Научный диалог = Nauchnyi dialog = Nauchnyy dialog
ISSN 2225-756X, eISSN 2227-1295

Бабенко Л. Г. Дискурс «Объяснение в любви» : проблема автоматического выявления (на материале произведений А. П. Чехова») / Л. Г. Бабенко, А. В. Эльстон-Бирон // Научный диалог. — 2021. — № 7. — С. 9—26. — DOI: 10.24224/2227-1295-2021-7-9-26.

Babenko, L. G., Elston-Biron, A. V. (2021). Discourse “Declaration of Love”: Problem of Automatic Identification (Works of A. P. Chekhov»). *Nauchnyi dialog, 7:* 9-26. DOI: 10.24224/2227-1295-2021-7-9-26. (In Russ.).

Журнал включен в Перечень ВАК

DOI: 10.24224/2227-1295-2021-7-9-26

Дискурс «Объяснение в любви»: проблема автоматического выявления (на материале произведений А. П. Чехова»)

Бабенко Людмила Григорьевна
orcid.org/0000-0002-9250-3232
доктор филологических наук
профессор
lgbabenko@yandex.ru

Эльстон-Бирон Александр Владимирович
orcid.org/0000-0002-5978-0852
аспирант
кафедра фундаментальной и прикладной лингвистики и текстоведения
eleston@mail.ru

Уральский федеральный университет имени первого Президента России Б. Н. Ельцина
(Екатеринбург, Россия)

Благодарности:
Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-312-90017/20

Discourse “Declaration of Love”: Problem of Automatic Identification (Works of A. P. Chekhov»)

Ludmila G. Babenko
orcid.org/0000-0002-9250-3232
Doctor of Philology
Professor
lgbabenko@yandex.ru

Alexander V. Elston-Biron
orcid.org/0000-0002-5978-0852
Post-graduate student
Department of Fundamental and Applied Linguistics and Textual Science
eleston@mail.ru

Ural Federal University named after the first President of Russia B. N. Yeltsin
(Yekaterinburg, Russia)

Acknowledgments:
The study was carried out with the financial support of the RFBR in the framework of a scientific project № 20-312-90017/20

© Бабенко Л. Г., Эльстон-Бирон А. В., 2021

ОРИГИНАЛЬНЫЕ СТАТЬИ

Аннотация:

Рассматривается проблема выявления дискурса «Объяснение в любви» в художественных произведениях. Представлена авторская разработка по автоматическому обнаружению ситуации *объяснения в любви*, апробированная на материале творчества А. П. Чехова. Поиск осуществлен на базе Национального корпуса русского языка. Всего выявлено свыше 200 текстов, содержащих текстовые репрезентации ситуации *объяснения в любви.* 40 из 200 текстов выделяются авторами статьи как наиболее полно репрезентирующие искомую ситуацию. Совокупность текстовых отрывков, описывающих ее, рассматривается в качестве когнитивно-дискурсивного множества *объяснения в любви*. Разработка алгоритма проводилась с опорой на выявление когнитивных схем писателя и статистический анализ лексического состава ситуации *объяснения в любви*. Среди частотных компонентов когнитивной модели, которой следует А. П. Чехов при описании процесса *объяснения в любви*, выявлено открытое пространство, заполненное растениями, птицами и др. В общем когнитивно-дискурсивном множестве *объяснения в любви* выявлены и проанализированы ядерная и спутниковая зоны, лексические компоненты которых относятся к определенным функционально-семантическим классам и функционально-семантическим группам.

Ключевые слова:
дискурс «Объяснения в любви»; когнитивно-дискурсивные множества; автоматический анализ; семантика художественных текстов; А. П. Чехов.

ORIGINAL ARTICLES

Abstract:

The problem of revealing the “Declaration of Love” discourse in works of art is considered. The author’s development on the automatic detection of the situation of *declaration of love*, tested on the material of the work of A. P. Chekhov, is presented. The search was carried out on the basis of the Russian National Corpus. In total, more than 200 texts have been identified containing textual representations of the situation of *declaration of love*. 40 out of 200 texts are identified by the authors of the article as the most fully representative of the desired situation. The set of textual passages describing it is viewed as a cognitive-discursive set of *declarations of love*. The development of the algorithm was carried out based on the identification of the cognitive schemes of the writer and the statistical analysis of the lexical composition of the situation of *declaration of love.* Among the frequency components of the cognitive model, which A. P. Chekhov follows in describing the process of *declaring love*, an open space filled with plants, birds, etc. was revealed. In the general cognitive-discursive set of *declarations of love*, the nuclear and satellite zones were identified and analyzed, the lexical components of which belong to certain functional-semantic classes and functional-semantic groups.

Key words:
discourse “Declaration of Love”; cognitive-discursive sets; automatic analysis; semantics of literary texts; A. P. Chekhov.

OPEN ACCESS

Научный диалог = Nauchnyi dialog = Nauchnyy dialog
ISSN 2225-756X, eISSN 2227-1295

УДК 81'42+821.161.1Чехов.07

Дискурс «Объяснение в любви»: проблема автоматического выявления (на материале произведений А. П. Чехова»)

© Бабенко Л. Г., Эльстон-Бирон А. В., 2021

1. Введение

Предметом теоретического изучения дискурс стал относительно недавно и в настоящее время представляет собой многозначное понятие. Одним из первых, кто попытался закрепить за дискурсом конкретное терминологическое значение, был Эмиль Бенвенист. Французский лингвист понимал под дискурсом речь, присваиваемую говорящим, и противопоставлял его объективному повествованию [Бенвенист, 1974, с. 292—300]. Широкое распространение термин *дискурс* получил во второй половине XX века в связи с работами М. Фуко, Р. Барта, Ж. Деррида, М. Мински, П. Серио и др. В отечественной лингвистике того периода термин *дискурс* как продукт функционально-коммуникативного подхода был близок понятию «функциональный стиль». Отечественные академики В. В. Виноградов и Г. О. Винокур под функциональным стилем понимали не только особый тип текстов (разговорных, бюрократических, газетных и т. п.), но и «соответствующую каждому типу лексическую систему и свою грамматику» [Цит. по: Горбунов, 2013, с. 11]. Выдающийся отечественный исследователь М. Н. Кожина отмечает, что отечественная функциональная стилистика во многом предвосхитила современный дискурсивный анализ, имея с ним один общий объект исследования — речь, взятую в функциональном аспекте. Под функционализмом она понимает «функционирование языка в целом в актах речевого общения с учетом социокультурных факторов» [Кожина, 2004, с. 11—24]. Зарубежный исследователь П. Серио видит истоки дискурса в теории речевых жанров М. М. Бахтина. Исследователь указывает на неверный перевод понятия «речевой жанр» на французский и другие европейские языки: «Бахтин писал не по-французски, а по-русски, а значит и говорил он не о “видах дискурса”, а о “жанрах речи” или “речевых жанрах”» [Серио, 2012, с. 8—21]. Таким образом, в современном понимании основы изучения дискурса проистекают из западноевропейской, американской и отечественной школ дискурсивного анализа. Сегодня термин *дискурс* востребован в исторических, философских, социальных, психоаналитических и лингвистических дисциплинах и, соответственно, имеет разные значения. Дополняющие друг друга представления о дискурсе в различных на-

учных школах и направлениях гуманитарного знания способствуют развитию общей дискурсивной теории.

В «Объяснительном словаре теории языка» А.-Ж. Греймаса и Ж. Курте под *дискурсом* понимается семиотический процесс, протекающий в различных видах дискурсивных практик: неязыковых (значимое поведение, манифестирующееся в формах, доступных чувственному восприятию) и языковых (способы словесного поведения) [Степанов, 1983, с. 488]. Во втором случае *дискурс*, с одной стороны, является синонимом *текста* и становится объектом изучения дискурсивной лингвистики. В то же время одним из важнейших компонентов современной дискурсивной теории является понятие «дискурсивные формации». По мнению О. Г. Ревзиной, у основоположника данного термина М. Фуко *дискурсивная формация* в общем виде «осуществляет спецификацию разных видов знаний через разное вербальное воплощение» [Ревзина, 2005, с. 66—78] и может относиться к различным авторам, сферам функционирования, эпохам, языкам. В таком прочтении релевантным для дискурса становится принцип «семейного сходства» Л. Витгенштейна [Микешина, 2005, с. 433]. Дискурс предстает в виде динамически открытого множества текстов или высказываний. В то же время Н. Д. Арутюнова дает другое хрестоматийное определение дискурса как речи, «рассматриваемой как целенаправленное социальное действие, как компонент, участвующий во взаимодействии людей и механизмах их сознания (когнитивных процессах)» [Арутюнова, 1990, с. 136].

Таким образом, опираясь на видение дискурса А.-Ж. Греймаса и Ж. Курте, на представление о дискурсивных формациях М. Фуко и на классическое определение дискурса Н. Д. Арутюновой, под дискурсом мы пониманием открытое и динамическое множество текстов, имеющих один денотат. В нашем исследовании таким денотатом является ситуация *объяснения в любви*. При этом грамматические, синтаксические, лексико-семантические и функционально-семантические средства выражения дискурса «Объяснение в любви» базируются на когнитивных структурах авторского сознания. По словам А. И. Новикова, «вычленяя реалии из континуума окружающей действительности, субъект воспринимает некую реальную ситуацию — денотативный источник текста» [Новиков, 1990, с. 68].

Т. А. ван Дейк указывает, что «дискурс дает представление о предметах или людях, об их свойствах и отношениях, о событиях или действиях или об их сложном сплетении, то есть о некотором фрагменте мира, который мы именуем ситуацией» [Дейк, 2000, 68—69]. Текст, считает исследователь, следует рассматривать неотрывно от когнитивных особенностей его автора. Для этого ученый использует понятие «ситуативной модели», под

которой понимает форму представления личного опыта. Причем данный опыт, по мнению Т. А. ван Дейка, не просто хранится в памяти субъекта и воспроизводится им при необходимости. Опыт является предписанием определенных действий в соответствии с динамическим изменением условий ситуации. Новые возникающие условия трансформируют ситуативную модель, в результате чего появляются ее варианты. Таким образом, ситуативные модели выступают как особые формы структуры организации знания о возможных формах взаимодействия с окружающим миром и могут отражаться автором в литературном художественном тексте [Там же, 68—110]. Создание художественного текстового описания ситуативной модели *объяснения в любви* обязательно затрагивает когнитивные структуры знания автора. Для того, чтобы учесть данный аспект, каждый вариант текстового воплощения типовой ситуации *объяснения в любви* воспринимается нами как когнитивно-дискурсивная единица (далее — КДЕ). Весь же дискурс «Объяснение в любви» представляет собой совокупность таких КДЕ и видится как соответствующее когнитивно-дискурсивное множество (далее — КДМ) [Бабенко, 2020, с. 34].

Целью нашей статьи является разработка алгоритма по автоматическому выявлению текстовых репрезентаций ситуации *объяснения в любви* в творчестве А. П. Чехова, а также анализ общих и специфических дискурсивных особенностей лексического состава общего дискурса «Объяснение в любви». Исследование проводилось на базе Национального корпуса русского языка (далее — НКРЯ). Материалом послужил подкорпус произведений А. П. Чехова, включенный в аннотированную базу НКРЯ. Общий объем подкорпуса составляет 280 текстов, которые принадлежат различным литературным жанрам и охватывают период творчества с 1883 по 1904 годы.

В процессе работы использованы различные методы исследования, среди которых метод корпусного анализа текста, метод статистического анализа лексического состава ситуации *объяснения в любви*, метод полевого структурирования дискурса «Объяснение в любви» с выявлением ядерных и спутниковых зон, метод когнитивно-дискурсивного анализа, метод интерпретации. В процессе разработки был решен комплекс задач, связанных с моделированием алгоритма по автоматическому выявлению ситуации *объяснения в любви* в произведениях А. П. Чехова. Среди них:

— определение лексико-семантического состава репрезентаций ситуации *объяснения в любви* в произведениях А. П. Чехова;

— определение функционально-семантических групп (далее — ФСГ) [Кузнецова, 1975] и функционально-семантический классов (далее — ФСК) выявленных лексических единиц;

— установление числового показателя для лексических маркеров, который отражает вероятность обнаружения ситуации *объяснения в любви* в конкретном произведении автора;

— формирование общего дискурса «Объяснение в любви» на основе полученного КДМ.

Подход к анализу текстовых репрезентаций ситуации *объяснения в любви* в творчестве А. П. Чехова с позиций дискурсивной теории является сегодня принципиально новым, актуальным и теоретически значимым.

2. Ситуативные модели в художественном тексте. Проблема репрезентативности любовного дискурса

В современной лингвистике глобальный литературно-художественный любовный дискурс еще далек от полного и всестороннего освещения. В нашей работе основным ориентиром в определении литературно-художественного дискурса является видение В. И. Карасика. Ученый полагает, что такой тип дискурса выступает формой художественного осмысления фундаментальных характеристик бытия индивидуальным сознанием, воплощающемся в литературном тексте [Карасик, 2002, с. 240—243]. Такое представление подчеркивает связь когнитивных и дискурсивных процессов в ситуативных моделях как формах знания, воплощенных в литературно-художественных текстах.

Исследователи, обращающиеся к различным ситуативным моделям в глобальном любовном литературно-художественном дискурсе, ограничивают материал рамками одного автора, одной ситуации или тем и другим. Например, выдающийся семиотик Р. Барт в работе «Фрагменты речи влюбленного» [Барт, 2002] создает структуральный портрет влюбленного на основе глобального художественного любовного дискурса, то есть концентрируется только на субъекте дискурса. Работа французского семиотика алогична по структуре и выдает себя за книгу о любви. Однако читать ее следует как металитературный текст. Р. Барт оформляет свой комментарий речи влюбленного «как хаотичное множество отрывков — “фигур”, “приступов речи”, более или менее случайно перетасованных» [Там же, с. 2]. Сам автор дает следующую характеристику: «Фигуры вычленяются, когда удается распознать в протекающем дискурсе что-то, что было прочитано, услышано, испытано. Фигура очерчена (как знак) и памятна (как образ или рассказ)» [Там же, с. 32]. Такая характеристика сближает «фигуру» с ситуационной моделью Т. А. ван Дейка. Представленное Р. Бартом собрание фигур отдаленно напоминает идеографический словарь, раскрывающий смыслы таких понятий, как «восхищение», «встреча», «ревность» и др., в контексте любовного дискурса. Хаотичность, отмеченная самим Р. Бар-

том, касается и выборки материала для исследования. В качестве примеров Р. Барт приводит цитаты из самых различных источников: Библии (Новый Завет), «Страдания юного Вертера» И. Гете, «Тошнота» Ж-П. Сартра, «Жюстина, или Несчастная судьба добродетели» маркиза де Сада, «Рождение трагедии из духа музыки» Ф. Ницше и др. С одной стороны, такая разнонаправленность оправдана формационным, синкретичным подходом к дискурсу М. Фуко, о котором речь шла выше. С другой — действительно вносит хаос, поскольку материал совершенно не систематизирован и отбирается на основе личного опыта Р. Барта. Безусловно, опыта богатейшего, однако ограниченного в силу объективных причин.

А. Греймас и Ж. Фонтани обращаются к проблемам анализа семиотики страстей и дискурса страсти (как любви) и трактуют эмоциональность и страсть как образ жизни, существующий за пределами рациональности и семантической компетенции. Исследование проводится на материале художественных произведений «Отелло» У. Шекспира, «Любовь Свана» и «Пленница» М. Пруста, «Ревность» А. Роб-Грийе, а также нескольких сцен из трагедий Ж. Расина. Объектом исследования становится «ревность» как эмоция, сопутствующая любовному дискурсу. Ученые разрабатывают дискурсивную схему «ревности», в которой ими обнаруживаются определенные последовательности в действиях героев литературных произведений на микро- и макроуровнях развития сюжета [Греймас, 2007, 271—322]. Отбор литературных произведений вызывает уважение, однако является далеко не исчерпывающим.

Большая конкретика в отборе материала наблюдается в литературоведческом дискурс-анализе, представленном в диссертационном исследовании любовного дискурса О. А. Велюго. Работа носит название «Художественное воплощение дискурса любви в прозе Джулиана Барнса и Иэна Макьюэна» [Велюго, 2019]. Как видно из названия, материалом послужили художественные тексты двух британских авторов. Исследователь приходит к выводу о том, что в творчестве этих писателей существует определенное противопоставление классического романтического (пасторального) дискурса и современного (жесткого и жестокого) дискурса любви. О. А. Велюго считает, что дискурс любви в творчестве данных британских писателей является рефлексией персонажей, «функции которой — создание образа “героя нашего времени” и поиск рефлексивных стратегий, необходимых человеку для анализа собственных чувств, взаимоотношений, ситуаций» [Там же, с. 2]. Однако исследователь в качестве материалов использует только отдельные произведения Дж. Барнса и И. Макьюэна и не охватывает их творчество целиком.

Более детальный подход к отбору материала представлен в работе М. В. Ползуновой. Исследование проводится с позиций теории речевых

жанров М. М. Бахтина и носит название «“Объяснение в любви” как сложный речевой жанр: лексика, грамматика, прагматика» [Ползунова, 2008]. Анализу подвергаются 84 текстовых фрагмента из трилогии Дж. Голсуорси «Сага о Форсайтах». Все фрагменты получены методом сплошной выборки. Ученый приходит к выводу, что «особенностью сложного речевого жанра “Объяснение в любви” является его внутрижанровая структура, представляющая собой комбинаторику простых речевых жанров» [Там же, с. 23]. Различия в комбинациях дают варианты текстовых репрезентаций речевого жанра «Объяснения в любви».

Перечисленные выше работы имеют одно общее свойство. Исследования проводились на материале художественных текстов, принадлежащих ограниченному числу произведений. Это обусловлено методом ручной сплошной выборки текстовых репрезентаций той или иной ситуативной модели. В случае с Р. Бартом такая выборка, охватывая огромный материал, порождает проблему систематизации. Выводы, к которым приходят ученые, касаются только конкретных (отобранных) литературно-художественных произведений, то есть далеко не всего творческого наследия автора, и тем более глобального любовного дискурса, имеющего динамическую структуру. Следом возникает проблема интерпретации, выражающаяся в вопросе, какие критерии необходимо применить к тексту, чтобы говорить о наличии в нем исследуемой ситуации. Современные методы корпусной лингвистики позволяют обработать практически неограниченную базу художественных текстов с целью автоматического выявления необходимых текстовых репрезентаций для их дальнейшего многоаспектного исследования. Статистические методы компьютерной лингвистики могут быть использованы для оценки формальных параметров текста. Подобные данные можно использовать в качестве аргумента при решении о включении конкретной текстовой репрезентации в общий дискурс.

Для того, чтобы обеспечить репрезентативность КДМ *объяснения в любви* — то есть указать на реальное распределение изучаемых КДЕ в глобальной совокупности художественных текстов А. П. Чехова — нами был учтен ряд обязательных условий:

— каждая КДЕ *объяснения в любви*, включенная в соответствующее КДМ, имела равную вероятность попадания в выборку;

— выборка КДЕ производилась независимо от изучаемого признака, то есть независимо от выводов о наличии репрезентации ситуации *объяснения в любви* в тексте конкретного произведения;

— отбор КДЕ производился из однородных совокупностей, то есть из всех художественных текстов А. П. Чехова, доступных в НКРЯ;

— число КДЕ в финальной выборке было достаточно большим по отношению к количеству текстов А. П. Чехова в подкорпусе НКРЯ.

Следование этим принципам позволило в дальнейшем сформировать КДМ *объяснения в любви* в произведениях А. П. Чехова.

3. Выявление структуры КДМ *объяснения в любви*: лексико-семантического ядра и спутниковой зоны

При попытке выявить ситуацию *объяснения в любви* в материалах НКРЯ пришлось столкнуться с рядом традиционных для лингвистических корпусов ограничений. Лингвистическая разметка художественных текстов А. П. Чехова, содержащихся в соответствующем подкорпусе, ориентирована на уровневую языковую модель. Функциональной, коммуникативной или речежанровой характеристики подкорпус А. П. Чехова не имеет. Фонетическая, морфологическая, синтаксическая разметки могут служить только дополнительными инструментами при анализе прагматической семантики ситуации. Хотя выделенный нами подкорпус обладает семантической разметкой, однако «семантическая омонимия в Корпусе не снимается: многозначным словам приписывается несколько альтернативных наборов семантических признаков» [НКРЯ, 2021]. Таким образом, семантическая аннотация НКРЯ в большинстве случаев не учитывает актуализацию значений языковых единиц, обусловленную прагматикой коммуникации. При попытке поиска по отдельным лексемам *объяснение* и *любить* количество случаев омонимии оказалось настолько высоким, что поиск стал нецелесообразным. С другой стороны, при поисковом запросе показать ситуацию *объяснения в любви* через соответствующую фразу, мы получили в результатах лишь 7 вхождений в 280 текстах автора, например: *пропускают описания природы и прочитывают* ***объяснения в любви*** *по пяти раз; встречи, наши поездки за город,* ***объяснение в любви*** *и погоду, которая …; лунного вечера на Волге, ни* ***объяснений в любви****, ни поэтической жизни; меня она … Отлично помню наше* ***объяснение в любви****. Она сидит …; Комплименты, любезности,* ***объяснения в любви — ****ваша речь; В маленькой зале, происходило, по-видимому,* ***объяснение в любви****; объяснялись их дочь …; В обязанность наблюдать: а) чтобы* ***объяснения в любви*** *производились высоким слогом …* [Там же].

Для того чтобы решить проблему малого количества репрезентаций, нам пришлось установить дополнительные лексико-семантические компоненты ситуации *объяснения в любви*. Автор описывает ситуацию с опорой на свои когнитивные стратегии и ментальные репрезентации. Обнаружение этих механизмов позволило выявить конкретный лексико-семантический состав КДЕ и способствовало решению поставленной нами задачи.

Анализ семи представленных выше примеров показал, что чеховская ситуация *объяснения в любви*:

— разворачивается в уединенном пространстве;

— данное пространство может быть открытым;

— описывая открытое пространство, автор указывает на близкое расположение природных водоемов, растений или присутствие птиц.

В результате мы смогли уточнить лексические границы КДЕ *объяснения в любви*, то есть охарактеризовать ту лексику, которая может гипотетически относиться к искомой ситуативной модели. При этом некоторые лексические компоненты обладают более устойчивой ситуативной семантикой (*объясняться в любви, объяснение в любви, любовь*), другие с относительно высокой вероятностью сопутствуют описанию ситуации *объяснения в любви* (различные наименования птиц, растений, типы водоемов). Следовательно, можно говорить о наличии у КДЕ *объяснения в любви* ядерной и спутниковой лексической зон. В связи с этим мы обратились к теории риторических структур [Taboada, 2006; Ковальчук, 2016]. В соответствии с данной теорией ядро каждого дискурса связано с его общей семантикой и выражается на уровне макротемы. Спутники дискурса, во-первых, конкретизируют ядро, а во-вторых, участвуют в процессе когерентности всего дискурса. Основываясь на этих теоретических предпосылках и результатах нашего анализа, мы предположили, что к ядерной зоне КДЕ *объяснения в любви* относится глагол *любить*. Однако дискурсивные связи становятся лучше различимы на уровне предложения. В соответствии с актантной теорией Л. Теньера [Теньер, 1988, с. 121—126] к предикативной компоненте *любить* мы добавили актанты *Я* и *Вас / тебя*. Глагол задал позиции актантов в качестве агентива и объектива. В результате мы получили 12 возможных синтаксических конфигураций, относящихся к ядерной зоне. Запрос на уровне высказывания *Я Вас / тебя люблю* позволил снизить число случаев с возможной омонимией ядерного глагола *любить*. Поиск в НКРЯ среди подкорпуса художественных произведений А. П. Чехова по всем возможным конфигурациям позволил выявить уже 43 текстовых отрывка, содержащих КДЕ *объяснения в любви*. Репрезентации относятся к таким текстам, как, например, «Вишневый сад» (1904), «Три сестры» (1901), «Рассказ неизвестного человека» (1893), «Моя жизнь» (1896), «О том, как я в законный брак вступил» (1883), «Злой мальчик» (1883), «Жалобная книга» (1884), «Черный монах» (1893) и др.

Статистический анализ частотных лексических единиц, находящихся в данных текстовых репрезентациях ситуации *объяснения в любви*, показал наличие новых ядерных и спутниковых лексических компонентов. Это позволило уточнить когнитивные схемы, которым следует А. П. Чехов при

описании ситуации *объяснения в любви* и установить конкретные лексические компоненты, что подтверждает выявленные ранее характеристики данной ситуации:

— автор регулярно помещает участников ситуации в уединенное пространство;

— это пространство чаще является открытым (объяснение происходит на природе или в садово-парковых сооружениях);

— в описании ситуации *объяснения в любви* присутствует высокочастотное описание растений и указание на близкое расположение различных птиц.

Лексические компоненты были объединены нами в определенные ФСГ и ФСК. В качестве примера приведем некоторые лексемы с указанием зоны, к которой они относятся. Ядерная зона представлена глаголами (*любить, полюбить, влюбиться, влюбляться / объясниться, объясняться, признаться*), существительными (*любовь, влюбленность / объяснение, признание*), прилагательными (*любимый, влюбленный*), модально ориентированной конструкцией (*Да или нет?*). В спутниковой зоне выявлены существительные, встречающиеся в хронотопических характеристиках КДЕ. Существительные спутниковой зоны относятся к различным денотативно-идеографическим классам (далее — ДИК): строения (*беседка, беседочка*), искусственные насаждения (*клумба, аллея, палисадник, цветник* и др.), цветы (*ландыш, сирень, тюльпан, роза, гелиотроп, ирис* и др.), деревья (*черемуха, акация*), птицы (*голубь, бекас, дрохва, утка, гоголь, журавль, воробей, галка, выпь, иволга, коршун* и др.). Также в спутниковой зоне выявлены частотные прилагательные (*лунный, ночной* — относятся к хронотопическим характеристикам / *ландышевый, сиреневый, черемуховый* — производные от наименования растений) и глагол (*пахнуть*). Поскольку данные денотативно-идеографические группы (далее — ДИГ) и ДИК оказались важны для автора при описании ситуации *объяснения в любви*, то мы обратились к частотным словарям языка А. П. Чехова. Это «Частотный грамматико-семантический словарь языка художественных произведений А. П. Чехова с электронным приложением» [Частотный …, 2012] и «Частотный словарь рассказов А. П. Чехова» [Частотный …, 1999]. В данных словарях методом сплошной электронной выборки были выявлены все лексические единицы, относящиеся к обозначенным нами частотным ДИГ и ДИК слов. К ДИК относятся: птицы, цветы, деревья, кустарники, астрономические объекты, водные объекты, рукотворные парковые сооружения и др. Всего в ядерной зоне нам удалось обнаружить 15 (рис. 1) высокочастотных лексических единиц, в спутниковой — 74 (рис. 2). Результаты статистического анализа мы приводим частично. Общее количество выяв-

ленных текстов, потенциально содержащих КДЕ *объяснения в любви*, составляет свыше 220 наименований. Всего А. П. Чеховым написано около 550 художественных произведений.

№	РАССКАЗ / ЛЕКСЕМА	1 любовь	2 любить	3 полюбить	4 объяснение	5 объясниться	6 объясняться	15 влюбляться	ИТОГО
1	Актерская гибель (1885-86)	1	4		1		1		7
2	Анна на шее (1895)								0
3	Антрепренер под диваном (1885-85)								0
4	Ариадна (1895)	13	12				1	2	32
5	Архиерей (1902)	2	1						3
6	Бабы (1891)		3						3
7	Бабье царство (1894)	11	12		1				25
215	Чайка (1896)	11	16	1					32
216	Человек в футляре (1898)	1	2		1				7
217	Черный монах (1893)	8	10				1		22
218	Шило в мешке (1885-86)								0
219	Экзамен на чин (1884-85)		1						1
220	Юбилей (1892)		5						7

Рис. 1. Статистика лексем ядерной зоны КДМ *объяснения в любви*

№	РАССКАЗ / ЛЕКСЕМА	1 ночь	2 соловей	3 скамеечка	4 запах	5 пахнет	6 цветок	74 луна	ИТОГО
1	Актерская гибель (1885-86)	2			2			1	15
2	Анна на шее (1895)	1			1	3	1	1	10
3	Антрепренер под диваном (1885-85)								3
4	Ариадна (1895)	7			1	1	1		51
5	Архиерей (1902)	3			1			5	19
6	Бабы (1891)	4	2			1		3	22
7	Бабье царство (1894)	4		1	9	4	2		51
215	Чайка (1896)	10					17	9	95
216	Человек в футляре (1898)	5						4	19
217	Черный монах (1893)	11	1		1		8	2	69
218	Шило в мешке (1885-86)								1
219	Экзамен на чин (1884-85)								1
220	Юбилей (1892)	1					1		11

Рис. 2. Суммарная статистика лексем КДМ *объяснения в любви*

Каждая лексема (в своей полной грамматической парадигме) была проанализирована на предмет вхождения в тексты А. П. Чехова. Например, существительное *любовь* искалось нами во всех падежах, глагол *любить* — во всех временных формах, прилагательное *любимый* — по роду, числу, падежу, краткости и т. д. В результате были выявлены тексты, в которых лексема встречается. При этом нами учитывался параметр количества вхождений лексемы в конкретный текст. Таким образом, мы смогли указать на концентрацию выявленных лексем в каждом тексте. Далее, путем сложения количества лексем, относящихся к ядерной и спутниковой зонам, мы установили общую сумму вхождений (см. на рис. 1. столбец

«итого»). Проведенный анализ дал несколько важных для разработки алгоритма результатов.

Во-первых, посредством спутниковых лексем нам удалось обнаружить новые текстовые отрывки, в которых присутствует КДЕ *объяснения в любви*. При этом КДЕ может быть выражено другими лексическими средствами, не учтенными нами ранее. Например, ядерное высказывание *Я Вас / тебя люблю* неоднократно заменяется или дополняется А. П. Чеховым ситуативно синонимичным *жить я не могу без вас*: ***Жить я не могу без вас**, моя дорогая!* [Чехов, 1975, с. 262] — или *я понял для чего живу*: *Увидев впервые вас, **я понял, для чего я живу**, я узнал цель моей жизни!* [Чехов, 1974, с. 103] и т. п. Все новые лексические единицы или фразовые единства, показавшие высокую статистику, использовались нами в новых поисковых запросах. Это давало возможность обнаружить новые репрезентации ситуации *объяснения в любви*. Данный метод можно обозначить метафорически как наматывание клубка из разных нитей. Каждая новая найденная высокочастотная лексема или фраза потенциально отсылала к текстам, в которых ранее на основе уже известной нам лексики не удавалось автоматически выявить ситуацию *объяснения в любви*. То есть поиск в НКРЯ давал не только сами репрезентации, но и, совершенствуясь, позволял найти новые лексические единицы для запроса.

Во-вторых, количественный учет пересекающихся ядерных и спутниковых лексем в одном произведении дал возможность выявить те художественные тексты, в которых вероятность обнаружения КДЕ выше. Например, для рассказа «Архиерей» (1902) число таких взаимных вхождений составляет 19. Тогда как для рассказа «Из записок вспыльчивого человека» (1887) это число равно 37. Следовательно, первое произведение с меньшей вероятностью содержит КДЕ *объяснения в любви*, чем второе, что легко подтверждается эмпирически. При этом отметим, что при переводе полученных результатов в процентную величину (отношение числа вхождений к числу слов всего текста в произведении) результаты сохраняются.

В-третьих, нам удалось частично решить вопрос об основаниях интерпретации наличия ситуации «объяснение в любви» в конкретных текстах. Спорные случаи стало возможным оценивать по статистическому параметру с учетом когнитивной стратегии автора по отбору лексики для описания ситуации.

Отметим, что первостепенное значение в работе с алгоритмом по автоматическому выявлению КДЕ *объяснения в любви* имеют лексемы ядерной зоны. Например, для рассказа «В Москве» (1891) число вхождений равно 48, в то время как для повести «Степь» (1888) оно составляет 111. Интерпретация этих результатов должна в первую очередь основываться

на учете ядерной лексической зоны КДЕ и ее семантике. Несмотря на то, что повесть «Степь» имеет больший показатель, вероятность содержания в ней КДЕ *объяснения в любви* ниже, поскольку свыше 100 лексем данного произведения относятся к спутниковым зонам (*цветок, ночь, запах, пахнет* и др.), а лексемы ядерной зоны имеют иную, отличную от искомой семантику. Рассказ «В Москве» содержит всего 9 вхождений. Однако 3 из них непосредственно относятся к макротеме КДЕ и выражены лексемами *объяснение* (1 вхождение) и *любовь* (2 вхождения). Таким образом, рассказ «В Москве» с большей долей вероятности содержит КДЕ *объяснения в любви*, что легко доказывается обращением к тексту произведения.

4. Определение порогового числового значения для КДЕ. Формирование КДМ *объяснения в любви*

На основании полученных статистических результатов мы смогли установить пороговое числовое значение, которое отражает вероятность обнаружения КДЕ *объяснения в любви* в произведениях А. П. Чехова. Пороговое значение устанавливается в отношении ядерной и спутниковой зон КДЕ суммарно. Оно говорит о минимально необходимой концентрации ядерных и спутниковых лексем в составе КДЕ *объяснения в любви*. Наш эмпирический анализ показал, что таким пороговым значением является число 20 (при условии распределения ядерной и спутниковой лексем в минимальном соотношении ⅓). В случае роста ядерной зоны вероятность содержания КДЕ *объяснения в любви* возрастает, в случае уменьшения — снижается. Однако встречаются такие исключения, как, например, «Жалобная книга» (1884). Здесь обнаруживается лишь одно вхождение, представленное ядерным глаголом *любить* с актантами *я, Вас / тебя* и почти целиком репрезентирующее всю КДЕ *объяснения в любви.* Полный текст данной КДЕ следующий: *Катинька, я вас люблю безумно!* [Чехов, 1975, с. 359]. Остальные компоненты относятся исключительно к экстралингвистическим факторам.

На основании полученных статистических данных можно констатировать, что дискурс *объяснения в любви* для А. П. Чехова является частотным. Всего автором создано свыше 550 произведений, из которых 220 потенциально содержат КДЕ *объяснения в любви*. Учитывая пороговое значение (20 единиц), мы можем с большой долей уверенности сказать, что ситуация *объяснения в любви* содержится в 40 произведениях автора, таких как «Ариадна» (1895), «Бабы» (1891), «Бабье царство» (1894), «Безотцовщина» (1878), «В родном углу» (1897), «Вишневый сад» (1904), «Дама с собачкой» (1899), «Дачные правила» (1884-85), «Драма на охоте» (1884),

«Дуэль» (1891), «Иванов» (1887), «Из записок вспыльчивого человека» (1887) и др.

Представив репрезентации ситуации *объяснения в любви* в вышеперечисленных рассказах как КДМ, мы можем указать, что центральными произведениями, более полно репрезентирующими ситуацию, можно считать «Драму на охоте» (1884), «Дуэль» (1891), «Три года» (1895), «Чайка» (1896). Помимо этого, статистика свидетельствует о более частом обращении автора к ситуации «объяснение в любви» после 1887 года. Здесь можно особо выделить 1891 год и повесть «Дуэль». В тексте этого произведения содержится 100 лексических единиц, относящихся к ситуации *объяснения в любви*. 1893 год представлен повестью «Черный монах» и рассказом «Из записок вспыльчивого человека», насчитывающими соответственно 69 и 37 лексических единиц. К 1895 и 1896 годам относятся повести «Три года» (104 лексических единицы) и «Моя жизнь» (117), пьеса «Чайка» (106) и т. д.

5. Заключение

Разработанная нами методика по автоматическому обнаружению репрезентаций ситуации *объяснения в любви* в произведениях А. П. Чехова позволила сформировать соответствующее КДМ. Всего нами выявлено свыше 200 произведений, потенциально содержащих исследуемую ситуацию. Помимо этого, мы можем указать на 40 конкретных художественных текстов А. П. Чехова, в которых ситуация *объяснения в любви* содержится обязательно. К сожалению, рамки статьи не позволяют перечислить их все, выше мы указали некоторые примеры.

Разработанный алгоритм прошел проверку на произведениях, отсутствующих в НКРЯ. В настоящее время нами формируется электронная база данных всех произведений А. П. Чехова. Алгоритм уже позволил выявить два дополнительных текста, содержащих ситуацию *объяснения в любви*. Это «Моя Нана» (1883—1884) и «Верочка» (1887). Дальнейшая работа в этом направлении позволит охватить все творчество автора и более полно исследовать ситуацию *объяснения в любви* с позиций дискурсивной теории. Многоаспектный анализ выявленных нами репрезентаций позволит указать на особенности проявления дискурса «Объяснение в любви» в творчестве А. П. Чехова. Помимо этого, сформированное КДМ, может способствовать изучению общего любовного дискурса в творчестве писателя.

Разработанный алгоритм может быть внедрен в компьютерные самообучающиеся системы по автоматическому анализу различных типов дискурсов и автоматическому анализу семантики текстов.

Источники и принятые сокращения

1. НКРЯ — *Национальный* корпус русского языка [Электронный ресурс]. — Режим доступа : https://ruscorpora.ru/new/ (дата обращения 11.05.2021).

2. *Частотный* грамматико-семантический словарь языка художественных произведений А. П. Чехова (с электронным приложением) / О. В. Кукушкина, Е. В. Суровцева, Л. В. Лапонина, Д. Ю. Рюдигер ; под общ. ред. А. А. Поликарпова. — М.: МАКС Пресс, 2012. — 571 с. ISBN 978-5-317-04249-3

3. *Частотный* словарь рассказов А. П. Чехова / авт.-сост. А. О. Гребенников ; под ред. Г. Я. Мартыненко. — Санкт-Петербург : Изд-во Санкт-Петербургского университета, 1999. — 171 с. ISBN 5-288-02166-X.

4. *Чехов А. П.* И то и се : (Поэзия и проза) / А. П. Чехов // Полное собрание сочинений и писем : В 30 т. Сочинения : В 18 т. / АН СССР. Ин-т мировой лит. им. А. М. Горького. — Москва : Наука, 1974—1982. Т. 1. [Рассказы. Повести. Юморески], 1880—1882. — Москва : Наука, 1974. — С. 103—105.

5. *Чехов А. П.* Знамение времени / А. П. Чехов // Полное собрание сочинений и писем : В 30 т. Сочинения : В 18 т. / АН СССР. Ин-т мировой лит. им. А. М. Горького. — Москва : Наука, 1974—1982. Т. 2. [Рассказы. Юморески], 1883—1884. — Москва : Наука, 1975. — С. 262.

Литература

1. *Арутюнова Н. Д.* Дискурс / Н. Д. Арутюнова // Лингвистический энциклопедический словарь. — Москва : Советская Энциклопедия, 1990. — С. 136—137.

2. *Бабенко Л. Г.* Типы лексических множеств в структурно-семантическом, когнитивно-дискурсивном и лексикографическом освещении : динамика интерпретаций / Л. Г. Бабенко // Научный диалог. — 2020. — № 9. — С. 9—47. — DOI: 10.24224/2227-1295-2020-9-9-47.

3. *Барт Р* Фрагменты любовной речи / Р. Барт. — Москва : Ад Маргинем, 2002. — 432 с. — ISBN 978-5-91103-208-1.

4. *Бенвенист Э.* Общая лингвистика / Э. Бенвенист. — Москва : Прогресс, 1974. — 448 с.

5. *Велюго О. А.* Художественное воплощение дискурса любви в прозе Джулиана Барнса и Иэна Макьюэна : автореферат диссертации … кандидата филологических наук : 10.01.03 / О. А. Велюго. — Минск, 2019. — 26 с.

6. *Горбунов А. Г.* Дискурс как новая лингвофилософская парадигма : учебное пособие / А. Г. Горбунов. — Ижевск : Изд-во «Удмуртский университет», 2013. — 56 с. — ISBN 978-5-4312-0215-5.

7. *Греймас А.* Семиотика страстей. От состояния вещей к состоянию души / А. Греймас, Ж. Фонтаний : Перевод с французского И. Г. Меркулова. — Москва : Издательство ЛКИ, 2007. — 334 с. — ISBN 978-5-382-00061-9.

8. *Дейк Т. А. ван.* Язык. Познание. Коммуникация / Т. А. ван Дейк. — Благовещенск : Типографкомплекс, 2000. — 308 с. — ISBN 5-80157-202-3.

9. *Карасик В. И.* Языковой круг : личность, концепты, дискурс / В. И. Карасик. — Волгоград : Перемена, 2002. — 477 с. — ISBN 5–88234–552–2.

10. *Кожина М. Н.* Речеведение : функциональная стилистика и дискурсный анализ / М. Н. Кожина // Стил.–Београд–Бањалука. — 2004. — № 3. — С. 11—24.

11. *Микешина Л. А.* Философия науки : учебное пособие для гуманитарных и негуманитарных направлений и специальностей вузов России / Л. А. Ми-

кешина. — Москва : Международный университет в Москве, 2006. — 463 с. — ISBN 5-9248-0121-7.

12. *Новиков А. И.* Семантические расстояния в языке и тексте / А. И. Новиков, Е. И. Ярославцева. — Москва : Наука, 1990. — 135 с. — ISBN 5-02-010986-X.

13. *Ползунова М. В.* Объяснение в любви как сложный речевой жанр : лексика, грамматика, прагматика : автореферат диссертации … кандидата филологических наук : 10.02.19 / М. В. Ползунова. — Екатеринбург, 2008. — 24 с.

14. *Ревзина О. Г.* Дискурс и дискурсивные формации / О. Г. Ревзина // Критика и семиотика. — 2005. — № 8. — С. 66—78.

15. *Серио П.* Почему Бахтин — не Пешё? Об одном большом недоразумении по поводу анализа дискурса / П. Серио // Язык, коммуникация и социальная среда. — 2012. — Выпуск 10. — С. 8—21.

16. *Степанов Ю. С.* Семиотика / Ю. С. Степанов. — Москва : Радуга, 1983. — 637 с.

17. *Теньер Л.* Основы структурного синтаксиса / Л. Теньер ; пер. с франц., под общ. ред. В. Г. Гака. — Москва : Прогресс, 1988. — 656 с.

18. *Taboada M.* Rhetorical Structure Theory : looking back and moving ahead [Electronic resource] / M. Taboada, W. C. Mann // Discourse Studies. — 2006. — № 8. — Pp. 423—459. — Access mode: http://dis.sagepub.com/content/8/3/423.

Material resources

Chekhov, A. P. (1974). And this and that: (Poetry and prose). *The complete collection of works and letters: In 30 vols. Essays: In 18 volumes, 1. [Stories. Stories. Humoreski]. AS of the USSR. Institute of World Literature named after A. M. Gorky.* Moscow: The science. Moscow: The science. 103—105. (In Russ).

Chekhov, A. P. (1975). The Sign of the Time. *The complete collection of works and letters: In 30 vols. Essays: In 18 volumes, 2. [Stories. Humoreski]. AS of the USSR. Institute of World Literature named after A. M. Gorky.* Moscow: The science. Moscow: The science. 262. (In Russ).

Grebennikov, A. O. (1999). *Frequency dictionary of stories by A. P. Chekhov.* St. Petersburg: Publishing House of St. Petersburg University. 171 p. ISBN 5-288-02166-X. (In Russ).

Kukushkina, O. V., Surovtseva, E. V., Laponina, L. V., Ryudiger, D. Yu. (2012). *Frequency grammatical and semantic dictionary of the language of literary works of A. P. Chekhov (with an electronic application).* Moscow: MAKS Press. 571 p. ISBN 978-5-317-04249-3. (In Russ).

NKRYA — *The National corpus of the Russian language.* Available at: https://ruscorpora.ru/new/ (accessed 11.05.2021). (In Russ).

References

Arutyunova, N. D. (1990). Discourse. *Linguistic encyclopedic dictionary.* Moscow: Soviet Encyclopedia. 136—137. (In Russ).

Babenko, L. G. (2020). Types of Lexical Sets in Structural-Semantic, Cognitive-Discursive and Lexicographic Coverage: Dynamics of Interpretations. *Nauchnyi dialog, 9:* 9—47. DOI: 10.24224/2227-1295-2020-9-9-47. (In Russ).

Bart, R. (2002). *Fragment of a love speech.* Moscow: Ad Marginem. 432 p. ISBN 978-5-91103-208-1. (In Russ).

Benveniste, E. (1974). *General linguistics.* Moscow: Progress. 448 p. (In Russ).

Deik, T. A. van. (2000). *Language. Poznań. Communications.* Blagoveshchensk: Printing House. 308 p. ISBN 5-80157-202-3. (In Russ).

Gorbunov, A. G. (2013). *Discourse as a new linguophilosophical paradigm: a textbook.* Izhevsk: Publishing house "Udmurt University". 56 p. ISBN 978-5-4312-0215-5. (In Russ).

Greimas, A, Fontania, J. (2007). *Semiotic of passions. From the state of things to the state of the soul.* Moscow: LKI Publishing House. 334 p. ISBN 978-5-382-00061-9. (In Russ).

Karasik, V. I. (2002). *Language circle: personality, concepts, discourse.* Volgograd: Smena. 477 p. ISBN 5-88234-552-2. (In Russ).

Kozhina, M. N. (2004). Speech production: functional stylistics and discursive analysis. *Steele.Beograd-Bangalore, 3:* 11—24. (In Russ).

Mikeshina, L. A. (2006). *Philosophy of science: a textbook for humanitarian and non-humanitarian areas and specialties of Russian universities.* Moscow: International University in Moscow. 463 p. ISBN 5-9248-0121-7. (In Russ).

Novikov, A. I., Yaroslavtseva, E. I. (1990). *Semantic distances in language and text.* Moscow: The science. 135 P. ISBN 5-02-010986-X. (In Russ).

Polzunova, M. V. (2008). *Explanation in love as a complex speech genre: vocabulary, grammar, pragmatics.* Author's abstract of the PhD Diss. Yekaterinburg. 24 P. (In Russ).

Revzina, O. G. (2005). Discourse and discursive formations. *Criticism and semiotics, 8:* 66—78. (In Russ).

Serio, P. (2012). Why is Bakhtin not a pedestrian? About one big misunderstanding about the approach to the analysis of discourse. *Language, communication and social environment, 10:* 8—21. (In Russ).

Stepanov, Yu. S. (1983). *Semiotics.* Moscow: Rainbow. 637 p. (In Russ).

Taboada M., Mann, W. C. (2006). Rhetorical Structure Theory: looking back and moving ahead. *Discourse Studies, 8*: 423—459. Available at: http://dis.sagepub.com/content/8/3/423.

Tenier, L. (1988). *Fundamentals of structural syntax.* Moscow: Progress. 656 p. (In Russ).

Velyugo, O. A. (2019). *The artistic embodiment of the discourse of love in the prose of Julian Barnes and Ian McEwan.* Author's abstract of PhD Diss. Minsk. 26 p. (In Russ).