

Научная статья

УДК 811.161.1

DOI 10.25205/1818-7935-2022-20-3-64-74

## К вопросу о частотностях глагольных конструкций в некоторых жанрах массовой литературы

Надежда Николаевна Буйлова<sup>1</sup>

Ольга Николаевна Ляшевская<sup>2</sup>

<sup>1</sup>Институт проблем передачи информации им. Харкевича РАН  
Москва, Россия

<sup>2</sup>Национальный исследовательский университет «Высшая школа экономики»  
Москва, Россия

Институт русского языка им. Виноградова РАН  
Москва, Россия

<sup>1</sup>[nbujlova@ruscorpora.ru](mailto:nbujlova@ruscorpora.ru), <https://orcid.org/0000-0002-0604-6603>

<sup>2</sup>[olesar@yandex.ru](mailto:olesar@yandex.ru), <https://orcid.org/0000-0001-8374-423X>

### Аннотация

Использование корпусов для изучения специфических областей языка – широко распространенная практика в современной лингвистике. В нашем исследовании мы обращаемся к такому лингвокультурологическому явлению, как массовая литература (бульварный роман). Клишированность такой литературы позволяет описывать ее при помощи так называемых литературных формул, в универсальной форме воплощающих культурные темы и стереотипы социума. В основном исследуемые культурологами и литературоведами, с лингвистической точки зрения литературные формулы изучены недостаточно. В работе мы рассматриваем четыре микрожанра<sup>1</sup> (любовный роман, детектив, научно-фантастический роман и фэнтези), которые примерно соответствуют литературным формулам «романтическая история», «тайна», «чуждые сущности и состояния» и «приключения». Поскольку низкоуровневые признаки (длина текста, длина слова, частеречные характеристики) не дают однозначного ответа относительно принадлежности текста к той или иной литературной формуле, в работе на материале корпуса текстов массовой литературы анализируются глагольные конструкции (комплексы вида «глагол + зависимые»), характерные для одного или нескольких микрожанров. Для выделения подобных комплексов используются методы машинного обучения. Конструкции, внесшие наибольший вклад в кластеризацию по жанрам, мы называем маркерными, т. е. характерными для определенного микрожанра. Конструкции рассматриваются с точки зрения заполненности валентностей глагола. В исследовании выделяются полные (словарные), неполные (с опущением субъекта, прямого или косвенного объекта) и расширенные (сирконстантные) глагольные конструкции. Проанализированы и описаны конкретные примеры реализации конструкций в каждом из микрожанров. На основе схожести конструкционного профиля сделаны выводы о преобладании в любовных романах и детективах глаголов и конструкций, сопутствующих прямой речи, и о преобладании полных конструкций в научно-фантастических романах и фэнтези. Предложены также внетекстовые объяснения подобного предпочтения. Предложенный в исследовании метод может быть использован для изучения синтаксических особенностей различного генеза: обусловленных авторской, временной или жанровой спецификой.

### Ключевые слова

массовая литература, литературная формула, формульная литература, глагольные конструкции, синтаксические маркеры

---

<sup>1</sup> В данной статье под микрожанром понимаются произведения, выделяющиеся в отдельные кластеры непосредственно внутри жанров массовой литературы, такие, как шпионский детектив, «алый» (любовный) роман, хронофантастику и т. п.

© Буйлова Н. Н., Ляшевская О. Н., 2022

## Для цитирования

Буйлова Н. Н., Ляшевская О. Н. К вопросу о частотностях глагольных конструкций в некоторых жанрах массовой литературы // Вестник НГУ. Серия: Лингвистика и межкультурная коммуникация. 2022. Т. 20. № 3. С. 64–74. DOI 10.25205/1818-7935-2022-20-3-64-74

## On Verbal Construction Frequency in Certain Genres of Popular Literature

Nadezhda N. Buylova<sup>1</sup>  
Olga. N. Lyashevskaya<sup>2</sup>

<sup>1</sup>The Institute for Information Transmission Problems RAS (Kharkevich Institute)  
Moscow, Russia

<sup>2</sup>HSE University  
Moscow, Russia

<sup>2</sup>The Russian Language Institute RAS (Vinogradov Institute)  
Moscow, Russia

<sup>1</sup>nbujlova@ruscorpora.ru, <https://orcid.org/0000-0002-0604-6603>

<sup>2</sup>olesar@yandex.ru, <https://orcid.org/0000-0001-8374-423X>

### Abstract

Corpora-based language studies is a widespread practice in modern linguistics. In our study, we address so-called mass literature (or paraliterature) via a corpus of texts. The standardization of mass literature allows us to describe its genres by applying literary formulas. In brief, a formula serves for the embodiment of cultural themes and stereotypes in a universal form. While mass literature is a common subject of literary and cultural studies, from a linguistic point of view, literary formulas have not been studied well enough. We suggest that differences between the microgenres of paraliterature may be in syntax as well as in the vocabulary. Our work is based on the mass literature corpora and provides analysis of verb constructions characteristic of microgenres (love story, detective story, science fiction novel, and fantasy). In order to identify the distinctive features of mass literature microgenres, we have conducted a series of machine learning experiments. As a dataset, we compiled a corpus of 1,200 texts belonging to four microgenres. In our previous studies we showed that statistical features (text length, sentence length, and parts of speech frequencies) were insufficient for successful classification. The usage of lexical features has improved the quality of machine learning, however, the classifier based on syntactic features has shown the best results. 81 constructions have been selected as the most important features of syntactic-based machine learning. A verb construction consists of a “verb + dependencies”. We consider several types of dependencies: arguments (subject, direct and indirect objects) and adjuncts (subordinate clause and adverbial classifier). The constructions are described in terms of the fullness of the verb valencies. Several types of constructions are distinguished: complete, incomplete (with omissions of direct or indirect complements), and extended (with adjuncts). Specific examples of verb constructions in each of the microgenres have been analyzed and described. Based on the similarity of the construction profile, romance novels and detective stories show the predominance of verbs and direct speech constructions, whereas science fiction novels and fantasy demonstrate the prevalence of full constructions. Possible non-textual explanations for our observations are also provided. The method proposed in the study can be used for the investigation of various kinds of syntactic features, for instance, those associated with the author’s, temporal or genre specificities.

### Keywords

paraliterature, formula literature, verb constructions, syntactic markers

### For citation

Buylova N. N., Lyashevskaya O. N. On Verbal Construction Frequency in Certain Genres of Popular Literature. *Vestnik NSU. Series: Linguistics and Intercultural Communication*, 2022, vol. 20, no. 3, pp. 64–74. DOI 10.25205/1818-7935-2022-20-3-64-74

## Введение

Поворот квантитативной лингвистики от исследования «чистого», строго нормированного языка классических жанров литературы к литературе «реального узуса» повлек за собой расширение объектно-предметных областей и включение в них так называемой массовой литературы, ранее рассматривавшейся в основном литературоведами и культурологами [Купина, Николина, 2009]. По мнению Д. Куэньяса, массовой литературе присущи одномерность и силуэтность персонажей, однозначность и типизированность идеологического высказывания, диктуемого жанровым канонам (как, например, в детективах), достаточно строго соблюдаемым авторами [Souégnas, 1992]. Эти базовые установки ее литературности привели затем к идее о формульности подобного литературного творчества. Эксплицитно понятие «формульная литература» (*formula stories*) было введено Дж. Г. Кавелти в работе «Изучение литературных формул» [Cawelty, 1976/1996], где формула определяется как способ, с помощью которого «конкретные культурные темы и стереотипы воплощаются в более универсальных повествовательных архетипах» [Cawelty, 1996, с. 34]. Кавелти предлагает достаточно разнородный набор формул, часть из которых объединяет схожие сюжеты, а часть описывает авторские представления о законах создания и функционирования вымышленного мира: например, наряду с формулой «романтика» у него существует формула «чуждые существа и состояния» [Cawelty, 1996, с. 35].

Среди лингвистических исследований русской формульной литературы можно выделить работы Т. Валдаевой [Валдаева, 2018], которая отмечает предельную схематичность главной героини любовного (или «розового») романа<sup>2</sup>, и Е. Ганапольской, опубликовавшей «Фразеологический словарь современного русского детектива» [Ганапольская, 2015].

В нашей работе мы обращаемся к синтаксическим особенностям формульной литературы, выбрав в качестве основного объекта изучения поверхностный синтаксис. Реализация поверхностного синтаксиса описывается теорией конструкций глагола, изложенной в работах по лексической семантике и синтаксису [Tesniere, 1953; Апресян, 1967; Падучева, 2004; Кустова, 2004] и грамматике конструкций [Fillmore, 1976; Goldberg, 1995; Stefanowitsch, Gries, 2003]. Нами поставлена задача выделить специфические синтаксические маркеры следующих четырех разновидностей массового романа: любовных («розовых») романов, женских детективов, космической научной фантастики и так называемого попаданческого<sup>3</sup> фэнтези. По Кавелти, эти жанры<sup>4</sup> примерно соотносятся со следующими литературными формулами: романтическая история, тайна, чуждые сущности и состояния, приключения.

Материалом послужил составленный авторами корпус из 1 200 произведений общим объемом около 100 млн словоупотреблений. Наша гипотеза состоит в том, что культурная вторичность этих произведений, связанная со следованием некоему сюжетному шаблону<sup>5</sup>, со стандартными ожиданиями читателей, и, следовательно, определенная клишированность языка позволит выявить не только тематическое своеобразие лексики, но и менее очевидные особенности авторской риторики в синтаксическом построении предложений.

В частности, нами исследуются конструкционные особенности глаголов, представляющих организующее начало фразы.

<sup>2</sup> Канон этого типа литературности не допускает повествования от первого лица, и, следовательно, в качестве его негативного грамматического маркера можно рассматривать отсутствие глаголов в этой форме.

<sup>3</sup> «Попаданческая» литература – поджанр фантастической или фэнтезийной литературы, в котором главный герой внезапно переносится в прошлое, будущее, на другую планету, в параллельный мир или в мир художественного произведения.

<sup>4</sup> Они иногда называются микрожанрами.

<sup>5</sup> Можно говорить, например, о единой сюжетно-событийной «морфологии» детектива в том смысле, в каком Пропп [1929] описал единую морфологию сказки.

## 1. Метод выделения конструкций-маркеров

Изучение конструкций глагола с точки зрения реализуемых им валентностей и их структурных ограничений позволяет делать выводы как о специфичности взаимодействия синтаксиса и семантики на уровне предложения, так и об организации текста в целом. Под конструкцией мы понимаем глагол и его непосредственные синтаксические зависимые. В ее состав могут входить как обязательные зависимые, соотносимые с ядерными актантами семантического фрейма, так и факультативные (сирконстанты). Изучение свойств глагольной конструкции строится на статистических корпусных данных с привлечением многоуровневой лингвистической разметки.

Использованный в работе корпус русскоязычной литературы массовых жанров был размечен лексически и синтаксически в соответствии с моделью синтаксиса зависимых Universal Dependencies [Буйлова, Ляшевская, 2021]. В итоге было выделено более четырех миллионов глагольных конструкций типа *<убрать nsubj obj obl>*<sup>6,7</sup>. При помощи метода машинного обучения «Случайный лес» для классификации текстов корпуса по формульным жанрам были выделены маркерные конструкции, характерные для одного, двух или трех микрожанров (подробное описание метода см. [Буйлова, Ляшевская 2021], список конструкций см. Github<sup>8</sup>).

В итоговую выборку вошла 81 маркерная конструкция, однако пять из них мы были вынуждены исключить из-за ошибок разметки (например, смешения прямого и косвенного дополнения у глагола *производить* или инфинитивной клаузы и косвенного дополнения у глагола *становиться*). Оставшиеся 76 конструкций принадлежат 74 глаголам: глаголы *касаться* и *хотеться* образуют по две различные маркерные конструкции: *<касаться obl>* и *<касаться nsubj obl>*, *<хотеться obl xcomp>* и *<хотеться nsubj xcomp>* для любовных романов и детективов соответственно.

## 2. Типы маркерных конструкций

Изучая полноту реализации валентностного потенциала глаголов среди маркерных конструкций литературных формул, мы выделили полные (с реализацией всех обязательных валентностей) и неполные (с опущением одной или нескольких из них) конструкции. Последние в свою очередь можно разделить по способу (не)выражения синтаксического актанта на бессубъектные (с опущением подлежащего) и безобъектные (с опущением прямого объекта и конструкции с конкуренцией переходной и непереходной реализации). Кроме того, были проанализированы расширенные конструкции, в которых, помимо обязательных актантов, выражается вовлеченность сирконстантных участников. Далее будут обобщены свойства этих групп маркерных конструкций во всех четырех жанрах.

### 2.1. Полные конструкции

Полные конструкции (т. е. совпадающие со словарной формулой вида «*видеть кого-либо <obj> где-либо <obl>*») более характерны для формул фэнтези и научной фантастики. Это может быть связано с их риторической спецификой, проявляющейся в большей описательности, а также с большей средней длиной предложений (СДП) в этих жанрах. В описанном ниже примере глагол *поднимать* из маркерной конструкции *<поднимать nsubj obj obl>* имеет три заполненные валентности: подлежащее *кто-то*, прямое дополнение *их* и не прямое дополнение *по очереди*.

<sup>6</sup> Здесь и далее используются следующие сокращения: *nsubj* – подлежащее, *obj* – прямое дополнение, *obl* – косвенное дополнение, *xcomp* – инфинитивная клауза, *advcl* – адverbальный модификатор.

<sup>7</sup> Эта конструкция описывает предложение *Он (nsubj) убрал тарелку(obj) со стола (obl)*.

<sup>8</sup> <https://github.com/nadin4123/RussianFormulaStories>.

FT<sup>9</sup> *Как будто кто-то поднимал их по очереди и тут же ставил на место.*

Отдельную группу полных конструкций составляют конструкции глаголов движения, физиологии, ментальных и бытийных глаголов, не допускающих прямого дополнения в принципе:

FT *Воронка направляется ко мне.*

Интересно отметить, что все собственно непереходные глаголы, конструкции которых стали маркерными, выступают только в роли финитной вершины клаузы с выраженным субъектом. Любовные романы лидируют по количеству таких конструкций (19 конструкций-маркеров любовных романов из 50 маркерных конструкций с непереходными глаголами).

LS *Губы покраснели и распухли от его поцелуев, грудь вздымалась и опадала с каждым вдохом и выдохом.*

## 2.2. Неполные конструкции

Неполные конструкции в целом чаще встречаются в любовных романах и детективах. Далее мы проанализируем их свойства, разделив по типу незаполненной валентности.

### 2.2.1. Бессубъектные конструкции

Бессубъектные конструкции присущи в большей степени любовному роману, детективу и фэнтези. Основная их часть приходится на употребление глагола в составе деепричастного оборота или сочиненной группы. Опускание актора – серьезная модификация глагольной конструкции, которая может свидетельствовать о сравнительной сложности синтаксиса: как минимум, в таких предложениях несколько глаголов, и поэтому первый участник часто синтаксически контролируется другим предикатом:

LS *Мой партнер, не убирая руки с моей талии, нагнулся к моей кисти и поцеловал.*

DS – *Это Рип, – сказал я, убирая руку с рукоятки «зиг-зауэра».*

FT – *И кто у нас тут самый глазастый, – буркнула лучница, убирая стрелу в колчан.*

### 2.2.2. Безобъектные конструкции

В отличие от субъекта, объект опускается гораздо реже. Соответствующая неполная конструкция практически не наблюдается среди конструкций-маркеров формульной литературы. Исключением стала конструкция глагола *глотать/глотнуть* (маркер детектива). Она представляет случай null instantiation – имплицитной передачи информации об обязательном участнике, заданной фреймом глагола либо восстанавливаемой из контекста [Fillmore, 1986]. Так, если *глотают (сглатывают) судорожно*, то, скорее, слюну, а если кто-то *глотнул еще* – скорее всего, подразумевается алкогольный напиток.

SF *Олег снова глотнул.*

#### 2.2.2.3. Конкуренция за валентности как объяснение опущения прямого дополнения

Невыраженность прямого объекта у переходных глаголов может объясняться конкуренцией разных способов обозначения второго семантического актанта. В первую очередь, это конкуренция конструкции с прямым объектом и конструкции с комплементарной клаузой, ср. *боялся насмешек* и *боялся показаться смешным / что его дядя-лесоруб посчитает племянника плаксой*. В русском у глагола прямой объект и комплемент крайне редко относятся к разным семантическим актантам или выражаются одновременно, чаще это два разных способа выраже-

<sup>9</sup> Здесь и далее используются следующие сокращения: LS – любовные романы, DS – детективные романы, SF – научная фантастика, FT – фэнтези.



ния одного участника. Так, прямой объект может сжато соответствовать пропозиции, которую в развернутом виде выражал бы комплемент (*решить вопрос*).

В наших данных прямое дополнение может быть опущено в трех случаях, из которых самый частый – прямая речь. В эту группу входят глаголы речи *буркнуть, говорить, вскрикнуть, отвечать*. Отсутствие объекта в этом случае объяснимо, несмотря на то, что в эту группу попали как переходные, так и непереходные глаголы: тема разговора была вынесена в реплику. Такие конструкции характерны для любовных романов и детективов, в которых персонажи много говорят (спрашивают, рассказывают, отвечают), в результате чего в тексте доминирует диалог как форма коммуникации:

DS – *Ничего особенного, уже уезжаем, – буркнул Родионов.*

LS – *Доминик! – вскрикнула она, когда он крепко схватил ее за запястье.*

Другой тип опущения дополнения представлен конструкциями с вытеснением объекта на синтаксическую периферию. В таких конструкциях объект существует, но, пользуясь терминами динамического представления глагольных диатез, имеет низкий коммуникативный ранг [Падучева, 2004], т. е. становится синтаксической периферией (маркерной конструкцией становится *рассказывать о происшествии / говорить о делах, но не рассказывать сказку / говорить речь*). Большая часть этих конструкций принадлежит глаголам речи, широко представленным в русскоязычной художественной литературе: валентность на содержание высказывания (уже выраженная прямой речью) вытесняется валентностью на тему разговора.

FT – *Он не упоминал о том, где он собирался там жить?*

Третий тип представлен конкуренцией моделей управления *хотр* и прямого дополнения. Теоретически количество зависимых глагола не ограничено, однако в массовой литературе глагол редко имеет более трех заполненных валентностей. Подобное ограничение приводит к вытеснению прямого дополнения комплементарной клаузой. Большая часть этой группы маркирует фэнтези.

FT *А то ждать заставляешь.*

FT *Ледяные струи дождя больно хлещут по спине, заставляя еще ниже пригнуться к седлу и спрятать лицо.*

#### 2.2.4. Опущение других зависимых

Чаще всего опускаются зависимые, связанные с локацией или состоянием (для глаголов движения или трансформации и бытийных глаголов), и зависимые, связанные с участниками и содержанием коммуникации (для глаголов речи). В маркерных конструкциях глаголов движения остается наиболее значимая валентность: «конечная точка», «начальная точка» либо «место». Кроме того, такие глаголы часто используются в переносных значениях и фразеологизмах:

FT *Уверена, теперь он оборачивается на каждую лошадь на улице.* (конечная точка)

SF *Астроному в школе проходил?* (место)

LS *Представления не имею, как можно весь день проходить на высоких каблуках.* (образ действия)

DS *Так, изредка доходили слухи...* (имплицитная конечная точка)

Схожее поведение демонстрируют бытийные глаголы и глаголы трансформации.

LS *Но Джейн много рассказывает мне об Америке, а я помогаю ей решать проблемы, возникающие у нее в Англии.* (место)

DS *А сами стремитесь к тому же, хотя постоянно оказываетесь в центре внимания.* (место)

DS *То, что повторяется изо дня в день, превращается в обыденность.* (конечное состояние)

LS *Молчание становилось невыносимым.* (конечное состояние)

Что касается глаголов речи, то им присущи валентности с семантикой «адресат», «тема» и «содержание высказывания», которые могут выражаться при помощи косвенного объекта. В таком случае можно было бы ожидать использование конструкции с тремя косвенными объектами, но в маркерных конструкциях их зачастую опускают. В большинстве случаев это объясняется тем, что семантические актанты «адресат» и «тема» уже выражены в реплике персонажа и не дублируются в комментариях к прямой речи, содержащей целевой глагол:

DS *Он просит, чтобы я встретился с ним с глазу на глаз.*

LS – *Отец, может, мы обсудим это позже, наедине? – сдерживаясь из последних сил, попросил Бромвелл.*

Также могут опускаться зависимые с семантикой «способ действия» (для *звонить*: *по телефону* или *в дверь*); «точка контакта» (для *подхватывать*: *на руки, под локоток*); «стимул» (*смеяться*: *над кем* или *от чего*, *краснеть*, *розоветь*: *от удовольствия*); «содержание пропозиции» (*подозревать в чем*). Отметим, однако, что системной картины статистика этих примеров не образует и служить надежными маркерами жанра они не могут.

### 2.3. Конструкции с сирконстантами

Расширенные конструкции предполагают выражение сирконстантного участника. В наших примерах он обозначается чаще всего как *obl* или *advcl*, однако необходимо иметь в виду, что у глаголов *просить* и *бояться* *advcl* кодирует и *содержание* (в придаточной клаузе с союзами *чтобы* и *если*), и *время*, и *цель*, и прочую категориальную семантику, которая может быть выражена непосредственно в реплике:

DS – *И главное: взломайте компьютер, там могут быть важные сведения! – попросила я, особенно не веря в удачу, и торопливо поведала все, о чем успела узнать.*

Аналогичным образом у глагола *вскрикнуть* обстоятельственная клауза выражает причину эмоциональной реакции (обязательный участник) или *состояние, образ действия героя* (сирконстант):

LS *Решиив, что это Линвуд, Венеция негромко вскрикнула, но потом узнала его.*

В заключение обратимся к трем конструкциям, в которых участник *advcl* всегда является сирконстантным: *<бродить nsubj obl advcl>*, *<подниматься nsubj obl advcl>*, *<стать nsubj xcomp advcl>*.

Эти глаголы становятся своеобразными входными точками для читателя, позволяя ему подключиться к эмоциональным переживаниям и интерпретировать действия героев. Автор формульной литературы вынужден балансировать между раскрытием описательной, эмоциональной или интеллектуальной составляющей, не слишком усложняя при этом восприятие произведения. Это приводит к многозначности глаголов, которые могут использоваться в переносном значении. Зачастую упрощение достигается следующим способом: физический глагол перемещается в синтаксический центр предложения, а глагол, обеспечивающий читательское «эмоциональное подключение», помещается в зависимую клаузу, что упрощает создание эффекта эмоциональной вовлеченности через механизм *embodiment* (воплощенность происходящего в физическом опыте).

DS *Загрузив провизию в машину, я стала придумывать меню на вечер.*

LS *Наверху Рэн бродила по комнате, заламывая руки и вытирая бегущие слезы.*

FT *Сердце словно вырвали у нее из груди, волна слез поднималась к глазам, грозя затопить печалью.*

### 3. Сравнительная характеристика микрожанров

Для дальнейшего обсуждения сходств и различий между четырьмя вышеперечисленными жанрами необходимо вспомнить отличительные черты, установленные ранее в парадигме компьютерной лингвистики для каждого из них в [Буйлова, Ляшевская, 2021]. Так, оказалось,

что они различаются не только по параметру *лексический состав* жанрового корпуса, что вполне очевидно и естественно объясняется предметно-тематической спецификой DS, LS, FT и SF, но и по параметрам *длина предложений в токенах*<sup>10</sup> и *длина текстов в токенах* (средние значения составляли 11 982 токена для DS, 13 170 – для LS, 21 999 – для FT, 22 495 – для SF). В нашем корпусе средняя длина произведения в жанре SF больше, чем в трех остальных, а в LS самая короткая. Средняя длина слова в фэнтези на две буквы длиннее<sup>11</sup>. В то же время частеречные характеристики четырех подкорпусов остаются практически идентичными [Bujlova, 2018, с. 109]. Возможно (и вероятнее всего), такие расхождения в базовой лингвостатистике вызваны не языковыми, а собственно литературными факторами, связанными с канонами жанров, с различиями в «технике и технологии» создания текста произведений.<sup>12</sup> Специфика научно-фантастических романов предполагает, к примеру, более развернутые объяснения, с одной стороны, и определенную сериальность, заключающуюся в тяготении к многотомным сагам, с другой. Любовные романы ориентированы на формат покетбуков, предполагающий определенный тип «потребления», и по размеру они не столько романы, сколько повести, что отражено в английском названии жанра словом *story*.

Стоит также отметить, что любовные романы и детективы, в большей степени ориентированные на описание внутреннего мира героев (интеллектуального в DS и эмоционального в LS), используют больше глаголов речи (таких как *буркнуть*, *сказать*, *вскрикнуть* и т. д.), чем фэнтези и научная фантастика для описания взаимодействия персонажей или передачи внутреннего монолога. Также ожидаемо у любовных романов выделяется класс глаголов, описывающих эмоциональные состояния персонажей (*розоветь*, *краснеть*), как маркер изменения цвета кожи вследствие внутренних переживаний.

Однако, как было нами показано, не все жанровые различия могут быть одинаково эффективно использованы для классификации текстов, что говорит о небольшом вкладе некоторых параметров и связанных с ними количественных показателей в определение литературно-языковой специфичности жанра. Так, низкоуровневые признаки (частотность различных частей речи, длина текстов в словах, и особенно в предложениях) не позволяют однозначно различить жанры. При этом использование глаголов с высоким коэффициентом логарифмического подобию существенно улучшает результаты классификатора, хотя наилучшие результаты он демонстрирует на наборе данных с включением синтаксических признаков [Bujlova, 2018].

При сравнении глагольных конструкций-маркеров важно отметить, что некоторые из них могут являться отличительным признаком более чем одного жанра. Поэтому на данный момент следует говорить не только о различиях между жанрами, но и о сходстве литературных приемов, используемых в них. Потому наши наблюдения носят, скорее, характер гипотезы, нежели однозначных утверждений; их необходимо тестировать на иных наборах данных.

Таким образом, мы видим, что для научной фантастики и фэнтези более характерны полные конструкции и что фэнтези характеризуется вытеснением прямого дополнения компонентной клаузой. С точки зрения литературной технологии и характера повествования это может быть объяснено, с одной стороны, тяготением этих двух жанров к некоторой клишированности в конструкциях (т. е. стремлением воспроизводить наиболее привычные из них), а также, с другой стороны, к построению описательных предложений большей длины, поскольку к особенностям текстов о вымышленных мирах можно отнести требование давать большую экспозицию.

Если сравнивать прогностический потенциал неполных конструкций, то для фантастики, например, наиболее характерны бессубъектные, что противопоставляет ее трем другим жан-

<sup>10</sup> Средняя длина предложения в токенах составляла для LS – 14, для DS – 13, для SF – 14, для FT – 15 [Bujlova, 2018, с. 109].

<sup>11</sup> Средняя длина слова в символах для LS – 5, DS – 5, SF – 14, FT – 7 [Bujlova, 2018, с. 109].

<sup>12</sup> По крайней мере, у нас пока нет оснований утверждать, что такое варьирование связано исключительно с субъективной авторской стилистикой, хотя этот фактор требует дополнительных исследований.



рам. Однако это также может свидетельствовать о сравнительной простоте синтаксиса массовой научной фантастики.

Вместе с тем выяснилось, что синтаксические характеристики детективов и любовных романов близки. Как мы уже упоминали выше, это может объясняться тем, что действие обоих происходит в «более реальном» мире, отражающем процессы, происходящие в реальном социуме с определенным местом и временем действия. Можно сказать, что в противовес подчеркнуто воображаемым мирам фэнтези и фантастики, оторванным (за редкими исключениями) от конкретного хронотопа, они связаны с ментальностью обычных людей, их жизненным опытом, и потому являются «более разговорными». Они «строятся и держатся» на диалогах: герои либо выясняют отношения (в любовных романах), либо (д)опрашивают друг друга (в детективах). Этот прагматизм жанра проявляется как в лексическом составе маркеров (глаголы речи чаще становятся маркерами именно этих жанров), так и в конструкционном: глаголы речи становятся маркерными в ремарках к репликам. Именно введение прямой речи обуславливает изменение конструкции: смещение актантов на синтаксическую периферию и превращение их в сирконстанты.

### Заключение

В заключение стоит сказать, что описанные конструкции-маркеры могут быть очень полезны в качестве признаков машинного обучения. Рассмотрение таких конструкций, позволяющих определять с большей или меньшей достоверностью жанр текста, выявило несомненное наличие синтаксических особенностей, характерных для литературных формул. В частности, можно считать доказанным, что детективы и любовные романы демонстрируют тяготение к употреблению конструкций с глаголом речи и, за счет прямой речи, смещению объекта на синтаксическую периферию, тогда как для научной фантастики и фэнтези характерны синтаксические паттерны описания (полные либо бессубъектные конструкции).

Это свидетельствует о том, что литературная формула накладывает определенные ограничения на синтаксические конструкции, подтверждая связь синтаксиса и семантики, однако взаимовлияние этих уровней построения текста должно быть исследовано более подробно. Кроме того, включение в анализ менее типизированных текстов позволит выделить не только жанровые особенности, но и авторский стиль, а также влияние текущего литературного окружения.

Поскольку микрожанр накладывает определенные ограничения на выбор синтаксических конструкций, использованный нами метод изучения текстов демонстрирует перспективность исследований взаимодействия семантики и синтаксиса на материале корпуса.

### Список литературы

- Апресян Ю. Д.** Лексические конверсивы в русском языке // *Linguistica Silesiana*, 1975. Т. 1. С. 256–283.
- Апресян Ю. Д.** Экспериментальное исследование семантики русского глагола. М.: Наука, 1967.
- Буйлова Н. Н., Ляшевская О. Н.** Лексико-синтаксические маркеры малых литературных жанров // *Вестник ПСТГУ. Серия III: Филология*, 2021. Вып. 66. С. 11–23.
- Валдаева Т. Н.** Жив ли розовый роман: читательская среда и популярный жанр // *Культурные инициативы: материалы 50 Всероссийской с международным участием научной конференции молодых исследователей*, 2018. С. 111–114.
- Гананольская Е. В.** Фразеологический словарь современного российского детектива: в двух томах. СПб.: Златоуст, 2015–2016.
- Купина Н. А., Литовская М. А., Николина Н. А.** Массовая литература сегодня: учеб. пособие. М.: Флинта, 2009.

- Кустова Г. И.** Типы производных значений и механизмы языкового расширения. М.: Языки славянской культуры, 2004.
- Падучева Е. В.** Динамические модели в семантике лексики. М.: Языки славянской культуры, 2004.
- Рахилина Е. В.** Лингвистика конструкций, М.: Азбуковник, 2010.
- Byjlova N.** Amateur Prose on the Web: Verb Construction as a Feature of Genre Classification. In: Proc. of ARANEA 2018. Web Corpora as a Language Training Tool. Slovakia, Bratislava, 2018.
- Cawelty J. G.** Adventure, Mystery and Romance. Chicago: University of Chicago Press, 1976.
- Couégnas D.** Introduction à la paralittérature. P.: Seuil, 1992.
- Fillmore Ch.** Frame Semantics and the Nature of Language // Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech. Berkeley, NY, USA, 1976. Pp. 20–32.
- Fillmore Ch.** Pragmatically Controlled Zero Anaphora // Proc. of the 12th Annual Meeting of the Berkeley Linguistics Society. Berkeley, NY, USA, 1986. Pp. 95–107.
- Goldberg A.** Constructions: A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press, 1995.
- Stefanowitsch A., Gries S. T.** Collocations: Investigating the Interaction of Words and Constructions // Int. J. of Corpus Linguistics, 2003. No. 8(2). Pp. 209–243.
- Tesnière L.** Esquisse d'une syntaxe structurale. P.: Klincksieck, 1953.

#### References

- Apresyan, Yu. D.** Lexical conversion in Russian. *Linguistica Silesiana*, 1975, 1, pp. 256–283. (in Russ.).
- Apresyan, Yu. D.** Experimental Study of the Semantics of the Russian Verb. Moscow. Nauka Publ, 1967, 251 p. (in Russ.)
- Bujlova, N. N., Lyashevskaya, O. N.** Lexical-Syntactic Markers of Minor Literary Genres. *Vestnik PSTGU*, 2021, 66, pp. 11–23. (in Russ.)
- Byjlova, N.** Amateur Prose on the Web: Verb Construction as a Feature of Genre Classification. In: Proc. of ARANEA 2018. Web Corpora as a Language Training Tool. Slovakia, Bratislava, 2018.
- Cawelty, J. G.** Adventure, Mystery and Romance. Chicago: University of Chicago Press, 1976.
- Couégnas, D.** Introduction à la paralittérature. P.: Seuil, 1992.
- Fillmore, Ch.** Pragmatically Controlled Zero Anaphora. In: Proc. of the 12th Annual Meeting of the Berkeley Linguistics Society. Berkeley, NY, USA, 1986, pp. 95–107.
- Fillmore, Ch.** Frame Semantics and the Nature of Language. In: Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech. Berkeley, NY, USA, 1976, pp. 20–32.
- Ganapol'skaya, E. V.** Phraseological Dictionary of Modern Russian Detective Literature: in 2 vol. Saint-Petersburg, Zlatoust Publ, 2015–2016, 256 p. (in Russ.)
- Goldberg A.** Constructions: A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press, 1995.
- Kupina, N. A., Litovskaya, M. A., Nikolina, N. A.** Mass Literature Today: A Tutorial. Moscow, Flinta: Nauka Publ, 2009, 423 p. (in Russ.)
- Kustova, G. I.** Types of Derivative Meanings and Mechanisms of Language Development. Moscow, Yazyki rossiskoi kultury Publ, 2004, 471 p. (in Russ.)
- Paducheva, E. V.** Dynamic Models in Vocabulary Semantics. Moscow, Yazyki rossiskoi kultury Publ, 2004, 607 p. (in Russ.)
- Rahilina, E.** Linguistics of Constructions. M.: Azbukovnik, 2010.

- Stefanowitsch, A., Gries, S. T.** Collostructions: Investigating the Interaction of Words and Constructions. *Inter. Journal of Corpus Linguistics*, 2003, no. 8(2), pp. 209–243.
- Tesnière, L.** Esquisse d'une syntaxe structurale. P.: Klincksieck, 1953.
- Valdaeva, T. N.** Is the Pink Romance Alive: Reading Environment and Popular Genre. In: *Kul'turnye iniciativy: materialy 50 Vserossijskoj s mezhdunarodnym uchastiem nauchnoj konferencii molodyh issledovatelej*, 2018, pp. 111–114. (in Russ.)

#### **Информация об авторах**

**Буйлова Надежда Николаевна**, младший научный сотрудник  
**Ляшевская Ольга Николаевна**, кандидат филологических наук, профессор

#### **Information about the Authors**

**Nadezhda N. Bujlova**, junior researcher  
**Olga N. Lyashevskaya**, Doctor of Philosophy, Professor

*Статья поступила в редакцию 14.04.2022;  
одобрена после рецензирования 05.07.2022; принята к публикации 27.07.2022  
The article was submitted 21.04.2022; approved after reviewing 05.07.2022; accepted for publication 27.07.2022*