или причина в большей «дипломатичности» финского дискурса, в стремлении избегать резких оценок? Или договоры являются отражением «финляндизации»<sup>3</sup>?

Чтобы сказать что-то более определенное, перейдем к конкретным примерам употребления: к результатам выборочного просмотра конкордансов.

#### 4. Анализ конкордансов

Как правило, если в одном тексте есть отрицание, то оно есть и во втором. Только в одном языке оно может быть имплицитным. Приведем несколько примеров замены эксплицитного отрицания на имплицитное.

<sup>2</sup> В список вошли только слова, имеющиеся в наших частотных списках, поэтому могут встречаться глаголы без видовых пар или отсутствовать очевидные соответствия.

<sup>3</sup> Финляндизация — термин политологии, применяющийся для описания ситуаций, при которых более слабое государство вынуждено идти на уступки более сильному соседу для сохранения суверенитета.

#### 4.1. В русском языке — эксплицитное отрицание, в финском — имплицитное

- (3) <...> автобус, то есть автотранспортное средство, предназначенное для перевозки пассажиров и имеющее **не менее** 9 мест для сидения, не считая места водителя.

<...> linja-autoa eli moottoriajoneuvoa, joka on tarkoitettu matkustajien kuljetukseen ja jossa on **vähintään** 9 istumapaikkaa kuljettajan istuimen lisäksi.

(= “автобус, или автотранспортное средство, которое предназначено для перевозки пассажиров и в котором **минимум** 9 мест для сидения в дополнение к месту водителя”).

(Соглашение о внесении изменений в соглашение об автомобильном сообщении, 2014)

В текстах соглашений *vähintään* (96) чаще всего соответствует русскому *не менее* (54) и *не позднее* (17), другие соответствия: как минимум, не позднее, и более, не более, по крайней мере, не ниже, не выше, по меньшей мере. Большая часть русских соответствий содержит эксплицитное отрицание.<sup>4</sup>

- (4) **Никакие** таможенные и иные, <...>, налоги или сборы **не могут взиматься** с граждан договаривающихся государств в связи с ввозом в воды другого договаривающегося государства или вывозом из последних, уловов, <...> и всего относящегося к промыслу.

Viedessään toisen sopimusvaltion vesille tai sieltä tuodessaan saalista, <...> ja kaikkia muita pyynnissä tarvittavia esineitä sopimusvaltioiden kansalaiset ovat vapautetut kaikista tulli- ja muista <...> maksuista ja suorituksista.

(= “Ввозя в воды другого договаривающегося государства или вывозя из них улов <...> и все другие необходимые для лова предметы, граждане договаривающихся государств **освобождены** от всех таможенных и других <...> налогов и платежей”).

(Конвенция о рыбном и тюленьем промысле, 1934)

В финском тексте отрицание выражается имплицитно, с помощью глагола *vapauttaa* ‘освободить’, а в русской версии два маркера отрицания: при именной группе и при модальном глаголе. В русской версии выражается запрет, в финской — разрешение, русский текст ограничивает произвол чиновника, финский — дает свободу гражданину.

#### 4.2. В русском языке отрицание имплицитное, в финском — эксплицитное

- (5) При перевозке автомашин, <...> отправителем должны быть закрыты на защелки или замки и опломбированы двери и люки кабины, <...>.

<sup>4</sup> Финско-русские словари указывают большое количество соответствий для этого финского слова, причем не менее не является первым соответствием. В финском языке возможны и другие соответствия приведенным русским словам, например, не позднее — viimeistään, более — enemmän, yli и т.д. (см. BFRS 2007).

а в случае **свободного** доступа и аккумулятор, запасное колесо, ящики с инструментами и запасными частями.

*Kuljetettaessa <...> autoja, <...> lähettäjän on lukittava ja sinetöitävä korin ovet ja kattoluukut, <...> sekä myös akku, varapyörä ja työkaluja varaosalaatikoit, jos niihin on esteetön pääsy.*

(=“При перевозке <...> автомобилей <...> отправитель должен закрыть и опломбировать двери и люки кабины, <...> а также аккумулятор, запасное колесо и ящики с инструментами и запасными частями, если к ним есть **беспрепятственный** доступ”)

(Условия перевозок в железнодорожном сообщении, 2005)

Финскому слову *esteetön* в русских текстах соответствует *беспрепятственный* во всех случаях, кроме приведенного выше (16). Соответствие *esteetön* — *беспрепятственный* представляется более логичным, поскольку эти слова не только совпадают по значению, но и являются аналогами с точки зрения словообразования.<sup>5</sup>

- (6) *Работы по устройству сплавного пути, <...>, допускаются лишь на основании особаго<sup>6</sup> заключенного о сем договаривающимся государствами соглашения.*

*Sellaisia lauttausväylän kuntoonpanotöitä, <...>, ei sallita ilman asianomaisen sopimusvaltion kanssa tehtävää eri sopimusta.*

(=“Такие работы по устройству сплавного пути <...> **не разрешаются без** заключенного с соответствующей договаривающейся стороной отдельного соглашения”)

(Конвенция о сплаве лесных материалов, 1922)

В финском тексте два отрицания, которые почти дословно могут быть переведены на русский язык, при этом смысл высказывания не меняется. Но в русском тексте выбрана более мягкая форма запрета. Русский текст не запрещает работы, а разрешает их выполнение при определенных условиях.

### 4.3. Случаи, когда в одном из языков отрицание пропущено

В корпусе есть случаи, когда в финском тексте слово с отрицанием пропущено.

- (7) *Если от отправителя, <...>, не поступит никаких выполнимых указаний <...>*

*Jos lähettäjältä, <...>, ei saada toteuttamiskelpoisia ohjeita <...>*

(= “Если от отправителя <...> **не получено** выполнимых указаний <...>”).

(Условия перевозок в железнодорожном сообщении, 2005)

<sup>5</sup> *Esteetön* — *беспрепятственный*; (*vara*) — *свободный*; (*aukea*) — *открытый*; ~ *näköala* — *открытый вид* (BFRS 2007: 78).

<sup>6</sup> В текстах некоторых договоров, вплоть до 1947 года, иногда встречаются описки, связанные с реформой орфографии 1918 года. Описки носят несистематический, случайный характер.

Местоимение *никакой* опущено в финском тексте, причем этот пропуск не единичный: из 31 контекста с *никакой* пропуск в финских версиях наблюдается в 10 случаях.<sup>7</sup>

Пропуск слов в одной из версий договоров не редкое явление. Как правило, это слова, не вносящие в текст ничего нового. Во всяком случае, такие пропуски не ведут к разночтениям. Но это еще раз доказывает нашу мысль о том, что тексты договоров не являются полностью идентичными по содержанию, хотя это и предполагается.

Следующий случай отсутствия отрицания, по-видимому, связан с терминологической лакуной в финском языке.

(8) *Неядерный материал.*

*Miut aineet.*

(= “Другие вещества”).

(Соглашение о сотрудничестве в области использования атомной энергии, 2014).

В текстах договора даны определения всем употребляемым понятиям и терминам, поэтому такое расхождение не является принципиальным.

## 5. Заключение

Проведенный анализ показал большую частоту отрицания в русских текстах по сравнению с финскими. Проверка параллельных конкордансов позволила найти конкретные случаи замены эксплицитного отрицания на имплицитное и полного исчезновения отрицания.

Различия могут быть в некоторой степени обусловлены особенностями языков. Так, у финского слова *vähintään* ‘не менее’ (пример 3), являющегося имплицитным средством выражения отрицания, большое количество русских соответствий, в которых есть эксплицитный маркер отрицания. Есть и другие аналогичные случаи. Не значит ли это, что в финском языке тенденция к имплицитному выражению отрицания сильнее, чем в русском языке?

В финско-русских соглашениях сообщения об отсутствии, несогласии, отказе и запрете не всегда выражаются симметрично, даже если в языке контрагента есть точное соответствие (пример 6). Пример 4 демонстрирует, что одна и та же ситуация может описываться с разных точек зрения (например, разрешение vs. запрет). Это может быть связано с различиями в дипломатической практике в России и Финляндии, однако для подтверждения этой гипотезы необходим фронтальный анализ всего корпуса.

В некоторых случаях выбор языкового средства в одном из языков нельзя объяснить ни языковыми различиями, ни контекстом, ни стремлением избежать резких оценок. Так, в примерах (5) и (7) при наличии стандартных

<sup>7</sup> Ср. в том же договоре: <...> и не возлагают никаких обязательств и никакой ответственности на железную дорогу. — ja jotka eivät aseta rautatielle mitään velvollisuuksia tai vastuuta (= «<...> и которые не возлагают на железную дорогу никаких обязательств и ответственности»)

лексических и грамматических соответствий приняты единичные нестандартные решения. Вполне возможно, что это лишь результат того, что в основу разных языковых версий легли разные редакции документа, и это не заметили при сверке окончательных версий.

В ходе дальнейшего исследования планируется выяснить, какой из факторов больше влияет на тенденцию к имплицитному выражению отрицания в финских версиях документов: особенности финского языка и дипломатического дискурса или перевод с русского языка.

## Исследовательский материал

1. PEST, Parallel Corpus of State Treaties. University of Tampere, 2016. URL: [mustikka.uta.fi/texthammer](http://mustikka.uta.fi/texthammer) (contact: [mikhail.mikhailov@staff.uta.fi](mailto:mikhail.mikhailov@staff.uta.fi)).

## Литература

2. BFRS. *Vahros I., Šerbakov A.* (2007). Finnish-Russian Dictionary [Bol'shoj finsko-russkij slovar']. Moskva: Živoj âzyk. — 816 p.
3. *Bondarenko V. N.* (1983). Negation as logical-grammatical category. [Otricanie kak logiko-grammatičeskaâ kategoriâ]. Moskva.
4. Grammatika (1958). Finnish Grammar. Phonetics and Morphology. [Grammatika finskogo âzyka. Fonetika i morfologiâ.] Moskva-Leningrad, Izdatel'stvo Akademii nauk SSSR.
5. *Mattila H. E. S.* (2002) Comparable Legal Linguistics [= Vertaileva oikeuslingvistiikka]. Kauppakaari, Helsinki.
6. *Miestamo M., Tamm A., Wagner-Nagy B.* (eds) 2015. Negation in Uralic Languages. Amsterdam/Philadelphia, John Benjamins.
7. *Padučeva E. V.* (2011) Negation. [Otricanie]. Moskva. Available at: <http://rusgram.ru/Отрицание>.
8. *Probirskaja S.* (2009). Across the Border: Finnish and Russian State Treaties in the Light of Keyword Analysis [= Rajankäyntiä: Suomen ja Venäjän kahdenväliset valtiosopimukset käänntieteellisen avainsana-analyysin valossa]. Tampere, Tampere University Press.
9. RKG 2011. Russian corpus grammar [Russkaâ korpusnaâ grammatika] Moskva. Available at: <http://rusgram.ru/>
10. *Sharoff S. & Nivre J.* (2011). The proper place of men and machines in language technology: Processing Russian without any linguistic knowledge, Computational Linguistics and Intellectual Technologies, Papers from the Annual International Conference “Dialogue” (2011), Issue 10 [Komputernaja lingvistika i intellektual'nye tekhnologii: Po materialam Mezhdunarodnoj konferencii “Dialog” (Bekasovo, 25–29 maja 2011)], pp. 591–604. Available at: <http://www.dialog-21.ru/media/1459/58.pdf>
11. VISK (2008) = The network version of the Large Grammar of Finnish [= Iso suomen kieliopin verkkoversio]. Kotus. Available at: <http://scripta.kotus.fi/visk/etusivu.php>.
12. WALs. The World Atlas of Language Structures. Available at: <http://wals.info/>

# ПОРЯДОК СЛЕДОВАНИЯ ПРИЛАГАТЕЛЬНЫХ РАЗНЫХ СЕМАНТИЧЕСКИХ КЛАССОВ В РУССКОМ ЯЗЫКЕ В СВЕТЕ КОРПУСНЫХ ДАННЫХ<sup>1</sup>

**Толдова С. Ю.** (toldova@yandex.ru),

**Муханова Р. В.** (muhanova.regina@mail.ru)

Национальный исследовательский университет  
Высшая Школа Экономики, Москва, Россия

Работа посвящена исследованию порядка следования прилагательных в русском языке. Исследование проводилось на материале Национального корпуса русского языка. За основу взята иерархия порядка прилагательных, предложенная в работе Чинкве [Cinque 1994]: 'посессивные' > 'количественные' > 'порядковые' > 'оценка' > 'размер' > 'форма' > 'цвет' > 'национальность'. В ходе исследования был разработан метод построения иерархии, описывающей допустимые порядки, который основывается на попарном сравнении семантических классов прилагательных по доминирующей частоте следования прилагательных в соответствующей паре. В итоге, исследование выявило тенденцию к тому, что линейаризация прилагательных в реальных текстах соответствует иерархии за исключением класса посессивных прилагательных. Этот класс прилагательных делится на два: 'конкретно-референтные' (на *-ин, -ов: папин, дедов*) и 'генерические' (на *-ий, -ский: лисий*). 'Генерические' располагаются ближе к существительному и ведут себя как относительные прилагательные. Конкретно-референтные прилагательные занимают не крайнюю левую позицию, как в иерархии Чинкве, а располагаются ближе к относительным прилагательным — между классом 'цвет' и классом 'национальность'. Такой результат служит подтверждением теоретических выводов относительно безартиклевых языков: в них прилагательные, соответствующие классу определителей (артиклей, посессивных конструкций и т.п.), приобретают детерминативную функцию в крайне-левой позиции в именной группе, в других позициях они ведут себя как относительные прилагательные. Анализ примеров, нарушающих ожидаемый порядок, показал, что такие нарушения могут быть связаны с семантическим сдвигом (коэрсией) в значении прилагательного, а также с образованием устойчивых словосочетаний. Контексты с нарушением порядка следования прилагательных могут быть использованы при оптимизации процедуры полуавтоматического снятия семантической неоднозначности, а в некоторых случаях могут служить одним из признаков, помогающих детектировать ошибки в семантической разметке.

**Ключевые слова:** порядок прилагательных, семантическая разметка, корпусные исследования

---

<sup>1</sup> В данной научной работе использованы результаты проекта «Точки роста в языковой системе: корпусные исследования и моделирование», выполненного в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» в 2017 году.

# MULTIPLE PRENOMINAL ADJECTIVES ORDERING IN RUSSIAN: A CORPUS STUDY

**Toldova S. Yu.** (toldova@yandex.ru),

**Mukhanova R. V.** (muhanova.regina@mail.ru)

National Research University Higher School of Economics, Moscow

The work deals with the multiple prenominal adjectives ordering within a DP/ NP in Russian. The aim is to examine the claim concerning word order hierarchy for adjectives suggested in [Cinque 1994] via quantitative analysis of corpus data from the Russian National Corpus (RNC). The following hierarchy was checked: possessive > quantity > order > quality > size > shape > colour > nationality. We use the RNC semantic annotation taxonomy for the semantic class of an adjective. We checked the pairwise orders for different semantic classes (e.g. possessive > colour vs. colour > possessive). Our data confirm the claim (suggested in the experimental studies) that the possessive adjectives should be divided into two classes: 'referential' possessives (e.g. adjectives with suffix *-in*, *Petin* 'Peter's') and so called 'generic' possessives (e.g. *chicken breast*). The latter occupy the position closer to the noun. There is one case of significant hierarchy violation. The referential possessive adjectives occupy the intermediate position between gradable adjectives (namely, after 'colour') and ungradable ones ('nationality' class). Hence, the determiner-like possessives occupy not the left-most position as in languages with articles. These findings serve the additional evidence in favor of theoretical claim that determiner-like adjectives in articleless languages acquire determiner function only in the left-most position, otherwise they function as other adjectives. The corpus data confirm the claim that the change of the adjective structural position induces its semantic type coercion. The detection of expected word-order violations can be helpful in word sense disambiguation and for error detection in RNC semantic annotation.

**Key words:** adjective order restrictions, word order hierarchy, corpus study, semantic annotation

## 1. Введение

Настоящая работа посвящена ограничениям, действующим при линеаризации неоднородных прилагательных в именной группе в русском языке. Для многих языков порядок следования нескольких прилагательных перед вершинной именной группы не является произвольным (ср., например, *красивый стеклянный шар* vs. *стеклянный красивый шар*). Проблема линеаризации прилагательных (в англоязычной литературе часто используется сокращение «AOR» — 'adjective ordering restrictions') имеет достаточно длительную историю изучения (ср. [Vendler 1968], [Quirk et al 1972]). Несмотря на большое количество исследований в данной области как в рамках отдельных языков, так и в типологической перспективе, эта проблема до сих пор вызывает споры. Общеизвестным является тот факт, что ограничения на порядок прилагательных связаны



с их семантическими свойствами (ср., например, иерархии, предложенные в [Quirk et al 1972], [Dixon 1977], [Scott 2002] и т.д.). Однако сам набор семантических классов, в терминах которых формулируются ограничения, а также порядок классов, варьируются в зависимости от исследуемых языков, теоретической базы исследования, методов исследования. Одним из центральных является вопрос о том, возможно ли предложить универсальную иерархию семантических классов прилагательных, а также в какой степени предлагаемые иерархии работают в языках со свободным порядком слов таких, как русский.

Существование ограничений на порядок прилагательных ставит перед синтаксической теорией следующий вопрос: если прилагательные занимают позицию адъюнктов в синтаксическом дереве, чем можно объяснить ограничения на их линейный порядок. Наличие таких ограничений способна объяснить гипотеза Г. Чинкве [Cinque 1994], согласно которой прилагательные занимают позиции спецификаторов соответствующих функциональных проекций в именной группе. Она претендует на универсальность. Согласно Чинкве, имена прилагательные с определенной семантикой соотносятся с соответствующей проекцией. Считается, что в языках со свободным порядком слов, в русском языке, например, порядок менее строгий. Однако целый ряд исследований на материале русского языка показывает, что порядок прилагательных несвободный.

До сих пор большинство исследований базировалось на экспериментальных данных. Настоящее исследование имело целью проверить, соответствуют ли выявленные в экспериментах ограничения на порядок прилагательных ограничениям, наблюдаемым в реальных текстах; действительно ли предпочтения или ограничения в реальных текстах соответствуют предлагаемым в работах иерархиям семантических классов прилагательных.

Нас интересовали следующие вопросы:

- выявляются ли на материале корпусных данных те же тенденции и ограничения, что и полученные экспериментальным путем (при конструировании примеров с фиксированным множеством прилагательных);
- каким образом можно по корпусным данным пытаться строить иерархию с учетом того, что цепочки длиннее трех прилагательных не очень частотны в корпусе;
- существуют ли какие-то принципиальные расхождения между теоретическими ожиданиями и реальными данными корпуса;
- если расхождения существуют, то чем их можно объяснить (многозначностью прилагательных, ошибками семантической разметки, объективными расхождениями между теоретическими ожиданиями и иерархией, релевантной для реальных текстов);
- дают ли эти расхождения что-то новое для теории и/или для решения каких-то практических задач.

Материалом послужили данные Национального корпуса русского языка (НКРЯ). За основу были взяты классы определений, рассматриваемые в работе Г. Чинкве [Cinque 1994], дополненные двумя морфосемантическими посесивными классами, предложенными в работе Москалевой [Москалева 2010].

Иерархия Г. Чинкве учитывала также порядковые и количественные числительные. Итоговый набор классов следующий: количественные, порядковые, размер, цвет, форма, национальность (происхождение), посессивные ‘конкретно-референтные’ и посессивные ‘генерические’. Эти классы были соотнесены с таксономическими классами прилагательных, выделяемых в семантической разметке НКРЯ. Порядок разных классов определялся по частоте их попарной линейной встречаемости в препозиции к существительному-вершине. Например, сравнивалась встречаемость прилагательных классов ‘цвет — размер’ и ‘размер — цвет’.

Мы учитывали следующие обстоятельства. Во-первых, при автоматическом поиске соответствующих пар прилагательных возможна достаточно высокая доля контекстов, ошибочно отнесенных к искомому множеству. Семантическая разметка НКРЯ далеко не всегда достаточно точна, возможно ошибочное приписывание прилагательного к некоторому классу. Для отсеивания подобных случаев мы использовали «ручную» фильтрацию. Во-вторых, в НКРЯ семантические теги приписаны не всем прилагательным. В-третьих, в НКРЯ не снята семантическая неоднозначность.

Таким образом, расчеты, полученные нами в ходе эксперимента, учитывая несовершенство корпусного поиска и несовершенство семантической разметки, не являются точными, а скорее отражают некоторые тенденции. Несмотря на все вышеперечисленные оговорки, в общем и целом, наше исследование подтвердило ожидаемую иерархию. Однако, выявленные в ходе эксперимента нарушения порядка оказались значимыми.

## 2. Семантические шкалы

В целом ряде работ правила линеаризации формулируются в терминах иерархии семантических классов прилагательных (ср., например, [Vendler 1968], [Quirk et al 1972], [Dixon 1977], [Dixon 1982] (на данных 17 языков), [Sproat&Shih 1991] и многие другие. Иерархии варьируются не только по выделяемым параметрам (ср. внутренне ориентированные vs. внешне ориентированные прилагательные), но и степенью дробности выделяемых классов.

В рамках генеративной грамматики существенным моментом является противопоставление именной группы (NP) группе определителя (DP): в артиклевых языках именная группа требует при себе элемента детерминации (например, артикля), т.е. явного указания на референциальные свойства составляющей. Также это могут быть посессивные местоимения и посессивные именные группы (например, в английском *Peter's*, в русском этой конструкции могут соответствовать посессивные прилагательные на *-ин Машин*). Во многих языках посессивные конструкции (включая прилагательные) располагаются в крайне-левой позиции в именной группе. Так, в работе Г. Чинкве ([Cinque 1994]) на материале итальянского языка для прилагательных при предметных существительных предлагается следующая иерархия (i):

- (i) ‘possessive’ > ‘cardinal number’ > ‘ordinal number’ > ‘quality’ > ‘size’ > ‘shape’ > ‘colour’ > ‘nation’

Что касается славянских языков, то большинство из них безартиклевые. До сих пор ведутся споры, следует ли в таких языках выделять группу определителя (DP). В них порядок прилагательных более свободный, однако, и в этих языках проявляются тенденции к определенному порядку прилагательных при их линеаризации. Существующие работы на материале русского языка, такие как [Pereltsvaig 2007], [Москалева 2010], а также на материале польского языка [Zielinska 2007] это подтверждают. В работе А. Перельцвайг [Pereltsvaig 2007] описывается эксперимент с носителями русского (34 респондента) и английского (26 респондентов) языков. Испытуемым предлагались словосочетания, где высокочастотные прилагательные из разных семантических классов располагались в разном порядке. Исследование показало, что в обоих языках есть ограничения на линейное расположение прилагательных. В работе Москалевой отдельно рассматривался класс посессивных прилагательных на *-ин*, *-ов* (конкретно-референтные прилагательные) и класс 'генерических' посессивных прилагательных, например, прилагательные на *-ий* (*птичий*, *кошачий* и т.п.)

Опираясь на результаты экспериментов, описанных в работе Перельцвайг и Москалевой, мы взяли за основу классификацию Чинкве, разделив класс посессивных прилагательных на два. Более детально процедура выделения различных семантических классов прилагательных будет описана ниже.

### 3. Данные. Постановка задачи

Как было сказано выше, в наши задачи входило проверить гипотезу соответствия иерархии классов прилагательных Г. Чинкве [Sinque 1994] на корпусных данных. Мы учли девять семантических классов модификаторов к существительному: 'конкретно-референтные', 'количественные', 'порядковые', 'оценка', 'размер', 'форма', 'цвет', 'национальность' (происхождение), 'генерические'. Исследование проводилось на материале Национального корпуса русского языка.

Цепочки из трех прилагательных и более встречаются в корпусе не так уж часто. Если примеры с какой-то последовательностью классов в длинных цепочках не представлены в корпусе, то это не дает гарантии того, что такой порядок недопустим. В связи с этим, мы рассмотрели различные попарные комбинации порядков семантических классов прилагательных. Например: 'форма' > 'цвет' vs. 'цвет' > 'форма'. Нас интересовали цепочки длиной в два прилагательных в препозиции к определяемому имени существительному:

- (1) *Я пошёл на звук. Обнаружил квадратную чёрную яму. Лёг на снег и осторожно заглянул вниз.* [Сергей Довлатов. Наши (1983)]
- (2) *При вечернем свете она казалась чересчур мрачной со своей выдавшей виды мебелью, черным круглым столом...* [Николай Тихонов. Белое чудо // «Огонек» №№ 18, 23, 1956] (НКРЯ)

Для каждой пары семантических классов определялся предпочтительный порядок прилагательных в цепочке. Конечная иерархия семантических классов строилась на основе таких предпочтений.

### 3.1. Семантическая разметка прилагательных

Для определения класса прилагательного мы использовали семантическую разметку НКРЯ. Классы 'цвет', 'форма', 'размер' и 'национальность' входят в таксономические классы прилагательных НКРЯ. Для классов 'количественный' и 'порядковый' мы использовали непосредственно семантическую разметку НКРЯ для числительных.

Что касается притяжательных прилагательных, то в семантической разметке НКРЯ в этот класс попадают прилагательные с разными референциальными свойствами: морфосемантические классы генерических и конкретно-референтных притяжательных прилагательных. Морфологическими признаками конкретно-референтных прилагательных являются суффиксы *-ин, -ов*, они отсылают к конкретному лицу, как, например: *Матрениными родными сестрами; бабушкин родной брат* (но: *пустое бекасиное гнездо*). Генерические прилагательные оканчиваются на *ий/ский*, имеют значение «свойственный / принадлежащий тому, кто назван мотивирующим словом»). Например, *с желтым верблюжьим лицом, чистый девичий голос*. Соответствия классов приведены в табл. 1.

**Таблица 1.** Соответствие классов НКРЯ семантическим классам прилагательных в иерархии Чинкве и исследовании Москалевой

тег	Класс	пример
card	количественное	<i>три</i>
ord	порядковое	<i>второй</i>
size	размер	<i>большой</i>
ev	оценка	<i>отвратительный</i>
form	форма	<i>круглый</i>
poss	конкретно-референтное (посессивное с суффиксом <i>-ин</i> )	<i>папин</i>
color	цвет	<i>белый, темный</i>
etn	национальность (происхождение)	<i>английский</i>
generic	генерическое (посессивное на <i>-ий, -ский</i> )	<i>кошачий</i>

### 3.2. Предобработка данных

Для каждой пары семантических классов прилагательных и каждого порядка этих классов (всего 72 пары) была получена выдача из корпуса, упорядоченная случайным образом. На основе выдачи формировался подкорпус примеров, проверенных вручную. Если в выдаче было более 500 примеров, то общее количество примеров, соответствующих данному порядку, оценивалось по контрольной выборке из 500–600 примеров. Число «хороших» примеров на 500 найденных проецировалось на общее количество найденных в НКРЯ пропорционально количеству «хороших» примеров в контрольной выборке. Если примеров было меньше 500, то просматривалось все множество примеров.

Каждое словоупотребление слова в корпусе размечено всеми тегами, которые приписаны слову в словаре. Например, прилагательное 'большой' имеет два тега:

- {«qualitative», «size»} 'большой мужчина'
- {«qualitative», «age»} 'большой мальчик (взрослый)'

Семантическая неоднозначность в корпусе не снята. При подсчетах мы ориентировались на первое значение прилагательного. Например, при поиске цепочки прилагательных с порядком следования: <'цвет', 'форма'> (Adj:form, Adj:colour) находится следующий пример (см. пример (3)):

(3) *Она нарезала круглый кремовый торт* // [Алексей Рыбин. Последняя игра (2000)]

Пример был найден поисковой системой НКРЯ, поскольку в одном из своих переносных значений прилагательного *кремовый* относится к классу 'цвет'. Такой пример не попадает в нашу выборку, поскольку в (3) *кремовый* употреблено в своем основном значении, оно обозначает материал, а не цвет объекта.

Примеры, когда прилагательное имело переносное значение, а запрос относился к его прямому значению, напротив, учитывались. Нам важно было проверить гипотезу о том, что, если прилагательное занимает не свойственное своему семантическому классу место, то происходит «коэрсия», т.е. оно приобретает переносное значение.

Мы не учитывали значения прилагательных в случае, если семантическая помета прилагательного соотносилась с совсем периферийным значением. Так, например, помета, соответствующая оценке (ev), приписана в НКРЯ помимо таких прилагательных как *хороший*, *отличный* и т.п., большому количеству прилагательных из других семантических классов: *зеленый*, *последний*, *высокий* и др. Такие прилагательные мы не рассматривали как оценочные.

## 4. Результаты

Результаты попарного сравнения количества примеров с разным порядками приведены в таблице 2. (100% составляет общее количество примеров для пары).

Таблица 2. Частота порядков внутри пары классов

	ord	card	size	ev	Form	color	poss	etn	generic
ord		66,59% (4213)	93,30% (763)	84,07% (2043)	88,24% (30)	97,68% (126)	69,23% (18) (p-val = 0,25)	99,04% (1031)	97,78% (44)
card	33,41% (2153)		96,20% (574)	91,68% (6322)	99,73% (731)	99,44% (2488)	94,40% (51)	99,46% (2213)	98,82% (501)
size	6,70% (55)	3,80% (306)		78,84% (108)	86,53% (3518)	88,81% (10847)	84,86% (118)	97,16% (1954)	96,92% (849)
ev	15,93% (574)	8,32% (387)	21,16% (29)		78,88% (560)	78,30% (4298)	78,79% (430)	94,13% (6747)	94,10% (1411)

	ord	card	size	ev	Form	color	poss	etn	generic
<b>form</b>	11,76% (2)	0,27% (4)	13,47% (548)	21,12% (161)		55,48% (1341)	80,13% (20)	59,52% (25) (p-val = 0,5)	96,00% (504)
<b>color</b>	2,32% (14)	0,566% (3)	11,19% (1367)	21,70% (1191)	44,52% (1076)		65,60% (50)	82,20% (673)	89,82% (635)
<b>poss</b>	30,77% (3)	5,60% (8)	15,14% (21)	21,21% (116)	19,86% (5)	34,41% (26)		100,00% (6) (p-val = 0,18)	78,57% (11) (p-val = 0,24)
<b>etn</b>	0,96% (12)	0,54% (10)	2,84% (57)	5,87% (421)	40,48% (17)	17,80% (146)	0,00% (0)		76,71% (56)
<b>generic</b>	2,22% (6)	1,18% (1)	3,08% (27)	5,90% (89)	4,00% (21)	10,18% (72)	21,43% (3)	23,29% (17)	

Как видно из таблицы 2, некоторые порядки определяются достаточно надежно. Так, например, для пары 'размер' — 'форма' общее количество примеров близко к 4000, при этом пар с таким порядком почти в 10 раз больше, чем с обратным. Некоторые пары семантических классов практически не встречаются в корпусе, например, пара 'порядковое числительное — форма'. Для определения статистической значимости разницы в порядке внутри пары мы использовали критерий Хи-квадрат и критерий Фишера (для малых чисел) с предположением о равномерном распределении порядков (в таблице приведены значения p-value больше 0.05). Как видно из таблицы, такое значение p-value соответствует случаям, когда данных слишком мало, т.е. в ситуациях, когда два семантических класса плохо сочетаются в паре.

В общем случае, могло оказаться так, что иерархию не удалось бы выстроить, поскольку мы рассматривали порядки в парах. В частности, для какой-то тройки классов порядок мог бы оказаться не транзитивным. Однако, такого не произошло (см. таблица 2). Все значения в процентах над диагональю выше значений для соответствующих «обратных» порядков (под диагональю).

Таким образом, если в паре фиксировать порядок, частота которого получилась выше, и для каждого семантического класса вычислить количество классов, которые расположены правее (ближе к существительному), то можно построить иерархию классов. Так, например, в сочетании с порядковыми числительными восемь классов чаще встречаются правее, для генерических — ни один.

Получившаяся иерархия выглядит следующим образом (ii):

- (ii) 'порядковые' > 'количественные' > 'размер' > 'оценка' > 'форма' > 'цвет'  
> 'конкретно-референтные' > 'национальность' > 'генерические'.

Для сравнения повторим здесь иерархию, предложенную Чинкве:

- (i) 'possessive' > 'cardinal number' > 'ordinal number' > 'quality' > 'size' > 'shape' > 'colour' > 'nation'

Таким образом, подтверждаются следующие закономерности: (1) крайнюю правую позицию занимают относительные прилагательные (классы 'генерические', 'национальность'), а градуированные прилагательные располагаются

левее; (2) классы 'размер', 'форма' и 'цвет' сохраняют относительный порядок; (3) числительные располагаются левее вышеупомянутых классов, что также соответствует иерархии Чинкве. Однако можно выявить следующие несоответствия: (4) класс 'конкретно-референтные притяжательные' (прилагательных на *-ин*) не соответствует крайней левой позиции класса 'possessive' в иерархии Чинкве; (5) класс оценочных прилагательных «вклинивается» между прилагательными размера и формы; (6) неожиданным является приоритетный порядок в паре 'порядковые — количественные числительные'. С позиций формального синтаксиса скорее ожидается обратный порядок.

Более подробно выявленные несоответствия, а также случаи колебаний в порядке прилагательных в парах, где мы бы ожидали более строгие ограничения (например, порядок 'национальность — форма' или 'цвет — оценка'), рассмотрим в следующем разделе.

## 5. Обсуждения

### 5.1. Конкретно-референтные прилагательные на *-ин*

Рассмотрим примеры (4)–(6). В примере (5) прилагательное на *-ин* следует после оценочного прилагательного (оценка является результатом семантического переноса), в (4) — после прилагательного из таксономического класса «время/возраст», в (6) — после прилагательного цвета:

- (4) *Первый раз — старинные бабушкины часы и нутриевый полушубок...*  
[Алла Сурикова. Любовь со второго взгляда (2001)]
- (5) *Мальчик, золотой бабушкин внук, вставал часов в одиннадцать, ел что придется...* [Маша Трауб. Замочная скважина (2012)]
- (6) *Поверх ковбойки — почему-то жёлтая Степашина куртка с орлом.*  
[Андрей Волос. Недвижимость (2000) // «Новый Мир», 2001]

Несоответствие позиции русских конкретно-референтных притяжательных прилагательных и притяжательных прилагательных в иерархии Чинкве неслучайно и имеет значимое с точки зрения синтаксической теории основание. Дело в том, что в артиклевых языках притяжательные прилагательные и другие посессивные конструкции играют в именной группе роль определителя (артикля). Они «приписывают» именной группе определенный конкретно-референтный статус. Как правило, они находятся в дополнительной дистрибуции с артиклями и указательными местоимениями (ср. *\*the Peter's house*, *\*Peter's the house* vs. *эта Машина* грядка более ровная, чем та). Согласно объяснению синтаксических свойств данного класса прилагательных, предложенному Е. Лютиковой [Лютикова 2012], они могут иметь функцию определителя в крайне-левой позиции, но могут и не иметь такой функции, занимая другие позиции. В языках с артиклями посессивные прилагательные занимают позицию определителя в синтаксической структуре и, соответственно, могут занимать только крайнюю левую позицию.

## 5.2. Реальная вариативность порядков

В ряде случаев наблюдается реальная вариативность порядков для некоторых классов. В (7) вполне допустимым является порядок 'цвет — форма' (ср. (7a) и (7b)), в (8) — 'национальность — форма':

- (7) а *При вечернем свете она казалась чересчур мрачной со своей выдавшей виды мебелью, черным круглым столом..* [Николай Тихонов. Белое чудо // "Огонек" №№ 18, 23, 1956] (НКРЯ)  
б *На круглом белом столике лежали изогнутые круассаны* [Валерий Попов. Будни гарема (1994)]
- (8) — *спрашивает Епкин, выпялив свое монгольское круглое лицо и выставив стриженую голову чуть вперед.* [Эдуард Лимонов. Подросток Савенко (1982)]

## 5.3. Порядковые и количественные числительные

Несоответствие порядка иерархии Чинкве также наблюдается в паре 'порядковые — количественные числительные':

- (9) *Я первые два года работы Снегурочкой перемещалась с места на место на метро.* [Юлия Вишневецкая. Большой дед следит за тобой // «Русский репортер», 2014]

Разница в порядках не столь существенная по сравнению, например, с парами 'порядковые/количественные — размер' (см. таблица 2). При более подробном рассмотрении примеров оказывается, что основная часть примеров с данным порядком — это пример с числительным *первый* (иногда *второй*) с малыми количественными числительными, модифицирующими временные отрезки, как в *первые три года*. Такие примеры составляют почти 80% всей выборки для данной пары классов.

## 5.4. Семантические переносы (коэрсия)

«Нарушения» ожидаемых порядков могут объясняться тем, что в реальном контексте прилагательное определенного семантического класса в позиции, несвойственной данному классу, имеет переносное значение. Это может быть закрепленный в словаре семантический перенос (в (10) представлен перенос 'материал -> оценка'), окказиональный семантический перенос и «метафоризация» всей именной группы (11). Достаточно часто такие «переносы» закреплены в устойчивых словосочетаниях (12), включая терминологические.

- (10) *Мальчик, золотой бабушкин внук, вставал часов в одиннадцать, ел что придется...* [Маша Трауб. Замочная скважина (2012)]
- (11) *Старший лейтенант, несмотря на свои румяные двадцать годков, в составе все той же Третьей Танковой успел перемахнуть и Вислу и Одер.* [Илья Бояшов. Танкист, или "Белый тигр" (2008)]



- (12) *голова бычка напоминает приплюснутый большой палец...*  
[С. Д. Кржижановский. Хорошее море (1937)]

## 6. Выводы

В результате проведенного исследования мы можем ответить на основные вопросы, которые стояли перед нами.

Во-первых, наш эксперимент показал, что возможно построить иерархию таксономических классов прилагательных, отражающую предпочтения в линеаризации группы прилагательных (несочиненных), основываясь на корпусных данных. Мы предложили метод такого построения: можно сравнить попарно таксономические классы с точки зрения частоты того или иного порядка классов в цепочках из двух прилагательных. Во-вторых, наш эксперимент с корпусными данными, в общем и целом, подтвердил соответствие порядка прилагательных в русском языке универсальной иерархии с точностью до притяжательных прилагательных, которые могут и не иметь в некоторых контекстах детерминирующей функции в безартиклевых языках.

В-третьих, корпусной подход также позволил выявить существенные синтаксические свойства класса посессивных прилагательных в русском языке, их «недетерминирующую» функцию в значительном количестве контекстов, и, следовательно, не крайнее, а скорее близкое к существительному положение в именной группе.

В-четвертых, анализ корпусных примеров, которые нарушают ожидаемый порядок, показал, что часть нарушений соответствуют случаям переносных значений, тем самым, могут быть полезными при полуавтоматическом режиме снятия неоднозначности, а также при выявлении непоследовательностей и ошибок в корпусной семантической разметке.

Таким образом, анализ расхождений между выявленными на материале корпуса тенденциями относительно порядка следования прилагательных и теоретическими ожиданиями, с одной стороны, позволяет сделать ряд теоретических выводов, с другой, — может помочь в практической задаче снятия семантической неоднозначности, а также в задаче выявления непоследовательностей и ошибок в семантической разметке.

## References

1. *Cinque, G.* (1994). On The Evidence for Partial N-Movement in the Romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini (Eds.), *Paths towards universal grammar*, Georgetown University Press, Georgetown, pp. 85–110.
2. *Dixon R. M. W.* (1977), Where have all the adjectives gone?, *Studies in Language*, 1: pp. 19–80.
3. *Dixon, R. M. W.* (1982), *Where Have All The Adjectives Gone? and other essays in semantics and syntax*, Mouton, Berlin.

4. *Dixon R. M. W.* (2005), *A Semantic Approach to English Grammar*, Oxford University press, New York.
5. *Lyutikova E. A.* (2012), Two types of inverse in noun phrase in Russian. Russian syntax in a scientific light [O dvuh tipah inversii v russloj imennoj gruppe. Russkij sintaksis v nauchnom osveschenii]. 2012: 65–106.
6. *Moskaleva I. S.* (2010), Linear order of Adjectival components in a Noun Phrase [Linearizacija adjektivnyh komponentov v imennoj gruppe] Graduated Thesis. Manuscript. MSU, Moscow.
7. *Pereltsvaig A.* (2007), The universality of DP: A view from Russian. *Studia Linguistica* 61(1), pp. 59–94.
8. *Quirk R., Greenbaum S., Leech G., Svartvik J.* (1972). *A Grammar of Contemporary English*. Longman, London.
9. *Scott G.* (2002), Stacked Adjectival Modification and the Structure of Nominal Phrases. In G. Cinque (Ed.), *Functional Structure in DP and IP*, Vol. 1 of *The Cartography of Syntactic Structures*, Chapter 4, Oxford University Press, New York, pp. 91–120.
10. *Sproat, R. and Shih, C.* (1991), The Cross-Linguistic Distribution of Adjective Ordering Restrictions in C. Georgopoulos, R. Ishihara (eds.) *Interdisciplinary Approaches to Language*, pp. 565–593. Kluwer Academic Publishers. Dordrecht.
11. *Vendler, Z.* (1968), *Adjectives and Nominalisations*, Mouton and Co., The Netherlands.
12. *Zielinska, D.* (2007), A Polish-English Contrastive Study of the Order of Pre-modifying Adjectives: A Procedural Model Account [<http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2007/5Paper.pdf>].

# КОГДА СТАТЬЯ — ЭТО НЕ СТАТЬЯ: «ОТРИЦАТЕЛЬНЫЕ ТАВТОЛОГИИ» В РУССКОМ ЯЗЫКЕ<sup>1,2</sup>

**Вилинбахова Е. Л.** (e.vilinbakhova@spbu.ru)

Санкт-Петербургский государственный  
университет, Санкт-Петербург, Россия

**Ключевые слова:** языковые тавтологии, противоречия, русский язык,  
микросинтаксис, семантика, прагматика

## WHEN AN ARTICLE IS NOT AN ARTICLE: “NEGATED TAUTOLOGIES” IN RUSSIAN

**Vilinbakhova E. L.** (e.vilinbakhova@spbu.ru)

St Petersburg State University, St Petersburg, Russia

The study is dedicated to contradictions  $A \neq A$  in Russian, also known in literature as ‘negated tautologies’. They are often viewed as derivations from equative tautologies  $A = A$ . Here I describe structural and semantic features of negated tautologies that are established in Russian with patterns  $X \text{ не } X$  ‘X is not X’,  $X \text{ не } \text{est}' X$  ‘X is not X’ and  $X - \text{eto } \text{не } X$  ‘X this is not X’. Such constructions show that the speaker is not able to use the corresponding tautology  $X \text{ is } X$  because (a) the referent of a linguistic expression  $X$  does not belong to the category  $x$ ; (b) characteristics of the referent of a linguistic expression  $X$  or an attitude towards it differ from the norm; (c) the linguistic expression  $X$  is not employed in its common, straight meaning. Besides, negated tautologies are compared to similar Russian constructions  $S_{\text{dat}} + X \text{ не } X_2$  ‘ $S_{\text{dat}} + X$  is not X’ and  $X \text{ не } \vee X_2$  ‘X is not in X’, and to tautologies  $X \text{ est}' X$  ‘X is X’ and  $X - \text{eto } X$  ‘X this is X’.

**Key words:** tautologies, contradictions, Russian language, microsyntax, semantics, pragmatics

*Если дважды два не равно четырем, то  
либо два — не два, либо «жды» — не «жды»,  
либо четыре — не четыре.*

[В. Ф. Турчин. Феномен науки. Ки-  
бернетический подход к эволюции (1970)]

---

<sup>1</sup> Автор выражает признательность анонимным рецензентам статьи за их замечания и предложения, а также Е. А. Бабаиной, А. В. Крюковой, А. О. Прусаковой за помощь в сборе материала.

<sup>2</sup> Работа поддержана грантом Президента Российской Федерации МК-713.2017.6.

## 1. Вводные замечания

В статье Ю. Д. Апресяна «Тавтологические и конрадикторные аномалии» выделяется два рода комбинаторных правил в языке: они отражают, с одной стороны, иррациональные формальные ограничения на сочетаемость языковых единиц, а с другой стороны, рациональные содержательные ограничения на их совместимость [Apresyan 1989 / 1995: 622]. К нарушениям правил второго рода относятся тавтологии и противоречия, причем в рамках обоих классов можно выделить как примеры, требующие знания семантики слов: *Этот холостяк неженат*; *Я его люблю и ненавижу* и др., так и случаи, когда структурные характеристики позволяют распознать их еще до обращения к содержанию: *Жизнь есть жизнь* ( $A = A$ ); *Я верил и не верил* ( $p \& \sim p$ ), *Дети — это не дети* ( $A \neq A$ ).

В данной работе на материале Национального корпуса русского языка (далее — НКРЯ), интернет-источников и данных интернет-опроса носителей русского языка рассматриваются именно такие противоречия с формулой  $A \neq A$  в русском языке, см. (1):

- (1) *Вдруг он увидел ясно: Арсюшка далеко уже не Арсюшка. <...> Арсений Иустинович Флоринский, действительный тайный советник, сенатор, вхож к государю, один из заправил департамента.* [Ю. В. Трифонов. Исчезновение (1981)]

В отличие от конрадикторных высказываний с формулой  $p \& \sim p$ , которые активно обсуждаются в литературе (см. библиографию в [Escandell-Vidal 1990; Snider 2015]), конструкции с формулой  $A \neq A$  пока не получили подробного описания. Как правило, они обозначаются как «отрицательные тавтологии» (далее — ОТ) и упоминаются в качестве производных от тавтологий  $A = A$ , см. [Smit 2003; Meibauer 2008; Snider 2015]; см. также [Horn 1989: 562], где ОТ приводится в качестве иллюстрации т.н. металингвистического отрицания (англ. *metalinguistic negation*)<sup>3</sup>.

Авторы отмечают, что второй повторяющийся элемент конструкции используется как предикат, а сами ОТ указывают на несоответствие общепринятым представлениям об объекте, см. пример (2) Й. Майбауэра и его комментарий: «в подобных примерах отрицается возможность отсылки к стереотипным свойствам референта» [Meibauer 2008: 448].

- (2) *Die Stones sind nicht die Stones* (пример из [Meibauer 2008: 448])  
'Роллинг Стоунз уже не те, что прежде' (букв. 'Роллинг Стоунз — это не Роллинг Стоунз')

На материале русского языка высказывания с формулой  $A \neq A$  не рассматривались в ряду языковых тавтологий, однако в литературе представлены описания ряда конструкций с лексическими повторами и отрицанием, см. (3–8). Из перечисленных конструкций на роль ОТ с идентичными повторяющимися элементами и семантикой несоответствия референта норме подходит (5), ср. толкования 'положительные ожидания, обычно связываемые с X-ом, не оправдываются'

<sup>3</sup> Приблизительно соответствует противопоставительному отрицанию в [Boguslavskiy 1985].

[Bulygina, Shmelev 1997: 446], 'несоответствие норме, эталонному X' [Shemanaeva 2010: 582], перекликающиеся с интерпретацией Й. Майбауэра.

- (3) *Дружба не дружба, а хорошие отношения* [Mel'chuk 1974]
- (4) *Дружба не дружба, а работать надо* [Mel'chuk 1974]
- (5) *Без детей человек не человек* [Bulygina, Shmelev 1997; Shemanaeva 2010]
- (6) *Без тебя мне праздник не в праздник* [Shemanaeva 2010]
- (7) *Я верил и не верил* [Apresyana et al. 2010]
- (8) *Читать я эту книгу не читал, но кое-что о ней слышал* [Iomdin 2013]

В настоящей работе демонстрируется, что формальные и содержательные характеристики ОТ далеко не исчерпываются уже существующими описаниями. Изложение построено по следующему плану. В разделе 2 описаны структура и семантика анализируемых конструкций. В разделе 3 ОТ рассматриваются в ряду близких к ним конструкций с лексическими повторами и отрицанием, а затем обсуждаются в сопоставлении с тавтологиями, устанавливающими тождество, в русском языке. В заключительном разделе 4 представлены итоги исследования и обозначен его практический потенциал.

## 2. Общие свойства отрицательных тавтологий в русском языке

### 2.1. Структурные свойства отрицательных тавтологий

По данным НКРЯ и интернет-источников, в русском языке для ОТ используются следующие структурные модели:  $X \text{ не } X$ ,  $X \text{ — это не } X$  и  $X \text{ не есть } X$ , причем, в отличие от тавтологий, устанавливающих тождество, ОТ регулярно употребляются с различными модификаторами (*уже, ещё, вовсе* и пр.), см. (9–11)<sup>4</sup>.

- (9) *Это, конечно, аллегория. Змей — вовсе не змей, а яблоко — не яблоко. Тут каждому ясно, что за аллегорией стоит...* [Владимир Скрипкин. Тинга // «Октябрь», 2002]
- (10) *Вот видите, тетя Даша, Таня не может про кактус рассказать, однако никто ее в обмане не подозревает. А вдруг Таня — это не Таня?* [Дарья Донцова. Личное дело женщины-кошки (2007)]

---

<sup>4</sup> Также встретились модели  $X \text{ не } X \text{ есть}$  и  $X \text{ есть не } X$ , см. примеры (а–б) из поэтических произведений, однако они не частотны, и потому были оставлены за рамками настоящей работы:

(а) *Если ты, такое око смерклось, / Значит, жизнь не жизнь есть, смерть не смерть есть* [Марина Цветаева. Новогоднее (1927)]

(б) *Смерть — есть не смерть, / То лишь перерожденье / Из старой формы бытия / В другую.* [Сергей Колосков. Смерть есть не смерть... (2005)]

- (11) Уинстон оказывается единственным человеком, который, кажется, помнит, что **«всегда» не есть всегда**, что четыре года назад, к примеру, расстановка сил была другая<sup>5</sup>.

Возможно выражение связки ОТ формами прошедшего и будущего времени с заменой именительного падежа на творительный или без нее, см. (12–14).

- (12) **И дождь был не дождь**, а нечто похожее на то, когда на тебя выливают из ведра воду чохом. [Владимир Толмасов. Парусная вахта // «Вокруг света», 1989]

- (13) **Человек так создан, что всякое увлечение... что без некоторого увлечения жизнь это будет не жизнь...**

- (14) **Отдых не будет отдыхом** без сытного и вкусного завтрака.

Далее, в ОТ, по-видимому, нет ограничений на заполнение переменных; в качестве повторяющихся элементов в конструкциях допустимы не только различные синтаксические группы, но и вообще любые языковые единицы, см. (15–17), а также эпиграф.

- (15) **А резюме Ваших слов, вот оно: главным в обезьяне всегда была душа. И убить не есть убить. И бог не создавал никого — мы потомки обезьян.**

- (16) **P.S.: переписка недельной давности, так что завтра это не завтра на самом деле.**

- (17) **Пожалуй, один из иллюминаторов я мог бы прочесть даже как прописное Ю. Ключик. Но мешало отсутствие впереди палочки, без которой Ю не Ю, не ключик, а всего лишь ноль, зеро, знак пустоты.** [В. П. Катаев. Алмазный мой венец (1975–1977)]

## 2.2. Семантические свойства отрицательных тавтологий

Переходя к семантике ОТ, важно отметить, что при их интерпретации, как и при интерпретации тавтологий, устанавливающих тождество, ключевую роль играет определенное знание о предмете речи, которое может быть *референциальным* (о соотношении языкового выражения с явлением действительности), *атрибутивным* (о характеристиках денотата языкового выражения), и *метаязыковым* (о его значении) [Vilinbakhova, Escandell-Vidal 2016].

В первом случае ОТ отрицают реальную принадлежность сущности к категории (даже если в нее входит только один элемент), т.е. языковое выражение на месте второй переменной используется в строгом, формальном, терминологическом понимании: например, в (18) референт предположительно не является биологическим отцом ребенка. Если на месте переменной имя

---

<sup>5</sup> Примеры взяты из Интернета, если не указано иное. Из соображений краткости адреса не указаны, но, разумеется, в материалах они присутствуют. Орфография и пунктуация оригинала почти полностью сохранены.

собственное, см. (19), то речь идёт именно об отсутствии реального тождества, а не о ситуации, когда человек меняет поведение, внешний вид, социальный статус<sup>6</sup>. Для данной разновидности ОТ можно предложить следующее толкование: 'говорящий подчеркивает, что референт языкового выражения X вообще не принадлежит к категории x'.

(18) *Так значит мой папа — вовсе не мой папа? Значит мой папа — вообще неизвестно кто...* [коллективный. Форум (2007)]

(19) *Хотите доказать, что Путин — это не Путин? Найдите фотографии, на которых можно увидеть детали радужной оболочки его глаз.*

Во втором случае речь идёт о несоответствии референта норме, стереотипу, что как раз и отмечалось в указанных выше работах. Можно добавить, что допустима отсылка как к дескриптивной, так и к прескриптивной норме, см. (20–21) и ниже.

(20) *И вообще Маша — это не Маша, если не влипнет в очередное приключение, на голову медведя, разумеется* (имеются в виду герои мультфильма «Маша и медведь» — Е.В.)

(21) *В общем, что и говорить, без террасы — дача не дача.* [Еще не дом, но уже не улица (2003) // «Сад своими руками», 2003.01.15]

Некоторые примеры допускают двоякую трактовку, как (22), который можно понять и как 'испытание нельзя признать полноценным, если тот, кто его проходит, заранее знает результат' (несоответствие прескриптивной норме), и как призыв вообще не считать описываемое событие испытанием (исключение из категории).

(22) <...> *если результат заранее известен, то испытание не есть испытание.* [Сергей Прокофьев: дневники (2003) // Российская музыкальная газета, 2003.02.12]

В рамках второго класса может обсуждаться не только сам референт, но и его восприятие кем-либо, что является важным отличием ОТ от тавтологий  $A = A$ . Дело в том, что при использовании в речи последних субъект восприятия чаще всего не обозначен, а высказанная точка зрения подается как объективная истина<sup>7</sup>. Напротив, в ОТ субъект восприятия играет важную роль, см. пример (23), где он выражен эксплицитно, а высказывание характеризует не людей (клиентов), а сотрудников компании, которые несправедливо к ним относятся. В примере (24) также речь идет о восприятии мира говорящим, но, в отличие от (23), здесь говорится не о постоянных личностных качествах,

<sup>6</sup> См. также следующий пример: *Гробница царя Давида — не гробница, не Давида и не царя.* [Анатолий Найман. No comment // «Октябрь», 2001]

<sup>7</sup> Исключение составляют метаязыковые тавтологии, см. *Он человек надежный в том смысле, что для него, как в Библии: «да» — есть да, а «нет» — есть нет,* хотя таких примеров тоже немного.

а о его состоянии. Отказ от привычного восприятия референта проявляется и в примерах вроде (25), где при сохранении референциального тождества происходит смена перспективы.

(23) Для них (сотрудники страховой компании — Е. В.) *люди это не люди* издеваются как могут.

(24) *И день не день, и ночь не ночь, когда о красоте думаю.* [Михаил Беляев. Пожар (1978)]

(25) *Дети — это не дети. Это люди.*

Значение второй разновидности ОТ можно описать следующим образом: 'говорящий подчеркивает, что характеристики референта языкового выражения *X* или его восприятие отличаются от нормы'.

Наконец, для метаязыковых ОТ возможно следующее толкование: 'говорящий подчеркивает, что произнесенная им языковая единица *X* не используется в наиболее общепринятом значении'. Как и в метаязыковых тавтологиях, устанавливающих тождество, на месте связки могут использоваться глаголы *значить, означать* и пр., см. (26–28).

(26) *В данном контексте — «впереди — это не впереди, а как бы сзади», хотя и впереди.*

(27) *А что значат её эти ответные смайлики: «Ни да, ни нет, но нет не так, чтобы прям нет, но и да вовсе не да...»? Девушки, объясните!*

(28) *Голова оратора оказалась огнеупорной — с танком не сгорела, а, напротив, ожила, осмелела и громко закончила речь: — Амнистия еще не значит амнистия!* [Евгений Весник. Дарю, что помню (1997)]

Следует отметить, что перечисленные значения ОТ могут быть выражены и высказываниями без тавтологии<sup>8</sup>, поэтому выбор формулы  $A \neq A$  в дополнение к передаче одного из трёх значений одновременно указывает на то, что соответствующее тавтологическое высказывание  $A = A$  в данном контексте неправомерно. Таким образом, общее толкование для ОТ могло бы выглядеть так: 'Говорящий подчеркивает, что альтернативное тавтологическое высказывание  $X \text{ сор } X$  не подходит для обозначения описываемой ситуации, т.к. (а) референт языкового выражения *X* вообще не принадлежит к категории *x*; (б) характеристики референта языкового выражения *X* или его восприятие отличаются от нормы; (в) языковая единица *X* не используется в наиболее общепринятом значении'.

---

<sup>8</sup> Автор благодарен за это наблюдение анонимному рецензенту.



### 3. Отрицательные тавтологии в ряду других конструкций

#### 3.1. Отрицательные тавтологии и близкие к ним конструкции

Как можно заметить из приведённых выше примеров, указанные структурные и семантические характеристики справедливы для любой из трех моделей ОТ  $X$  не  $X$ ,  $X$  — *это не X* и  $X$  не *есть X*.

Единственное ограничение касается модели  $X$  не  $X$ : если субъект восприятия эксплицитно выражен дательным падежом, то  $S_{\text{дат}} + X$  не  $X$  может отрицать референциальное тождество только при одновременном отрицании соответствующего восприятия референта. Так  $S_{\text{дат}} + X$  не  $X$  возможно в примере (29), когда референт теряет свой формальный статус (начальство перестаёт быть таковым после увольнения) для субъекта восприятия, и одновременно происходит изменение отношений между ними, но не в (30), когда изменение отношений между начальником и бывшей подчинённой не происходит<sup>9</sup>.

(29) [Андрей выиграл в лотерею миллион долларов и сразу уволился с нелюбимой работы] Бывший коллега: *Повезло Андрею! Теперь ему начальство — не начальство.*

(30) [Таню из нашего отдела повысили, и теперь формально она не подчиняется нашему шефу, хотя по-прежнему относится к нему с большим уважением и готова ему помочь, выполняя различные поручения] ?? / \*Таню повысили, и теперь ей **начальство** — **не начальство.**

Поскольку модель  $X$  не  $X$  единственная допускает выражение субъекта восприятия в дательном падеже (это недопустимо для  $X$  не *есть X* и  $X$  — *это не X*), естественно предположить, что речь идёт о двух разных конструкциях:  $X$  не  $X$  и  $S_{\text{дат}} + X$  не  $X_2$ , по аналогии с конструкциями *мне негде спать* и *у меня негде спать*, описанными в [Aresyan, Iomdin 1989 / 2010].

По-видимому, значение  $S_{\text{дат}} + X$  не  $X_2$  — 'несоответствие восприятия референта дескриптивной / прескриптивной норме', — полностью входит в возможный спектр значений  $X$  не  $X$ : характеристика *постоянных* физических, духовных и т.п. качеств субъекта, или его *временного* эмоционального, физического и пр. состояния может быть выражена обеими конструкциями. Так для примеров конструкции  $S_{\text{дат}} + X$  не  $X_2$  (31) и (32) можно подобрать аналогичные примеры конструкции  $X$  не  $X$ , см. (33) и (34).

(31) *Так что, я думаю, если человек находится в государственном аппарате, то ему Родина не Родина, народ не народ, а живет он где-нибудь за границей* ('безразличное отношение к Родине и народу — постоянное свойство политика')

<sup>9</sup> Оба примера проверялись в рамках интернет-опроса 20 человек. Первый пример сочли приемлемым 90% испытуемых, второй — только 15%.

(32) *Я не выспалась... <...>/ Что же делать теперь буду / Как денёк я проведу... / Мне работа не работа / Я качаюсь на ходу...* ('неспособность сосредоточиться на работе — временное состояние')

(33) *А для него мать не мать, политические идеи выше родства, поэтому ее репутацию и любовь можно принести в жертву* ('пренебрежение к матери — постоянное свойство героя')

(34) *Побыть вашим личным психологом, который чувствует себя как в тюрьме, и для которого жизнь — не жизнь?* ('неспособность ценить жизнь — временное состояние')

Здесь же следует упомянуть о конструкции  $X$  не в  $X$ , описанной в [Shemapaeva 2010], см. (35), значение которой также входит в спектр значений ОТ, но имеет ещё более строгие ограничения, указывая на временный, а не на постоянный признак: ср. (31а).

(35) *Если его друг в беде, он снимет с себя последнее, он поедет за сотни километров, ночь не в ночь, если чем-то сможет помочь* [Наталья Бестемьянова и др. Пара, в которой трое (2000–2001)]

(31а) *\*Так что, я думаю, если человек находится в государственном аппарате, но ему Родина не в Родину, народ не в народ, а живет он где-нибудь за границей*

Таким образом, хотя конструкции  $S_{дам} + X$  не  $X_2$  и  $X$  не в  $X$  не способны передать все значения, доступные для ОТ<sup>10</sup>, их семантика позволяет считать их наиболее близкими к ОТ из представленных в русском языке конструкций с повторами и отрицанием.

### 3.2. Отрицательные тавтологии и тавтологии, устанавливающие тождество

Как было отмечено выше, ОТ считаются производными от тавтологий, устанавливающих тождество. В русском языке выделяется две модели тавтологий:  $X$  есть  $X$  и  $X$  — это  $X$  [Bulygina, Shmelev 1997], и логично предположить, что с первой соотносятся ОТ  $X$  не  $X$  и  $X$  не есть  $X$ , а со второй, в свою очередь, ОТ  $X$  — это не  $X$ . Согласно предложенному выше толкованию, ОТ отрицают соответствующие исходные тавтологии. Поскольку все три модели ОТ синонимичны и способны отсылать к любому из трех типов знаний — референциальному, атрибутивному и метаязыковому (см. раздел 2.2), этого можно ожидать и от тавтологий  $X$  есть  $X$  и  $X$  — это  $X$ .

Эти ожидания, однако, не вполне оправдываются. В реальном языковом узусе  $X$  — это  $X$  передает гораздо более широкий спектр значений, чем  $X$  есть  $X$ . В то время как для обеих конструкций возможна отсылка к атрибутивному

<sup>10</sup> Кроме того,  $X$  не в  $X$  имеет структурные отличия.

знанию о характеристиках объекта, доступному обоим участникам общения, см. (36), при передаче других значений использование *X есть X* ограничено.

- (36) Компания «Боско»? Хорошая марка, но **Карден есть / это Карден**. Он обладает абсолютным вкусом, великолепно чувствует линию, красоту. [Долгополов Николай. Пьер Карден против «Боско» // Труд-7, 2007.03.02]<sup>11</sup>

Так, *X есть X*, как правило, не используется для отсылки к референциальному знанию, а метаязыковые тавтологии со связкой *есть* встречаются преимущественно в поэтических и религиозных текстах или в сопоставительных тавтологиях *X сор X*, а *Y сор Y*; см. (37–38), где замена связки *это* на *есть* делает примеры менее естественными.

- (37) И. ЛИНДЕР: Ну, тогда этот вопрос надо задавать Юлиану Семенову, правда ли, **что Штирлиц это Штирлиц / ??Штирлиц есть Штирлиц**<sup>12</sup>

- (38) — **Исчез — это исчез / ??Исчез есть исчез**, и все тут, — вздохнул Болсохоев. [Вячеслав Рыбаков. Гравилет «Цесаревич» (1993)]

Также употребление *X есть X* затруднено, если знание о характеристиках объекта доступно только говорящему, но не адресату.

- (39) А: Что за человек этот Иванов из бухгалтерии?  
Б: Как бы тебе объяснить... **Иванов — это Иванов / ??Иванов есть Иванов**. <Скоро сам познакомишься с ним поближе>  
[модифицированный пример из Meibauer 2008: 447]

Следовательно, получается, что производные модели *X не X* и *X не есть X* как будто обладают большим спектром значений, чем исходная конструкция *X есть X*. Значит ли это, что существует другая исходная конструкция? Рассмотрим возможные пути решения.

Прежде всего, стоит отметить, что семантические ограничения для *X есть X* снимаются не только при отрицании: появление модификаторов действует таким же образом, см. (40), где идет отсылка к референциальному знанию, и (41), где знание о предмете речи доступно только говорящему.

- (40) Если еще **Путин и есть Путин** и не заменен на двойника...

<sup>11</sup> Примеры (36) и (38) взяты из [Vilinbakhova, Kopotev 2017].

<sup>12</sup> Также по результатам интернет-опроса 25 носителей русского языка фраза *В этом фильме Хичкок есть Хичкок* не может интерпретироваться как 'в этом фильме Хичкок играет самого себя' (100% испытуемых), в отличие от аналогичной фразы *В этом фильме Хичкок — это Хичкок*, для которой такую интерпретацию сочли приемлемой 80% испытуемых.

(41) *Мама, ты знаешь, что такое сволочь? <...> спросил Иван.*

— *Сволочь и есть сволочь, — мрачно ответила она* [Валентин Распутин.

Дочь Ивана, мать Ивана (2003) // «Наш современник», 2003.11.15]<sup>13</sup>

Таким образом, можно предположить, во-первых, существование дополнительной тавтологической конструкции с нулевой связкой  $X \emptyset X$  (или же омонимичной конструкции  $X \text{ есть } X_2$ ), которая обладает теми же значениями, что  $X$  — это  $X$  и (почти) не используется в настоящем времени без модификаторов; тогда именно она является исходной для соответствующих ОТ, а примеры (40–41) являются её манифестациями. В этом случае  $X \text{ есть } X$  (или  $X \text{ есть } X_1$ ) является отдельной лингвоспецифичной псевдотавтологией.

Во-вторых, если оказаться от введения третьей тавтологической конструкции, можно говорить о дефолтном значении  $X \text{ есть } X$  (' $X$  обладает типичным для него свойством  $Y$ '), которое блокирует все остальные возможные интерпретации; если же данное значение входит в противоречие с контекстом, то примеры воспринимаются как менее естественные. С другой стороны, при появлении модификаторов, отрицании и пр. «блокировка» снимается, что позволяет отсылку к другим типам знаний.

Хотя окончательное решение данного вопроса остается за рамками настоящей работы, намеченные здесь трактовки могут быть использованы в дальнейших исследованиях тавтологий в русском языке.

#### 4. Выводы

В настоящей статье на материале русского языка была предпринята попытка внести вклад в описание контрадикторных высказываний с формулой  $A \neq A$ , которые в литературе обозначаются термином «отрицательные тавтологии» (ОТ) в параллель тавтологиям, устанавливающим тождество,  $A = A$ . На материале НКРЯ и интернет-источников были выделены три модели ОТ  $X \text{ не } X$  и  $X \text{ не есть } X$  и  $X$  — это не  $X$ , которым было предложено следующее толкование: 'Говорящий подчеркивает, что альтернативное тавтологическое высказывание  $X \text{ сор } X$  не подходит для обозначения описываемой ситуации, т.к. (а) референт языкового выражения  $X$  вообще не принадлежит к категории  $x$ ; (б) характеристики референта языкового выражения  $X$  или его восприятие отличаются от нормы; (в) языковая единица  $X$  не используется в наиболее общепринятом значении'. Далее ОТ были рассмотрены в ряду близких к ним конструкций  $S_{\text{дам}} + X \text{ не } X_2$  и  $X \text{ не в } X$  и сопоставлены с соответствующими тавтологиями, устанавливающими тождество,  $X \text{ есть } X$  и  $X$  — это  $X$ .

Обращаясь к прикладному потенциалу исследования, представляется, что полученные данные окажутся полезными в области автоматической обработки текста. В частности, значения отрицательных тавтологий можно учесть при описании соответствующих синтаксических моделей, повышая вероятность их правильной интерпретации.

---

<sup>13</sup> Пример минимально модифицирован: исходный текст *Сволочь она и есть сволочь* является примером пролептической конструкции, см. об этом, например, [Argesyan et al. 2010: 113–129].

## References

1. *Apresyan Ju. D.* (1989 / 1995), Tautological and contradictory anomalies [Tavtologicheskie i kontradiktornye anomalii], in Selected works: In 2 vol. Vol. II: Integrated description of language and systematic lexicography [Izbrannye trudy. Vol. 2: Integral'noe opisaniye yazyka i sistemnaya leksikografiya], Yazyki Russkoy Kul'tury, Moscow, pp. 622–629.
2. *Apresyan Ju. D., Boguslavskiy I. M., Iomdin L. L., Sannikov V. Z.* (2010), Theoretical issues in Russian syntax: interaction of grammar and dictionary [Teoreticheskie problemy russkogo sintaksisa: vzaimodeystvie grammatiki i slovarya], Jazyki slavyanskikh kul'tur, Moscow.
3. *Apresyan Ju. D., Iomdin L. L.* (1989 / 2010) Constructions of the type *negde spat'*: syntax, semantics, lexicography [Konstruktsii tipa *negde spat'*: sintaksis, semantika, leksikografiya], in *Apresyan Ju. D., Boguslavskiy I. M., Iomdin L. L., Sannikov V. Z.*, Theoretical issues in Russian syntax: interaction of grammar and dictionary [Teoreticheskie problemy russkogo sintaksisa: vzaimodeystvie grammatiki i slovarya], Jazyki slavyanskikh kul'tur, Moscow, pp. 59–112.
4. *Boguslavskiy I. M.* (1985), Studies in syntactic semantics: Scopes of logical words [Issledovaniya po sintaksicheskoy semantike: sfery deistviya logicheskikh slov], Nauka, Moscow.
5. *Bulygina T. V., Shmelev A. D.* (1997), Linguistic conceptualization of the world (on the material of the Russian grammar) [Yazykovaya kontseptualizatsiya mira (na materiale russkoy grammatiki)], Shkola "Jazyki russkoy kul'tury", Moscow.
6. *Escandell-Vidal V.* Strategies in the interpretation of contradictory utterances [Estrategias en la interpretación de enunciados contradictorios] in Acts of the Congress of Spanish Society of Linguistics, XX Anniversary [Actas del Congreso de la Sociedad Española de Lingüística, XX Aniversario], Gredos, Madrid, pp. 923–936.
7. *Horn L. R.* (1989), A natural history of negation, The University of Chicago Press, Chicago and London.
8. *Iomdin L. L.* (2013), Chitat' ne chital, no...: on a Russian Construction with Repeated Lexical Elements [Chitat' ne chital, no...: ob odnoy russkoy konstruktsii s povtoryayushchimisya slovesnymi elementami], Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2013" [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii "Dialog 2013"], Moscow, pp. 309–322.
9. *Meibauer J.* (2008), Tautology as presumptive meaning, Pragmatics and Cognition, Vol. 16, pp. 439–470.
10. *Mel'chuk I. A.* (1974). Experience in theories of Meaning-Text linguistic models [Opyt teorii lingvisticheskikh modeley Smysl–Tekst], Moscow, Nauka.
11. Russian National Corpus [Natsional'nyy korpus russkogo yazyka], available at: [www.ruscorpora.ru](http://www.ruscorpora.ru)
12. *Smit O. V.* (2003) Normative component in the structure of lexical meaning [Normativnyy component v structure leksicheskogo znacheniya], PhD thesis. St Petersburg, Herzen State Pedagogical University.
13. *Snider T.* (2015), Using tautologies and contradictions, *Proceedings of Sinn und Bedeutung*, 19, pp. 590–607.

14. *Shemanaeva O. Yu.* (2010), Konstruktsiya “Prazdnik ne v prazdnik” na fone drugih konstruktsiy [The peculiarities of the Russian idiomatic construction “Prazdnik ne v prazdnik”] Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2009” [Komp’yuternaya Lingvistika i Intellektual’nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2009”], Moscow, pp. 578–583.
15. *Vilinbakhova E. L., Escandell-Vidal V.* (2016), Linguistic tautologies: interpretation and use in discourse [Tautologías lingüísticas: interpretación y uso en el discurso], talk given at the UNED (Madrid, Spain) on 01.12.2016.
16. *Vilinbakhova E. L., Kopotev M. V.* (2017), Does «X est’ X» mean «X eto X»? Looking for an answer in synchrony and diachrony [«X est’ X» znachit «X eto X»? Ishchem otvet v sinhronii i diahronii], Topics in the Study of Language [Voprosy yazykoznaniiya], № 3, pp. 110–124.

# ИНТОНАЦИЯ ПЕРЕЧИСЛЕНИЯ В РУССКОМ ЯЗЫКЕ<sup>1</sup>

**Янко Т. Е.** (tanya\_yanko@list.ru)

Институт языкознания РАН, Москва, Россия

В работе [Брызгунова 1963: 223–228] показано, что интонация перечисления имеет целый ряд интонационных реализаций, которые зависят от дискурсивного контекста. Цель нашей работы — детализировать типы контекстов и показать, что дискурсивный смысл перечисления комбинируется с иллокутивными (создающими речевой акт, такими, как тема и рема) и дискурсивными (создающими связный дискурс) значениями, а также рассмотреть конкретные типы интонационных конструкций (по [Брызгунова 1982]), которые соответствуют этим комбинациям значений. Перечисление создается за счет наложения на ряд однородных членов предложения или однородных предложений разнообразных интонационных конструкций, которые изначально выражают другие значения: тему, рему, незавершенность, «рассказ по порядку». Для анализа на основе мультимедийного подкорпуса Национального корпуса русского языка создан интонационно размеченный рабочий массив записей дискурсивных стратегий перечисления. Рассмотрено семантическое поле перечисления и средства выражения ядерных и периферийных значений. Работа иллюстрирована графиками изменения частоты основного тона, полученными с помощью системы анализа устной речи Praat.

**Ключевые слова:** интонация, перечисление, дискурсивные стратегии, незавершенность, тема, рема, русский язык, семантическое поле, звучащая речь, НКРЯ, Praat

## WHAT IS INTONATION OF ENUMERATION IN RUSSIAN?

**Yanko T. E.** (tanya\_yanko@list.ru)

Institute of Linguistics, Moscow, Russia

This paper is aimed at answering the question what intonation of enumeration in Russian is. Our main result is that there is no universal intonation of enumeration in isolation from illocutionary meanings (i.e. the meanings

---

<sup>1</sup> Работа выполнена в ФГБУН «Институт языкознания РАН», грант РНФ №14-28-00130.

forming a speech act such as either topics or foci), or discourse meanings (i.e. the meanings forming the coherent discourse). Prosodic patterns of enumeration are based on a variety of pitch accents (which originally express the meanings of the topic, the focus, the discourse continuity) aligned with either homogeneous parts of the sentence or homogeneous sentences. The most common (but not the only ones) prosodic patterns of enumeration are represented by pitch accents IK-6, IK-2, IK-3, and IK-4 in the terminology of E.A. Bryzgunova. We argue that enumeration has a field semantic structure. The nuclear of the semantic field of enumeration consists of the meanings expressed by pitch accents IK-6, IK-2, IK-3, and IK-4. The semantic periphery of enumeration consists of such meanings as a roll call, or a specific Russian strategy of splitting a complex sentence into autonomous segments of information. The peripheral meanings as well as the nuclear meanings are expressed by a variety of pitch accents. For investigation, a minor working corpus of the Russian speech recordings was set up. The corpus comprises records retrieved in the Multimodal subcorpus of the Russian National Corpus. This paper is illustrated throughout with the frequency tracings of the sounding patterns of enumeration. The software program Praat was used in the process of analyzing the sounding data.

**Key-words:** pitch accents, intonation, intonation of enumeration, enumeration, Praat, frequency tracings, semantic field, spoken language, Russian

В работе [Брызгунова 1963: 223–228] показано, что перечисление имеет целое множество интонационных реализаций, которые формируются в зависимости от дискурсивного контекста. Цель нашей работы — детализировать типы контекстов и показать, что значение перечисления комбинируется с иллокутивными (создающими речевой акт, такими, как тема и рема) и дискурсивными (создающими связный дискурс) значениями, а также рассмотреть конкретные типы интонационных конструкций (по [Брызгунова 1982]), которые соответствуют этим комбинациям значений. Заметим также, что в лингвистической учебной и методической литературе широко используется понятие интонации перечисления [Пешковский 1928: 510], [Канакина 2005]; [Гвоздев 2007], где авторы связывают понятие перечисления с понятиями однородных членов предложения и сочиненных предложений [Пешковский 1928: 510], [Черемисина-Ениколопова 1999: 128], логического ударения [Канакина 2005], монотонной интонацией<sup>2</sup> («однотонной» по [Пешковский 1928: 509–510], жестом «загибания пальцев» и т.д., однако какие конкретно просодии стоят за понятием интонации перечисления ничего не говорится.

В этой работе ставится задача детализации контекстов, в которых функционирует интонация перечисления, и выделения дискурсивных значений, с которыми комбинируется перечисление. В разделе 1 обсуждается процедура анализа и методы отбора примеров в рабочий корпус. В разделе 2 дается обзор выделенных на основе корпуса дискурсивных стратегий перечисления.

<sup>2</sup> «...однородные члены стремятся к однотонной интонации и... вследствие этого заключительное понижение наступает... в предложении позже, чем в обычном... они [паузы] являются главным средством выражения (наряду с усилительными ударениями и однотонной «перечисляющей» интонацией)» [Пешковский 1928: 509–510].



В разделе 3 рассматриваются периферийные модели перечисления и обсуждается полевая структура перечисления в русском языке.

## 1. Процедура анализа

Методы решения задачи мы видим в следующем. На первом этапе создается рабочий массив звучащих примеров, предположительно содержащих интонацию перечисления. Затем интонация этих примеров анализируется в терминах интонационных конструкций Е. А. Брызгуновой. Исследование интонационных конструкций производится слуховым и инструментальным методом: вначале применяется анализ на слух, который впоследствии должен получить инструментальную верификацию. При инструментальном анализе используется компьютерная система анализа устной речи Praat [Boersma, Weenink 2012]). В дальнейшем дается интерпретация полученных данных.

Остановимся на критериях отбора примеров в рабочий массив. Для отбора примеров мы используем звучащий подкорпус Национального корпуса русского языка «Мультимедийный», в котором звучащая речь синхронизирована с графическим текстом. Таким образом, структура корпуса позволяет вести поиск по графической записи звучащих текстов. Поиск же интонационных моделей ориентирован на явление близости значений (или даже синонимии) лексических и дискурсивных единиц. Иначе говоря, подбор примеров интонации перечисления заменяется поиском текстов, включающих ключевые слова перечисления: *перечислить*, *пересчитать*, *назвать (их)*, *список*, *наконец*, *следующий*. Эти слова (а также пунктуационный знак «двоеточие») мы интерпретируем как выражение дискурсивного смысла, называемого перечислением. Наша гипотеза состоит в том, что присутствие в тексте ключевых слов перечисления в контексте однородных членов предложения (или однородных предложений) в количестве более двух, может сопровождаться искомыми образцами интонации перечисления. Таким образом, доступ к интонации перечисления дают лексические единицы, включающие компонент перечисления в свое толкование.

Не все употребления лексемы *перечислить*, а также других лексем из списка выше и тем более двоеточия дают искомые звучащие примеры, содержащие интонацию перечисления. Так, когда в качестве единиц перечисления используются пространные описания перечисляемых событий, фактов, умозаключений, включающие сложные предложения или пассажи, состоящие из более чем одного предложения, выделить образцы интонации перечисления не удастся. Таким образом, вопрос, соответствует контекст, содержащий лексемы перечисления, интуитивно понимаемой интонации перечисления или не соответствует, — решается методом экспертной оценки.

Приведем примеры фрагментов текстов, содержащих значение перечисления; ключевые слова перечисления выделены полужирным шрифтом: *Современная семья / вообще семья / подвержена многим недугам. Усталость друг от друга / расхождения во взглядах / вкусах / обиды / стрессы. Это хорошо / что вы всё **перечислили**; Это четыре организации — МПО Лавочкина/ ИКИ/*

Геохим/ институт вот Виноградова бывший/ и/ **наконец**/ наш Институт прикладной математики; На конференции были поставлены и решены **следующие** вопросы / разгром Германии / оккупация Германии и контроль над ней / репарации с Германии / вопрос о конференции объединённых наций; **Список** / товарищ Сухов. Зарина! Джамиля! Гюзель! Саида! Хафиза! Зухра! Лейла! Зульфия! Гюльчатая!; Ну тут же всё **перечислено**. Квартира однокомнатная в Химках / шестьсот пятьдесят рублей на сберкнижке / телевизор «Горизонт» / тут кое-какая мебель / пудель Несси. [НКРЯ].

В результате произведенной процедуры был получен массив звучащих примеров, предположительно содержащих интонацию перечисления. В примерах имеется сорок три встречаемости различных интонационных моделей, отвечающих значению перечисления. Этого количества, безусловно, недостаточно для того, чтобы получить полноценный статистический анализ интонационных моделей перечисления, однако определенная картина имеющегося разнообразия интонаций перечисления тем не менее возникает. Массивы интонационных данных, связанных с выражением некоторого наперед заданного значения, как правило, не бывают слишком представительными. Реальные образцы воплощения значений, выражающихся интонационно, — это, как правило, «штучные», драгоценные находки.

## 2. Интонационные модели перечисления

Сопоставление примеров из полученного рабочего массива со списком интонационных конструкций Е. А. Брызгуновой говорит о том, что в модели перечисления вписываются практически все интонационные конструкции, кроме ИК-7 (об ИК-7 см. [Брызгунова 1982: 118]). Из 43 интонационных конструкций перечисления (при подсчете повторяющиеся в одном примере конструкции считаются за одну) 20 встречаемостей приходится на ИК-6 (об ИК-6 см. [Брызгунова 1982: 118]), 7 — на ИК-2 (об ИК-2 см. [Брызгунова 1982: 109]), 7 — на ИК-4 (об ИК-4 см. [Брызгунова 1982: 114]), 6 — на ИК-3 [Брызгунова 1982: 111]) и 3 — на ИК-5 [Брызгунова 1982: 115].

Обратимся к примерам.

- (1) *Сейчас я тебе зачитаю список. Убийство, изнасилование, вооруженный грабёж, шпионаж я отмел сразу.* [НКРЯ]



Рис. 1. Тоннограмма примера (1)

На рис. 1 изображена тонограмма примера (1). Тонограмма — это кривая изменения частоты основного тона звучащего голоса. По оси ординат на графике откладывается частота в герцах, по оси абсцисс — время, прошедшее с начала речи в секундах. Тонограмма — один из основных способов передать на бумаге значимые интонационные параметры речи. Лексическим показателем перечисления в примере (1) служит словоформа *список*. В ряд перечисления входят четыре единицы: *убийство, изнасилование, вооружённый грабёж, шпионаж*. Присутствие в данном фрагменте текста дискурсивной ситуации перечисления сомнений не вызывает. С точки зрения коммуникативной структуры каждая из именных групп играет роль одной из множественных тем предложения, включающего перечисление. Пятая тема — это цепочка, состоящая из подлежащего и сказуемого данного предложения *я отмел*. Венчает предложение рема *сразу*. Темы рассматриваемого ряда несут акцент ИК-3, по Е. А. Брызгуновой. Последняя в ряду тем тема *я отмел* не входит в перечислительный ряд. Пример (1) говорит о том, что интонация перечисления формируется той коммуникативной функцией, которую имеют соответствующие группы в предложении.

В примере (2) лексическим показателем перечисления служит словоформа *перечислил*, а в качестве интонации перечисления служит ряд нисходящих акцентов типа ИК-1 и ИК-2. Наша гипотеза состоит в том, что говорящий формирует перечислительный ряд как ряд последовательных рем.

- (2) *Саша только что перечислил. Золовы процессы/ дюны/ вулканы/ вулканические потоки/ кратеры.* [НКРЯ].

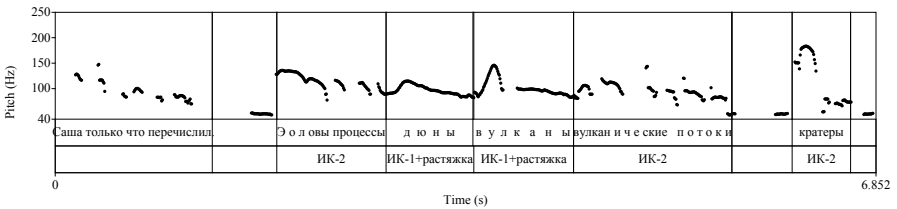


Рис. 2. Тонограмма примера (2)

В первом элементе перечисления *золовы процессы* нисходящий акцент типа ИК-2 фиксируется на словоформе *золовы*: мы наблюдаем рельефное падение на ударном слоге этой словоформы, которое завершается рядом нисходящих движений тона на финальных слогах ремы *золовы процессы*. Следующая рема в ряду рем — она соответствует словоформе *дюны* — имеет менее рельефное падение типа ИК-1 на ударном слоге. Падение сопровождается удлинением времени звучания. Аналогично оформлен и следующий элемент в ряду перечисления *вулканы*. Элемент *вулканические потоки* оформляется без растяжки и с достаточно рельефным падением на ударном слоге словоформы *вулканические*. Наибольший диапазон частот в падении отличает словоформу *кратеры*. Это заключительный элемент перечислительного ряда. Здесь падение в больших диапазонах частот, которому предшествует небольшой подъем в начале

ударного слога, необходимый для набора высоты, соответствует особой иллюкативной силе: здесь перед нами не просто рема, значение которой состоит в передаче сообщаемого от говорящего к слушающему, но и призыв к слушающему обратить внимание на то, что перед нами особый (последний, но важный) элемент в перечислительном ряду: 'не забудем о кратерах, они венчают наш список'.

Таким образом, пример (2) иллюстрирует перечислительный список, в котором перечисляемые элементы формируют ряд последовательных рем. Ремы маркируются достаточно разнообразно: это «плоское» падение типа ИК-1 с растяжкой, более рельефное падение ИК-2, и наконец, падение типа ИК-2 с наибольшим перепадом частот, которое берет начало на относительно высоком уровне. При общем подсчете интонационных конструкций, используемых при перечислении, мы для простоты заменили разнообразие нисходящих акцентов, представленных, в частности, примером (2), только одним представителем падений: ИК-2.

Пример (3) построен на оформлении перечислительного ряда интонационной конструкцией ИК-4. При ИК-4 на ударном слоге словоформы-акцентоносителя фиксируется нисходящий или ровный низкий тон, который завершается рельефным подъемом на заударных слогах, если они есть. Если заударных слогов нет, характерное для ИК-4 нисходяще-восходящее движение тона фиксируется на единственном или последнем слоге акцентоносителя.

Одной из функций ИК-4 в русском дискурсе служит формирование т.н. «рассказа по порядку». «Рассказ по порядку» — это одна из стратегий связного нарратива, представленная не во всех языках (скажем, в русском и в немецком стратегия рассказа по порядку представлена, а в английском — нет, и там, где можно было бы ожидать рассказа по порядку, в английском используется простая незавершенность). Рассказ по порядку дает возможность выстроить предметы или события в соответствии с их упорядоченностью во времени, в пространстве или в соответствии с их внутренней логикой. ИК-4 оформляет расписания, повестки дня, кулинарные рецепты (которые подразумевают строгую упорядоченность действий кулинара), отчеты о командировках и экскурсиях, воссоздающие маршрут поездки. Стратегия рассказа по порядку может использоваться и в рассказе о сущностях или событиях, для которых порядок не служит определяющей чертой, а только указывает на стремление оратора говорить по порядку, поэтому ИК-4 в режиме незавершенности служит характерной чертой речи преподавателей, профессиональных лекторов, судей и других официальных лиц. Стратегия рассказа по порядку, в отличие от немаркированных стратегий незавершенности, подразумевает, что не только текущий шаг повествования не последний, но также и то, что шаг, следующий за текущим (кроме последнего и предпоследнего), тоже не последний.

В примере (3) говорящий очень последовательно использует возможности не только ИК-4, но и другой интонации незавершенности, которая, когда перечисление заканчивается, приходит на смену ИК-4, а именно — интонационной конструкции ИК-6.

- (3) *Я еще раз перечислю: это желание быть свободным, независимым, не продаваться никому, непредсказуемость и, конечно, индивидуализм и эгоизм*<sup>3</sup>.

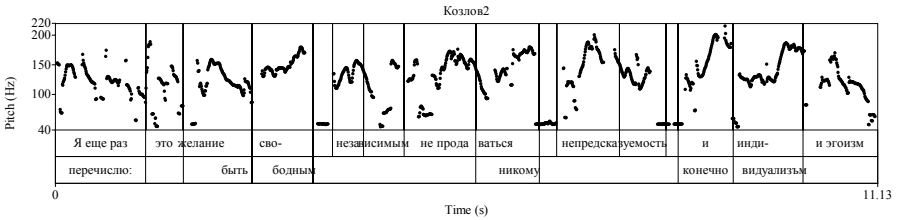


Рис. 3. Тонограмма примера (3)

Перечисление открывается обобщающим словом *перечислю*. Далее, на акцентоносителях групп (кроме последней и предпоследней), которые соответствуют перечисляемым сущностям, — *свободным*, *независимым*, *продаваться*, и *непредсказуемость* — единообразно фиксируется акцент ИК-4. Так, на словоформе *непредсказуемость* на ударном слоге –зу- фиксируется падение, а на заударных –*е*мость рельефный заударный подъем. Здесь перед нами типичная реализация акцента типа ИК-4: ИК-4 говорит о том, что перечисление продолжается. На предпоследней же в ряду перечисления словоформе *индивидуализм* (произносится с гласным звуком на -зм [зъм]) фиксируется уже другая просодическая фигура: подъем на ударном слоге -ли- и ровный тон на заударном – зъм. Это ИК-б, по Е. А. Брызгуновой, типичный акцент предпоследнего члена сочинительной конструкции. На последнем члене перечисления фиксируется интонация конца: интонационная конструкция ИК-1. Перечисление закончилось, говорящий поставил точку.

Итак, в примере (3) структура перечислительного ряда основана на одной из дискурсивных стратегий, формирующих связный текст, — стратегии рассказа по порядку: все шаги перечисления несут интонационную конструкцию, предусматривающую продолжение не только на текущем шаге, но и на шаге, следующем за текущим, предпоследний шаг перечисления несет одну из немаркированных интонационных конструкций незавершенности ИК-б, последний — завершающий — шаг перечисления оформляется интонационной конструкцией финала — ИК-1.

В примере (4) основой для перечисления служит стратегия указания на дискурсивную незавершенность, которая в данном случае имеет в качестве средства выражения интонационную конструкцию ИК-3. Графическим маркером перечисления в примере (4) служит двоеточие.

<sup>3</sup> Примеры, не имеющие здесь ссылки на сайт НКРЯ, взяты из рабочего массива, подготовленного автором с опорой на различные источники звучащей речи: радио- и телеинтервью, т.н. разговорные радиопередачи, записи дружеских бесед.

- (4) *По восшествии вашего величества подверглись опале: семь маршалов, триста генералов, две тысячи двести офицеров...* [НКРЯ].

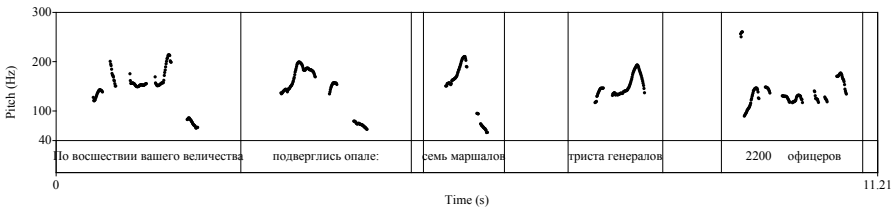


Рис. 4. Тонограмма примера (4)

Тонограмма демонстрирует рельефные подъемы на ударных слогах словоформ *маршалов*, *генералов* и *офицеров*; заударные слоги несут резкое падение частоты тона.

Пример (5) демонстрирует интонационную конструкцию ИК-5 в контексте перечисления. Ключевое слово перечисления: *назовем их*.

- (5) *Но сначала назовём некоторых из них. Славянофилы — это Алексей Хомяков. Это Иван Киреевский. Это братья Аксаковы.* [НКРЯ]

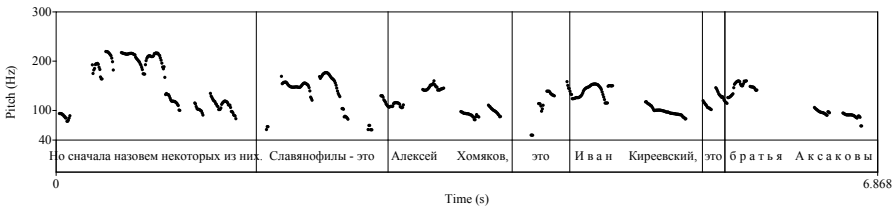


Рис. 5. Тонограмма примера (5)

Интонационная конструкция ИК-5 (по Брызгуновой) имеет два т.н. центра (или, иначе, две словоформы-акцентоносителя). Так, в примере (5) первая встречаемость ИК-5 реализована на сегментном материале имени *Алексей Хомяков*, где на ударном слоге словоформы *Алексей* фиксируется подъем, а на ударном слоге словоформы *Хомяков* — падение. Аналогичное распределение движений тона можно наблюдать на сочетаниях *Иван Киреевский* и *братья Аксаковы*: на ударных слогах словоформ *Иван* и *братья* наблюдается подъем, *Киреевский* и *Аксаковы* — падение. Здесь перед нами особая — индивидуальная — манера конкретного говорящего (который остается, тем не менее, в рамках русской литературной нормы): при перечислении он использует интонационную стратегию целого предложения с подъемом в начале и завершающим падением в конце каждой фигуры перечисления. Перечисляя впоследствии философов-западников в той же речи, говорящий использует ту же фигуру перечисления. Заметим, что в качестве альтернативного описания вместо ИК-5 здесь можно было бы предложить интерпретацию в виде

последовательности «подъем-падение» (или «ИК-6 плюс ИК-1»), однако выбор альтернатив в данном случае не принципиален: существенно, что говорящий в перечислительном ряду неизменно использует рассмотренную здесь интонационную последовательность.

Для перечисления в целом характерен повтор интонационных конструкций, формирующих объекты перечисления, но в принципе, повтор не служит необходимым параметром перечисления: в речи свободно возникают перечислительные ряды с разнообразными способами интонационного оформления перечисляемых сущностей. Ср. об этом [Миловидова 1960: 207]. Приведем пример.

- (6) *На конференции были поставлены и решены следующие вопросы / разгром Германии / оккупация Германии и контроль над ней / репарации с Германии / вопрос о конференции объединённых наций.* [НКРЯ]



Рис. 6. Тонограмма примера (6)

Перед нами речь Ю. Левитана. Ключевое слово перечисления: *следующие*. Далее перечисляемые сущности имеют весьма разнообразное интонационное оформление. Это последовательность ИК-5, ИК-6, ИК-2, ИК-2. Наша гипотеза состоит в том, что именно разнообразие, вносимое в монотонную, как правило, реализацию перечисления, придает этой речи пафосное звучание. На первом шаге перечисления, которому соответствует именная группа *разгром Германии*, реализуется как ИК-5 с рельефным подъемом на ударном слоге словоформы *разгром* и падением, которое также реализовано в большом диапазоне частот, на ударном слоге словоформы *Германии*. Затем диктор делает паузу. Следующие два шага перечисления объединены в конъюнктивную группу с союзом *и*: на первом члене конъюнкции конструкция типа ИК-6, сопровождающаяся растяжкой, на втором члене — рельефное падение типа ИК-2. Затем опять следует эффектная пауза. На ударном слоге словоформы *репарации* в именной группе *репарации с Германии* рельефное падение ИК-2. Пауза. *Вопрос о конференции объединённых наций* реализуется с подъемами незавершенности типа ИК-6 на *вопрос*, *конференции* и *объединённых*, и, наконец, пологое завершающее падение на словоформе *наций*. Здесь перед нами особая стилистика дикторского чтения, при которой диктор отходит от монотонной манеры перечисления, присущей спонтанной речи. Эта стилистика характеризуется большими диапазонами частот при подъемах и падениях тона, эффектными паузами и разнообразием интонационной реализации однородных членов. Нельзя, однако, не признать, что отход от монотонной манеры перечисления «ослабляет» выразительную силу перечисления, заменяя ее на другие оттенки смысла.

### 3. Периферия семантического поля перечисления

Среди значений, близких перечислению, мы выделяем дискурсивные значения «переключки» (*Зарина! Джамиля! Гюзель!*) и стратегию эмфатического членения речи на сопоставимые кванты информации (см. Янко 2016). Обратимся к примерам «дальней» периферии. (При подсчете частотности средств выражения перечисления значения, относящиеся к «дальней» периферии, не учитывались).

В рабочем массиве имеется несколько фрагментов из фильмов, содержащих кадры «переключки». Так, в фильме «Белое солнце пустыни» красноармеец Сухов регулярно устраивает построения личного состава вверенного ему гарема. В этой ситуации используются как минимум две стратегии: 1) стратегия с использованием интонационной конструкции ИК-6, основанная только на дискурсивном значении незавершенности (пример (7)), и 2) специальная стратегия, имеющая иллокутивную функцию обращения (зова), которая, кроме интонации зова, несет указание на незавершенный характер списка лиц, к которым обращается говорящий, или, иначе, стратегия «переключки» (пример (8)). Сравним пару примеров (7) и (8).

- (7) *Ну / Джамиля / Зарина / Гюзель / Саида / Хафиза / Зухра / Лейла / Зульфия. До свиданья / барышни. Извините / коль что не так.* [НКРЯ]

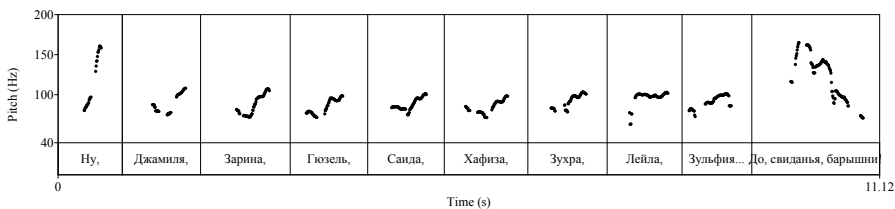


Рис. 7. Тонаграмма примера (7)

В примере (7) говорящий зачитывает список. На всех элементах перечисления фиксируется относительно пологий подъем плюс ровный тон на последнем и ударном слоге каждого из имен в списке, кроме словоформы *Лейла*. Имя *Лейла* имеет ударение на первом слоге, и такая структура слова наглядно демонстрирует подъем на ударном слоге плюс ровный тон на заударном, характерный для ИК-6. В примере (7) список так и остается незавершенным, потому что имя *Гюльчатай*, которое на переключках всегда звучало в конце списка, остается произнесенным: *Гюльчатай* погибла. Это заключительная переключка сценария. Первая же переключка в тот момент, когда Сухов только получает список девушек, представлена примером (8):

- (8) *Список / товарищ Сухов. Зарина! Джамиля! Гюзель! Саида! ...* [НКРЯ]



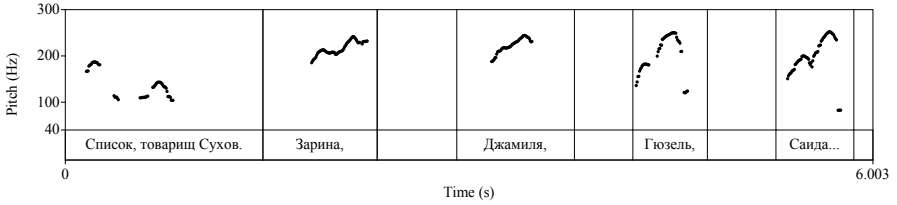


Рис. 8. Тоннограмма примера (8)

Здесь перед нами комбинация значений ‘звон в условиях относительной удаленности говорящего от слушающего’ (Сухов находится на некотором удалении от строя женщин, как командир перед строем солдат) и ‘незавершенность списка’. Перед нами особая интонация зова; как сугубо частная интонационная конструкция в инвентаре интонационных конструкций Е. А. Брызгуновой она не значит. Мы наблюдаем подъем на высоком уровне частот на предупредительном слоге плюс продолжение подъема на ударном. Назовем условно эту интонационную модель «интонацией переключки». Интонация переключки — одна из периферийных интонационных стратегий перечисления.

И наконец, к семантическому полю перечисления также можно отнести особую дискурсивную стратегию, которую мы называем стратегией членения информации на однородные кванты. Эта стратегия всегда реализуется в режиме повтора и, в контексте однородных членов предложения или однородных предложений, она становится одним из способов перечисления аргументов, которые должны убедить слушающего в точке зрения говорящего:

- (9) *Когда человек или группа людей слишком долго задерживаются на властных вершинах, исчезает обратная связь, исчезает уровень экспертного анализа...*

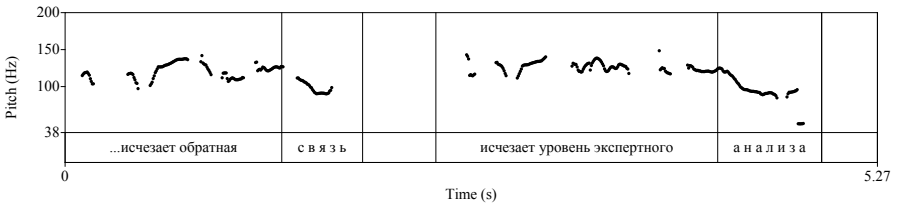


Рис. 9. Тоннограмма фрагмента примера (9)

В примере (9) носителями рассматриваемого акцента служат словоформы *связь* и *анализа*. Эта стратегия рассмотрена нами в [Янко 2016], где приводится подборка других примеров и обсуждаются проблема интонационной транскрипции для документирования особой интонационной конструкции, которая формирует эту стратегию. Эта проблема возникает в связи с тем, что в списке Е. А. Брызгуновой соответствующая модель отсутствует.

Для рассматриваемой интонационной конструкции выделяются следующие различительные признаки: 1) рельефное падение частоты на ударном

слоге словоформы-акцентоносителя; 2) подъем в небольшом диапазоне частот на заударных слогах акцентоносителя, если заударные есть (как в словоформе *анализа*); если заударных нет (как в словоформе *связь*) финальный подъем тона сохраняется, но смещается на финал конечного или единственного слога словоформы-акцентоносителя; финальный подъем не превышает максимума частот, которые формируют предшествующее подъему падение; 3) удлинение артикуляции, или растяжка; 4) повтор конструкции. Дискурсивная функция анализируемого акцента состоит в подчеркнутом членении речи на фрагменты, соответствующие отдельным квантам информации: говорящий «дробит» речь, чтобы придать отдельным квантам смысла автономное звучание.

## Заключение

Поле перечисления составляет целый спектр значений и соответствующих интонационных средств, в котором мы предлагаем выделять ядерный тип понятия перечисления и периферийные.

Ядро понятия перечисления образуют значения, которые выражаются интонационными конструкциями ИК-6, ИК-2, ИК-4 и ИК-3. Эти интонационные конструкции дают доступ к большинству значений, которые в контексте однородных членов предложения и однородных предложений, создают дискурсивную ситуацию перечисления. К «периферии» мы относим значения, близкие перечислению: это «переключка» и «стратегия членения речи на автономные кванты информации».

Заключим. Универсальной интонации, которая выражала бы значение перечисления в отвлечении от иллокутивных (создающих речевой акт, таких, как тема и рема) и дискурсивных (создающих связный дискурс) значений, не существует. Перечисление создается за счет наложения на ряд однородных членов предложения или однородных предложений разнообразных интонационных конструкций (необязательно повторяющихся), которые изначально выражают другие значения: тему, рему, незавершенность, «рассказ по порядку».

## Литература

1. Брызгунова Е. А. (1963) Практическая фонетика и интонация русского языка. М.: Издательство Московского университета, 1963.
2. Брызгунова Е. А. (1982) Интонация // Русская грамматика. Том 1. М. Наука, 1982.
3. Гвоздев А. Н. (2007) Вопросы изучения детской речи. М.: Детство-Пресс, 2007.
4. Канакина В. П. (2005) Русский язык. Методическое пособие. М.: Просвещение, 2005.
5. Миловидова Р. В. (1960) Интонация перечисления в русском языке сравнительно с русским языком // Экспериментальная фонетика и психология речи. Изд-во МГУ, 1960.
6. Пешковский А. М. (1928) Русский синтаксис в научном освещении. Изд. 3. М., Ленинград: Государственное издательство.

7. Черемисина-Ениколопова Н. В. (1999) Законы и правила русской интонации. М.: Флинта, Наука.
8. Янко Т. Е. (2016) Новые интонационные конструкции русского языка: разработка транскрипции // Компьютерная лингвистика и интеллектуальные технологии. По материалам ежегодной Международной конференции Диалог 2016. Вып. 15 (22), М.: Издательство РГГУ.
9. Boersma P., Weenink D. (2012). Praat: Doing phonetics by computer. Version 5.3.04. Online: <http://www.praat.org/>.

## References

1. Bryzgunova E. A. (1963) Practical phonetics and intonation of the Russian language [Prakticheskaja fonetika i intonatsija russkogo yazyka]. Moscow: Практическая фонетика и интонация русского языка. М.: Издательство Московского университета, 1963.
2. Bryzgunova E. A. (1982) Intonation [Intonatsija], Russian Grammar [Russkaja grammatika]. Vol. 1, Nauka, Moscow, pp. 98–118.
3. Cheremisina-Enikolopova N. V. (1999) Laws and rules of the Russian Intonation [Zakony i pravila russkoj intonatsii]. Moscow: Flinta, Nauka, 1999.
4. Gvozdev A. N. (2007) Voprosy izucheniya detskoj rechi. [Problems of infants speech]. Moscow: Detstvo-Press.
5. Kanakina V. P. (2005) Russian language. Methodological manual. [Russkij yazyk: Metodicheskoe posobie]. Moscow: Prosveshchenie.
6. Milovidova R. V. (1960) Intonation of enumeration in Russian as contrasted to German [Intonatsija perechislenija v russkom jazyke sravniteljno s nemetskim] // Experimental phonetics and psychology of speech [Eksperimentaljnaja fonetika i psihologija rechi]. Moscow: MGU, 1960.
7. Peshkovskij A. M. (1928) The Russian syntax in scientific interpretation [Russkij sintaksis v nauchnom osveshchenii], Leningrad: Gosudarstvennoe izdatel'stvo.
8. Yanko T. E. (2016) Prosodic transcription for the new Russian prosodic constructions [Novye intonatsionnye konstruksii russkogo yazyka: razrabotka transkriptsii] // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialog 2016” [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii: Trudy Mezhdunarodnoj Konferentsii “Dialog 2016”], Issue. 15 (22). Moscow.
9. Boersma P., Weenink D. (2012). Praat: Doing phonetics by computer. Version 5.3.04. Online: <http://www.praat.org/>.

# РУССКИЕ ПРЕДИКАТИВЫ В ЗЕРКАЛЕ ЭКСПЕРИМЕНТА И КОРПУСНОЙ ГРАММАТИКИ<sup>1</sup>

**Циммерлинг А. В.** (fagraey64@hotmail.com)

Московский педагогический государственный университет / Институт языкознания РАН, Москва, Россия

В статье описывается двойной эксперимент по определению объема и структуры активного словаря предикативов дативно-предикативной конструкции (ДПС). Вначале представляется онтология ситуаций, покрываемых данной конструкцией в русском языке, и тестовый словарь, состоящий из 422 стимулов, разбитых на 15 тематических классов. Словник тестируется на группе из 18 испытуемых и составляется выборка предикативов на основе рейтинга одобрения Socio. Тот же набор из 422 стимулов тестируется на НКРЯ и составляется вторая выборка на основе показателя  $m$ , отражающего число клауз ДПС в контексте мне Z-во. Объем и структура словаря ДПС определяются с помощью соотношения  $m$  и Socio. В ходе эксперимента строится модель конструкции ДПС, с выделением ее словарного ядра  $\{Lex_i\}$  и грамматического расширения  $\{Gram_i\}$ , предсказывающая, что носители языка могут использовать одинаковые принципы семантической селекции и добавления элементов в  $\{Gram_i\}$  при несовпадающем инвентаре лексических единиц.

**Ключевые слова:** грамматика конструкций, словарь, предикативы, дативно-предикативные структуры, идиолект, лексическое расширение, лексическая редукция, тематический класс

## RUSSIAN PREDICATIVES IN THE PERSPECTIVE OF SOCIOLINGUISTIC EXPERIMENT AND CORPUS GRAMMAR

**Zimmerling A. V.** (fagraey64@hotmail.com)

Moscow State University of Education / Institute of Linguistics,  
Russian Academy of Sciences, Moscow, Russia

This paper aims at measuring the active vocabulary of Russian predicatives licensing dative-predicative-structures (DPS) and introduces the results of the sociolinguistic experiment paired with the corpus study for the same set of stimuli. The sociolinguistic experiment aimed at measuring the

---

<sup>1</sup> Работа выполнена при поддержке гранта РФФ 16-18-00203а «Структура значения и ее отображение в системе лексических и функциональных категорий русского языка», реализуемого в МПГУ. Я благодарю анонимных рецензентов конференции «Диалог», а также А. Ч. Пиперски и Л. Л. Иомдина за ценные замечания. Ответственность за все недочеты лежит на авторе.

active DPS vocabulary in the idiolects of 18 native speakers of Russian and produced a sample of 422 stimuli ranked according to their acceptance rate. The same set of stimuli was tested on data from the Russian National Corpus, which produced the second sample based on the frequency ranks of DPS clauses in the diagnostic context *мне Z-во*, with the priority dative argument placed contiguously to the predicative at the distance <-1:1> in the same clause. The correlation of two ranked samples predicts the volume of the DPS vocabulary in an average idiolect of Russian and helps to establish the proportion between the lexical nucleus of the DPS construction and its extension. The analysis verifies the hypothesis that native speakers of Russian apply to the same underlying principles of semantic selection for DPS elements while using non-identical sets of lexical elements. The paper introduces a formal model of lexical extension with the entire DPS being grouped onto 15 thematic classes. The results prove that each thematic class includes high-, middle- and low-frequency DPS elements. Low-frequency DPS predicatives are modelled after high-frequency elements from the same thematic class.

**Key words:** construction grammar, lexicon, predicatives, dative-predicative-structures, idiolect, lexical extension, lexical reduction, thematic class

## 1. Конструкции ДПС

Дативно-предикативная-структура (ДПС), предполагающая маркировку субъекта несогласуемого неглагольного предикатива дат. п., — одна из наиболее известных черт русской грамматики [Щерба 2008; Поспелов 1955]. Вместе с тем, точно не установлено, какие лексемы обслуживают конструкцию ДПС. Теоретические соображения свидетельствуют о том, что валентность на дат. п. лица — характеристика семантического класса предикатов состояния в русском языке [Циммерлинг 1998; 1999; Бонч-Осмоловская 2009], но условия его пополнения неясны [Циммерлинг 2010; Say 2013].

## 2. Русские предикативы в теоретической и корпусной лингвистике

Доступного справочного ресурса, с помощью которого можно проверить состав класса предикативов ДПС, нет. Лакуна частично восполняется синтаксическим словарем системы «Этап 3» [Этап-3; Aresjan et alii 2003], который насчитывает около 800 единиц, для которых на основе регулярных или окказиональных употреблений типа *?мне было не очень эротично*, *?X-у было не эргономично делать Z* и т. п. постулируется валентность на дат. п. лица. Словари данного типа ориентированы на парсинг, поэтому учет окказионализмов, подтверждающих, что ЭРГОНОМИЧНО, ЭРОТИЧНО имеют валентность на дат. п. лица в речи отдельных носителей, необходим. Но это не доказывает совместимости подобных употреблений друг с другом в речи одного носителя русского языка. В «Грамматике 1980» [Шведова 1982: 151] ИГ в дат. п. лица при предикативе трактуется не как приоритетный актант, но как т. н.

детерминант. Предикативы ДПС изучались на материале НКРЯ, ср. [Маркова 2015], но прогресс в этом направлении сдерживается тем, что «с предикативами на -о само составление ... списка и определение того, насколько тот или иной предикат имеет валентность на дативный актант, является непростой исследовательской задачей» [Бонч-Осмоловская 2015: 83]. Еще хуже изучены неадеквативные предикативы (X-у) НЕ ПО СЕБЕ, ТАК И НАДО, ПО СИЛАМ и т. п. [Циммерлинг 2016: 362]. Такие выражения обычно не снабжаются пометой «предикатив» и не распознаются в качестве единых элементов.

## 2.1. Семантический инвариант конструкции ДПС

Если семантический инвариант конструкции ДПС существует, вероятным кандидатом является значение внутреннего состояния, соотношенного с одушевленным субъектом в конкретной референтной ситуации<sup>2</sup> [Циммерлинг 2014: 52].

- (i) валентность на дат. п. лица отражает способность предикатива ДПС обозначать актуальные положения, соотношенные с референтным одушевленным субъектом в течение некоторого отрезка времени.

Гипотеза (i) является модификацией тезиса Л. В. Щербы и Н. С. Поспелова [Поспелов 1955: 65] о том, что конструкция ДПС выражает в русском языке значение состояния. Она не требует представлять все предикативы ДПС как метафоры физического состояния или аффекта, ср. критику в [Селиверстова 1982: 131]. Конструкция ДПС покрывает некоторый спектр лексических значений, которые можно представить в виде онтологии ситуаций.

## 2.2. Семантическая селекция основ признаков слов

Употребление конструкции ДПС в русском языке поддерживается существованием большого числа признаков слов, которые в данной конструкции употребляться не могут, причем этот признак не скоррелирован с невозможностью предикативного употребления в структуре без согласования. Этот факт впервые отмечен в [Циммерлинг 1999: 224]. Предикаты типа *желтый* всегда обозначают свойства предмета или лица и не имеют синтаксических валентностей. Предикаты типа *солнечный* формируют нульместную схему, ср. *Сегодня солнечно, здесь пыльно*, но не конструкцию ДПС: *\*мне солнечно, \*нам пыльно*. Признак «быть СОЛНЕЧНО/ПЫЛЬНО» осмысливается как свойство внешней ситуации, одинаково воспринимаемое всеми потенциальными наблюдателями, и не может быть представлен как внутреннее состояние конкретного одушевленного субъекта в некоторый момент времени. Аналогичным образом, оценочные предикаты типа *аморальный*, которые допускают употребления *Z-во*,

---

<sup>2</sup> О других дативных схемах русского языка, где семантический субъект может быть неодушевленным, в том числе, модальной конструкции *пирогу надо остыть*, см. [Циммерлинг 2010; 2016: 358].

что *P*, *Z*-во *делать P*, отсеиваются конструкцией ДПС. Хотя квалификация поступка *U* как *аморального алогичного* может быть личным мнением говорящего, навязываемая языком концептуализация заставляет считать, что любой *X* оценит *U* как аморальный или алогичный. Наконец, предикаты, обозначающие активные свойства одушевленного субъекта, ср. *злобный, гневный, бешеный*, в отличие, например, от древнеисландского или болгарского [Циммерлинг 2010], тоже отсеиваются конструкцией ДПС.

В [Циммерлинг 1998: 73; 2010: 553] производящие основы прилагательных, от которых образуются только предикаты постоянного свойства, отнесены к типу т. н. актантно-поляризованных. Производящие основы другого типа порождают предикативы ДПС, ср. *X-у грустно*, но параллельно используются и для порождения предикатов свойства: *грустный X*. Данные основы определяются как амбивалентные. Основы третьего типа, которые определяются как ситуативно-поляризованные, используются только для порождения предикативов ДПС, а коррелятивные формы прилагательного утрачены: *X-у стыдно, совестно, \*стыдный X, \*совестный X*. Данную классификацию можно расширить, учтя предикативы неадеквативной морфологии, что показано в табл. 1.

**Таблица 1.** Типы производящих основ предикативов ДПС в русском языке

	Тип I: актантно-поляризованные основы	Тип II: амбивалентные основы	Тип III: ситуативно-поляризованные основы
А. Адеквативные основы	<i>желт-; пыльн-; аморальн-, алогичн-; гневн-, зло-; стыдлив-;</i>	<i>нужн-, грустн-;</i>	<i>стыдн-; совестн-; можн-;</i>
В. Неадеквативные основы	(...)	<i>по силам, по душе, некстати</i>	<i>жаль, пора, надо, нельзя, наплевать; невдомек; все равно; нечего, некогда</i>

### 2.3. Пополнение класса ДПС

Основной динамической тенденцией системы, представленной в табл. 1, является пополнение типов II и III. Главным ресурсом служит переход A1 → A2. Если в 1800–1850 гг. основной корпус НКРЯ фиксирует ок. 200 предикативов ДПС, к 1900 г. их уже более 300 [Циммерлинг, Астахова 2017], а в современном языке — около 500. Переход A1 → A2 иллюстрируется распространившимися в XX в. употреблениями *X-у слабо, жирно, лениво* и т. п., которые характерны для речи определенных поколений [Циммерлинг 2010: 554]. Переход A2 → A3, т. е. утрата производящей основы атрибутивной формы, редок, ср. *стыдно, совестно, боязно при \*???стыдный, \*???совестный, \*боязный* [Циммерлинг 1998: 73]. Переход B2 → B3 более обычен. В XIX в. *невдомек, невтерпеж, невмоготу* употреблялись не только в ДПС *X-у было Z-во*, но и в дативно-номинативной схеме *Y-и X-у были Z-ы*, что в современной норме не допускается. Под рубрику B2 → B3 можно также

подвести случаи размежевания предикатива и его субстантивного источника, отражающиеся в смене модели согласования: *X-у была (крайняя) неохота делать P* → *X-у было очень неохота делать P* и т. п.<sup>3</sup>

## 2.4. Модель лексического расширения конструкции ДПС

Гипотезу (i) можно проверить для ситуации, когда лексическое заполнение класса предикативов ДПС у разных носителей языка неодинаково, например, при синхронном варьировании или при диахроническом сдвиге. Построим модель расширения класса ДПС, при которой принцип тернарного распределения основ, иллюстрированный выше на рис. 1, т. е. семантическая избирательность конструкции ДПС, не нарушается.

Примем за точку отсчета употребление  $I_0^n$ , где имеется  $n$  предикативов ДПС. Содержательно  $I_0^n$  можно интерпретировать как идиолект-эталон, либо как обобщенное употребление некоторого периода, восстанавливаемое по корпусу. Множество  $\{n\}$  должно состоять из непустых подмножеств  $\{Lex\}$  и  $\{Gram\}$ , куда входят, соответственно, элементы, воспроизводимые как словарные единицы со своей моделью управления, и элементы, принадлежность которых к классу ДПС, т. е. наличие у них диагностической валентности на дат. п. лица, устанавливается на уровне речи, в соответствии с некоторыми правилами. Пусть  $I_1^{n+m}$  — идиолект, в котором имеется на  $m$  больше предикативов ДПС, чем в эталоне  $I_0^n$ . По условию, хотя бы часть предикативов множества  $\{Lex\}$ , элементы которого наследуются из словаря, должна быть общими в  $I_0^n$  и  $I_1^{n+m}$ . Расширение класса предикативов ДПС происходит за счет пополнения  $\{Gram\}$ . Возможны два сценария.

А.  $\{Gram_0\}$  и  $\{Gram_1\}$  имеют непустое пересечение, и сохраняются признаковые основы типа I, которым не соответствуют никакие элементы в  $\{Gram_0\} \cup \{Gram_1\}$ <sup>4</sup>.

В этом случае семантическая избирательность конструкции сохранится.

В.  $\{Gram_0\}$  и  $\{Gram_1\}$  имеют пустое пересечение, и сохраняются признаковые основы типа I, которым не соответствуют никакие элементы в  $\{Gram_0\} \cup \{Gram_1\}$ .

Дополнительное распределение элементов  $\{Gram_0\}$  и  $\{Gram_1\}$  совместимо с семантической избирательностью, если конструкция ДПС ориентирована не на конкретные лексемы, а на ряды значений. В этом случае тот факт, что, например, для выражения значения «X-у безразлично, P или не P» в  $I_0$  и  $I_1$  используются разные единицы, не нарушит принципов функционирования конструкции.

<sup>3</sup> Аналогично — при *охота, жаль, пора, грех, досуг, недосуг,лень*. О квазипоссессивных конструкциях с дат. п. и предикатным именем в истории русского языка см. [Июдин & Июдин 2014: 221].

<sup>4</sup> Данное условие является дефиниционным. Основы типа I являются источником пополнения класса предикативов ДПС. Если при расширении  $I_0^n \rightarrow I_1^{n+m}$  не останется основ типа I, принцип семантической селекции основ перестанет действовать.



Для основного сценария можно уточнить прогноз о пополнении класса ДПС:

- (ii) В ситуации расширения класса предикативов ДПС  $I_0^n \rightarrow I_1^{n+m}$  главным источником пополнения  $\{Gram_0\} \rightarrow \{Gram_1\}$  является переход основ признаков слов из типа I (актантно-поляризованные) в тип II (амбивалентные):  $A1 \rightarrow A2, A2 \rightarrow B2$ .

## 2.5. Семантическая избыточность рядов предикативов ДПС

В [Циммерлинг 2000: 199] утверждается, что предикативы ДПС в русском языке образуют ограниченно открытые ряды синонимов. Уточним эту гипотезу применительно к ситуации варьирования. Пусть в идиолекте  $I_1$  для выражения значения  $M$  в конструкции ДПС используется лексема  $a_1 \in \{a\}$ , а в идиолекте  $I_2$  — лексема  $a_2 \in \{a\}$ , и существует идиолект  $I_k$ , где для выражения данного значения используются  $a_1$  и  $a_2$ . Тогда в  $I_k$  имеет место семантическая избыточность, которая устранена в  $I_1$  и  $I_2$ .

## 3. Словник социолингвистического эксперимента

Для проверки гипотез (i) и (ii) был составлен словник предикативов ДПС, использующий принцип разбиения на тем. классы. Словник тестировался на группе носителей русского языка от 20 до 63 лет, уроженцев Москвы или Санкт-Петербурга, с высшим образованием. Всего было выделено 15 тем. классов. В основной список вошло 342 стимула. 80 стимулов дополнительного списка добавлялись в те же 15 классов, т. е. представляли собой их расширение.

### 3.1. Структура словника

Словник эксперимента представляет синтезированный идиолект. В тем. классы включались и предикативы ДПС, допустимые в литературном языке, *Х-у лень, охота*, и субстандартные формы, ср. *Х-у лениво, влом* и т. п. Словник соответствует модели идиолекта  $I_k$  с семантической избыточностью, где для выражения значения  $M^a$  в конструкции ДПС используется ряд лексем  $a_1, a_2 \dots a_k \in \{a\}$ . Каждому тем. классу соответствует набор значений  $M^a, M^b \dots M^k$ . Тем самым, используется представление о том, что конструкции ДПС в русском языке соответствует некоторый закрытый и строго конечный набор значений, при том, что в идиолектах этим значениям могут соответствовать нетождественные ряды лексем. Одобрение испытуемым элемента доп. списка может интерпретироваться либо как добавление конкретной лексемы к ряду, закреплённому за уже имеющимся значением, либо как добавление нового значения к набору значений, присущих тем. классу, либо как отступление от принципов функционирования всей конструкции ДПС в эталонном идиолекте.

**Таблица 2.** Тематические классы предикативов ДПС в словнике.

Прописными буквами выделены предикативы неадъективной морфологии

	Тем. класс	Примеры	Число стимулов осн. списка	Число стимулов доп. списка
1	Физические состояния	<i>Х-у холодно, плохо<sub>1</sub>, НЕ ПО СЕБЕ</i>	27	10
2	Модальности	<i>Х-у НАДО, неприлично, ПОРА</i>	44	5
3	Эмоциональные состояния	<i>Х-у грустно, НЕ ПО ДУШЕ.</i>	57	24
4	Моральные оценки	<i>Х-у стыдно, ГРЕХ, ТАК И НАДО</i>	16	0
5	Удобство исполнения	<i>Х-у удобно<sub>1</sub>, сподручно, ПО ПУТИ</i>	8	0
6	Уместность/неуместность	<i>Х-у уместно, В САМЫЙ РАЗ</i>	13	2
7	Внутренняя потребность	<i>Х-у НЕВМОГОТУ, ПОТРЕБНО.</i>	7	2
8	(не) соответствие задаче	<i>Х-у ВПОРУ, жирно, слабó, ЛЕНЬ</i>	11	0
9	Трудность выполнения	<i>Х-у сложно, трудно, нелегко</i>	10	4
10	(Не) желание выполнять	<i>Х-у ОХОТА, НЕОХОТА, ВЛОМ</i>	9	1
11	Общая оценка	<i>Х-у хорошо, плохо<sub>2</sub>, НЕ ПО СЕБЕ</i>	41	7
12	(Не) релевантность	<i>Х-у важно, НАПЛЕВАТЬ</i>	16	4
13	Эффективность	<i>Х-у вредно, полезно, БЕЗ ТОЛКУ</i>	6	1
14	Сенсорные и интеллектуальные реакции	<i>Х-у видно, слышно, ясно, известно, любопытно</i>	25	12
15	Параметризуемый признак	<i>Х-у темно, великó, ПО ПОЯС.</i>	52	8
<b>ВСЕГО:</b>			<b>342</b>	<b>80</b>

### 3.2. Тестовое задание

Испытуемые должны были оценить 422 стимула в плане употребления в одном из четырех контекстов: 1) *Мне было очень приятно.* 2) *Мне было очень приятно сделать Р.* 3) *Мне приятно, что Р.* 4) *Мне это очень приятно* (где слово *это* не отсылает к существительному ср. р.). Инструкция предписывала ответ «ДА» в случае, если испытуемый может сам употребить данный стимул хотя бы в одном из 4 перечисленных диагностических контекстов, и «НЕТ» во всех остальных случаях, включая знакомство с соответствующими употреблениями в речи других людей. Разрешалось менять формы лица, местоимения на существительные, порядок слов и лексическое наполнение.

### 4. Результаты эксперимента

В таблице 3 представлены данные об употреблении предикативов ДПС в идиолектах  $I_1$  —  $I_{18}$ . Возраст указан на момент эксперимента (март 2016 г.). В нижней строчке указан процент от общего числа стимулов осн. списка (342) в идиолекте каждого информанта.

**Таблица 3.** Заполнение тем. классов предикативов ДПС в идиолектах испытуемых. Зеленым выделены идиолекты с лексическим расширением, красным — идиолекты с лексической редукцией

		ИДИОЛЕКТЫ ИНФОРМАНТОВ																	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		F25	M51	F37	F61	M63	F61	F32	F59	F35	M25	M27	F56	F25	F34	F28	F22	M55	M20
Осн.		190	266	205	282	268	288	210	169+7	260	244	292	210	205	236	275	139	306	313
Доп.		5	0	9	14	9	20	6	7	12	10	25	6	10	6	41	1	27	38
		55,5%	77,7%	59,9%	82,4%	78,4%	84,2%	61,2%	49,4%	76%	71,3%	85,4%	61,4%	59,9%	69%	80,4%	40,6%	89,5%	91,5%

#### 4.1. Грамматика идиолектов

В 16 идиолектах из 18 объем словаря ДПС уместился в диапазон от 49,3% ( $I_8$ ) до 89,5% ( $I_{17}$ ). Стоит исключить аномально низкий результат  $I_{16}$  (40,5%, 139 предикативов), и аномально высокий результат  $I_{18}$  (91,5%, 313 предикативов). В  $I_{16}$  и  $I_8$  большинство стимулов представляют {Lex}, т.е. носители данных идиолектов не пользуются или мало пользуются правилами расширения {Gram}. У  $I_{18}$  налицо проблемы с семантической селекцией, он забраковал всего 29 стимулов осн. списка и 42 элемента доп. списка. Тем самым,  $I_{18}$  с трудом вычлняет основы класса I. У  $I_{15}$  аномально высок процент одобренных элементов доп. списка (41 из 80), но процент элементов осн. списка (80,4%, 275 стимулов) находится в пределах основного диапазона. Тем самым,  $I_{15}$  — информант со стандартной для русского языка грамматикой ДПС, но смещенной границей между элементами осн. и доп. списка.

##### 4.1.1. Редукционные идиолекты и идиолекты с расширением

Чтобы интерпретировать синхронное варьирование словаря ДПС как случай расширения/редукции при общих принципах селекции стимулов, нужно выбрать в качестве эталона  $I_0$  значение в нормальном диапазоне. Если взять значение нижней медианы выборки  $I^m = I_{10} = 71,3\%$ , редукционными будут  $I_{16}$ ,  $I_8$  и  $I_1$ ,  $I_3$ ,  $I_{13}$ ,  $I_7$ ,  $I_{12}$ ,  $I_{14}$ .

#### 4.2. Ранжирование тестовых стимулов по рейтингу Socio

18 положительных ответов из 18 соответствуют высшему рангу 1, а 0 положительных ответов из 18 — низшему рангу 19. 48 предикативов (14% осн. списка) имеют ранг 1, что можно расценить как подтверждение того, что они воспроизводятся как словарные единицы и принадлежат {Lex<sub>k</sub>}. 4 элемента получили 0 баллов, т.е. они отсутствуют в активном словаре ДПС у всех информантов. Далее количество баллов называется 'рейтингом Socio'.

#### 4.2.1. Наборы стимулов

Из 80 стимулов доп. списка 68 (85%) относятся к 6 низшим рангам ( $Socio < 6$ ), всего 7 предикативов (*X-у обязательно, некомфортно, чуждо, симпатично, дешево, твердо, неплохо*) получило более 50% положительных ответов ( $Socio \geq 9$ ). Напротив, в осн. списке более 50% положительных ответов получило 245 стимулов (71,6%). Тем самым, разделение стимулов на осн. и доп. списки подтвердилось, а исключения касаются небольшого числа единиц, где оценки приемлемости в эталоне  $I_0$  и усредненном идиолекте выборки  $I_k$  разошлись.

#### 4.2.2. Дополнительные стимулы

В  $I_0$  перечисленные 7 предикативов не входят в  $\{Lex_0\}$  и отсеиваются из  $\{Gram_0\}$ , так их основы в  $I_0$  относятся к типу I (актантно-поляризованные). В  $I_k$  данные 7 предикативов либо входят в  $\{Lex_k\}$ , либо помещаются в  $\{Gram_k\}$  благодаря переходу в тип амбивалентных ( $A1 \rightarrow B1$ ) по принципу (ii).

#### 4.2.3. Основные стимулы

74 стимула, т. е. 21,7% осн. списка, получили меньше 50% положительных ответов. 64 стимула, т. е. 18,7% осн. списка относятся к диапазону  $9 \leq Socio \leq 12$ . Их естественно отнести к  $\{Gram_k\}$ . Логично принять, что стимул не воспроизводится как словарная единица, если он получает треть и более отрицательных ответов.

### 4.3. Интерпретация данных

Эксперимент подтвердил, что класс предикативов ДПС в речи испытуемых не наследуется из словаря целиком, а достраивается по правилам грамматики. Правила селекции у большинства информантов оказались сходными при значительном варьировании лексического наполнения класса предикативов ДПС. Модель расширения  $\{Gram_k\}$  применима к большинству идиолектов. Нижняя медиана выборки  $I^m = 71,3\%$  соответствует уровню в 244 стимула осн. списка. Почти идентичный показатель (245) дает рейтинг Socio: этому числу соответствует та часть стимулов осн. списка, которую одобрили более 50% испытуемых, т. е. элементы с  $Socio \geq 9$ .

## 5. Корпусная грамматика предикативов ДПС

Второй этап исследования был связан с тестированием 422 стимулов на основном корпусе НКРЯ<sup>5</sup>.

### 5.1. Аппроксимативная мне-мера

Использовалась выборка употреблений предикатива ДПС с 1 л. ед.ч. личного мест. на расстоянии  $1 <-1: 1>$ , т. е. контексты *мне Z-во ~ Z-во мне*. Обоснование такого решения состоит в том, что большинство предикативов

---

<sup>5</sup> 23803881 предложение, дата обращения 05.03.2017.

ДПС — автореферентные слова, ориентированные на употребление в 1 л. ед. ч. наст. вр. [Циммерлинг 2014]. Использовалась функция точного поиска, с последующей очисткой выборки аннотатором. Реализации ДПС и все прочие входящие элемента в корпус трактовались как омонимы.

## 5.2. Выборка по мне-мере и общее число вхождений

Между числом клауз с ДПС в выборке по *мне-мере* ( $m$ ), и общим числом вхождений элемента (с учетом омонимии) в НКРЯ ( $N$ ), нет явной корреляции. При больших значениях  $n$  коэффициент  $m/N$  зависит не столько от абсолютной величины  $m$ , сколько от продуктивности употреблений, не связанных с ДПС.

## 5.3. Частотные классы предикативов

Выборка делалась для тех же 342 стимулов осн. и 80 стимулов доп. списка. Кроме того, проверялись 56 стимулов, которые не вошли в анкету: данная группа указана в таблице 4 в строке «компенсированный список». Стимулы были распределены по 11 частотным классам: для класса I,  $m > 1000$ , в классе XI,  $m = 0$ , т. е. выборка дает 0 клауз ДПС.

**Таблица 4.** Распределение предикативов ДПС по частотным классам в выборке НКРЯ по *мне-мере*

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	> 1000	400–1000	200–399	100–199	50–99	20–49	10–19	5–9	2–4	1	0
Основ.	7	17	17	18	24	25	37	33	50	35	79
Доп.	0	0	0	0	0	1	0	3	9	9	58
Компенс.	1	0	0	2	3	6	1	9	7	4	23
	8	17	17	20	27	32	38	45	66	48	160

## 5.4. Соотношение ранжированных выборок

422 стимулов анкеты, ранжированных по Socio, были поделены на 7 групп в диапазонах: 16–18 баллов, 13–15 баллов, 9–12 баллов, 6–8 баллов, 3–5 баллов, 1–2 балла, 0 баллов. Соотношение этой выборки с распределением предикативов ДПС по частотным классам, показывает таблица 5.

**Таблица 5.** Соотношение ранжированных выборок эксперимента и выборки НКРЯ по *мне-мере* для предикативов ДПС (по осн. и доп. спискам)

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI		
	> 1000	400–1000	200–399	100–199	50–99	20–49	10–19	5–9	2–4	1	0		
16–18 баллов (137)	7	16	14	13	19	15	17	9	12	7	8	137	
13–15 баллов (70)	0	0	2	3	3	6	14	8	17	1	17	70	
9–12 баллов (70)	0	1	1	1	1	3	2	11	12	17	21	70	
6–8 баллов (42)	0	0	0	1	1	1	1	4	5	10	19	42	
3–5 баллов (44)	0	0	0	0	0	0	2	3	9	3	27	44	
1–2 балла (39)	0	0	0	0	0	1	1	0	3	5	29	39	
0 баллов (20)	0	0	0	0	0	0	0	1	1	1	17	20	
422	7	17	17	18	24	26	37	36	59	44	137		
	Высокочастотные предикативы (146)							Среднечастотные предикативы (95)		Низкочастотные предикативы (181)			

Высокочастотные предикативы почти всего являются высокорейтинговыми, в первых 6 частотных классах ( $m \geq 20$ ) всего 4 предикатива с  $Socio < 9,3$  из которых являются архаизмами. Обратное не всегда верно: в частотном классе XI ( $m = 0$ ) есть предикативы всех групп, хотя низкорейтинговые ( $Socio < 9$ ) преобладают (68, 14%).

#### 5.4.1. Высокочастотные предикативы

Емкость словаря ДПС в среднем идиолекте находится на уровне 244–245 стимулов. Такое количество единиц соответствует суммарному объему частотных классов I–VIII ( $m \geq 5$ ) и части класса IX ( $m \geq 2$ ). Примем, что большая часть этих элементов принадлежит  $\{Lex_k\}$ . Тогда понижение рейтинга Socio интерпретируется по модели редукции. Нижнюю границу  $\{Lex_k\}$  стоит ограничить классом VII ( $m \geq 10$ ), после него в выборке возрастает доля предикативов с  $Socio < 13$ . Тем самым, в группе высокорейтинговых/высокочастотных оказывается 145 стимулов.

#### 5.4.2. Среднечастотные предикативы

К среднечастотным относятся стимулы классов VIII и IX, где  $2 \leq m \leq 9$ . В данной зоне 84 элемента, они могут относиться либо к  $\{Lex_k\}$ , либо к  $\{Gram_k\}$ . Ключ дает разбиение употреблений на два периода: 1760–1974 гг. vs. 1975–2016 гг. Знание вектора развития, значений Socio и  $m/N$  дает достаточные сведения для проверки гипотез о том, реализует ли употребление элемента сценарий редукции или расширения.

#### 5.4.3. Низкочастотные предикативы

К данной группе относятся 113 элементов классов X ( $m = 1$ ) и XI ( $m = 0$ ). 33 из них имеют высокий рейтинг Socio (от Socio = 13 до Socio = 18). Низкий показатель  $m$  в данном случае отражает то, что эти предикативы описывают онтологически ограниченные ситуации. Информанты легко восстанавливают

контекст, где уместны высказывания *X-у малó* <о платье>, *X-у горько* <во рту>, <в ручье> *X-у было по щиколотку*, но соответствующие контексты редко фиксируются корпусом.

### 5.5. Тематические классы

Соотношение частотных и тем. классов показано в таблице 6.

**Таблица 6.** Распределение 342 предикативов осн. списка по частотным классам НКРЯ в выборке по мне-мере

Частотные классы			Тематические классы														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Высокочастотные предикативы	I	>1000		3	1						1			1		1	
	II	400–999	1	5	5	1							2			2	1
	III	200–399		1	5	2					1	1	1	2		2	2
	IV	100–199	3	3	3	1					1		3			3	1
	V	50–99	5	4	2		1			1	2	2	1	1		1	4
	VI	20–49	1	1	6	2			1	1				3	2	5	3
	VII	10–19	2	5	10	2	3	1	2			1	2	2	2	1	4
				12	22	32	8	4	1	3	2	5	4	9	9	4	15
Среднечастотные	VIII	5–9	4	4	3	2	1	0	2	3	2	0	7	1	0	2	3
	IX	2–4	5	7	9	3	0	4	0	2	1	2	8	0	1	2	6
			9	11	12	5	1	4	2	5	3	2	15	1	1	4	9
Низкочастотные	X	1	2	2	5	3	1	3	0	0	1	0	2	2	0	3	10
	XI	0	4	9	8	0	2	5	2	4	1	3	15	4	1	3	18
			6	11	13	3	3	8	2	4	2	3	17	6	1	6	28
Емкость класса			27	44	57	16	8	13	7	11	10	9	41	16	6	25	52

145 высокочастотных предикативов ДПС ( $m > 9$ ) интерпретируются как ядро словаря ДПС, т. е. элементы  $\{Lex_k\}$ . 113 низкочастотных предикативов ДПС ( $m \leq 1$ ), интерпретируются как расширения  $\{Gram_k\}$ , либо как элементы, исключенные из  $\{Lex_k\}$  по сценарию лексической редукции. 84 среднечастотных предикатива ДПС ( $1 < m < 10$ ) образуют промежуточную зону между ядром и расширением конструкции.

## 6. Выводы и перспективы

Проведенный двойной эксперимент позволил установить объем класса предикативов ДПС в идиолектах русского языка и смоделировать усредненный идиолект  $I_k$  с подмножествами  $\{Lex_k\}$  (словарь) и  $\{Gram_k\}$  (грамматика) по данным НКРЯ. Гипотеза о том, что одна часть предикативов ДПС принадлежит общерусскому словарю  $\{Lex_k\}$  и является общей для всех носителей,

а другая часть предикативов ДПС порождается информантами в речи по правилам семантической селекции и добавляется в {Gram<sub>i</sub>} в их идиолектах, подтвердилась. Средний объем класса ДПС в идиолекте носителя современного русского языка составил 244 стимулов осн. списка (71,3%), что почти совпало с числом стимулов осн. списка, одобренных более 50% испытуемых — 245 стимулов, 71,6%.

Целью проведенного эксперимента было тестирование словника конструкции ДПС, а не изучение варьирования в разных социальных и возрастных группах. В то же время, для оценки устойчивости результатов полезно уточнить требования, предъявляемые к испытуемым и к корпусам текстов. От информанта требуется способность строить проверочный контекст для каждого стимула. Заранее неизвестно, обладают ли все испытуемые способностью к семантической селекции стимулов ДПС (в проведенном эксперименте такая способность была подтверждена для 16 из 18 человек), но для проверки гипотезы о лексическом расширении {Gram<sub>k</sub>} нужно, чтобы большинство испытуемых группы такой способностью обладали. Для создания ранжированной корпусной выборки предикативов ДПС подходит любой корпус, который отражает все 15 тем. классов словника (т.е. полную онтологию конструкции ДПС) и включает предикативы, зафиксированные в разные периоды существования языка.

## Литература

1. *Апресян Ю. Д., Богуславский И. М., Иомдин Л. Л., Санников В. З.* (2010) Теоретические проблемы русского синтаксиса: Взаимодействие лексики и грамматики / отв. ред. Ю. Д. Апресян. М.: 408 с.
2. *Бонч-Осмоловская А. А.* (2009) Дативные субъектные конструкции с предикативами на — о/-е // Корпусные исследования по русской грамматике. Ред.-сост. К. Л. Киселева, В. А. Плунгян, Е. В. Рахилина, С. Г. Татевосов. М. С. 157–183.
3. *Бонч-Осмоловская А. А.* (2015) Квантитативные методы в диахронических корпусных исследованиях // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2015», вып. 14 (21). М. С. 80–94.
4. *Иомдин Б. Л., Иомдин Л. Л.* (2014) Валентности русских предикатных существительных и микросинтаксические конструкции // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции «Диалог-2014», вып. 13 (20). М. С. 219–231.
5. *Маркова В. Р.* (2015) Предикативы состояния с финалью -о в русском языке // Типология морфосинтаксических параметров. Материалы международной конференции «Типология морфосинтаксических параметров 2015» / Под ред. Е. А. Лютиковой, А. В. Циммерлинга, М. Б. Коношенко. Вып. 2. М.: С. 266–278.
6. *Поспелов Н. С.* (1955) В защиту категории состояния // Вопросы языкознания. 1955. № 2. С. 55–65.



7. Селиверстова О. Н. (1982) Второй вариант классификационной сетки и описание Некоторых предикатных типов русского языка // Семантические типы предикатов / Отв. ред. О. Н. Селиверстова. М. С. 88–157.
8. Циммерлинг А. В. (1998) История одной полемики // Язык и речевая деятельность. 1998. № 1. С. 63–87.
9. Циммерлинг А. В. (1999) Субъект состояния и субъект оценки // Логический анализ естественного языка. Образ человека в культуре и языке. П. ред. Н. Д. Арутюновой, И. Б. Левонтиной. М. С. 221–228.
10. Циммерлинг А. В. (2000) Этические концепты и семантические поля // Н. Д. Арутюнова, Т. Е. Янко, Н. К. Рябцева (ред.). Логический анализ языка. Языки этики. М. С. 190–199.
11. Циммерлинг А. В. (2010) Именные предикативы и дативные предложения в европейских языках // Компьютерная лингвистика и интеллектуальные технологии. Вып. 9 (16). Материалы международной конференции «Диалог 2010». М. С. 549–558.
12. Циммерлинг А. В. (2014) Предикаты состояния и фактивность // Вопросы филологии. 2014. № 3–4. С. 52–60.
13. Циммерлинг А. В. (2016) Предикативы параметрического признака в русском языке // Труды ИРЯ РАН. Вып. 10. М. С. 358–369.
14. Циммерлинг А. В., Астахова Т. С. (2017) Корпусная грамматика и язык русских писателей XIX в. М., 2017. В печати.
15. Щерба Л. В. (2008). О частях речи в русском языке // Щерба Л. В. Языковая система и речевая деятельность. Изд. 4-е. М. С. 77–100.
16. Шведова Н. Ю. (1982). Распространение предложения в целом (детерминация) // Русская грамматика, том 2, М. С. 149–180.
17. Apresjan Ju., Boguslavsky I., Iomdin L. et al. (2003) ETAP-3 Linguistic Processor: a Full-Fledged NLP Implementation of the MTT // First International Conference on Meaning-Text Theory (MTT'2003). June 16–18, 2003. Paris, 2003. pp. 279–288. (Соавторы: Ю. Д. Апресян, И. М. Богуславский, А. В. Лазурский, В. З. Санников, В. Г. Сизов, Л. Л. Цинман)
18. Say S. (2013). On the nature of dative arguments in Russian constructions with “predicatives” // Current studies in Slavic linguistics. Amsterdam, Philadelphia, 2013. pp. 225–246.
19. Лингвистический процессор ЭТАП-3: <http://proling.iitp.ru/ru/etap3>

## References

1. Apresjan Ju., Boguslavsky I., Iomdin L. et al. (2003) ETAP-3 Linguistic Processor: a Full-Fledged NLP Implementation of the MTT // First International Conference on Meaning-Text Theory (MTT'2003). June 16–18, 2003. Paris: Ecole Normale Supérieure, 2003, pp. 279–288. (Soavtory: Ju. D. Apresjan, I. M. Boguslavskij, A. V. Lazurskij, V. Z. Sannikov, V. G. Sizov, L. L. Cinman)
2. Say S. (2013). On the nature of dative arguments in Russian constructions with “predicatives” // Current studies in Slavic linguistics. Amsterdam, Philadelphia, 2013, pp. 225–246.

3. *Apresjan Ju. D., Boguslavskij I. M., Iomdin L. L., Sannikov V. Z.* (2010), Theoretic problems of Russian syntax: interaction of lexicon and grammar, Ju. D. Apresjan (ed.) [Teoreticheskie problemy russkogo sintaksisa: Vzaimodejstvie leksiki i grammatiki / otv. red. Ju. D. Apresjan]. — Moscow, 408 p.
4. *Bonch-Osmolovskaja A. A.* (2009), Dative subject constructions with the -o/-e final [Dativnye sub#ektnye konstrukcii s predikativami na — o/-e], *Corpus studies in Russian grammar*, K. L. Kiseleva, V. A. Plungian, E. V. Rahilina, S. G. Tatevosov (eds.) [Korpusnye issledovanija po russkoj grammatike. Red.-sost. K. L. Kiseleva, V. A. Plungjan, E. V. Rahilina, S. G. Tatevosov]. — Moscow, pp. 157–183.
5. *Bonch-Osmolovskaja A. A.* (2015), Quantitative methods in diachronic corpus studies [Kvantitativnye metody v diahronicheskikh korpusnyh issledovanijah], *Computational linguistics and intellectual technologies: Proceedings of the international conference “Dialogue 2015”*, Vol. 14 (21) [Komp’juternaja lingvistika i intellektual’nye tehnologii: Trudy mezhdunarodnoj konferencii «Dialog-2015», Vyp. 14 (21)], Moscow, pp. 80–94.
6. *Iomdin B. L., Iomdin L. L.* (2014), Valencies of Russian predicate nouns and microsyntactic constructions [Valentnosti russkikh predikatnyh sushhestvitel’nyh i mikrosintaksicheskie konstrukcii], *Computational linguistics and intellectual technologies: Proceedings of the international conference “Dialogue-2014”*, Vol. 13 (20) [Komp’juternaja lingvistika i intellektual’nye tehnologii: Trudy mezhdunarodnoj konferencii «Dialog-2014», Vyp. 13 (20)], Moscow, pp. 219–231.
7. *Markova V. R.* (2015), Predicates of State with the –o –final in Russian [Predikativy sostojanija s final’ju -o v russkom jazyke], *Typology of morphosyntactic parameters: Proceedings of the international conference “Typology of morphosyntactic parameter 2015”*, E. A. Lyutikova, A. V. Zimmerling, M. B. Konoshenko (eds.), Vol. 2 [Tipologija morfosintaksicheskikh parametrov. Materialy mezhdunarodnoj konferencii «Tipologija morfosintaksicheskikh parametrov 2015» / Pod red. E. A. Lyutikovej, A. V. Zimmerlinga, M. B. Konoshenko. Vyp. 2. ], Moscow, pp. 266–278.
8. *Pospelov N. S.* (1955), Defending the Category of State [V zashhitu kategorii sostojanija], *Issues in linguistics [Voprosy jazykoznanija]*, 1955. № 2, pp. 55–65.
9. *Seliverstova O. N.* (1982), A second variant of the classification of the selected predicate types in Russian [Vtoroj variant klassifikacionnoj setki i opisanie nekotoryh predikatnyh tipov russkogo jazyka] // *Semantic predicate types*. O. N. Seliverstova (ed.) [Semanticheskie tipy predikatov / Otv. red. O. N. Seliverstova], Moscow, pp. 88–157.
10. *Shvedova N. Ju.* (1982), Extension of the whole sentence (determination) [Rasprostranenie predlozhenija v celom (determinacija)], *Russian grammar*, Vol. 2 [Russkaja grammatika, tom 2], Moscow, pp. 149–180.
11. *Shcherba L. V.* (2008), On parts of speech in Russian [O chastjah rechi v russkom jazyke], *Shcherba L. V. Language system and speech activity [Shcherba L. V. Jazykovaja sistema i rechevaja dejatel’nost’*. Izd. 4-e.], Moscow, pp. 77–100.
12. *Zimmerling A. V.* (1998), The history of a polemic [Istorija odnoj polemiki], *Language and speech activity [Jazyk i rechevaja dejatel’nost’]*, 1998. № 1, pp. 63–87.

13. *Zimmerling A. V.* (1999), The Subject of State and the Subject of Evaluation [Sub#ekt sostojanija i sub#ekt ocenki], Logical analysis of a natural language: the Concept of a Human Being in language and culture, N. D. Arutyunova, I. B. Levontina (eds.) [Logicheskij analiz estestvennogo jazyka. Obraz cheloveka v kul'ture i jazyke. P.red. N. D. Arutyunovoj, I. B. Levontinoj], Moscow, pp. 221–228.
14. *Zimmerling A. V.* (2000), Ethic concepts and semantic fields [Eticheskie koncepty i semanticheskie polja], Logical analysis of a natural language: the languages of ethics, N. D. Arutyunova, T. E. Yanko, N. K. Rjabceva (eds.) [N. D. Arutjunova, T. E. Yanko, N. K. Rjabceva (red.). Logicheskij analiz jazyka. Jazyki etiki]. Moscow, pp. 190–199.
15. *Zimmerling A. V.* (2010), Nominal predicatives and dative sentences in European languages [Imennye predikativy i dativnye predlozhenija v evropejskih jazykah], Computational linguistics and intellectual technologies, Vol. 9 (16): Proceedings of the international conference “Dialogue 2010” [Komp'juternaja lingvistika i intellektual'nye tehnologii. Vyp. 9 (16). Materialy mezhdunarodnoj konferencii «Dialog 2010»], Moscow, pp. 549–558.
16. *Zimmerling A. V.* (2014). Predicates of State and factivity [Predikaty sostojanija i faktivnost'], Issues in philology [Voprosy filologii], 2014, № 3–4, pp. 52–60.
17. *Zimmerling A. V.* (2016), Predicatives of parametric features in Russian [Predikativy parametriceskogo priznaka v russkom jazyke], Proceedings of the Vinogradov Institute of Russian language [Trudy IRJa RAN], Vyp. 10, 2016, pp. 358–369.
18. *Zimmerling A. V., Astahova T. S.* (2017). Corpus grammar and language of Russian authors of the XIX century [Korpusnaja grammatika i jazyk russkih pisatelej XIX v], Moscow, 2017. In press.
19. *Shherba L. V.* (2008), On parts of speech in Russian [O chastjah rechi v russkom jazyke], Shherba L. V. Language system and speech activity [Shherba L. V. Jazykovaja sistema i rechevaja dejatel'nost'. Izd. 4-e.], Moscow, pp. 77–100.
20. *Shvedova N. Ju.* (1982), Extension of the whole sentence (determination) [Rasprostranenie predlozhenija v celom (determinacija)], Russian grammar, Vol. 2 [Russkaja grammatika, tom 2], Moscow, pp. 149–180.
21. Linguistic processor ETAP-3 [Lingvisticheskij processor ETAP-3]:<http://proling.iitp.ru/etap3>

## Abstracts

### POSITIVE AND NEGATIVE POLARITY: SEMANTIC UNDERPINNINGS

**Apresjan V. Ju.** (valentina.apresjan@gmail.com), National Research University Higher School of Economics; Vinogradov Russian Language Institute of the Russian Academy of Sciences; Moscow, Russia

Negative and positive polarity items (NPIs and PPIs) are one of the well-explored topics in formal semantics and typology. However, the phenomenon of polarization is only addressed on a very limited linguistic material, such as indefinite pronouns (*some vs. any*), temporal adverbs (*yet vs. already*), certain idioms (*not to lift one's finger*), expressions of attitude (*would rather, unfortunately*). The main focus of polarity research are licensing contexts, or contexts that allow to use NPIs or PPIs, respectively. The current paper introduces considerable new body of NPIs and PPIs, explores semantic underpinnings of polarization, demonstrates that polarization is an all-pervasive property of language and shows that it is a scalar feature. The paper provides evidence for the fact that polarization is a regular phenomenon in polysemy, especially in verbs, where it characterizes single senses of polysemous words rather than words as a whole. It further establishes semantic connections between the phenomena of negative and positive polarization. The degree of polarization depends on semantic structure: the more "weighty" is the modal frame, the more polarized an item is.

### "XENO" MARKERS IN THE LIGHT OF THE DATA OF PARALLEL CORPORA AND CONTEMPORARY MASS MEDIA: THE CASE OF THE RUSSIAN WORD *JAKOBY*

**Apresjan V. Ju.** (valentina.apresjan@gmail.com)<sup>1,2</sup>,  
**Shmelev A. D.** (shmelev.alexei@gmail.com)<sup>3,2,4</sup>

<sup>1</sup>National Research University Higher School of Economics; <sup>2</sup>Vinogradov Russian Language Institute of the Russian Academy of Sciences, Moscow, Russia; <sup>3</sup>Moscow Pedagogical State University, Moscow, Russia; <sup>4</sup>St Tikhon's Orthodox University

The paper considers the less known aspects in the functioning of Russian lexical "xeno" markers, in particular, of the particle *jakoby* 'allegedly, ostensibly'. Traditionally described as expressing the falsity of a proposition contained in somebody's utterance, in conjunction with a negative assessment of the utterer as aware of its falsity, *jakoby* displays very different usages in the language of contemporary mass media. Namely, it is frequently used as a mere marker of evidentiality, without an obligatory assessment of the proposition as false or of its source as untruthful. In fact, it can even be used to refer to statements that are treated as true within the very same text, only to indicate that the source of this information is not the writer herself but somebody else (e.g., a different news agency), in what might be termed as "safety" strategy.

Besides, *jakoby* in its mass media usages demonstrates unusual syntactic behaviors, namely shifts in scope, where it is placed before the speech verb rather than before the challenged proposition: *jakoby utverzhdat', chto P 'jakoby claim that P'* instead of *utverzhdat', chto jakoby P 'claim that jakoby P'*.

However, the study of the Russian-English parallel corpus reveals that these usages are not as unusual as they may appear. In Russian translations of English texts *jakoby* sometimes functions as a translation of the English *supposedly, allegedly, ostensibly* or other (e.g., verbal) markers of uncertainty, but more frequently occurs with no apparent stimulus in the source, merely to mark indirect quotation.

It appears therefore that there is a certain need in the Russian language for a neutral evidentiality marker. It is occasionally filled with *jakoby*, which in this case displays a tendency for grammaticalization: it expresses that the source of information is other than the speaker herself (but contains no other semantic components), and takes syntactic scope over the speech verb instead of the proposition it challenges.

## A CORPUS-BASED STUDY OF FACTORS AND MODELS METAPHORICITY; EVIDENCE FROM RUSSIAN VERBS

**Badryzlova Yu. G.** (yuliya.badryzlova@gmail.com), National Research University Higher School of Economics (HSE), Moscow, Russia

The paper presents a tentative scheme for the analysis and annotation of the factors that contribute to metaphoricity of verbs in real-world data. The need for a comprehensive metaphor annotation scheme was demonstrated in the comparative study of four state-of-the-art automatic metaphor identification systems (Dunn, 2013c), which showed low agreement on a gold standard. The proposed scheme aims to capture a variety of factors involved in metaphoricity, their cumulative effect, and the gradual nature of metaphoricity.

The proposed inventory of metaphoricity factors includes nonbasic meanings (the less concrete, body-related and precise meaning(s) of a word); semantic shifts (additional semantic components of the meaning), newly attested meanings; morphologically motivated metaphoricity; different types of personification (a switch in the semantic class of the argument from animate to inanimate); a switch in the semantic class of the argument between concrete and abstract; the use of a metonymic argument, participation in idiomatic expressions; and direct metaphors – the direct use of a word involved in a cross-domain mapping.

We also attempt to represent the dynamic interactions between these factors using graph-like visualization. Personifications and direct metaphors are described in terms of conceptual sources and targets. Secondary metaphoricity is attested for words that participate in extended metaphors. Metaphorically related words can be signaled by different types of cues: graphically, or by specific grammatical constructions, or by lexical cues.

The evidence is drawn from 2,036 occurrences of 100 Russian verbs sampled from the Russian syntactically annotated corpus SynTagRus.

## RUSSIAN IDIOMS OF DRUNKENNESS: CORPUS-BASED ANALYSIS

**Baranov A. N.** (baranov\_anatoly@hotmail.com),

**Voznesenskaja M. M.** (voznesh-masha@yandex.com), Institute of Russian Language (Vinogradov's Institute), Russian Academy of Sciences, Moscow, Russia

The paper provides a corpus-based analysis of Russian idioms of drunkenness from the Thesaurus of Modern Russian Idioms (Moscow, 2007). As a starting point of a study three text corpora differentiated in accordance with different spheres of functioning of Russian were selected: Corpus of Russian Literature (texts from 1960 to 2000 years, 35 millions of typetokens), Corpus of Mass-Media (texts from 1990 to 2010 years, 29 millions of typetokens) and Corpus of Drama (plays from 1960 to 2010 years, 23 millions of typetokens).

The main idea consists in comparing frequency of idioms of drunkenness under consideration in every corpus. The analysis shows that Russian idioms of drunkenness form four different groups in accordance with their frequency characteristics. The first group is formed by low-frequency idioms from all corpora. The second group consists of idioms with the same relative high-frequency in all corpora. The third group includes idioms which demonstrate a high variations in frequencies in all corpora. The fourth group consists of idioms with a high frequency rank within every single corpus. The paper describes each of these groups providing the interpretation of the statistic data from the field of stylistics. The high frequency range of an idiom in different corpora corresponds with its stylistic features. The paper argues that there is a strict correlation between the discourse frequency of an idiom and its stylistic properties.

## INTERACTION AND EMPATHY AS ELEMENTS OF NARRATIVE STRATEGIES IN THE RUSSIAN CLIPS CORPUS

**Bergelson M. B.** (mirabergelson@gmail.com), **Khudyakova M. V.**, National research University Higher School of Economics, Moscow, Russia

This paper focuses on evaluation of discourse abilities of speakers with brain damage: people with dynamic aphasia (PWA(d)) and right hemisphere damage (RHD) as compared to healthy speakers of Russian language. The study is based on the material from the Russian CliPS corpus that contains retellings of the Pear Film produced by PWA and RHD, as well as neurologically healthy controls.

The nature of the narratives in the corpus allows for a comparative investigation of discourse on the level of micro-structure (grammatical and lexical phenomena) and on the macro level: narrative structure, coherence and cohesion, interactional patterns and narrative discourse strategies. In this paper we present results of the comparative analysis of some macro level discourse strategies: the way interaction and empathy are realized in the stories by PWA(d), RHD and healthy speakers.

We have found significantly higher numbers of attitude expression markers, as well as significantly lower numbers of cognitive difficulties markers, in healthy speakers as compared to PWA(d). These results support what is known about difficulties that PWA(d) demonstrate in discourse production tasks. While PWA(d) use interactive markers to get a break from keeping with the story plan, they avoid using epistemic predicates whose subjects are the story characters.

We also present qualitative analysis of the discourse strategies of healthy speakers.

## DIRECTIVE UTTERANCE AS AN ELEMENT OF SPOKEN DIALOGUE: THE CASE OF IMPERATIVE 1PL

**Blinova O. V.** (o.blinova@spbu.ru), Saint Petersburg State University, Saint Petersburg, Russia

The article deals with Russian imperative constructions 1PL with particles *davay(te)* / *day(te)*. It is argued that conversation-oriented approach in analysis of these imperative utterances can be very productive. On the basis of the neighboring utterances and specifics of dialogue organization one can distinguish four semantic subtypes of 1PL imperative: propositive (which is an expression of 'suggestion' and presupposes a verbal reaction from the listener), hortative (which expresses an 'mild order' and no verbal reaction of the listener is expected), factitive permissive (expresses a request to remove obstacles for action, no verbal reaction of the listener is expected), permissive (expresses 'permission', the utterance appears as a reaction in the final part of the microdialogue). Imperative constructions are studied in two speech corpora. The difference in frequency for the four imperative subtypes in the two corpora can be explained by the difference in the amount of everyday and institutional talks in the sampling. In the corpus where institutional talks between «sellers» and «customers» are predominant most of the 1PL imperative usages (70%) are propositive. This is determined by the fact that such are specific strategies used by the «sellers». The research results can be used for the modeling of Russian dialogue.

## SOME INVARIANT FEATURES OF RUSSIAN EVERYDAY SPEECH: PHONOLOGY, MORPHOLOGY, SYNTAX

**Bogdanova-Beglarian N. V.** (n.bogdanova@spbu.ru), **Blinova O. V.** (o.blinova@spbu.ru), **Martynenko G. Ya.** (g.martynenko@spbu.ru), **Sherstinova T. Yu.** (t.sherstinova@spbu.ru)  
Philological Faculty of St. Petersburg State University, St. Petersburg, Russia

The presented research was carried out on the material of the ORD speech corpus in the framework of the project, dedicated to study sociolinguistic variation of Russian speech and aimed at identifying diagnostic features characterizing everyday speech of major social groups (age-, gender-, status-, professional-related, etc.). The obtained results showed that practically on each linguistic level one may observe the features exhibiting a very high similarity between different sociolects. In particular, the coincidence is observed in the distribution of phonemes, distribution of parts of speech, and the frequency of some syntactic structures. The distribution of phonemes was determined on the subcorpus of 172,000 allophones. The following ten phonemes are the most frequent in speech of all social groups: /a/ (18,18%), /i/ (9,04%), /t/ (6,36%), /o/ (5,43%), /u/ (4,49%), /n/ (4,11%), /j/ (3,82%), /e/ (3,57%), /k/ (3,35%), /s/ (3,01%). The distribution of parts of speech in everyday speech was obtained on the linguistically annotated subcorpus of 125,437 tokens and has the following breakdown: V (17,43%), S (15,29%), S-PRO (14,13%), PART (13,35%), CONJ (9,47%), PR (7,09%), ADV-PRO (5,30%), ADV (4,51%), A-PRO (4,30%), A (3,73%), PRAEDIC (1,84%), INTJ (1,41%), NUM (1,29%), PARENTH (0,56%), ANUM (0,27%), PRAEDIC-PRO (0,01%). At the syntactic level, one-element structures are prevailing in everyday speech of all social groups, the most frequent among them being D (particle / discursive word) (3,73%), S (2,26%), and V (1,88%). Statistical analysis of the left-branching and right-branching verb groups has showed that the first ones significantly prevail

in speech of all sociolects. The revealed features reflect some constant, universal properties of everyday spoken Russian and can be used for adjustment and improvement of speech synthesis and recognition systems.

## DEVELOPMENT OF POLYSEMY IN RUSSIAN NUMERALS AND ORDINALS

**Boguslavskaya O. Ju.** (boguslavskaya.olga@gmail.com), **Iomdin B. L.** (iomdin@ruslang.ru)

The Russian Language Institute of the RAS, Moscow, Russia

Words denoting numbers (cardinal and ordinal numerals, or adjectives) represent a small (although potentially infinite) lexicographic type. In this article we deal with the polysemy structure of these two lexical classes. We propose a lexicographic pattern and study standard types of semantic shifts, including regular metaphors and metonymies. The words of both classes normally develop special senses with conversion into other parts of speech. Additional senses, different for different words, can appear due to cultural conventions.

## DISCOURSE PARTICLES AND THEIR TRANSLATION:

### HY IN VLADIMIR SOROKIN'S NOVEL *THE QUEUE* (*OCHERED'*)

**Dobrovolskij D. O.** (dobrovolskij@gmail.com), **Levontina I. B.** (irina.levontina@mail.ru)

Russian Language Institute of the Russian Academy of Sciences, Moscow, Russia

The representation of oral speech in literature differs sharply from real spoken language. It is therefore interesting to study Vladimir Sorokin's novel *Ochered'* (*The Queue*), which consists entirely of dialogue and gives the impression of conveying authentic orality.

The present study examines the use of the particle *hy*, which occurs some 300 times in the novel. The aim is to analyze various uses of its German and English equivalents, which will allow us to better describe the functioning of the Russian particle. On the other hand, comparison with translations will lead to a better understanding of this particle in the structure of literary texts.

Lexicographical descriptions indicate that the standard equivalent of Russian *hy* in German is the particle *na*. As is clear from our analysis, however, in many cases translations select a different variant. In a context involving coaxing (*Nu, skazhite, ne vrednichayte*), the equivalents *aber* and *schon* are found more often, while the particles *na, naja, ja, also, so, eben* or the interjections *ach, tja* are used more frequently to express overcoming embarrassment (*Nu chto vy...*). In English this play on discourse particles is not conveyed. This indicates that we need to refine the contrastive and lexicographic description of these particles on the basis of a representative corpus of parallel texts. Our analysis shows that Sorokin uses *hy* not merely as a colloquial marker but also as a keyword, as this term is understood in linguistic poetics. In translation this function is not preserved or preserved only partially.

## THE CONTRIBUTION OF THE KINETIC COMPONENT TO THE MULTIMODAL COMMUNICATION, OR TANGRAM DESCRIPTION STRATEGIES REVISITED

**Fedorova O. V.** (olga.fedorova@msu.ru), Lomonosov Moscow State University, RANEPa,

Moscow, Russia

Multimodal communication — words, prosody, gesticulation, postures, eye gaze — is a fundamental part of language use. It is widely thought that a communication is more effective when it is accompanied by gesticulation than when it is not. In search of systematic evidence how the interface “speech–gesticulation” works we turn to the referential communication task originally devised by Krauss and specified by Clark. In our experiment, two participants were seated on tatami separated by a low or high plastic screen (the within-participants factor “visibility”), in front of each were twelve cards of so-called Tangram figures. For the Director the cards were already arranged in a target sequence, and for the Matcher the same figures lay in a random sequence. The Director's job was to get the Matcher to rearrange her figures to match the target ordering. Thirty two participants carried out the task in four trials. All descriptions (over three hours long) were annotated, in Praat at the vocal channel and in ELAN at the kinetic channel. We consider six concrete hypotheses: (1) any kinetic activity benefits communication; (2) iconic gestures and postures benefit communication; (3) iconic postures benefit communication; (4) visible kinetic activity benefit communication; (5) visible iconic gestures and postures benefit

communication; (6) iconic gestures and postures of the 1-st trial benefit communication. However, none of these hypotheses was confirmed. The most likely explanation of the obtained results is associated with the complex interface between speech and gesticulation: kinetic channel demonstrates not the additive effect, but the interaction one.

## SEMANTIC CLASSES AND GOVERNMENT OF ADJECTIVES

**Grashchenkov P. V.** (pavel.gra@gmail.com), MSU, IOS RAS, MSUE, Moscow, Russia;  
**Kobozeva I. M.** (kobozeva@list.ru), MSU, Moscow, Russia

The paper aims at establishing the list of semantic classes of Russian adjectives. Another goal is to enumerate semantic roles relevant for the participants of Russian adjectives and the surface tools for coding such adjectival participants. The discussion of adjectival classes starts from the classifications proposed in Western and Russian linguistics. Such classifications are often based on pure semantic grounds and sometimes have rather speculative character. To overcome such shortcomings, the paper takes into consideration participants introduced by situations described by adjectives. Such participants can be of two types. Participants of the first type are relevant for specific semantic classes of adjectives. The second type participants can be introduced by adjectives from different semantic classes (and under certain circumstances—by the adjectives that lack government at the first sight). The overall conclusion is that adjectival government is more idiosyncratic and “more semantic” in nature than the verbal one.

## PRINCIPLES OF HOW TO DETERMINE THE DEGREE OF LANGUAGE-SPECIFICITY OF CONNECTIVES

**Inkova O. Yu.** (Olga.Inkova@unige.ch), Institute of Informatics Problems, FRC CSC RAS, Moscow, Russia; University of Geneva, Geneva, Switzerland

The paper deals with the question of how to determine the degree of language-specificity of connectives. As connectives are non-descriptive linguistic units and their morphological nature may vary, the author suggests to apply functional semantic criteria and distinguishes five types of interlingual isomorphism violation in Russian/French and Russian/Italian.

The author proposes to establish five categories of language-specific connectives:

1. The relation R expressed by the connective  $K_A$  of language A is conveyed in language B by a linguistic unit  $U_B$  not belonging to the functional class of connectives (for instance, the relation of simultaneousness expressed by *pri etom* in Russian is frequently expressed in French or Italian by means of a present gerund);
2. The relation R expressed by the connective  $K_A$  of language A is conveyed in language B by the signal  $K_B$  of a relation R' which is semantically contiguous, but nevertheless possesses a more specific equivalent in language A (for instance, the *pritom* of illocutive simultaneousness is rendered in French with *en outre*, *en plus*, *de plus* disposing in Russian of the more precise equivalents *k tomu že* and *krome togo*);
3. The connective  $K_A$  of language A possesses in language B a set of equivalents  $K_A^1$ , all occurring with the same frequency (it's the case of Russian *vpročem*, which is mainly translated with *d'ailleurs* and *du reste*);
4. The connective  $K_A$  of language A has in language B a systemic equivalent  $K_B$  which is the most frequent in translation, but doesn't reproduce all the values and uses of  $K_A$  (for instance, the connective *a to* of negative alternative has the French equivalent *sinon*, but the Russian connective is polysemous: it can express both the relations of cause and alternative, as well as a relation of addition — what is not the case of *sinon*);
5. There is in language B a mostly used translation variant  $K_B$  of the connective  $K_A$  of language A, but this 'equivalent'  $K_B$  doesn't possess the combinatory variety of  $K_A$  (thus the Russian connective *hotya* introducing a subordinate clause combines in the main sentence with the conjunction *no*, but *hotya ... no* is translated in French either by a single *bien que*, the equivalent of *hotya*, or by a single *mais*, the equivalent of *no*).

The author also describes features helping to determine the degree of language-specificity of connectives that would take into account the translation direction (from Russian to French/Italian and vice versa) and might be subject to a quantitative analysis. These features are: 1) frequency of occurrence of a “zero” functionally equivalent fragment (respectively, translation stimulus); 2) type of functionally equivalent fragment or translation stimulus (congruent or



divergent); 3) number of possible translation patterns (respectively, translation stimuli). In conclusion, the author will comment on statistical data confirming the described types of interlingual isomorphism violation in case of the functional class of connectives. The author points out that statistics requires adequate interpretation and should be accompanied by a semantic analysis, as the latter also serves as a tool for the data's review. The study was carried out by using the Russian-French and Russian-Italian supracorpora database of connectives.

## WHAT TO DO ABOUT CONSTRUCTIONS LIKE *WHAT TO DO*?

**Iomdin L. L.** (iomdin@iitp.ru), Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Moscow, Russia

The paper discusses a number of challenging microsyntactic units of Russian formed with functional and broad semantics words, like pronouns, particles, conjunctions and auxiliary verbs. One microsyntactic unit, “kak byt” (≈ “what to do”), is described in full detail. Issues of adequate presentation of such units in linguistic resources and their treatment in natural language processing tasks are considered. Special emphasis is laid on two types of resources: the Microsyntactic dictionary, developed as a specific lexicographic resource that can be used in theoretical linguistics and natural language processing, and a deeply annotated corpus of Russian, Syntagrus, supplied with microsyntactic tagging. The Microsyntactic dictionary consists of lexical entries compiled according to a uniform pattern that comprises 9 zones: (1) name of microsyntactic unit; (2) type of unit (syntactic idiom or nonstandard construction); (3) lexical composition of the unit (individual words of the syntactic idiom or classes of words forming the nonstandard construction, if any); (4) analytical lexicographic definition of the unit; (5) government pattern; (6) syntactic properties of the unit presented in terms of dependency grammar, including the identification of syntactic relations that hold between the unit's constituent elements, and relations that incorporate this unit into the sentence; (7) synonyms and analogues; (8) comments and (9) illustrations. The microsyntactic markup of SynTagRus shows what units are present, how they fit in the structure and which of the polysemous entities occur in the tagged sentence.

## COARTICULATION OF NASALS AS A CUE FOR DETECTING PROSODIC BREAKS AND RULE ORDER IN STANDARD RUSSIAN

**Knyazev S. V.** (svknia@gmail.com), **Utesheva A. S.** (Utesheva102@gmail.com)  
Moscow State University<sup>(1,2)</sup>, Higher School of Economics<sup>(1)</sup>; Moscow, Russia

The paper reports some results of the research, aimed at finding out 1) whether place coarticulation occurs in clusters of [labial or dental nasal + labiodental obstruent] within the phonological word and in external sandhi position in Modern Standard Russian, and 2) whether it may serve as a cue for detecting the presence of prosodic breaks and the phonological rules order.

The results obtained show that the F2 value of nasal before labiodental obstruent is significantly higher for bilabial obstruent and significantly lower for coronal one in comparison with their F2 values in the position before gomorganic stops. This type of place coarticulation is found only within the phonological word and is not available in an external sandhi position; thus the absence of this type of coarticulation may serve as a cue for detecting the presence of prosodic break.

In the case of clusters with final palatalized labiodental obstruent, the F2 value of bilabial nasal is found to be noticeably higher than those of the coronal one because a palatalization coarticulation, which exists in Modern Standard Russian for bilabials but not coronals before labiodentals. Thus, we argue that the phonological rule of palatalization operates before the rule for place assimilation in Standard Russian.

## TYPOLOGY OF RUSSIAN GAITS

**Kreydlin G. E.** (gekr@iitp.ru), Russian State University for the Humanities, Moscow, Russia

The paper continues a series of research on gaits as a particular type of nonverbal complex sign units. It presents the detailed description of various types of gaits peculiar to Russian people. The research is conducted in the framework of the feature-based approach, which is a main tool for constructing a semiotic conceptualization of the somatic objects or human body in general. Using this approach we analyze physical features of gaits, such as width of pace, speed of walk-

ing, type of contact with the surface, noise, which is produced while walking, shape of the body or particular somatic objects (i.e. legs, feet), etc. We describe how the particular manner of walking reflects various characteristics of the human being. These are social characteristics, such as gender and age; traits of character and psychological features, such as being nervous, depressed or happy; various types of dysfunctions, such as being halting. Such parallels between the specific features of the gait and elements of the emotional and psychological states of the person highlight the connections between various systems of the human being.

## PHONETIC CHARACTERISTICS OF BREATHING PAUSES WITH DIFFERENT TEXT LOCALIZATION

**Krivnova O. F.** (okrivnova@mail.ru), Lomonosov Moscow State University, Moscow, Russia

In spite of its fundamental role in the processes of speech production, speech breathing is studied experimentally quite insufficiently. The main task of this paper was to obtain phonetic-acoustic data on the breathing pauses (BP) with different text localization in oral Russian texts. In the introduction we formulate the problems being discussed with the main goal to analyze the correlation of BP's acoustics with the boundaries of principal text units—paragraphs, sentences, clauses, taking into account the interspeaker variability in reading of the same text. In our previous studies it was supposed that the distinctive function of BP as prosodic markers of macrosegmentation is realized by the stable differences in their general phonetic patterns and such acoustic parameters as duration, intensity and noise spectrum of the inhalation phase. The quantitative analysis of these differences comprises the main part of the paper. The second section describes the material and the methods of research. In the third section the analysis of the general acoustic-physiological pattern of BP with different text localization is given, which allows to set the regular objective differences between them. The fourth section contains the data of BP's duration in connection with their text localization. It is shown that BP's text localization is the main factor to control its duration. In the fifth section the similar information is presented for the average BP's intensity. The sixth section contains an analysis of the spectral characteristics of the inhalation noise in BP. It is shown that it is the most invariant feature for all BP types, almost independent of its text localization. The conclusion summarizes the study's results emphasizing the fact that the presence of BP in any oral text can serve as a sufficient indication of its structure, but interspeaker variability shows that BP's realization is not a necessary feature of the text boundary.

## LINGUISTIC TENDENCIES IN ENGLISH TO RUSSIAN TRANSLATION: THE CASE OF CONNECTIVES

**Kunilovskaya M. A.** (mkunilovskaya@gmail.com), University of Tyumen, Tyumen, Russia

Lexical elements, which signal text cohesion (*though, moreover, etc*), are a convenient and attractive research object in Translation Studies. Unlike other words, connectives are structurally optional, more context-independent and, therefore, more revealing in terms of motivations behind translators' linguistic choices. Their frequencies are used to establish differences between translations and non-translations and are interpreted as a linguistic indicator of several tendencies in translation process such as explicitation, simplification and convergence. Particular patterns of translator's choices result in different degrees of 'being a translation' and can be related to translation quality and translational norms. We set out to reveal tendencies in translational behavior at different competence levels by describing frequency distributions of a single functional type of discourse markers (connectives) in English to Russian translation of mass-media texts. To this end, we compared data from a parallel translational learner corpus and a corpus of professional translations to customized selections from English and Russian national corpora. Using independent predefined lists of connectives for each language, we explored cross-linguistic differences and their influence over students' and professional translations. We hypothesized three possible tendencies: translations follow source language pattern (interference); translations follow target language pattern (normalization) or translations demonstrate independent idiosyncratic (over)use of connectives (explicitation). The observations were done with regard to the overall frequencies of the list items, their semantic groups and individual frequencies. Beside obvious pedagogical implications, the findings are useful in understanding cognitive processes behind translation, applicable in detecting translationese for a given language pair and assessing textual quality of translations.

## THE TYPES OF ADVERBIALS

**Kustova G. I.** (galinak03@gmail.com), Vinogradov Russian Language Institute of the Russian Academy of Sciences; Moscow State Pedagogical University; Moscow, Russia

The article discusses adverbial expressions (adverbials), cf. *v ozhidanii* ('in expectation'), *pod okhranoj* ('guarded by smb'), *po priglasheniju* ('by invitation'), *v blagodarnost'* ('being thankful for') as one of the stages of the process of verbal nouns grammaticalization — namely, as a stage of their transformation into derivative prepositions.

An adverbial is regarded as a reduced predications related to the main predication. There are two types of adverbials: included adverbials and circumstantial ones. Circumstantial adverbials correlate with adverbial subordinate clauses. The article discusses semantic-syntactic properties of included adverbials.

Among included adverbials two types can be distinguished:

- 1) coreferential (having the same subject) adverbials that correlate with an adverbial participle: the subject of the adverbial is identical to the subject of the main predication; and the subject (agent) of the adverbial is not expressed and can't be expressed syntactically, cf. (*Passazhiry khodili po perronu v ozhidanii poezda* 'Passengers were walking along the platform **waiting for the train to arrive**');)
- 2) non-coreferential (having different subjects) adverbials that correlate with a passive participle: the subject of the adverbial is not identical to the subject of the main predication; and the subject (agent) of the adverbial is necessarily expressed (*Prishel v soprovozhdenii advokata* 'came **accompanied by a lawyer**'). In contrast to such adverbials, nominalizations can express all their arguments.

## 'TENSED' AND 'NON-TENSED' PREDICATIVES IN RUSSIAN

**Letuchiy A. B.** (alexander.letuchiy@gmail.com), National Research University Higher School of Economics, Russia, Moscow

In the article, I focus on tense marking in Russian constructions with predicatives, such as *xolodno* '(it is) cold' and *ploxo* '(it is) bad'. Statistical data from the Russian National Corpus show that the frequency of past tense forms (e.g., combinations with the form *bylo*) is much greater for some predicatives than for others. This difference results both from semantic and formal factors. On the one hand, some predicatives denote evaluation (e.g. *ploxo* 'bad'). Evaluation can be applied to events that have finished or have never been realized. What is relevant is that the evaluation is made at the moment of speech, and this is why the present tense (= the zero copula verb) is used. On the other hand, it is important that the present tense is unmarked with predicatives, while with verbs, it is marked with special verbal affixes. The unmarked present tense form of a predicative can get the temporal meaning from the embedded verb. Interestingly, this phenomenon is in a sense opposite to the well-known phenomenon of relative tense marking. While the latter presupposes that the tense assignment in the embedded event is anchored to the tense meaning of the main event, the tense value of the construction with evaluation predicatives is assigned by 'agreement' with the embedded verb.

## ITALIAN *MAGARI* AND ITS RUSSIAN EQUIVALENTS: DIFFERENT DISCOURSE STRATEGIES

**Levontina I. B.** (irina.levontina@mail.ru), RLI RAS, Moscow, Russia;

**Denissova G.** (galina.denissova@unipi.it), Pisa State University, Italy

This paper is aimed to analyze the Russian equivalents of the Italian focus particle "magari". This lexeme has attracted much attention among linguists from different countries due to its especially intriguing polyfunctionality that sometimes knows no parallel in other languages. Russian equivalents of "magari" (extracted from the Russian-Italian subcorpus of the Russian National Corpus) clearly demonstrate that "magari" corresponds in Russian to a wide range of lexemes/units with different modality—equipotential non exclusion of factuality, concessivity, weakening of the illocutionary force of imperative, optative modality. The set of functions held by "magari" (non factual, non factual concessive, imperative, optative) also recalls in Russian the semantic network developed by several irrealis markers of non factuality. Moreover, in some contexts in the Italian translations from Russian "magari" appears while there is no concrete equivalent in the source language. In other words, the connotative range of "magari" is mostly

achieved in Russian by different semantic mechanisms. Cross-linguistic analysis helps in clarifying the set of interlinguistic Russian-Italian correspondences of the lexeme “magari” and in circumscribing the different constructions and contexts where it occurs.

## **ANNOTATION OF RUSSIAN MANUAL GESTURES: THEORETICAL AND PRACTICAL ISSUES**

**Litvinenko A. O.** (allal1978@gmail.com)<sup>1</sup>, **Nikolaeva Ju. V.** (julianikk@gmail.com)<sup>2,1</sup>,  
**Kibrik A. A.** (aakibrik@gmail.com)<sup>1,2</sup>

<sup>1</sup>Institute of Linguistics RAS, Moscow, Russia;

<sup>2</sup>Lomonosov Moscow State University, Moscow, Russia

This paper contributes to the field of multimodal linguistics and discusses annotation of manual gestures in multichannel (multimodal) discourse, using a corpus of Russian discussions of the Pear Film (Chafe ed. 1980); the corpus is currently under construction ([www.multidiscourse.ru](http://www.multidiscourse.ru)). The study is based on a fully annotated subcorpus of 3 recordings, total time about one hour, with three participants in each recording. One of the key issues in gesture annotation is the importance of systematic distinction between communicatively significant gestures and general kinetic background. To address this issue, we introduce the notion of a speaker's gesticulation portrait, involving properties such as (dis)inclination to stillness and to frequent use of physiologically motivated self-adaptors, as well as a set of habitual postures (“neutral hand positions”), serving as starting points for significant movements. An annotation scheme for the ELAN environment is proposed, allowing one to systematically annotate hierarchically organized gestural units as annotations in independent ELAN tiers: gesture chains, individual gestures, and gesture phases. Gesture phases are annotated separately for the left and the right hand; an individual gesture (sometimes called “gesture phrase”) is treated as a combined unit that can be one-handed or two-handed. To divide an ongoing stream of gesticulation into a series of gestures, annotators identify the points where kinetic features change. A simultaneous change of two or more parameters (effort, velocity, trajectory, movement direction, hand shape and orientation, location in gesture space, etc.) marks a phase boundary; a change of several features points to a gesture boundary. A set of subordinate ELAN tiers is used to describe each gesture, including handedness, phase structure, multi-strokes, gesture overlaps, gesture repetitions, etc. As left and right hands movements are often asynchronous, a set of formal rules is proposed for such instances. On the basis of the annotated subcorpus evidence, quantitative observations were made concerning speakers' preferences in gesture handedness. Three out of nine participants prefer (>60%) two-handed gestures, while six participants prefer either right- or left-handed gestures most of the time. While all the speakers identified themselves as right-handed persons, two of them show preference for the left hand (>63%); the degree of right hand preference varies strongly (from 60% to 97%). This data suggests that adding the “dominant hand” parameter for two-handed gestures may be quite useful. The proposed annotation scheme is oriented towards comparing manual gestures with units of other communication channels.

## **A COMPUTER SYSTEM OF TEACHING INTONATION PATTERNS OF RUSSIAN SPEECH**

**Lobanov B. M.** (Lobanov@newman.bas-net.by), **Zhitko V. A.** (zhitko.vladimir@gmail.com),  
**Kharlamov A. A.** (kharlamov@analyst.ru), United Institute of Informatics Problems NAS  
Belarus, Minsk, Belarus

A software system is presented which is designed to train learners in producing the basic intonation patterns of Russian speech. The system is based on comparing the melodic portraits of a reference sentence and a sentence pronounced by the learner and involves active interaction of the learner with the system. While parametric representation of intonation features of the speech signal faces fundamental difficulties, the paper shows how they can be overcome. The basic algorithms of analyzing and comparing intonation features, used in the proposed learning system, are presented. The features of the acoustic database composed of reference sentences and used in the learning system are presented. The set of reference sentences represents intonation patterns of Russian speech (IP1 to IP7) and their basic varieties. The system's interface is presented and the results of system operation are illustrated.

## SYNTAX AND SEMANTICS FOR NOMINAL COUNTERFACTUALS

**Ljutikova E. A.** (ljutikova2008@gmail.com), **Tatevosov S. G.** (tatevosov@gmail.com),  
Lomonosov Moscow State University; Moscow State University of Education, Moscow, Russia

The paper explores syntax and semantics of nominal counterfactuals, in which the antecedent contains a DP instead of a clause (“if BY not DP”). We address a number of issues related to their internal structure and semantic interpretation, including obligatoriness of negation and its inability to license negative polarity items. We argue that despite the apparent nominal character of the antecedent, its interpretation is not different from that of regular counterfactuals whereby the antecedent provides the restrictor for the universal quantifier in the domain of possible worlds. We propose that the nominal constituent undergoes re-interpretation along the following path. A pronominal element *pro* adjoins to DP, which takes its reference in the domain of events. This element finds an antecedent event in the standards way regular pronominals do. It is related to the denotation of the DP by an unpronounced thematic relation. This derives a proposition of the form “that there is an event identical to the referent of *pro*, which || DP || is a participant of”. Since the referent of *pro* is an eventuality in the evaluation world, the counterfactual only get a coherent interpretation if the antecedent is negated. For the same reason, the antecedent is not an environment where NPIs can be licensed.

## TOLERANCE IN RUSSIAN AND IN ENGLISH: A CORPUS STUDY

**Malafeev A. Y.** (amalafeev@yandex.ru), **Orehkova E. A.** (kate.orekhova@gmail.com),  
National Research University Higher School of Economics, Nizhny Novgorod, Russia

Tolerance is a complex and partly contradictory concept that can be understood differently not only in different cultures, but also within the same culture. This paper presents a comparative study of the perception of tolerance by Russian and English speakers based on analysis of corpus data. At the initial stage of the study, the authors semi-automatically compiled a pilot web-based corpus of texts about tolerance. The corpus consists of a Russian-language subcorpus of 199,607 words and an English-language subcorpus of 210,898 words. After the mini-corpus was analyzed, the results were verified on the data from the general corpora ruTenTen11 and enTenTen13 using the Sketch Engine platform. The authors compared the word sketches for *толерантность* (*tolerantnost'*), *tolerance*, *толерантный* (*tolerantnyi*) and *tolerant*. In particular, this implied analyzing various lexical-semantic fields and thematic groups of collocations, as well as the following patterns: *X толерантности* (*tolerantnosti*) and *X of tolerance*, *толерантность к* (*tolerantnost' k*) *X* and *tolerance towards X*, *толерантность и/или* (*tolerantnost' i/ili*) *X* and *tolerance and/or X*. In addition, various derivatives of *толерантность* (*tolerantnost'*) / *tolerance* were discovered in the corpora and analyzed, including numerous nonce words. The corpus analysis enabled a deep insight into the way tolerance is perceived by Russian and English speakers.

## EGOCENTRICITY IN THE SEMANTICS OF THE RUSSIAN WORD NIKAK

**Paducheva E. V.** (elena.paducheva@yandex.ru), Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow, Russia

E. Benvenist (Benvenist 1974) proposed a new approach to the study of egocentricity in a language. He delimitates two contexts of language use: the **context of discourse** and the **context of narration** (*plan de discourse* and *plan de récit*). Thereafter the notion **register of interpretation** came into being. Two registers are distinguished: **dialogical register** and **narrative register**, see Apresjan 2003, Paducheva 1986, 1996. This distinction is of utmost importance for the egocentric entities of language (egocentrals) whose semantics appeals to the implied speaker. The opposition of **primary** (rigid) and **secondary** (shiftable) egocentrals is essential: secondary egocentrals undergo projections (when the role of the speaker is fulfilled by another subject), while primary egocentrals don't allow projections.

In the paper it is demonstrated that the parenthetical *nikak* expresses an uncertain statement made on the basis of directly observed situation. The following properties of *nikak* are distinguished. *Nikak* is often used in an interrogative sentence, but in an utterance that doesn't require an answer; different particles are often used in a *nikak*-sentence and the second person pronoun; *nikak*-sentence can express an explanation or interpretation of the situation before the eyes of the observer; *nikak*-sentence can express surprise; *nikak* may have *kazhetsja* ‘it seems’ as a synonym. *Nikak* is used only in an independent clause (thus differing from *neuzheli*, which is appropriate in a subordinate context). It is used only in the dialogical register and is a primary egocentric: it can only have a speaker as an implied subject and doesn't allow hypotactical or narrative projection.

## SEMANTIC HALO OF A METER: A KEYWORD-BASED APPROACH

**Piperski A. Ch.** (apiperski@gmail.com), Russian State University for the Humanities; National Research University Higher School of Economics, Moscow, Russia

The semantic halo of a meter (*semantičeskij oreol metra*) is a notion that was introduced by Mikhail Gasparov to describe semantic invariants of poetic texts composed using the same metrical scheme. Most studies addressing this phenomenon have been based on expert knowledge of the text corpus. In this study, I propose an automated approach to analyzing the semantic halo of various meters based on keyword extraction, using a simple measure of keyness as developed by Adam Kilgarrieff. The method is applied to texts from the Poetic subcorpus of the Russian National Corpus. It allows us to discern basic motifs that are very close to those identified by literary scholars, which proves that it is a promising way to analyze the semantic halo of a meter. Some novel associations can also be inferred from the keyword lists. Clearly, keyword extraction cannot replace profound expert knowledge, but it can serve as a useful first step of the analysis.

## “JA SKAZHU TEBE S POSLEDNEJ PRJAMOTOJ”: DIRECT AND INDIRECT SPEECH VIEWED THROUGH THE PRISM OF PROSODICALLY ANNOTATED CORPUS DATA

**Podlesskaya V. I.** (vi\_podlesskaya@il-rgggu.ru), Russian State University for the Humanities; Russian Academy of National Economy and Public Administration; Moscow, Russia

The paper focuses on how prosody complements grammar in differentiating between two main strategies of quoting—direct and indirect speech (*Peter said: “I am a linguist”* vs. *Peter said that he was a linguist*). Basing on corpus data from spontaneous spoken Russian, the paper addresses two research questions: (1) how do grammatical, lexical and prosodic parameters correlate in canonical cases of direct and indirect speech in Russian, and (2) what are typical prosodic deviations from canonical patterns. Answering the first question, we indicated prosodic features (pitch movements, pitch reset, localization of phrasal accents, pauses etc.) that help to identify a prosodic break between the reporting frame (*Peter said*) and the following reported utterance in the case of direct deixis (when personal and spatiotemporal indexes are oriented towards the reported situation, cf. *I am a linguist*) and the lack of such a break in the case indirect deixis (when personal and spatiotemporal indexes are shifted towards the reporting situation, cf. *that he was a linguist*). Answering the second question, we indicated particular prosodic patterns used (i) to integrate otherwise prosodically disintegrated constructions with direct deixis, and (ii) to disintegrate otherwise prosodically integrated constructions with indirect deixis.

## ORTHOEPY IN AN ERA OF CORPUS LINGUISTICS: SOME CASE STUDIES IN VARIATION IN STANDARD RUSSIAN

**Pozharickaja S. K.** (sofkonst@yandex.ru), Moscow State University, Moscow, Russia;  
**Dobrushina E. R.** (edobrush@gmail.com), St. Tikhon's Orthodox University (PSTGU), Moscow, Russia

We discuss the status of orthoepy as a linguistic discipline; its scope and limitations; the nature of orthoepic prescriptions and the evolution of dictionary marks in its connection to the evolution of the stress and grammar of standard Russian. We consider the use of Russian National Corpus as an instrument of prediction of further orthoepic changes and envisaging prescriptions resulting from this change.

1. Prescription can not apply to contextual modifications of phonemes ([сm']ex/[с'м']ex), because they are not perceived by a “lay speaker”, are not subject to his or her conscious choice and can not lead to communicative failures.

2. On the other hand, prescription may apply to variants that are different in ‘sound types’ (звукотип, e.g. ж[а]леть/ж[ы]леть, [сейф]/[с'ейф], т[и]х[ъй]/т[и]х[йй], [ч]то/[ш]то, е[жж]у/е[ж'ж']у); these variants may be suggested to represent newer vs. older norms.

3. Stress variants are the most dynamic domain of orthoepy. Special focus is on the change in verbal stress as reflected in the difference between the prescribed vs. real usage and in the chronology of dictionary marks (*дружит*→*дружит*, *родился*→*родился* etc.).

4. The use of the data of Russian National Corpus and, more specifically, comparative analysis of the main corpus and the corpus of newspapers is an important method of analysis of grammatical variation. We consider the competition between such forms as *поезжай/езжай, одеть/надеть, тьчу/тыкаю, мучу/мучаю, их/ихний* and provide charts showing positive dynamics of the more recent variants.

In conclusion, we stress the difference in the nature of orthographic and orthoepic marks in the dictionaries. Orthographic marks may be prescriptive and prohibitive. Orthoepic marks may only, in most cases, have a status of recommendations; they can not (be regarded as means to) block language change.

## ГОВОЮ Я, КАРЛ...

**Rakhilina E. V.** (rakhilina@gmail.com), National Research University Higher School of Economics, Institute for Russian Language / Russian Academy of Sciences, Moscow, Russia

The paper describes the diachronic development of Russian construction with the verb *govorit* 'to say', the main marker introducing the direct speech. The analysis relies heavily on the data from XIX century Russian as presented in a special annotated corpus of Early Modern Russian.

## REVERSE TRANSLATION AS A TOOL FOR ANALYSIS OF DISCOURSE WORDS

**Shmelev A. D.** (shmelev.alexei@gmail.com), Moscow Pedagogical State University; Vinogradov Institute of Russian Language; Russian Academy of Sciences; St Tikhon's Orthodox University, Moscow, Russia; **Zalizniak Anna A.** (anna.zalizniak@gmail.com), Institute of Linguistics, RAS; Institute of Informatics Problems FRC CSC RAS, Moscow, Russia

The paper demonstrates the possibilities of "one-direction analysis" in contrastive studies based on the parallel corpora. We assume that one may regard translation equivalents and paraphrases of a linguistic unit extracted from real translated texts as a source of information about its semantics; translations into Russian may be even more revealing in this respect. Our methods are based on the above hypothesis: we correlate the explications of the discourse words given in earlier studies with the "stimuli to translation", that is, fragments of the original text that may cause the appearance of these discourse words in a Russian translation as a "reaction" to those "stimuli". Using this methodology, we seek to validate, disprove or improve the semantic analysis of these words made without recourse to electronic corpora. The results of the analysis of the Russian discourse words *eshche, vidimo, po-vidimomu* and *vidno* are set forth.

## SUBSTITUTION NEIGHBORS IN THE STUDY OF LEXICAL ACCESS

**Slioussar N. A.** (slioussar@gmail.com), HSE, Moscow, and St. Petersburg State University, St. Petersburg, Russia; **Alexeeva S. V.** (mail@s-alexeeva.ru), St. Petersburg State University, St. Petersburg, Russia

The role of orthographic neighbors (e.g. bank—tank) in word processing has been discussed in many experimental studies. However, these studies have been conducted on a limited pool of languages, and many important questions are still unresolved. After creating a lexical database StimulStat that contains various neighborhood parameters for Russian, we conducted the first experiment with substitution neighbors in Russian. We used lexical decision task with priming, and manipulated the following factors: whether the prime is more or less frequent than the target, whether the prime is a nominative singular (primary) form or an oblique form, and whether the substituted letter is word-final or in the middle of the word. The results suggest that noun forms undergo morphological decomposition at a very early stage and shed new light on the process of activating candidates during lexical. The results also have practical significance because it is well known that spelling errors are influenced by neighborhood effects.

## NEGATION IN RUSSIAN AND FINNISH VERSIONS OF THE STATE TREATIES BETWEEN RUSSIA AND FINLAND: A CORPUS-BASED RESEARCH

**Souma Ju. V.** (iuliia.souma@uta.fi), **Kudashev I. S.** (igor.kudashev@uta.fi),  
**Mikhailov M. N.** (mikhail.mikhailov@uta.fi), University of Tampere, Tampere, Finland

Different language versions of bilateral state treaties are supposed to be equally authentic and have identical meaning. This belief looks problematic from the linguistic point of view, at least on the sentence level and below. In this article, we discuss the use of negation markers in the Russian and Finnish versions of the state treaties between Russia and Finland. The research material is the Russian-Finnish subcorpus of the PEST corpus (Parallel Electronic Corpus of State Treaties) that includes all treaties starting from 1917, the year of the Russian revolution and Finland's independence. It was found out that in the Russian versions of the documents explicit negation markers (particles, prepositions, etc.) are used much more frequently than in the Finnish ones, while implicit negation markers (verbs and nouns expressing prohibition, refusal etc.) are more typical for the Finnish versions. This phenomenon was discovered by means of comparing frequencies of translation equivalents in the Russian and Finnish versions. The results were confirmed by studying parallel concordances: numerous examples of replacing explicit negative markers by implicit ones were found. This phenomenon can be explained by the differences between the two languages, the differences in diplomatic discourse conventions, and possible prevalence of translating the treaties from Russian into Finnish over composing texts directly in Finnish. At the current stage of the research, it is not yet clear which of these factors have an impact on the choice of negation markers and to what degree.

## MULTIPLE PRENOMINAL ADJECTIVES ORDERING IN RUSSIAN: A CORPUS STUDY

**Toldova S. Yu.** (toldova@yandex.ru), **Mukhanova R. V.** (muhanova.regina@mail.ru),  
National Research University Higher School of Economics, Moscow

The work deals with the multiple prenominal adjectives ordering within a DP/NP in Russian. The aim is to examine the claim concerning word order hierarchy for adjectives suggested in [Cinque 1994] via quantitative analysis of corpus data from the Russian National Corpus (RNC). The following hierarchy was checked: possessive > quantity > order > quality > size > shape > colour > nationality. We use the RNC semantic annotation taxonomy for the semantic class of an adjective. We checked the pairwise orders for different semantic classes (e.g. possessive > colour vs. colour > possessive). Our data confirm the claim (suggested in the experimental studies) that the possessive adjectives should be divided into two classes: 'referential' possessives (e.g. adjectives with suffix *-in*, *Petin* 'Peter's') and so called 'generic' possessives (e.g. *chicken breast*). The latter occupy the position closer to the noun. There is one case of significant hierarchy violation. The referential possessive adjectives occupy the intermediate position between gradable adjectives (namely, after 'colour' and ungradable ones ('nationality' class)). Hence, the determiner-like possessives occupy not the left-most position as in languages with articles. These findings serve the additional evidence in favor of theoretical claim that determiner-like adjectives in articleless languages acquire determiner function only in the left-most position, otherwise they function as other adjectives. The corpus data confirm the claim that the change of the adjective structural position induces its semantic type coercion. The detection of expected word-order violations can be helpful in word sense disambiguation and for error detection in RNC semantic annotation.

## WHEN AN ARTICLE IS NOT AN ARTICLE: "NEGATED TAUTOLOGIES" IN RUSSIAN

**Vilinbakhova E. L.** (e.vilinbakhova@spbu.ru), St Petersburg State University, St Petersburg, Russia

The study is dedicated to contradictions  $A \neq A$  in Russian, also known in literature as 'negated tautologies'. They are often viewed as derivations from equative tautologies  $A = A$ . Here I describe structural and semantic features of negated tautologies that are established in Russian with patterns *X ne X* 'X is not X', *X ne est' X* 'X is not X' and *X — eto ne X* 'X this is not X'. Such constructions show that the speaker is not able to use the corresponding tautology *X is X* because (a) the referent of a linguistic expression *X* does not belong to the category *x*; (b) characteristics of the referent of a linguistic expression *X* or an attitude towards it differ from the norm; (c) the linguistic expression *X* is not employed in its common, straight meaning. Besides, negated tautologies are compared to similar Russian constructions  $S_{dat} + X ne X_2$  'S<sub>dat</sub> + X is not X' and *X ne v X\_2* 'X is not in X', and to tautologies *X est' X* 'X is X' and *X — eto X* 'X this is X'.



## WHAT IS INTONATION OF ENUMERATION IN RUSSIAN?

**Yanko T. E.** (tanya\_yanko@list.ru), Institute of Linguistics, Moscow, Russia

This paper is aimed at answering the question what intonation of enumeration in Russian is. Our main result is that there is no universal intonation of enumeration in isolation from illocutionary meanings (i.e. the meanings forming a speech act such as either topics or foci), or discourse meanings (i.e. the meanings forming the coherent discourse). Prosodic patterns of enumeration are based on a variety of pitch accents (which originally express the meanings of the topic, the focus, the discourse continuity) aligned with either homogeneous parts of the sentence or homogeneous sentences. The most common (but not the only ones) prosodic patterns of enumeration are represented by pitch accents IK-6, IK-2, IK-3, and IK-4 in the terminology of E. A. Bryzgunova. We argue that enumeration has a field semantic structure. The nuclear of the semantic field of enumeration consists of the meanings expressed by pitch accents IK-6, IK-2, IK-3, and IK-4. The semantic periphery of enumeration consists of such meanings as a roll call, or a specific Russian strategy of splitting a complex sentence into autonomous segments of information. The peripheral meanings as well as the nuclear meanings are expressed by a variety of pitch accents. For investigation, a minor working corpus of the Russian speech recordings was set up. The corpus comprises records retrieved in the Multimodal subcorpus of the Russian National Corpus. This paper is illustrated throughout with the frequency tracings of the sounding patterns of enumeration. The software program Praat was used in the process of analyzing the sounding data.

## RUSSIAN PREDICATIVES IN THE PERSPECTIVE OF SOCIOLINGUISTIC EXPERIMENT AND CORPUS GRAMMAR

**Zimmerling A. V.** (fagraey64@hotmail.com), Moscow State University of Education / Institute of Linguistics, Russian Academy of Sciences, Moscow, Russia

This paper aims at measuring the active vocabulary of Russian predicatives licensing dative-predicative-structures (DPS) and introduces the results of the sociolinguistic experiment paired with the corpus study for the same set of stimuli. The sociolinguistic experiment aimed at measuring the active DPS vocabulary in the idiolects of 18 native speakers of Russian and produced a sample of 422 stimuli ranked according to their acceptance rate. The same set of stimuli was tested on data from the Russian National Corpus, which produced the second sample based on the frequency ranks of DPS clauses in the diagnostic context *мне Z-во*, with the priority dative argument placed contiguously to the predicative at the distance <-1:1> in the same clause. The correlation of two ranked samples predicts the volume of the DPS vocabulary in an average idiolect of Russian and helps to establish the proportion between the lexical nucleus of the DPS construction and its extension. The analysis verifies the hypothesis that native speakers of Russian apply to the same underlying principles of semantic selection for DPS elements while using non-identical sets of lexical elements. The paper introduces a formal model of lexical extension with the entire DPS being grouped onto 15 thematic classes. The results prove that each thematic class includes high-, middle- and low-frequency DPS elements. Low-frequency DPS predicatives are modelled after high-frequency elements from the same thematic class.

## Авторский указатель

Алексеева С. В. ....	407	Летучий А. Б. ....	249
Апресян В. Ю. ....	2, 17	Литвиненко А. О. ....	271
Бадрызлова Ю. Г. ....	30	Лобанов Б. М. ....	287
Баранов А. Н. ....	45	Лютикова Е. А. ....	302
Блинова О. В. ....	68, 82	Малафеев А. Ю. ....	321
Богданова-Бегларян Н. В. ....	82	Мартыненко Г. Я. ....	82
Богуславская О. Ю. ....	96	Михайлов М. Н. ....	419
Вилинбахова Е. Л. ....	441	Муханова Р. В. ....	429
Вознесенская М. М. ....	45	Николаева Ю. В. ....	271
Гращенков П. В. ....	134	Орехова Е. А. ....	321
Денисова Г. ....	261	Падучева Е. В. ....	333
Добровольский Д. О. ....	106	Пиперски А. Ч. ....	342
Добрушина Е. Р. ....	372	Подлеская В. И. ....	355
Житко В. А. ....	287	Пожарицкая С. К. ....	372
Зализняк Анна А. ....	394	Рахилина Е. В. ....	384
Инькова О. Ю. ....	150	Слюсарь Н. А. ....	407
Иомдин Б. Л. ....	96	Соума Ю. В. ....	419
Иомдин Л. Л. ....	161	Татевосов С. Г. ....	302
Кибрик А. А. ....	271	Толдова С. Ю. ....	429
Князев С. В. ....	175	Утешева А. С. ....	175
Кобозева И. М. ....	134	Федорова О. В. ....	118
Крейдлин Г. Е. ....	196	Харламов А. А. ....	287
Кривнова О. Ф. ....	207	Циммерлинг А. В. ....	466
Кудашев И. С. ....	419	Шерстинова Т. Ю. ....	82
Куниловская М. А. ....	222	Шмелев А. Д. ....	17, 394
Кустова Г. И. ....	234	Янко Т. Е. ....	453
Левонтина И. Б. ....	106, 261		

## Author Index

Alexeeva S. V. ....	407	Litvinenko A. O. ....	271
Apresjan V. Ju. ....	3, 17	Lyutikova E. A. ....	302
Badryzlova Yu. G. ....	30	Malafeev A. Y. ....	322
Baranov A. N. ....	45	Martynenko G. Ya. ....	83
Blinova O. V. ....	68, 83	Mikhailov M. N. ....	419
Bogdanova-Beglarian N. V. ....	83	Mukhanova R. V. ....	430
Boguslavskaya O. Ju. ....	96	Nikolaeva Ju. V. ....	271
Denissova G. ....	261	Orekhova E. A. ....	322
Dobrovol'skij D. O. ....	107	Paducheva E. V. ....	333
Dobrushina E. R. ....	373	Piperski A. Ch. ....	342
Fedorova O. V. ....	119	Podlesskaya V. I. ....	355
Grashchenkov P. V. ....	134	Pozharickaja S. K. ....	373
Inkova O. Yu. ....	150	Rakhilina E. V. ....	384
Iomdin B. L. ....	96	Sherstinova T. Yu. ....	83
Iomdin L. L. ....	161	Shmelev A. D. ....	17, 394
Kibrik A. A. ....	271	Slioussar N. A. ....	407
Knyazev S. V. ....	175	Souma Ju. V. ....	419
Kobozeva I. M. ....	134	Tatevosov S. G. ....	302
Kreydlin G. E. ....	196	Toldova S. Yu. ....	430
Krivnova O. F. ....	207	Utesheva A. S. ....	175
Kudashev I. S. ....	419	Vilinbakhova E. L. ....	441
Kunilovskaya M. A. ....	221	Voznesenskaja M. M. ....	45
Kustova G. I. ....	234	Yanko T. E. ....	453
Letuchiy A. B. ....	249	Zalizniak Anna A. ....	394
Levontina I. B. ....	107, 261	Zimmerling A. V. ....	466

*Научное издание*

## **Компьютерная лингвистика и интеллектуальные технологии**

По материалам ежегодной  
международной конференции «Диалог»

Выпуск 16 (23). 2017

Том 2

Ответственный за выпуск **А. А. Белкина**  
Вёрстка **К. А. Климентовский**  
Дизайн обложки **А. А. Светличная**

Подписано в печать 12.05.2017  
Формат 152 × 235  
Бумага офсетная  
Тираж 350 экз. Заказ № 79

Издательский центр «Российский  
государственный гуманитарный университет»  
125993, Москва, Миусская пл., д. 6  
Тел.: +7 499 973 42 06

Отпечатано с готового оригинал-макета в типографии  
ООО «Издательско-полиграфический центр Маска»  
117246, Москва, Научный пр-д, д. 20, стр. 9