

Компьютерная лингвистика и интеллектуальные технологии

По материалам ежегодной Международной
конференции «Диалог» (2015)

Выпуск 14

В двух томах

Том 1. Основная программа конференции

Computational Linguistics and Intellectual Technologies

Papers from the Annual International
Conference "Dialogue" (2015)

Issue 14

Volume 1 of 2. Main conference program

УДК 80/81; 004
ББК 81.1
К63

Программный комитет конференции выражает
искреннюю благодарность Российскому фонду
фундаментальных исследований за финансовую поддержку,
грант № 15-07-20554 Г

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, Й. Нивре,
Г. С. Осипов, В. Раскин, Э. Хови, С. А. Шаров, Т. Е. Янко*

Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Москва, 27–30 мая 2015 г.). Вып. 14 (21): В 2 т. Т. 1: Основная программа конференции. — М.: Изд-во РГГУ, 2015.

Сборник включает 69 докладов международной конференции по компьютерной лингвистике и интеллектуальным технологиям «Диалог 2015», 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), 2015

Предисловие

14-й выпуск ежегодника «Компьютерная лингвистика и интеллектуальные технологии» содержит избранные материалы 21-й Международной конференции «Диалог». На основании мнений наших рецензентов для публикации в ежегоднике Редсоветом было отобрано 69 докладов из числа примерно ста работ, которые были рекомендованы по результатам рецензирования для представления на конференции в этом году.

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

- Формальные модели языка и их применение в компьютерной лингвистике
- Модели и методы семантического анализа текста
- Лингвистические онтологии и извлечение знаний
- Теоретическая и компьютерная лексикография
- Методики тестирования технологий и верификации результатов лингвистических исследований (Dialogue Evaluation)
- Компьютерные лингвистические ресурсы и их связывание (Linked Data)
- Корпусная лингвистика: создание, разметка, методики применения и оценка корпусов
- Анализ Social Media
- Машинный перевод текста и речи
- Лингвистический анализ речи
- Модели общения. Коммуникация, диалог и речевой акт
- Мультимодальная лингвистика
- Компьютерный анализ документов: классификация, поиск, тематический анализ, оценка тональности и т. д.

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

В этом году продолжается и традиция проведения в рамках направления Dialogue Evaluation тестирования технологий решения отдельных задач компьютерного анализа русского языка. Значимость таких мероприятий трудно переоценить, поскольку их результаты создают основу для сравнительной оценки эффективности результатов в соответствующих областях исследований.

В этом году было проведено два таких тестирования: сравнивались различные подходы к анализу т.н. аспектного сентимента и оценке семантической близости слов.

Sentiment Analysis является важным самостоятельным прикладным направлением компьютерной лингвистики, особенно в той постановке, которая была предложена участникам: не только определение общей тональности документа, но и выделение и оценка в нем отдельных аспектов выражения мнения.

Тестирование методов определения семантической близости слов является важным для понимания сложной картины в современной вычислительной семантике, где конкурируют и взаимодействуют традиционные словарные и новые дистрибуционные методы исследования лексических значений.

Наиболее значимые работы, представленные участниками этих тестирований, выделены в отдельный второй том ежегодника. Там же опубликованы и итоговые статьи организаторов.

Программный комитет конференции выражает особую признательность Наталье Лукашевич и Александру Панченко за особую роль в организации и проведении этих тестирований.

Среди особых направлений «Диалога» в этом году — исследования в области русского мультимодального дискурса. Интерес к языковой коммуникации как целому всегда был характерным для нашей конференции, выросшей, напомним, из семинара, носившего название «Модели общения». Доклады мультимодального направления составляют важную часть этого сборника.

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 привлечь 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 языке).

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, на котором представлены обширные электронные архивы «Диалогов» последних лет и все результаты проведенных тестирований.

*Программный комитет «Диалога»
Редколлегия ежегодника «Компьютерная лингвистика
и интеллектуальные технологии»*

Организаторы

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

Учредителями конференции являются:

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

Конференция проводится при поддержке Российской ассоциации искусственного интеллекта.

Международный программный комитет

Байтин Алексей Владимирович	Компания Yandex
Богуславский Игорь Михайлович	Политехнический университет Мадрида
Буате Кристиан	Гренобльский университет
Гельбух Александр Феликсович	Национальный политехнический институт, Мехико
Иомдин Леонид Лейбович	Институт проблем передачи информации РАН
Кобозева Ирина Михайловна	Филологический факультет МГУ
Козеренко Елена Борисовна	Институт проблем информатики РАН
Корбетт Гревил	University of Surrey, UK
Кронгауз Максим Анисимович	Институт Лингвистики РГГУ
Лукашевич Наталья Валентиновна	НИВЦ МГУ
Маккарти Диана	Lexical Computing Ltd., UK
Мельчук Игорь Александрович	Монреальский университет
Нивре Йоаким	Уппсальский университет
Ниренбург Сергей	Университет Нью-Мексико
Осипов Геннадий Семёнович	Институт программных систем РАН
Попов Эдуард Викторович	РосНИИ информационной техники и САПР
Раскин Виктор	Purdue University, USA
Селегей Владимир Павлович	Компания АBBYУ
Хови Эдуард	University of Southern California
Шаров Сергей Александрович	University of Leeds, UK

Организационный комитет

Селегей Владимир Павлович, <i>председатель</i>	Компания ABBYY
Байтин Алексей Владимирович	Компания Yandex
Беликов Владимир Иванович	Институт русского языка им. В. В. Виноградова РАН
Браславский Павел Исаакович	Kontur Labs; Уральский федеральный университет
Добров Борис Викторович	НИВЦ МГУ
Иомдин Леонид Лейбович	Институт проблем передачи информации РАН
Кобозева Ирина Михайловна	Филологический факультет МГУ
Козеренко Елена Борисовна	Институт проблем информатики РАН
Лауфер Наталия Исаевна	ООО «проФан Продакшн»
Ляшевская Ольга Николаевна	Universitet i Tromsø, Norway
Сердюков Павел Викторович	Компания Yandex
Соколова Елена Григорьевна	РосНИИ искусственного интеллекта
Толдова Светлана Юрьевна	Филологический факультет МГУ
Шаров Сергей Александрович	University of Leeds, UK

Секретариат

Белкина Александра Андреевна, <i>секретарь оргкомитета</i>	Компания ABBYY
Атясова Анастасия Леонидовна, <i>координатор</i>	Компания ABBYY

Рецензенты

Августинова Тая
Азарова Ирина Владимировна
Апресян Валентина Юрьевна
Байтин Алексей Владимирович
Баранов Анатолий Николаевич
Беликов Владимир Иванович
Богданов Алексей Владимирович
Богданова-Бегларян Наталья
Викторовна
Богуславский Игорь Михайлович
Бонч-Осмоловская Анастасия
Александровна
Браславский Павел Исаакович
Васильев Виталий Геннадьевич
Гельбух Александр Феликсович
Гершман Анатолий
Гращенко Павел Валерьевич
Гриненко Михаил Михайлович
Гришина Елена Александровна
Губин Максим Вадимович
Даниэль Михаил Александрович
Добров Борис Викторович
Добровольский Дмитрий Олегович
Добрынин Владимир Юрьевич
Зализняк Анна Андреевна
Захаров Виктор Павлович
Захаров Леонид Михайлович
Иванов Владимир Владимирович
Иомдин Борис Леонидович
Иомдин Леонид Лейбович
Кибрик Андрей Александрович
Кобозева Ирина Михайловна
Козеренко Елена Борисовна
Коротаев Николай Алексеевич

Котельников Евгений Вячеславович
Котов Артемий Александрович
Крейдлин Григорий Ефимович
Кронгауз Максим Анисимович
Леонтьев Алексей Петрович
Лобанов Борис Мефодьевич
Лукашевич Наталья Валентиновна
Ляшевская Ольга Николаевна
Маккарти Диана
Минлос Филипп Робертович
Недолужко Анна Юрьевна
Новицкий Валерий Игоревич
Пазельская Анна Германовна
Панченко Александр Иванович
Паперно Денис Аронович
Пиперски Александр Чедович
Подлеская Вера Исааковна
Савельев Василий Евгеньевич
Селегей Владимир Павлович
Смирнов Иван Валентинович
Сокирко Алексей Викторович
Соколова Елена Григорьевна
Сорокин Алексей Андреевич
Старостин Анатолий Сергеевич
Тихомиров Илья Александрович
Толдова Светлана Юрьевна
Турдаков Денис Юрьевич
Урысон Елена Владимировна
Федорова Ольга Викторовна
Хови Эдуард
Хорошевский Владимир Федорович
Циммерлинг Антон Владимирович
Шаров Сергей Александрович
Янко Татьяна Евгеньевна

Contents*

Основная программа конференции

Апресян В. Ю. Связь семантических и коммуникативных свойств языковых единиц ...	2
Баранов А. Н. Справедливость versus несправедливость: метафорические осмысления в современном российском дискурсе (по материалам центральных печатных изданий)	19
Aleksandrs Berdičevskis, Hanne Eckhoff Automatic Identification of Shared Arguments in Verbal Coordinations	30
Бергельсон М. Б., Акинина Ю. С., Драгой О. В., Искра Е. В., Худякова М. В. Затруднения при порождении слов в дискурсе и их формальные маркеры: норма и патология, или о недискретности нормы в языке и речи	41
Bogdanov A. V., Gorbunova I. M. The Case of Russian Subject Pro in Machine Translation System	52
Igor Boguslavsky, Vyacheslav Dikonov, Leonid Iomdin, Alexander Lazursky, Victor Sizov, Svetlana Timoshenko Semantic Analysis and Question Answering: a System Under Development	62
Бонч-Осмоловская А. А. Квантитативные методы в диахронических корпусных исследованиях: конструкции с предикативами и дативным субъектом	80
Daniel M. A. Stem Initial Alternation in Russian Third Person Pronouns: Variation in Grammar	95
Добровольский Д. О., Левонтина И. Б. Модальные частицы и идея актуализации забытого (на материале параллельных корпусов)	104
Добрушина Н. Р. Показатель сослагательного наклонения как часть союза	118

* The reports of each section are ordered by the surname of the first author in compliance with the English alphabet.

Федорова О. В. Интродукция референта в русских устных пересказах (на материале «Рассказов о грушах» У. Чейфа)	131
Galitsky B. A. Document vs. Meta-Document: are Their Rhetoric Structures Different?	141
Галяшина Е. И. Лингвистический анализ в системах идентификации диктора: интегративный комплексный подход на базе экспертологии	156
Goncharova M. B., Kozlova E. A., Pasyukov A. V., Garashchuk R. V., Selegey V. P. Model-Based WSA as Means of New Language Integration into a Multilingual Lexical-Semantic Database with Interlingua	169
Гришина Е. А. Кванторные слова, жестикуляция и точка зрения	183
Grozin V. A., Dobrenko N. V., Gusarova N. F., Ning Tao The Application of Machine Learning Methods for Analysis of Text Forums for Creating Learning Objects	202
Июмдин Б. Л. Что такое <i>орехи</i>?	214
Kibrik A. A. The Problem of Non-Discreteness and Spoken Discourse Structure	231
Kipyatkova I. S., Karpov A. A. Development of Factored Language Models for Automatic Russian Speech Recognition	241
Yuri Kiselev, Andrew Krizhanovsky, Pavel Braslavski, Ilya Menshikov, Mikhail Mukhin, Nataly Krizhanovskaya Russian Lexicographic Landscape: a Tale of 12 Dictionaries	254
Киселева К. Л., Татевосов С. Г. Гулял — нагулял — нагулялся. Заметки о структуре префиксально-постфиксальных глаголов	272
Князев С. В. Редукция гласного как показатель его ударности в современном русском литературном языке	284
Коротаев Н. А. Коммуникативно-просодический подход к выявлению элементарных дискурсивных единиц в устном монологическом тексте ...	294

Котов А. А., Зинина А. А. Функциональный анализ невербального коммуникативного поведения	308
Крейдлиן Г. Е., Хесед Л. А. Тело в диалоге и некоторые проблемы мультимодальной коммуникации: признак «размер соматического объекта»	321
Кривнова О. Ф. Глубина просодических швов в звучащем тексте (экспериментальные данные)	338
Крылова Т. В. Частицы «ВОТ» И «ВОН»: механизмы формирования переносных значений на основе исходных дейктических значений	352
Kudinov M., Piontkovskaya I. Automatic Update of the Named Entities Database Based on the Users Queries	369
Кустова Г. И. Валентности абстрактных существительных: Редукция vs. Конкретизация	376
Kuzmenko E. A., Mustakimova E. G. Automatic Disambiguation in the Corpora of Modern Greek and Yiddish	388
Lagutin M. B., Katinskaya A. Y., Selegey V. P., Sharoff S., Sorokin A. A. Automatic Classification of Web Texts Using Functional Text Dimensions	398
Лобанов Б. М. Опыт создания мелодических портретов сложных повествовательных предложений русской речи	414
Olga Lyashevskaya, Egor Kashkin Inducing Verb Classes from Frames in Russian: Morpho-Syntax and Semantic Roles	427
Malafeev A. Yu. Exercise Maker: Automatic Language Exercise Generation	441
Мустайоки А., Вепрева И. Т. Метаязыковой портрет модных слов	453
Muzychka S., Piontkovskaya I. Graph-Based Approach in the Dependency Parsing Task for Russian Language	468

Anna Nedoluzhko, Svetlana Toldova, Michal Novák Coreference Chains in Czech, English and Russian: Preliminary Findings	474
Николаева Ю. В., Кибрик А. А., Федорова О. В. Структура устного дискурса: взгляд со стороны мультимодальной лингвистики	487
Падучева Е. В. Глаголы <i>быть</i> и <i>бывать</i>: история и современность	500
Piperski A. Ch. To Be or not to Be: Corpora as Indicators of (Non-)Existence	515
Подлеская В. И. «И не друг, и не враг, а так...»: дистрибуция и просодия маркеров нерелевантности по данным мультимедийного корпуса МУРКО	523
Ponomarev S. V. Learning by Analogy in a Hybrid Ontological Network	536
Protopopova E., Antonova A., Misyurev A. Acquiring Relevant Context Examples for a Translation Dictionary	548
Shelmanov A. O., Smirnov I. V., Vishneva E. A. Information Extraction from Clinical Texts in Russian	560
Sheremetyeva S. O. On Summarization Supporting Readability and Translatability	573
Шмелев А. Д. Русские лингвоспецифичные лексические единицы в параллельных корпусах: возможности исследования и «подводные камни»	584
Tarasov D. S. Natural Language Generation, Paraphrasing and Summarization of User Reviews with Recurrent Neural Networks	595
Урысон Е. В. Выбор союза <i>и</i> vs. <i>но/а</i> при сочинении двух предложений (уточнение понятий «ожидание» и «норма»)	603
Ustalov D. A. Russian Thesauri as Linked Open Data	616
Вилинбахова Е. Л. Статья значит статья: об одном классе тавтологических конструкций в русском языке,	626

Яковлева И. В. «Давай ронять слова»: метафора каузации перемещения по воздуху в семантической зоне глаголов речи в русском языке в контрастивном аспекте	638
Янко Т. Е. К проблеме сопоставительного анализа просодии: одесский региональный вариант русского языка vs. русская разговорная норма	650
Захаров В. П. Сочетаемость через призму корпусов	667
Зализняк Анна А. Лингвоспецифичные единицы русского языка в свете контрастивного корпусного анализа	683
Anton Zimmerling, Ekaterina Lyutikova Approaching V2: Verb Second and Verb Movement	696
Abstracts	711
Авторский указатель	730
Author's Index	732

Раздел I.

Основная программа конференции

СТАТЬЯ ЗНАЧИТ СТАТЬЯ: ОБ ОДНОМ КЛАССЕ ТАВТОЛОГИЧЕСКИХ КОНСТРУКЦИЙ В РУССКОМ ЯЗЫКЕ^{1, 2}

Вилинбахова Е. Л. (elenavilinb@yandex.ru)

Санкт-Петербургский государственный университет,
Санкт-Петербург, Россия

Ключевые слова: тавтологии, метаязыковые конструкции, русский язык, интернет

ARTICLE MEANS ARTICLE: ON ONE PATTERN OF TAUTOLOGIES IN RUSSIAN

Vilinbakhova E. L. (elenavilinb@yandex.ru)

St. Petersburg State University, St. Petersburg, Russia

The study based on Internet and corpus data [RNC] deals with a special pattern of tautological constructions in Russian called metalinguistic tautologies. The notion was briefly introduced in 1996 by E. Miki as a label for a set of quite heterogeneous examples, but was not further developed. While other tautologies describe entities in the real world, metalinguistic tautologies refer to the use of a linguistic expression. Such constructions show that the speaker is employing a word or an expression in its common, straight meaning. Therefore, they are most often used when context allows other possible interpretations of the linguistic expressions (such as euphemisation, irony, or hyperbole), and sometimes such alternatives are explicitly spelled out: *Inexpensive means inexpensive, not poor quality*. Metalinguistic tautologies are established in Russian with patterns *X znachit X*, *X oznachaet X* 'X means X', *X eto X* 'X is X', and are distinguished from homonymous constructions by their semantic and pragmatic features.

Key words: tautologies, metalinguistic constructions, Russian language, Internet

¹ Автор выражает признательность анонимным рецензентам статьи за их замечания и предложения.

² Работа поддержана грантом Президента РФ для государственной поддержки ведущих научных школ РФ «Школа общего языкознания Ю. С. Маслова» НШ-1778.2014.6.

*Если дипломат говорит «да», это значит «может быть»,
если дипломат говорит «может быть», это значит «нет»,
если дипломат говорит «нет», значит, это не дипломат.*

*Если женщина говорит «нет», это значит «может быть»,
Если женщина говорит «может быть», это значит «да»,
если женщина говорит «да», значит, это не женщина.*

*Если военный говорит «да», это значит «да»,
если он говорит «нет», это значит «нет»,
если военный говорит «может быть», значит, это не военный.*

(Анекдот)

1. Вводные замечания

Тавтологические конструкции вида *Жизнь есть жизнь* неоднократно становились объектом лингвистических исследований. Особенно активно обсуждались вопросы, связанные с механизмом их интерпретации, возможностью их перевода с одного языка на другой без потери смысла, влиянием контекста и лексического наполнения конструкций на возможные инференции и т.д., см. [Grice 1975], [Levinson 1983], [Wierzbicka 1987], [Gibbs, McCarrell 1990], [Ward, Hirschberg 1991], [Miki 1996], [Bulhof, Gimbel 2001], [Падучева 2004], [Meibauer 2008], [Rhodes 2009] и др.

На материале русского языка Т. В. Булыгина и А. Д. Шмелёв выделяют два ряда тавтологических конструкций: *X (и) есть X, X всегда / везде X* и т.п., с одной стороны, и *X — это X*, с другой стороны: «значение конструкций первого ряда: все манифестации X-а, т.е. члены класса X или «инстанты» индивида X, в общем, одинаковы, и нет оснований ждать от манифестации, с которой мы имеем дело, чего-то другого. <...> Для X — это X общее значение состоит в «выделении X-а среди прочих индивидов или классов; констатация его особого места; особенности отдельных манифестаций могут не приниматься во внимание» [Булыгина, Шмелёв 1997: 506–509].

Можно заметить, что Т. В. Булыгина и А. Д. Шмелёв связывают значения обоих рядов тавтологий с определёнными синтаксическими моделями, однако первичными являются именно семантические признаки: идея равноценности членов внутри класса для первого ряда и идея исключительности класса для второго ряда. Отметим, что конструкции обоих рядов используются говорящим для выражения мнения по поводу объектов или явлений окружающей действительности.

Тавтологии, о которых пойдёт речь, образуют третий ряд и могут быть проиллюстрированы примерами (1) и (2). В настоящей работе они обозначаются как *метаязыковые тавтологии* (далее — МТ), термин заимствован из [Miki 1996].

- (1) А: *Лео, тут надо разобраться, что в вашем понимании есть рыбалка!!!*
Б: *Рыбалка значит рыбалка.... РЫБУ ЛОВИТЬ.....и народ ухой угостить...*³

³ Примеры взяты из Интернета, если не указано иное. Из соображений краткости адреса не указаны, но, разумеется, в материалах они присутствуют.

- (2) А: Что значит «завтра»? Он должен взять билет.
Б: Завтра означает завтра. Успеет купить.

Статья построена следующим образом: после краткой истории вопроса (раздел 2) будут описаны особенности МТ, отличающие их от других тавтологических конструкций (раздел 3); разновидности МТ и их общее значение (раздел 4); репрезентации МТ в русском языке и формально сходные конструкции, которые МТ не являются (раздел 5). В разделе 6 представлены итоги исследования.

2. История вопроса

Первое упоминание о МТ встречается в статье [Miki 1996], посвящённой значению тавтологических конструкций. Э. Мики исследует, каким образом «бессмысленные» тавтологические высказывания обретают значение в коммуникации, и приходит к выводу, что в тавтологиях набор известных собеседникам характеристик объекта идентифицируется в контексте текущей ситуации с самим объектом. Автор указывает на существование особых случаев, которые отсылают к языковому коду участников коммуникации, и обозначает их как «метаязыковые» тавтологии (англ. *metalinguistic tautologies*). В статье даны следующие примеры:

- «(33) *When I said kid, I meant K-I-D, kid!* ‘Когда я сказал ребёнок, я имел в виду Р-Е-Б-Ё-Н-О-К, ребёнок’.
- (35) *When he said three o'clock, he meant three o'clock* (P. G. Wodehouse, *The Romance of an Ugly Policeman*) ‘Когда он сказал «три часа», он имел в виду три часа’.
- (37) *Partial means partial* ‘Неполный значит неполный’.
- (41) *I'm glad to meet you, I said, meaning the words.* (J. Braine. *Room at the top*) ‘Очень рад познакомиться с вами, — сказал я, и это были не пустые слова: я действительно был рад’ (пер. Т. Озерской и Т. Кудрявцевой).
- (42) *I don't have one. / — What do you mean you don't have one? / — I mean I don't have one* (C. Webb. *The Graduate*) ‘У меня его нет / Что значит, у тебя его нет?/ Это значит, у меня его нет’
- (43) *We are getting sick and tired of it. When I say he's in Chicago, he is in Chicago.* (F. S. Fitzgerald. *The Great Gatsby*) ‘Покою нет от вашего брата. Раз я говорю, он в Чикаго, значит, он в Чикаго’ (пер. Е. Калашниковой)»

[Там же: 641].

Можно заметить, что приведённые Э. Мики примеры разнородны по форме и особенно по содержанию: ср. (35) и (43). В первом случае говорящий указывает,

что языковое выражение *три часа* в устах героя действительно имело буквальное значение ‘три часа’, а не ‘2.50’ или ‘пол четвёртого’, т. е. он имел в виду ровно то, что сказал, и этот частный смысл совпадает с общепринятым значением выражения *три часа*. Во втором случае говорящий подчёркивает достоверность предоставляемой слушателю информации и убеждает посетителя в том, что босса *действительно* нет на месте, т. е. в данном контексте языковое выражение *он в Чикаго* является истинным. Отметим, что для случаев, аналогичных примеру (35), соответствие содержания языкового выражения и реальной действительности не обязательно:

(35') *Когда он сказал «три часа», он, конечно, имел в виду три часа, но, к сожалению, из-за пробок смог добраться только к половине четвёртого.*

(43') **Если я говорю, он в Чикаго, значит, он в Чикаго, но на самом деле он Калифорнии.*

По-видимому, эти противоречия можно было бы устранить при более тщательной разработке понятия МТ, однако идея не получила дальнейшего развития ни у самого автора, ни у его коллег: класс МТ не упоминается в обзорных частях дальнейших исследований, посвящённых тавтологиям, см. [Bulhof, Gimbel 2001, [Meibauer 2008], [Rhodes 2009].

На русском материале вопрос о МТ был затронут в [Вилинбахова 2013], однако с тех пор удалось сделать некоторое количество дополнений и уточнений.

3. Метаязыковые тавтологии в ряду других тавтологических конструкций

Может показаться, что в определённых контекстах МТ очень близки по смыслу к конструкциям первого и второго ряда, описанным в [Булыгина, Шмелёв 1997]: ср. конструкцию первого ряда (3) и МТ (4) с общей имплицатурой ‘необходимость выполнения обязательств’.

(3) *Но обещание — есть обещание, слово следует держать.*

(4) А: *Что значит — обещание?*

Б: *Обещание значит обещание, то, что нарушить нельзя.*

Тем не менее, МТ имеют свои особенности, выделяющие их в ряду остальных тавтологий. Наиболее очевидное отличие — это связка ‘обозначать, иметь значение’, с помощью которой в МТ даётся характеристика языковому знаку, а не объектам действительности, как в тавтологиях первого и второго ряда. Кроме того, отличия касаются языковых выражений, используемых в качестве повторяющихся элементов МТ: это их автонимный, а не термовый,

референциальный статус⁴ и их употребление в первом, «наименее переносном» значении.

Как было указано выше, тавтологии первого и второго ряда описывают *внеязыковую действительность*. Рассмотрим следующие примеры: в (5) говорящий признаёт, что семья занимает важное место в жизни человека, и в частности — его собеседника, который пропускает дружеские встречи; в (6) подразумевается, что семья — это надёжный тыл, т. е. в обоих высказываний речь идёт о семье как о социальном явлении.

- (5) *Семья есть семья, но общаться надо и с друзьями, выкидывать нельзя людей из жизни!*
- (6) *Даже если не получится, ты всегда успеешь вернуться домой, ведь семья — это семья.*

МТ, напротив, дают инструкцию к восприятию *языкового выражения*: см. (7), где речь идёт о слове *семья*, толкование которого говорящий даёт в следующей реплике.

- (7) А: *Что для Вас значит семья?*
Б: *Семья — значит семья! Семья — это дорогие люди, которые всегда будут рядом, что бы ни случилось в жизни.*

Таким образом, в тавтологиях, описанных Т. В. Булыгиной и А. Д. Шмелёвым, первый из повторяющихся элементов имеет термовый референциальный статус, а в МТ — автономный референциальный статус⁵.

Различия в референциальном статусе элементов определяют формальные ограничения на языковые выражения, допустимые в тавтологических конструкциях. Если в тавтологиях первого и второго ряда преобладают именные группы и невозможны, например, союзы или частицы, которые не относятся к референциальным выражениям, то МТ не имеют таких ограничений, см. (8). Также в МТ в качестве повторяющихся элементов могут выступать предикативные конструкции, см. (9) и (10).

⁴ Референциальный статус именных групп, вводящих в рассмотрение ту или иную внеязыковую сущность, называется **термовым**. <...> Референциальный статус имени, которое отсылает не к той или иной сущности в мире и не к свойству, а к самому себе, называется **автономным** [Кобозева 2009: 229].

⁵ С этой точки зрения МТ напоминают «нетавтологические» примеры (см. ниже), где для читателей расшифровываются потенциально незнакомые слова. Однако, в отличие от реальных толкований, МТ не дают новых данных, а отсылают к уже известной собеседнику информации.

Это слово происходит от итальянского «дилетто», означающее «удовольствие» <...>. Дилетант — значит неспециалист, точнее, не получивший специального образования в той отрасли науки, где он отваживается что-то сказать. [НКРЯ]

- (8) А: Ой. / Б: Что ой? / А: **Ой значит ой** и глаза закрой.
- (9) Ну у нас всё просто... **хочу есть значит хочу есть...** а не своди меня в ресторан)))
- (10) Что значит «не люблю»? — Экий ты, брат, непонятливый, — ответил за Лену Майоров, делая еще один шаг в их сторону. — **«Не люблю» значит «не люблю», ни больше, ни меньше.**

Следующая особенность МТ связана с вопросом о характере общих фоновых знаний собеседников и нормативности употребления языковой единицы. Неоднократно отмечалось, что при использовании тавтологических конструкций говорящий даёт отсылку к некоторой информации, хорошо знакомой слушателю, см. [Gibbs, McCarrell 1990], [Miki 1996], [Bulhof, Gimbel 2001], [Meibauer 2008]. Это могут быть общекультурные знания, принятые в языковом коллективе, убеждения, распространённые в кругу отдельной семьи, друзей или коллег, и, наконец, факты, которые только что обсуждались участниками коммуникации. Для тавтологий первого и второго ряда речь идёт о свойствах объекта, поэтому подразумеваемая информация могут быть очень «закрытой» (знать о конкретных свойствах объекта могут только собеседники), а нормативность употребления языкового выражения не имеет значения, ср. (11), где существительное *вышка* используется в значении ‘высшая лига’.

- (11) **Вышка это вышка**, не нужно путать её с Украинской лигой.

В случае с МТ идёт отсылка к языковой компетенции адресата, которая также входит в фонд общих знаний собеседников, но уже не может быть «локальной», т.к. это противоречит семантике конструкции. Дело в том, что говорящий, используя МТ, представляет значение языкового выражения как общеизвестное, очевидное, простейшее, поэтому языковые выражения должны не просто присутствовать в сознании носителей языка, но и использоваться в своём первом, «наименее переносном» значении.

Устойчивый, общепринятый характер значения языковых единиц может быть подчёркнут апелляцией к словарю как «единому источнику нормативных сведений о литературном языке, который не пополняется новыми данными и, вообще говоря, содержит информацию, которая заведомо известна говорящему как достаточно образованному человеку» [Июмдин 2014: 130], см. пример (12) из данной работы:

- (12) **Новый это значит новый.** (Открой словарь русского языка). Если есть какие то сомнения то не покупай дешевле всё равно не найдёшь [Июмдин 2014: 119]

В примере (13) у глагола *спать* имеются оба рассматриваемых значения, но второе является переносным и потому отмечается говорящим.

- (13) **ВНИМАНИЕ!** Это реальное предложение, а не приколы. И, да, когда я говорю «спать» — это не завуалированный намек на интим-услуги. **Спать** — означает спать.

Поскольку в значение МТ входит апелляция к наиболее общепринятому значению, изучение данных конструкций могло бы пролить свет на то, какие значения многозначных слов говорящие считают главными, исходными, подобно тому как тавтологии первого и второго ряда используются для изучения стереотипных представлений об объектах и ситуациях, см., например, [Кобозева 2009: 185–197].

4. Разновидности метаязыковых тавтологий

Как уже отмечалось выше, МТ указывают на соответствие частного употребления языкового выражения общепринятому значению. Такое речевое поведение имеет смысл, когда возможна альтернативная интерпретация, см. тип апелляций к словарю в интернет-дискуссиях и его толкование в уже упомянутой работе Б. Л. Иомдина: «Говорящий употребляет слово X в общепринятом значении, но предполагает, что адресат может неправильно понять значение слова X, и на этом основании заранее оценивает поведение адресата отрицательно» [Иомдин 2014: 118].

Ср. также рассуждения Г. Уорда и Д. Хиршберга, которые считают, что наличие альтернативных вариантов является важнейшим условием употребления в речи любых, а не только метаязыковых, тавтологий: «используя в речи тавтологии, говорящий нарушает постулат количества (информативности), позволяя слушателю сделать вывод, что намеренно не были выбраны определённые альтернативные высказывания. <...> Конкретные варианты отвергнутых альтернативных высказываний зависят от контекста» [Ward, Hirschberg 1990: 510].

В случае с МТ контексты, допускающие альтернативные варианты понимания языкового выражения, делятся, как минимум, на три группы, и, таким образом, можно говорить о трёх разновидностях конструкции.

4.1. МТ, отрицающие эвфемизацию

Говорящий подчеркивает, что произнесенная им языковая единица употреблена в наиболее общепринятом значении и не является эвфемизмом, как неправильно предполагает адресат⁷. Сюда относятся контексты, где языковая единица могла бы заменять собой другое, социально менее приемлемое выражение, как в (13) и (14), где *спать* и *кофе* означает вполне определённый досуг. В эту же группу попадают МТ, отрицающие различного рода намёки, см. (9).

- (14) [обсуждение на тему: «Если мужчина приглашает на чашечку кофе»]
А с чего ты взяла, что секс вообще будет? **Кофе** — значит кофе.

4.2. МТ, отрицающие иронию

Говорящий подчеркивает, что произнесенная им языковая единица употреблена в наиболее общепринятом значении, потому что адресат может заподозрить иронию и понять его неправильно'. В данных контекстах возможно шутовское, ироническое и пр. употребление языкового выражения, меняющее его смысл на противоположное, как в (14), где собеседник понимает слово *подарок* как 'товар за деньги'.

- (15) — *Ладно, а я тебе, шотландо-фриз, между прочим, подарочек приготовил. — Он нагнулся и достал из-под прилавка толстенный том. — Кто-то из университетских профессоров заказал — и не выкупил. Словарь британских фамилий. Я думал, какому придурку мне толкнуть эту дурацкую книгу? И вспомнил про тебя.*
— *Спасибо, <...>... Сколько сдерешь за подарок?*
— *Подарок — значит, подарок. Наслаждайся, пока я добрый!*
Фабель поблагодарил Отто еще раз, уже не насмешливо...

4.3. МТ, отрицающие гиперболу

Говорящий подчеркивает, что произнесенная им языковая единица употреблена в наиболее общепринятом значении, потому что адресат может предположить преувеличение / преуменьшение и т. п. и понять его неправильно'. Речь идёт о контекстах, где языковое выражение в силу различных обстоятельств может использоваться неточно, с допустимой долей погрешности, как в (15), где автор предполагает, что читатели могут понять выражение *чуть-чуть* (*согнуть*) как, скажем, 'до некоторой степени', а не 'совсем немного'.

- (16) [Всероссийский виндсерфинг форум] *И чтобы доска хорошо ребрилась заднюю ногу действительно нужно чуть-чуть сгибать. Но чуть-чуть значит «чуть-чуть», т. е. только чтобы загнуть плавник.*

Можно заметить, что предложенные разновидности МТ соотносятся с коммуникативными постулатами Г. П. Грайса [Grice 1975]: говорящий настаивает, что нарушение, например, постулата истинности не имело места, поскольку, по его мнению, адресат может быть уверен в обратном. МТ указывают, что в данном контексте возможно другое, не буквальное восприятие языкового выражения, однако это не соответствует замыслу говорящего. Таким образом, можно описать общее значение МТ: 'Говорящий подчеркивает, что произнесенная им языковая единица употреблена в наиболее общепринятом значении и не содержит обычных отклонений от него'⁶ (эти отклонения — эвфемизация, ирония и гипербола — и соответствуют трем приводимым разновидностям конструкций).

⁶ Автор благодарен за эту формулировку анонимному рецензенту.

5. «Истинные» и «ложные» метаязыковые тавтологии

Как видно из примеров в [Miki 1996], на материале английского языка класс МТ представлен различными синтаксическими моделями. В русском языке распространены модели *X значит X*, *X означает X*, а также модель *X — это X*, см. (17).

- (17) *Воспитанный в строгих канонах западной дипломатии, когда «да» — это «да», а «нет» — это «нет», он не понимал всех хитростей и тонкостей политики на Востоке* [НКРЯ]

Пунктуационное оформление МТ в текстах разнообразно и определяется автором сообщения: возможны кавычки или курсив для одного или обоих повторяющихся элементов, запятая, тире, многоточие до и после связки и т. д.

Наиболее «удобной» для исследователя является модель *X означает X*, поскольку полученные при запросах данной модели примеры в большинстве случаев оказываются МТ. Возможность замены связки (*значит / это*) в остальных конструкциях на связку *означает* без изменения смысла зачастую может служить тестом на принадлежность их к МТ. Примеры по модели *X — это X* могут оказаться как МТ, так и тавтологиями второго ряда, выражающими идею исключительности класса (см. раздел 1). При запросе модели *X значит X*, помимо МТ, выдаются примеры омонимичной конструкции *X значит X* со значением неконтролируемости выбора и импликатурой 'ситуацию изменить невозможно, и переживать по этому поводу бессмысленно', сходную с конструкцией *X так X*, описанную в [Копотев, Файнвейц 2007].

- (20) *Двойка значит двойка! Это мелочи жизни.*

В указанных выше случаях «истинные» МТ отграничиваются от омонимичных конструкций достаточно легко: «ложные» МТ обладают другим значением, а языковые выражения в них имеют термовый, а не автономный референциальный статус.

Более сложным являются случаи вроде (*Если / раз*) *У говорит X, значит (,) X*, *Сказано X — значит X* и т. п., которые Э. Мики относит к МТ, см. известный рекламный текст, аналогичный приведённому выше примеру (43) из [Мики 1996].

- (18) *Муж: Лично я всё решаю сам. Как скажу, так и будет. Если я говорю на футбол, значит на футбол.
Жена: А не стоит ли нам поехать к маме?
Муж: Если я говорю к маме, значит, к маме.*

В разделе 2 уже было отмечено, что для подобных примеров важно соответствие значения языкового выражения (в данном случае, *на футбол* и *к маме*) действительности, которое для остальных МТ не играет роли. Более того, языковое выражение является лишь одним из компонентов ситуации, которая, как показывает

маркер *значит*, «служит говорящему основанием для того, чтобы он сделал речевой акт логического вывода» [Храковский 1998 цит. по: Апресян и др. 2010: 220]. Строго говоря, намеренность общения вообще не является обязательным при порождении языкового выражения в данных конструкциях, см. (17), где тесты показывают определённый результат, который соответствует действительности.

(19) *Ненене, они [тесты] никогда не врут! говорят добрый — значит добрый. Я им теперь знаешь как верю.*

Значение данных конструкций можно сформулировать следующим образом: 'В случае, если языковое выражение X порождается в заданном контексте, можно сделать вывод, что имеет место ситуация, которая совпадает с планом содержания X', что не соответствует рассмотренному выше значению МТ и, следовательно, позволяет не относить их к классу МТ.

6. Выводы

В русском языке, наряду с описанными Т. В. Булыгиной и А. Д. Шмелёвым двумя рядами тавтологических конструкций со значениями равноценности членов внутри класса и исключительности класса, представлен также класс метаязыковых тавтологий вида *Нет значит нет*. Используя МТ в речи, говорящий подчеркивает, что произнесенная им языковая единица употреблена в наиболее общепринятом значении и не содержит обычных отклонений от него, а именно: эвфемизации, иронии и гиперболы. В отличие от других тавтологий, языковые выражения в МТ имеют автономный референциальный статус, что позволяет использовать в данных конструкциях любые языковые единицы. В русском языке МТ представлены синтаксическими моделями *X значит X*, *X означает X* и *X — это X*, которые отграничиваются от омонимичных конструкций по семантико-прагматическим признакам.

Литература

1. Апресян Ю. Д., Богуславский И. М., Иомдин Л. Л., Санников В. З. (2010), Теоретические проблемы русского синтаксиса: взаимодействие грамматики и словаря, Языки славянских культур, Москва.
2. Булыгина Т. В., Шмелёв А. Д. (1997), Языковая концептуализация мира (на материале русской грамматики), Школа «Языки русской культуры», Москва.
3. Вилинбахова Е. Л. (2013), Доклад значит доклад: метаязыковые тавтологические конструкции в русском языке, Тезисы конференции «Русский язык: конструкционные и лексико-семантические подходы», Санкт-Петербург, с. 9–10, режим доступа: www.iling.spb.ru/confs/rusconstr2013/pdf/abstracts.pdf
4. Кобозева И. М. (2009), Лингвистическая семантика, Книжный дом «ЛИБРОКОМ», Москва.

5. *Коптев М. В., Файнвейц А. В.* (2007), Изучать так изучать: синхрония и диахрония, Научно-техническая информация, Серия 2, Информационные процессы и системы, Вып. 9, с. 29–37.
6. *Иомдин Б. Л.* (2014), Загугли в Дале. Словари в интернет-дискуссиях, в Современный русский язык в интернете, «Языки славянской культуры», Москва, с. 114–132.
7. Национальный корпус русского языка, режим доступа: www.ruscorpora.ru
8. *Падучева Е. В.* (2004), Динамические модели в семантике лексики, «Языки славянской культуры», Москва.
9. *Bulhof J., Gimbel S.* (2001), Deep tautologies, Pragmatics and Cognition, Vol. 9–2, pp. 279–291.
10. *Gibbs R. W., McCarrell N. S.* (1990), Why boys will be boys and girls will be girls: understanding colloquial tautologies, Journal of Psycholinguistic Research, Vol. 19, pp. 125–145.
11. *Grice H. P.* (1975), Logic and conversation, in Syntax and Semantics, Vol. 3: Speech Acts, Academic Press, New York, pp. 41–58.
12. *Levinson S.* (1983), Pragmatics, Cambridge University Press, Cambridge.
13. *Meibauer J.* (2008), Tautology as presumptive meaning, Pragmatics and Cognition, Vol. 16, pp. 439–470.
14. *Miki E.* (1996), Evocation and tautologies, Journal of Pragmatics, Vol. 25, pp. 635–648.
15. *Rhodes R.* (2009), A Cross-linguistic comparison of tautological constructions with special focus on English, available at: www.linguistics.berkeley.edu/~russellrhodes/pdfs/taut_qp.pdf
16. *Ward, G. L., Hirschberg J.* (1991), A pragmatic analysis of tautological utterances, Journal of Pragmatics, Vol. 15, pp. 507–520.
17. *Wierzbicka A.* (1987), Boys will be boys: ‘Radical semantics’ vs. ‘Radical pragmatics’, Language, Vol. 63, pp. 95–114.

References

1. *Апресьян Ю. Д., Богуславский И. М., Иомдин Л. Л., Саников В. З.* (2010), Theoretical issues in Russian syntax: interaction of grammar and dictionary [Teoreticheskie problemy russkogo sintaksisa: vzaimodeystvie grammatiki i slovarya], Jazyki slavyanskikh kul'tur, Moscow.
2. *Булыгина Т. В., Шмелев А. Д.* (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.
3. *Bulhof J., Gimbel S.* (2001), Deep tautologies, Pragmatics and Cognition, Vol. 9–2, pp. 279–291.
4. *Gibbs R. W., McCarrell N. S.* (1990), Why boys will be boys and girls will be girls: understanding colloquial tautologies, Journal of Psycholinguistic Research, Vol. 19, pp. 125–145.
5. *Grice H. P.* (1975), Logic and conversation, in Syntax and Semantics, Vol. 3: Speech Acts, Academic Press, New York, pp. 41–58.

6. *Iomdin B. L.* (2014), Google in Dal. Dictionaries in Internet discussions [Zagugli v Dale. Slovari v internet-diskussiyakh], in *Modern Russian language in Internet* [Sovremennyy russkiy yazyk v internete], “Jazyki slavyanskoy kul’tury”, Moscow, pp. 114–132.
7. *Kobozeva I. M.* (2009), *Linguistic semantics* [Lingvisticheskaya semantika], Knizhnyy dom “LIBROKOM”, Moscow.
8. *Kopotev M. V., Faynveyts A. V.* (2007) If we study we study: both synchrony and diachrony [Izuchat’ tak izuchat’: sinkhroniya i diakhroniya], *Scientific and technical information. Series 2, Informational processes and systems* [Nauchno-tekhnicheskaya informatsiya, Seriya 2, Informatsionnye protsessy i sistemy], Vol. 9, pp. 29–37.
9. *Levinson S.* (1983), *Pragmatics*, Cambridge University Press, Cambridge.
10. *Meibauer J.* (2008), Tautology as presumptive meaning, *Pragmatics and Cognition*, Vol. 16, pp. 439–470.
11. *Miki E.* (1996), Evocation and tautologies, *Journal of Pragmatics*, Vol. 25, pp. 635–648.
12. *Rhodes R.* (2009), A Cross-linguistic comparison of tautological constructions with special focus on English, available at: www.linguistics.berkeley.edu/~russellrhodes/pdfs/taut_qp.pdf
13. Russian National Corpus [Natsional’nyy korpus russkogo yazyka], available at: www.ruscorpora.ru
14. *Paducheva E. V.* (2004) *Dynamic models in lexical semantics* [Dinamicheskie modeli v semantike leksiki], *Jazyki slavyanskoy kul’tury*, Moscow.
15. *Vilinbakhova E. L.* (2013), Presentation means presentation: metalinguistic tautologies in Russian [Doklad znachit doklad: metayazykovye tautologicheskie konstruksii v russkom yazyke], *Proceedings of the Conference “The Russian language: constructional and lexico-semantic approaches”* [Materialy konferentsii “Russkiy yazyk: konstruksionnye i leksiko-semanticheskie podkhody], St. Petersburg, available at: www.iling.spb.ru/confs/rusconstr2013/pdf/abstracts.pdf
16. *Ward, G. L., Hirschberg J.* (1991), A pragmatic analysis of tautological utterances, *Journal of Pragmatics*, Vol. 15, pp. 507–520.
17. *Wierzbicka A.* (1987), Boys will be boys: ‘Radical semantics’ vs. ‘Radical pragmatics’, *Language*, Vol. 63, № 1, pp. 95–114.