




<https://doi.org/10.22363/2687-0088-27394>

Research article / Научная статья

Aspectual pairs: Prefix vs. suffix way of formation

Valery SOLOVYEV  , Vladimir BOCHKAREV 
and Venera BAYRASHEVA 

Kazan Federal University, Kazan, Russia

 maki.solovyev@mail.ru

Abstract

In linguistic theory, there is no common point of view on the question of whether verbs in aspectual pairs are in inflectional or derivational relations. At the same time, the prefix and suffix methods of forming aspectual pairs are contrasted in this respect. The publications (e.g. Janda & Lyashevskaya 2011) pointed out the need to develop new quantitative approaches to this aspect of the text corpus. We propose two new approaches that compare the quantitative characteristics of aspectual pairs of both types. One approach is based on the Google Books Ngram corpus and analyzes the dynamics of the frequency of the use of words in pairs. The aspectual pairs from the databases created by Janda and Lyashevskaya are considered. For a numerical assessment of the degree of proximity of the frequency graphs, the Pearson correlation coefficients were used. The second approach introduces a numerical characteristic of the semantic proximity of verbs in pairs using modern computer methods. Semantic proximity of verbs is calculated as a standard cosine measure between vectors representing the compatibility of the considered verbs in the corpus. Several computer models and text corpora are considered. Both proposed approaches did not reveal significant numerical differences in semantic proximity between verbs in aspectual pairs with prefix and suffix pairing. This is in good agreement with the results of an early study by Janda and Lyashevskaya (2011). Together with the results of this work, our research shows that the suffixal and affixal ways of forming aspectual pairs have an equal status in terms of their classification as inflectional or derivational.

Keywords: *aspectual pairs, inflection, word formation, quantitative analysis, corpora, Russian*




For citation:

Solovyev, Valery, Vladimir Bochkarev & Venera Bayrasheva. 2022. Aspectual pairs: Prefix vs. suffix way of formation. *Russian Journal of Linguistics* 26 (4). 1114–1135. <https://doi.org/10.22363/2687-0088-27394>

**Аспектуальные пары:
префиксальный vs. суффиксальный способ образования**

В.Д. СОЛОВЬЕВ  , **В.В. БОЧКАРЕВ** , **В.Р. БАЙРАШЕВА** 

Казанский федеральный университет, Казань, Россия

 maki.solovyev@mail.ru

Аннотация

В литературе нет общей точки зрения по вопросу, находятся ли глаголы в аспектуальных парах в словоизменительных или словообразовательных отношениях, при этом префиксальный и суффиксальный способы образования аспектуальных пар часто противопоставляются в этом отношении. В публикациях (например, Janda & Lyashevskaya 2011) указывалось на необходимость развития в этом вопросе новых квантитативных подходов на корпусах текстов. В данной статье мы предлагаем два новых подхода, сопоставляющих квантитативные характеристики аспектуальных пар русских глаголов обоих типов. Один подход основан на корпусе Google Books Ngram и анализирует динамику частот употребления слов в парах. Рассматриваются аспектуальные пары из баз данных, созданных Л. Яндой и О. Ляшевской. Для численной оценки степени близости графиков частот использованы коэффициенты корреляции Пирсона. Второй подход вводит числовую характеристику семантической близости глаголов в парах с использованием современных компьютерных методов. Семантическая близость глаголов рассчитывается как стандартная косинусная мера между векторами, репрезентирующими сочетаемость рассматриваемых глаголов в корпусах. Рассмотрено несколько компьютерных моделей и корпусов текстов. Ни один из предложенных подходов не выявил существенных численных различий в степени семантической близости между глаголами в аспектуальных парах при префиксальном и суффиксальном способах образования пар. Это находится в хорошем согласии с результатами раннего исследования Л. Янды и О. Ляшевской (Janda & Lyashevskaya 2011). Совместно с результатами этой работы наши исследования показывают, что суффиксальный и аффиксальный способы образования аспектуальных пар имеют равный статус с точки зрения их классификации как словоизменительных или словообразовательных.

Ключевые слова: *аспектуальные пары, словоизменение, словообразование, квантитативный анализ, корпуса, русский язык*

Для цитирования:

Соловьев В.Д., Боцкарев В.В., Байрашева В.Р. Аспектуальные пары: префиксальный vs. суффиксальный способ образования. *Russian Journal of Linguistics*. 2022. Т. 26. № 4. С. 1114–1135. <https://doi.org/10.22363/2687-0088-27394>

1. Введение

Русский вид (в зарубежной терминологии «аспект») является одной из наиболее сложных грамматических категорий, вызывающей много споров среди лингвистов и трудной для усвоения иностранцами. Значительный вклад

в формализацию и понимание природы этого понятия внесен И.А. Мельчуком в фундаментальной монографии “Курс общей морфологии” в 4-х томах. В настоящей работе мы опираемся на изложенные в ней представления.

Несмотря на длительную историю изучения русского вида остается неясным, как следует его трактовать с точки зрения грамматического статуса – находятся ли глаголы в аспектуальных парах в словоизменительных или словообразовательных отношениях. Основные точки зрения следующие.

1. Вид – словоизменительная категория (это классический подход В.В. Виноградова (Виноградов 1972, Мельчук 1997, Тихонов 1998).

2. Вид – словоклассифицирующая категория (Русская грамматика 1980, Петрухина 2000, Janda et al. 2013, Плунгян 2000).

3. Вид – словообразовательная категория (Всеволодова 1997, Filip 1999).

4. Префиксальное образование глаголов совершенного вида (перфективов) от базовых глаголов несовершенного вида (имперфективов) – это словообразование, а суффиксальное образование вторичных имперфективов от перфективов – словоизменение (пожалуй, наиболее часто представляемая точка зрения (Маслов 1978/2004, 1984, Бондарко 1971, Исаченко 1954/2003, Горбова 2017)).

Разумеется, эти работы различаются нюансами трактовки перечисленных понятий.

Таким образом, эти два способа видовой трансформации тесно связаны с грамматическим статусом аспектуальных пар. Отметим, что видовые пары образуются разным способом. Перфективы образуются из базовых имперфективов¹ с помощью приставок², а вторичные имперфективы образуются из перфективов добавлением суффиксов). Целью статьи является проверка квантитативными методами последней из перечисленных 4-х точек зрения. В последнее десятилетие проблема, является ли русский вид словоизменительной или словообразовательной категорией, активно исследовалась в работах Е.В. Горбовой и В.С. Храковского (Горбова 2014, 2015а, 2015б, 2017, Храковский 2015, 2018), однако к консенсусу они так и не пришли. Проблематика русского вида подробно освещена в монографии (Зализняк и др. 2015). По рассматриваемому вопросу точка зрения, представленная в этой монографии, следующая: «вопрос вообще не имеет однозначного решения... что отражает реально двойственную природу категории вида в русском языке» (с. 20).

К сожалению, общепризнанного формального (операционного) критерия разграничения словообразования и словоизменения тоже нет. В различных работах (Dressler 1989, Plank 1991, Перцов 1996, Мельчук 1997) обсуждались разные признаки этих категорий. Авторы работы (Singh 1980) вообще настаивают на отмене этого разграничения. В (Мельчук 1997) предлагается ввести промежуточную категорию квазиграммем. Все же есть общее интуитивное понимание различия между ними, хотя бы в прототипических случаях: при

¹ В последнее время используется также термин ‘симплексы’ (Gorbova 2022).

² Другое название – префигированные (Gorbova & Chuikova 2022).

словоизменении само слово не меняется, его лексическая семантика сохраняется, меняется лишь его форма, а при словообразовании меняется слово, т.е. его значение. Эта позиция представлена и в формализованном определении в (Мельчук 1997). Таким образом, мы приходим к вопросу о том, меняется ли семантика слова (кроме чисто видовой) при префиксальной и суффиксальной трансформации.

Однако, как известно, семантика не формализуема или крайне сложно формализуема. Предложенная А. Вежбицкой теория семантических примитивов (Вежбицкая 1999) к данной задаче плохо применима, прежде всего, в силу крайней сложности формул представления семантики, а также в силу сохраняющегося субъективизма в построении таких формул. Противоречивая оценка многих пар глаголов как аспектуальных или нет в работах (Зализняк, Микаэлян 2012, Кузнецова, Янда 2013) показывает, что интуиция даже ведущих лингвистов-аспектологов недостаточна для установления семантического тождества. Это приводит к сделанному (Janda & Lyashevskaya 2011: 201) выводу, что “At the level of intuitive analysis, it has not been possible to resolve this debate” и, следовательно, для решения этих проблем нужно применять статистические методы к корпусным данным.

Итак, исследовательский вопрос, решаемый в данной статье: в какой степени меняется семантика слов при суффиксальном и префиксальном способах образования аспектуальных пар? Проверяемая гипотеза: семантическая близость между глаголами в аспектуальных парах перфектив – вторичный имперфектив будет больше, чем между глаголами в парах базовый имперфектив – перфектив. В статье мы ограничиваемся рассмотрением основных способов образования аспектуальных пар: приставочного для образования перфектива и суффиксального для образования вторичного имперфектива.

В данной статье мы предлагаем два подхода к получению объективной численной оценки степени семантической близости слов, составляющих аспектуальные пары. Оба подхода применены к аспектуальным парам, получаемым префиксацией и суффиксацией. В итоге мы не нашли заметных количественных различий между этими двумя способами, которые позволили бы их четко дифференцировать как словообразовательные или словоизменительные. Это хорошо согласуется с работой (Janda & Lyashevskaya 2011), в которой также не найдено таких различий при применении еще одного количественного метода. В этой работе сравнивались грамматические профили – распределения форм глагола в Национальном корпусе русского языка – и был получен следующий результат: “we do not find reportable differences between the grammatical profiles of aspectual partners formed with prefixes as opposed to suffixes” (Janda & Lyashevskaya 2011: 213).

В этой статье мы используем предложенную в (Janda et al. 2013) терминологию: перфективы в аспектуальных парах с базовыми имперфективами будем называть естественными, а иные перфективы – специализированными. Например, *прочитать* – естественный перфектив, составляющий аспектуальную пару с *читать*, а *дочитать* – специализированный перфектив.

2. Теоретические предпосылки

В этом разделе мы кратко представим необходимый теоретический материал. В соответствии с заявленными целями уточним принципиально важные для данного исследования понятия словообразования и словоизменения. Мы отталкиваемся от классического труда И.А. Мельчука «Курс общей морфологии» в 4-х томах. Хотя эти понятия и хорошо известны, все же И.А. Мельчук в своей монографии подробно их рассматривает. Это связано, в частности, с тем, что между словоизменением и словообразованием нет четкой границы, эти понятия, скорее, градуированные.

Словоизменение и словообразование уместно рассматривать с двух точек зрения: семантико-лексикографической и формальной. Под семантико-лексикографической понимается вопрос: являются ли два слова, находящихся в словоизменительных/словообразовательных отношениях, вариантами одной лексемы или это две разных лексемы. Это определяет в том числе и представление слов в словарях – должны ли они описываться в одной словарной статье или в двух разных.

В (Мельчук 1997) автор исходит из следующих постулатов: «Дериватемы различают разные лексемы» (с. 274), где под дериватемой понимается выражаемое морфологическими средствами словообразовательное значение. В этой монографии используется также понятие граммема как значения, принадлежащего к словоизменительной категории. При этом постулируется (с. 258): «граммемы различают **формы** “одного и того же слова”» (выделение И.А. Мельчука). Таким образом, по Мельчуку, при словоизменении мы остаемся в пределах «одного и того же слова», т.е. лексическая семантика не меняется. А при словообразовании получаются две разные лексемы с разной семантикой (учитывая, что абсолютная синонимия практически невозможна).

И.А. Мельчук предлагает также формальное определение словоизменительной категории. Чтобы не усложнять изложение в данной статье, мы отсылаем читателя к первоисточнику (Мельчук 1997: 249). Чуть уточненный вариант этого определения приведен в свежей статье И.А. Мельчука с соавторами (Иорданская и др. 2020). Содержательно это определение сводится к двум требованиям – обязательности и регулярности словоизменения. Словообразование не обязано обладать этими свойствами (Мельчук 1997: 271). Русский вид (под названием ‘Аспект V’) описывается в этой монографии в главе II Семантические словоизменительные значения (Мельчук 1998).

Свойства обязательности и регулярности словоизменения рассматривались и в других работах. Так, в (Сичинава 2011) при обсуждении грамматического статуса вида приводятся следующие доводы (далее – цитата):

- выражение при приставочной перфективации совместно со словообразовательными значениями (в пользу словообразовательной трактовки);
- отсутствие общих морфологических показателей вида (против словоизменительной трактовки);

- однозначная охарактеризованность основного массива словоформ по виду (в пользу словоизменительной трактовки);
- высокая регулярность образования вторичных имперфективов (в пользу словоизменительной трактовки).

К сожалению, применение только формальных параметров не позволяет однозначно классифицировать вид. Так, регулярность является градуальным понятием, и не вполне очевидно, трактовать ли применение набора приставок при префиксальном образовании перфектива как регулярное или нет. Ряд более тонких формальных характеристик рассмотрены в работе (Горбова 2017).

Вернемся к семантико-лексикографическому подходу. Как уже отмечалось, критическим является вопрос, совпадает ли лексическая семантика у членов аспектуальной пары. Для глаголов пар ‘перфектив – вторичный имперфектив’ большинство исследователей считают, что их семантика идентична. Известный пример Ю.Д. Апресяна (Апресян 1995: 112) *пить – выпить – выпивать* демонстрирует наличие дополнительной лексической семантики у вторичного имперфектива. Более того, это отнюдь не единичный случай. В (Сичинава 2011) отмечено, что вторичный имперфектив «семантически сдвинут в сторону итератива». Несовпадение лексической семантики в этих парах трактуется в (Горбова 2017) как «следствие лексикализации (деграмматикализации) грамматической формы».

Соотношение семантики базового имперфектива и приставочного перфектива является предметом давних дискуссий. Совпадение их лексической семантики эквивалентно отсутствию собственной семантики у приставок в составе перфектива. Отправной является точка зрения (Виноградов 1972, Тихонов 1998), что в аспектуальных парах приставки этого типа являются «чистовидовым», то есть не несут собственной семантики, а выражают только вид. Однако есть и альтернативные точки зрения. Гипотеза Вея–Схоневелда (Schooneveld 1978) гласит, что эти приставки имеют собственную семантику, но она накладывается на семантику корня. В последнее десятилетие идея наличия собственной семантики у приставок в аспектуальных парах имперфектив – приставочный перфектив активно отстаивалась в работах Л. Янды (Янда 2012, Janda et al. 2013, Кузнецова, Янда 2013). Л. Янда трактует вид как классифицирующую категорию.

В.А. Плуноян также считает русский вид словоклассифицирующей категорией, хотя и с оговорками (Плуноян 2000: 125). В этой работе приводятся интересные наблюдения типологического и диахронического характера. По данным В.А. Плунояна, во многих языках мира аспект носит словоизменительный характер с тенденцией превращения в словоклассифицирующую категорию (Плуноян 2000: 294). Суффиксальные аспектуальные пары В.А. Плуноян трактует как словоизменительные, однако отмечает, что такие пары не являются доминирующими. Отмечается, что в последние 200–300 лет число таких пар возрастает (Плуноян 2000: 126). Подход В.А. Плунояна к словоизменению/словообразованию в целом близок к подходу И.А. Мельчука, но менее формализован.

Отметим интересную работу (Иорданская и др. 2020), в которой пары глаголов движения типа *лететь/летать* трактуются как находящиеся в словоизменительном отношении, несмотря на лексикографическую традицию их описания в разных словарных статьях. В работе обосновывается точка зрения, что различия в семантике глаголов в таких парах можно описать как грамматические, определяемые граммемами *однонаправленное/ненаправленное* движение. Это пример того, как, казалось бы, разные слова могут быть представлены как формы одного слова. Таким образом, разграничение словоизменения и словообразования является совершенно нетривиальной проблемой.

3. Данные и методы

3.1. Подход 1: динамика частотности употребления слова

Первый из двух предлагаемых нами подходов основан на использовании корпуса Google Books Ngram (в дальнейшем сокращенно GBN). Корпус содержит для русского языка более 80 миллиардов слов (в 100 раз больше, чем Национальный корпус русского языка) и охватывает более 2 веков. Сервис визуализации Ngram Viewer (<https://books.google.com/ngrams/>) позволяет строить графики частот встречаемости слов и словосочетаний по годам.

Коллекция GBN содержит только тексты книг, и в этом смысле она не является всеохватывающей. С другой стороны, методика создания этого корпуса была такова (Michel et al. 2011): полное сканирование книг (более 15 % от всех изданных в мире книг) более чем из 40 крупнейших библиотек по всему миру, включая библиотеки университетов Оксфорда, Гарварда и др. (ENA, September 20, 2022)³. Это обеспечило репрезентативное представление всех аспектов человеческой культуры, в той мере, в которой они представлены в изданных книгах, хранящихся в ведущих библиотеках. Вопросы, связанные со сбалансированностью и репрезентативностью GBN, обсуждаются в (Solovyev 2019, Richey & Taylor 2020).

Наша основная идея состоит в следующем. Если два слова находятся в словоизменительном отношении и, соответственно, имеют тождественную лексическую семантику, то частота употребления этих двух слов будет меняться синхронно, в зависимости от степени востребованности слов с этой семантикой в тот или иной период времени. Если же два слова находятся в словообразовательных отношениях, т.е. представляют собой два разных слова, то частоты их употребления могут меняться не синхронно, в зависимости от востребованности каждого из этих слов. Это следует, в частности, из дистрибутивной гипотезы (Sahlgren 2008). Таким образом, для того, чтобы оценить префиксальный и суффиксальный способы образования аспектуальных пар с этой точки зрения, требуется сопоставить динамику частот словоупотребления соответствующих глаголов. В GBN употребление слова репрезентируется временным рядом частот в последовательные годы. Мы

³ <http://www.google.com/intl/en/googlebooks/library/partners.html>

используем интервал времени с 1920 г., чтобы избежать влияния старой орфографии (до реформы 1917 г.), не всегда точно отражаемой в GBN.

Степень синхронности изменения частот употребления определяется близостью форм кривых – временных рядов и может быть подсчитана. Общей проблематике близости временных рядов и методам его оценки посвящена работа (Batyrshein et al. 2014). Для численной оценки степени близости будем использовать коэффициенты корреляции Пирсона, как часто используемый для оценки близости временных рядов. Данный подход был впервые анонсирован на конференции 15th Slavic Cognitive Linguistics Conference (Соловьев 2017). Обоснование подхода изложено в статье (Соловьев 2022).

Будут использованы два набора данных. Аспектуальные пары с префиксальным способом видообразования взяты из базы данных Exploring Emptiness (Janda et al. 2013), (ENA, September 20, 2022)⁴. Пары с суффиксальным типом видообразования – из базы данных видовых пар (О. Ляшевская 2016 (ENA, September 20, 2022)⁵).

При работе с временными рядами частот из GBN следует иметь в виду следующую особенность. Для низкочастотных слов возрастает влияние на форму графиков случайных факторов. Поэтому целесообразно анализировать по отдельности слова с разной частотностью. Будем сравнивать только слова, отличающиеся по частоте не более, чем в 10 раз. Т.е. делим все слова на группы с частотами $f: f > 10^{-4}$, $10^{-4} > f > 10^{-5}$ и т.д. Частота подсчитывается как средняя на интервале 1920–2008. Рассматриваются лишь аспектуальные пары, оба глагола которых принадлежат к одной частотной группе. В табл. 1 приведены усредненные коэффициенты корреляции по Пирсону временных рядов частот глаголов в аспектуальных парах.

Таблица 1. Средний коэффициент корреляции Пирсона частот глаголов в аспектуальных парах/
Table 1. Average Pearson correlation coefficient of verb frequencies in aspectual pairs

Частота, log10	Префиксальное видообразование	Суффиксальное видообразование
-4	0,65	0,61
-5	0,66	0,68
-6	0,53	0,58
-7	0,31	0,39
-8	0,19	0,27
-9	-	0,29
Всего пар	673	7800

Как видно из таблицы, для наиболее высокочастотных глаголов с $f > 10^{-4}$ корреляция оказывается несколько выше для префиксального способа видообразования. Для следующих двух частотных групп – для суффиксального способа видообразования, причем тоже не намного выше. Для низкочастотных глаголов в следующих группах коэффициент корреляции имеет небольшие значения в силу фактора, о котором было сказано выше. Таким образом, при таком подсчете существенной разницы не обнаруживается.

⁴ <http://emptyprefixes.uit.no/>

⁵ <http://ru-eval.ru/go/resources.html>

Одним из возможных влияющих на результат факторов может быть многозначность глаголов. Это снижает значения коэффициентов корреляции, так как график частот словоупотреблений строится для всех значений слов, усредняя частоты для всех его значений. Отберем некоторое число эталонных пар – достаточно высокочастотных глаголов, не являющихся омонимами или полисемичными словами с заведомо различными значениями, и рассчитаем коэффициенты корреляции для них. В табл. 2 приведены рассматриваемые пары слов. Список подготовлен Л. Яндой.

Расчет проводится так же, как и для полного набора аспектуальных пар. Результаты приведены в табл. 3.

Таблица 2. Пары базовый имперфектив – естественный перфектив и специализированный перфектив – вторичный имперфектив /
Table 2. Pairs of basic imperfective – natural perfective and specialized perfective – secondary imperfective

Базовый имперфектив	Естественный перфектив	Специализированный перфектив	Вторичный имперфектив
делать	сделать	переделать	переделывать
казаться	показаться	оказаться	оказываться
просить	попросить	выпросить	выпрашивать
ставить	поставить	приставить	приставлять
хранить	сохранить	охранить	охранять
прятать	спрятать	припрятать	припрятывать
плевать	наплевать	оплевать	оплевывать
растить	вырастить	отрастить	отращивать
жрать	сожрать	пожрать	пожирать
нюхать	понюхать	вынюхать	вынюхивать
щупать	пощупать	нащупать	нащупывать
играть	сыграть	выиграть	выигрывать
верить	поверить	проверить	проверять
звонить	позвонить	перезвонить	перезванивать
рисовать	нарисовать	срисовать	срисовывать
влечь	повлечь	привлечь	привлекать
жечь	сжечь	зажечь	зажигать
мерзнуть	замерзнуть	промерзнуть	промерзать
копать	выкопать	раскопать	раскапывать
красть	украсть	обокрасть	обкрадывать

Таблица 3. Средний коэффициент корреляции Пирсона частот глаголов в аспектуальных парах из таблицы 2 /

Table 3. Average Pearson correlation coefficient of verb frequencies in aspectual pairs from table 2

Частота, log ₁₀	Префиксальное видообразование	Суффиксальное видообразование
-5	0,94	-
-6	0,75	0,72
-7	0,71	0,72
-8	-	0,53

Здесь минус означает отсутствие пар в данной частотной группе. Как и ожидалось, для эталонных пар коэффициент корреляции заметно выше, чем

в табл. 1. Аспектуальные пары обоих типов трансформации присутствуют только в двух частотных группах: $10^{-5} > f > 10^{-6}$ и $10^{-6} > f > 10^{-7}$. В обеих группах коэффициенты корреляции почти равны.

Таким образом, при таком подходе различия в степени семантической близости при суффиксальном и префиксальном способах образования вида для эталонных пар также не обнаружены.

3.2. Подход 2: семантическое сходство в векторной семантике

Хорошо известной является дистрибутивная гипотеза, согласно которой лингвистические единицы, встречающиеся в схожих контекстах, имеют близкие значения. Основываясь на ней в компьютерной лингвистике разработаны методы векторной семантики, позволяющие формализовать семантику слова как вектор частот слов, с которыми оно вместе встречается в текстах корпуса. Семантическое сходство между словами легко вычислить, одним из стандартных математических методов, например, как косинусную меру сходства между векторами слов. Предложены различные методы построения векторов: skipgram, fasttext и др. (Mikolov et al. 2013, Wojanowski et al. 2017, Pennington et al. 2014). Данный подход и его последующие модификации широко используются в различных задачах обработки текстов, таких как классификация текстов, вопросно-ответные системы, суммаризация текстов (Sivakumar et al. 2020, Yang et al. 2019). Для русского языка существуют системы, в которых уже построены вектора для сотен тысяч слов (Grave et al. 2018, Kutuzov & Andreev 2015). Мы используем эти методы для расчета семантического сходства слов в аспектуальных парах.

В таблице 4 приведена величина сходства (в интервале от 0 до 1; 1 – максимальное сходство) для нашего набора эталонных пар. Для расчетов применялся инструментарий Rusvectors (ENA, September 20, 2022)⁶, использующий модель skipgram, обученную на корпусе Тайга (Shavrina & Shapovalova 2017). В нее включены также значения сходства между базовым имперфективом и специализированным перфективом. Сходство приведено в столбцах, расположенных между соответствующими словами.

В приложении приведены две таблицы, в которых использованы другие математические модели и другие корпуса. Численные данные в них немного отличаются от данных таблицы 4, но демонстрируют те же эффекты. В таблице S1 сходство рассчитывалось по предобученной модели fasttext без дополнительного дообучения, а в таблице S2 по skipgram, обученной на корпусе НКРЯ и Википедия. Различие в корпусах состоит в том, что Тайга состоит почти полностью (95%) из художественных текстов. Прочерк в таблицах ниже означает, что в данной вычислительной модели отсутствуют необходимые слова.

⁶ <https://rusvectors.org/ru/misc/>

Таблица 4. Величина сходства между словами в векторной семантике, модель skipgram, корпус Тайга /

Table 4. The value of the similarity between words in vector semantics, skipgram model, Taiga corpus

№	Естественный перфектив	Сходство	Базовый имперфектив	Сходство	Специализированный перфектив	Сходство	Вторичный имперфектив
1	сделать	0,73	делать	0,20	переделать	0,73	переделывать
2	показаться	0,64	казаться	0,21	оказаться	0,69	оказываться
3	попросить	0,82	просить	0,45	выпросить	0,60	выпрашивать
4	поставить	0,81	ставить	0,30	приставить	0,78	приставлять
5	сохранить	0,52	хранить	0,38	охранить	0,50	охранять
6	спрятать	0,81	прятать	0,68	припрятать	0,82	припрятывать
7	наплевать	0,84	плевать	0,41	оплевать	0,50	оплевывать
8	вырастить	0,73	растить	0,27	отрастить	0,69	отращивать
9	сожрать	0,62	жрать	0,60	пожрать	0,48	пожирать
10	понюхать	0,76	нюхать	0,34	вынюхать	0,59	вынюхивать
11	пощупать	0,70	щупать	0,46	нащупать	0,48	нащупывать
12	сыграть	0,65	играть	0,27	выиграть	0,81	выигрывать
13	поверить	0,76	верить	0,19	проверить	0,81	проверять
14	позвонить	0,77	звонить	0,71	перезвонить	0,75	перезванивать
15	нарисовать	0,80	рисовать	0,42	срисовать	0,55	срисовывать
16	повлечь	0,63	влечь	0,39	привлечь	0,76	привлекать
17	сжечь	0,58	жечь	0,38	зажечь	0,82	зажигать
18	замерзнуть	0,63	мерзнуть	0,52	промерзнуть	0,74	промерзать
19	выкопать	0,70	копать	0,50	раскопать	0,71	раскапывать
20	украсть	0,50	красть	0,46	обокрасть	0,67	обкрадывать
Среднее		0,70		0,41		0,67	

4. Обсуждение результатов

Анализируя данные таблиц, следует отметить следующее.

1. Сравнение сходства в парах «базовый имперфектив – естественный перфектив» с парами «базовый имперфектив – специализированный перфектив».

Как и следовало ожидать, сходство между словами в парах «базовый имперфектив – естественный перфектив» (столбец 3) значительно выше по сравнению с парами «базовый имперфектив – специализированный перфектив» (столбец 5). Различие в семантике слов в парах «базовый имперфектив – естественный перфектив» обусловлено только различием в виде, а в парах «базовый имперфектив – специализированный перфектив» кроме вида семантика глаголов различается еще и семантикой приставки. Это подтверждает правильность предложенного метода вычисления семантической схожести в данной задаче. Отметим, что не только в среднем, но и в каждой строке значение в третьем столбце выше, чем в пятом (исключение в строке 9 табл. S2 приложения со словом *жрать*, где имеет место равенство значений).

2. Сравнение сходства в парах «базовый имперфектив – специализированный перфектив» с парами «специализированный перфектив – вторичный имперфектив».

Аналогичное соотношение мы видим и при сопоставлении пар «базовый имперфектив – специализированный перфектив» с парами «специализированный перфектив – вторичный имперфектив». Степень сходства в парах «специализированный перфектив – вторичный имперфектив» (столбец 7) почти во всех случаях больше, чем в парах «базовый имперфектив – специализированный перфектив» (столбец 5). Обратим внимание на интересное исключение.

В табл. 4 и S1 в паре *пожрать* – *пожирать* степень близости слов меньше, чем в паре *жрать* – *пожрать*. Это указание на то, что *пожрать* – *пожирать* не является чистой аспектуальной парой, *пожирать* имеет какую-то дополнительную семантику по сравнению с *пожрать*. Проведем лингвистический анализ этих слов по их толкованиям, чтобы выявить возможные расхождения в семантике. Возьмем толкования в словаре Малом академическом словаре (Евгеньева 1981–84). *Пожрать* означает то же, что и *поест* с дополнительным значением ‘с жадностью’ (о животных) или грубое протонародное выражение, если говорится о человеке. Слово *пожирать* кроме этого имеет еще и переносные значения: 1) Быстро, с жадностью и интересом прочитывать (книги, журналы и т. д.). Сочетание ^{??} *пожрал книги* маргинально. В Google это словосочетание встречается 27 раз, в то время как *пожирал книги* – 6600 раз. 2) Поглощать, требовать для себя что-л. в больших количествах. 3) Уничтожать, сжигая (об огне, пламени). 4) Захватывать целиком (о чувствах). Таким образом, между словами *пожрать* и *пожирать* действительно есть значительные семантические различия, кроме видовых.

То, что суффиксальное образование вторичного имперфектива не всегда является чисто видовой трансформацией ранее отмечалось, например, в работе (Апресян 1995: 112), в которой указано, что глагол *выпивать* имеет дополнительное значение кратности, а чисто видовой парой с *выпить* является именно *пить*.

3. Сравнение средних значений сходства. Средние значения близости в парах «базовый имперфектив – естественный перфектив» и «специализированный перфектив – вторичный имперфектив» очень близки и существенно превышают среднее значение в парах «базовый имперфектив – специализированный перфектив».

При расчетах по модели skipgram с корпусом Тайга (табл. 4) в 10 случаях значение близости выше для пар «базовый имперфектив – естественный перфектив», в 9 случаях – для пар «специализированный перфектив – вторичный имперфектив» (в одном случае равенство). При расчетах по модели fasttext (табл. S1) в 8 случаях значение близости выше для пар «базовый имперфектив – естественный перфектив», в 10 случаях – для пар «специализированный перфектив – вторичный имперфектив» (в двух случаях равенство).

При расчетах по модели skipgram с корпусами НКРЯ и Википедия (табл. S2) в 10 случаях значение близости выше для пар «базовый имперфектив – естественный перфектив», в 8 случаях – для пар «специализированный перфектив – вторичный имперфектив».

Таким образом, хотя точные значения величины близости для пар слов несколько различаются для разных вычислительных моделей, общие закономерности сохраняются. Оба способа образования аспектуальных пар – префиксальный и суффиксальный – в одинаковой степени меняют семантику слов. В случаях значительного расхождения значений близости пар слов анализ значений слов по словарям позволяет объяснить причины расхождения. Например, в паре *зажечь* – *зажигать* семантическое сходство значительно выше, чем в паре *жечь* – *сжечь*. Семантик слов *зажечь*, *зажигать* практически совпадает, в то же время между словами *жечь*, *сжечь* имеется заметная разница: *жечь* имеет значение ‘вызывать ожег’, которого нет у глагола *сжечь*.

Полученные численные данные могут быть использованы для получения и других результатов. Рассчитаем коэффициенты корреляции по Пирсону между столбцами 3, 5, 7; результаты приведены в табл. 5.

Таблица 5. Корреляция между разными способами образования пар /
Table 5. Correlation between different ways of pairing

	Модель fastText	Модель skipgram+Тайга	Модель skipgram+НКРЯ+Wiki
Префиксальная аспектуальная пара – суффиксальная аспектуальная пара	0,173	0,018	0,332
Префиксальная аспектуальная пара – способ глагольного действия	0,065	0,097	0,040
Суффиксальная аспектуальная пара – способ глагольного действия	–0,441	–0,150	–0,371

Между префиксальным и суффиксальным способами образования аспектуальных пар корреляция близости пар глаголов слабая (по Чеддоку). Это же касается и сравнения: префиксальные аспектуальные пары – способ глагольного действия. В то же время при сопоставлении данных по суффиксальному способу образования аспектуальных пар с парами, классифицируемыми как способ глагольного действия, неожиданно обнаруживается значительно более высокий уровень корреляции, причем с отрицательным знаком. Для модели fastText эта корреляция характеризуется как умеренная, близкая к средней: –0,441. Качественно это означает, что чем меньший вклад вносит приставка в семантику специализированного перфектива по сравнению с базовым имперфективом, тем в большей степени вторичный имперфектив будет семантически отличаться от специализированного перфектива.

Например, в таблице 4 средняя степень близости в парах «базовый имперфектив – специализированный перфектив» равно 0,41, а в парах «специализированный перфектив – вторичный имперфектив» – 0,67. При этом у

рассмотренного выше глагола *жрать* в паре *жрать* – *пожрать* степень близости 0,60, заметно выше средней, что указывает на высокую степень семантической близости, хотя *пожрать* и специализированный перфектив, а в паре *пожрать* – *пожирать* только 0,48 – значительно ниже среднего, что указывает на меньшую семантическую близость, обусловленную дополнительным значением слова *пожирать*, как было указано выше. В то же время у пары *делать* – *переделать* степень близости всего 0,20 (меньше средней), обусловленная ясным различием в семантике, а у пары *переделать* – *переделывать* степень близости 0,73 (больше средней). Эта пара является чисто аспектуальной (Ушаков 1994). Насколько нам известно, такой эффект ранее не описывался. Он требует дополнительного изучения.

5. Заключение

С появлением больших корпусов текстов появилась надежда на то, что применение статистических, компьютерных методов позволит прояснить вопрос о том, находятся ли глаголы в аспектуальных парах в словоизменительных или иных грамматических отношениях. Для постулирования словоизменительного или словообразовательного (+словоклассифицирующего) статуса аспектуальных пар, получаемых суффиксальным или префиксальным способом, принципиальным является определение того, в какой степени глаголы в паре являются семантически близкими. В значительном числе работ предполагалось, что суффиксальный способ образования аспектуальных пары является словоизменительным, а префиксальный – словообразовательным (+словоклассифицирующим). Соответственно, можно было бы ожидать, что суффиксальный способ не меняет семантики глаголов, он меняет только вид, а префиксы вносят определенный вклад в глагольную семантику, помимо видовой.

В статье проверялась эта гипотеза. Предложены два различных подхода для получения точных количественных оценок семантической близости глаголов в аспектуальных парах. Оба подхода реализованы в двух вариантах. При этом нигде не обнаружено заметной разницы в степени семантической близости глаголов в префиксальных и суффиксальных аспектуальных парах. Таким образом, ответ на сформулированную в начале статьи гипотезу: «семантическая близость между глаголами в аспектуальных парах перфектив – вторичный имперфектив будет больше, чем между глаголами в парах базовый имперфектив – перфектив» отрицательный.

Аналогичный результат при применении совершенно иного подхода – а именно сопоставлении грамматических профилей глаголов – был получен и в работе (Janda & Ljashchinskaya 2011). Найти количественные различия в свойствах суффиксальных или префиксальных аспектуальных пар не удастся.

Это означает, что основной признак отнесения грамматической категории к словоизменению или словообразованию – сохранение или не сохранение лексической семантики – не позволяет разграничить префиксальный и

суффиксальный способы видообразования с точки зрения их грамматического статуса. Многолетнюю дискуссию о противопоставлении грамматического статуса аспектуальных пар, образованных префиксальным и суффиксальным способами, по нашему мнению, можно считать практически завершённой.

Обратим внимание на то, что второй из предложенных подходов – с вычислением векторной близости между словами на основе больших корпусов текстов – широко используется в компьютерной лингвистике (Solovyev et al. 2022), но лишь в последнее время начал применяться для решения собственно лингвистических задач (Ryzhova & Paperno 2022, Shavrina 2021). Между тем он может оказаться перспективным в исследованиях, связанных с семантической близостью слов.

Благодарности

Авторы благодарят Л. Янду и М. Исламова за вклад в подготовку материалов. Исследование поддержано РНФ, грант № 20-18-00206. Конфликт интересов отсутствует.

Вклад авторов

Соловьев В.Д. – основная идея, выполнение расчетов (табл. 4, 5, S2), квантитативный анализ результатов, написание текста статьи. Бочкарев В.В. – выполнение расчетов (табл. 1, 3). Байрашева В.Р. – лингвистический анализ примеров. Янда Л. – подготовка набора слов (табл. 2), Исламов М.И. – выполнение расчетов (табл. S1).

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Апресян Ю.Д. Избранные труды. Т. 2. М.: Школа “Языки русской культуры”, 1995. [Apresjan, Jury D. 1995. *Izbrannye Trudy (Elected Works)*. V. 2. M.: Shkola “Yazyki russkoi kul'tury”].
- Бондарко А.В. Вид и время русского глагола. М.: Просвещение 1971. [Bondarko, Alexander V. 1971. *Vid i vremya russkogo glagola (Type and tense of the Russian verb)*. M.: Prosveshchenie (In Russ.)].
- Виноградов В.В. Русский язык. М.: Высш. шк., 1972. [Vinogradov, Victor V. 1972. *Russkii yazyk (Russian language)*. M.: Vyssh. shk. (In Russ.)].
- Вежбицкая А. Семантические универсалии и описание языков. М.: Языки русской культуры, 1999. [Wierzbicka, Anna. 1999. *Semanticheskie universalii i opisanie yazykov (Semantic universals and the description of languages)*. M.: Yazyki russkoi kul'tury. (In Russ.)].
- Всеволодова М.В. Семный состав глагольного слова. (К вопросу типологии вида) // *Типология вида: проблемы, поиски, решения*. Тезисы международной научной конференции. Московский государственный университет имени М.В. Ломоносова. Филологический факультет. М.: Изд-во Моск. ун-та, 1997. [Vsevolodova, Maiya V. 1997. *Semnyi sostav glagol'nogo slova. (K voprosu tipologii vida)*. (The semantic composition of the verb word. (On the question of the typology of species)) *Tipologiya vida: problemy, poiski, resheniya*. Tezisy mezhdunarodnoi nauchnoi konferencii. Moscow State University named after M. V. Lomonosov. Philological department. M.: Izd-vo Mosk. un-ta. (In Russ.)].

- Горбова Е.В. Заметки о видообразовании русского глагола и словоизменительной vs. словоклассифицирующей трактовке вида // *Acta Linguistica Petropolitana*. Вып. X. Ч. 3. СПб.: Наука, 2014. С. 181–211. [Gorbova, Elena V. 2014. Notes on aspectual formation of the Russian verb and on inflectional vs. word-classifying interpretation of aspect. *Acta Linguistica Petropolitana* 10 (3). 181–211. St. Petersburg: Nauka. (In Russ.)].
- Горбова Е.В. Видообразование русского глагола: префиксация и/или суффиксация? // *Вопросы языкознания*. 2015а. № 1. С. 7–37. [Gorbova, Elena V. 2015а. Aspectual formation of the Russian verb: Prefication or suffixation? *Voprosy Yazykoznaniiya* 1. 7–37. (In Russ.)].
- Горбова Е.В. Словоизменение или словообразование, или является ли русский вид «видом славянского типа»? // *Аспектуальная семантическая зона: типология систем и сценарии диахронического развития*. Сборник статей V Международной конференции Комиссии по аспектологии Международного комитета славистов. Киото, 13–15 ноября 2015 / Китадзэ М. (сост.). Киото: Tanaka Print. С. 62–72. 2015б. [Gorbova, Elena V. 2015б. Inflection or derivation, or Is the Russian aspect «Slavic style aspect»? *Aspektual'naya semanticheskaya zona: tipologiya sistem i stsenarii diakhronicheskogo razvitiya*. Sbornik statei V Mezhdunarodnoi konferentsii Komissii po aspektologii Mezhdunarodnogo komiteta slavistov. Kioto, 13–15 noyabrya 2015. Kitajo M. (comp.). Kyoto: Tanaka Print. 62–72. (In Russ.)].
- Горбова Е.В. Русское видообразование: словоизменение, (слово)классификация или набор квазиграммем? (еще раз о болевых точках русской аспектологии) // *Вопросы языкознания*. 2017. № 1. С. 24–52. [Gorbova, Elena V. 2017. Aspectual formation of Russian verbs: Inflection, derivation, or a set of quasigrammemes? («sore points» of Russian aspectology revisited). *Voprosy Yazykoznaniiya* 1. 24–52. (In Russ.)].
- Евгеньева А.П. (ред.) Словарь русского языка: в 4 т. / АН СССР, Ин-т рус. яз. М.: Русский язык, 1981–1984. [Evgenyeva, Anastasia P. (ed.). 1981–1984. Slovar' russkogo yazyka: V 4-h t. (The dictionary of the Russian language in 4 volumes) / AN SSSR, In-t rus. yaz. M.: Russkii yazyk. (In Russ.)].
- Зализняк А.А., Микаэлян И.Л. О некоторых дискуссионных моментах аспектологической концепции Лоры Янды // *Вопросы языкознания*. 2012. №6. С. 48–65. [Zalizniak, Anna A. & Irina L. Mikaelian. 2012. O nekotorykh diskussionnykh momentah aspektologicheskoi konceptii Lory Jandy (On some discussion points in Laura Yanda's aspectological concept). *Voprosy yazykoznaniiya* 6. 48–65. (In Russ.)].
- Зализняк А.А., Шмелев А.Д., Микаэлян И.Л. Русская аспектология. В защиту видовой пары. М.: Языки славянских культур, 2015. [Zalizniak, Anna A., Aleksey D. Shmelev & Irina L. Mikaelian. 2015. Russkaya aspektologiya. V zashhitu vidovoi pary (Russian Aspectology. In defense of the species pair.). М.: Yazyki slavyanskih kul'tur. (In Russ.)].
- Иорданская Л.Н., Крылосова С.Г., Мельчук И.А., Михель П.А. Категория направленности у русских парных глаголов перемещения (на примере глаголов ЛЕТЕТЬ / ЛЕТАТЬ) // *Вопросы языкознания*. 2020. № 1. 27–64. [Iordanskaja, Lidija N., Svetlana G. Krylosova, Igor A. Melcuk & Polina A. Mihel'. 2020. Kategoriya napravlennosti u russkih parnyh glagolov peremeshcheniya (na primere glagolov LETET' / LETAT') (Directionality in Russian paired verbs of displacement (by the example of the verbs FLY / FLY)). *Voprosy yazykoznaniiya* 1. 27–64. (In Russ.)].
- Исаченко А.В. 1954/2003. *Грамматический строй русского языка в сопоставлении со словацким. Морфология*. Т. I–II. М.: Языки славянской культуры, 2003. (Исаченко А.В. Грамматический строй русского языка в сопоставлении со словацким: Морфология. Часть первая. Братислава, 1954.) [Issatschenko, Aleksandr V. 1954/2003. *Grammaticheskii stroi russkogo yazyka v sopostavlenii so slovatskim. Morfologiya* (Grammatical system of Russian versus Slovak. Morphology). Vol. I–II. Moscow: Yayki

- Slavyanskoj Kul'tury (Issatschenko, Aleksandr V. 1954. Grammaticheskii stroi russkogo yazyka v sopostavlenii so slovatskim: Morfologiya (Grammatical system of Russian versus Slovak. Morphology). Part 1. Bratislava. (In Russ.)).
- Кузнецова Ю.Л., Янда Л.А. Приставки в свете когнитивной лингвистики и типологии (отклик на статью А.А. Зализняк и И.Л. Микаэлян) // *Вопросы языкознания*. 2013. №4. С. 87–96. [Kuznecova, Julija L. & Laura A. Janda. 2013. Pristavki v svete kognitivnoi lingvistiki i tipologii (otklik na stat'yu Anna A. Zalizniak i Irina L. Mikaelian) (Prefixes in the light of cognitive linguistics and typology (response to the article of Anna A. Zaliznyak & Irina L. Mikaelian). *Voprosy yazykoznanija* 4. 87–96. (In Russ.)].
- Маслов Ю.С. 1978/2004. Избранные труды. Аспектология. Общее языкознание. М.: Языки славянских культур, 2004. С. 305–364. (Маслов Ю.С. К основаниям сопоставительной аспектологии // Маслов Ю.С. (отв. ред.) Проблемы современного теоретического и синхронно-сопоставительного языкознания. Вып. 1: Вопросы сопоставительной аспектологии. Л.: Изд-во ЛГУ, 1978. С. 4–44.) [Maslov, Yury S. 1978/2004. Izbrannye trudy. Aspektologiya. Obshchee yazykoznanie (Selected works. Aspectology. General linguistics). Moscow: Yazyki Slavyanskikh Kul'tur. 305–364. (Maslov, Yury S. 1978. Towards the fundamentals of comparative aspectology. In Yury Maslov (ed.), *Problems of Modern Theoretical and Synchronic Comparative Linguistics. Vol. 1: Issues of Comparative Aspectology*, 4–44. Leningrad: Leningrad State Univ. Publ. (In Russ.)].
- Маслов Ю.С. Очерки по аспектологии. Л.: ЛГУ, 1984. [Maslov, Yury S. *Ocherki po aspektologii* (Essays on Aspectology). L.: LGU. (In Russ.)].
- Мельчук И.А. 1997. Курс общей морфологии. Т. 1. М.: Прогресс, Языки русской культуры, Вена: WSA. [Mel'chuk, Igor A. 1997. Kurs obshhei morfologii (Course in General Morphology). V. 1. М.: Progress, Yazyki russkoi kul'tury, Vena: WSA. (In Russ.)].
- Мельчук И.А. 1998. Курс общей морфологии. Т. 2. М.: Языки русской культуры, Вена: WSA. [Mel'chuk, Igor A. 1998. Kurs obshhei morfologii (Course in General Morphology). V. 2. М.: Yazyki russkoi kul'tury, Vena: WSA. (In Russ.)].
- Перцов Н.В. 1996. Грамматическое и обязательное в языке // *Вопросы языкознания*. № 4. С. 39–61. [Percov, Nikolay V. 1996. Grammaticheskoe i obyazatel'noe v yazyke (Grammatical and obligatory in language). *Voprosy yazykoznanija* 4. 39–61. (In Russ.)].
- Петрухина Е.В. Аспектуальные категории глагола в русском языке в сопоставлении с чешским, словацким, польским и болгарским языками. М.: МГУ, 2000. [Petruhina, Elena V. 2000. Aspektual'nye kategorii glagola v russkom yazyke v sopostavlenii s cheshskim, slovatskim, pol'skim i bolgarskim yazykami (Aspectual Categories of the Verb in Russian Compared to Czech, Slovak, Polish, and Bulgarian). М.: MGU. (In Russ.)].
- Плунгян В.А. Общая морфология. Введение в проблематику. М.: УРСС, 2000. [Plungian, Vladimir A. 2000. Obshchaya morfologiya. Vvedenie v problematiku (General Morphology. Introduction to the Problematics). Mju: URSS. (In Russ.)].
- Русская грамматика: научные труды. В 2 томах / Н.Ю. Шведова (Гл. ред.). — М.: Институт русского языка им. В.В. Виноградова РАН, 2005. [Russkaya grammatika: nauchnye trudy. V 2 tomakh (Russian grammar: Scientific works. In 2 volumes) / Natalia Shvedova (ed.). 2005. М.: Institut russkogo yazyka im. V.V. Vinogradova RAN. (In Russ.)].
- Сичинава Д.В. Вид. Материалы для проекта корпусного описания русской грамматики (<http://rusgram.ru>). На правах рукописи. М., 2011. [Sichinava, Dmitrii V. 2011. Vid. Materialy dlya proekta korpusnogo opisaniya russkoi grammatiki (<http://rusgram.ru>). Na pravah rukopisi (Materials for the project of corpus description of Russian grammar (<http://rusgram.ru>). On the rights of manuscript). М. (In Russ.)]. (accesed 13 October 2012).

- Соловьев В., Бочкарев В., Янда Л. Аспектуальная система в русском языке: семантический вклад приставок и суффиксов // *15th Slavic Cognitive Linguistics Conference, Book of Abstracts*. С. 75–76. 2017. [Solovyev, Valery, Vladimir Bochkarev & Laura Janda. 2017. Aspektual'naya sistema v russkom yazyke: semanticheskii vklad pristavok i suffiksov (The aspect system in Russian: Semantic contribution of prefixes and suffixes). *15th Slavic Cognitive Linguistics Conference, Book of Abstracts*. 75–76. (In Russ.)].
- Соловьев В., Бочкарев В. Дело об аспектуальных парах открывается вновь // *Вестник Томского государственного университета. Филология*. Вып. 78. 2022. [Solovyev, Valery & Vladimir Bochkarev. 2022. Case for aspectual pairs reopened. *Vestnik Tomskogo gosudarstvennogo universiteta. Filologiya* 78. (In Russ.)].
- Тихонов А.Н. Русский глагол: проблемы теории и лексикографирования. М.: Academia, 1998. [Tikhonov, Alexander N. 1998. *Russkii glagol: problemy teorii i leksikografirovaniya* (The Russian verb: Problems of theory and lexicographic description). Moscow: Academia. (In Russ.)].
- Ушаков Д.Н. (ред.) *Tolkovyi slovar' russkogo yazyka: V 4 t.* М.: Сов. Энцикл, 1994. [Ushakov, Dmitry N. (ed.). 1994. *Tolkovyi slovar' russkogo yazyka: V 4 t.* (The Dictionary of the Russian language: In 4 vols.). М.: Sov. encik. (In Russ.)].
- Храковский В.С. Категория вида в русском языке: болевые точки // *Аспектуальная семантическая зона: типология систем и сценарии диахронического развития*. Сборник статей V Международной конференции Комиссии по аспектологии Международного комитета славистов. Киото, 13–15 ноября 2015 г. / Китадзэ М. (сост.). Киото: Tanaka Print. С. 321–334. 2015. [Xrakovskij, Viktor S. 2015. The category of aspect in Russian: Problematic issues. *Aspektual'naya semanticheskaya zona: tipologiya sistem i stsenarii diakhronicheskogo razvitiya*. Sbornik statei V Mezhdunarodnoi konferentsii Komissii po aspektologii Mezhdunarodnogo komiteta slavistov. Kioto, 13–15 November 2015 g. Kyoto: Tanaka Print. 321–334. (In Russ.)].
- Храковский В.С. И опять о болевых точках категории вида // *Вопросы языкознания*. 2018. № 1. С. 105–118. [Xrakovskij, Viktor S. 2018. Aspect's "sore points" revisited once more. *Voprosy yazykoznaniiya* 1. 105–118. (In Russ.)].
- Янда Л.А. Русские приставки как система глагольных классификаторов // *Вопросы языкознания*. 2012. № 6. С. 3–47. [Janda, Laura A. 2012. Russkie pristavki kak sistema glagol'nyh klassifikatorov (Russian prefixes as a system of verb classifiers). *Voprosy yazykoznaniiya* 6. 3–47. (In Russ.)].
- Batyrshin, Ildar, Valery Solovyev & Vladimir Ivanov. 2014. Time series shape association measures and local trend association patterns. *Neurocomputing* 175. 924–934
- Bojanowski, Piotr, Edouard Grave, Armand Joulin & Tomas Mikolov. 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics* 5. 135–146.
- Dressler, Wolfgang. 1989. Prototypical differences between inflection and derivation. *STUF – Language Typology and Universals* 42 (1). 3–10.
- Gorbova, Elena V. & Oksana Iu. Chuikova. 2022. Suffixal imperfectivation of prefixed verbs: Record breakers and outsiders (based on the dictionary, corpus and Runet). *Computational Linguistics and Intellectual Technologies*. Papers from the Annual International Conference "Dialogue". 21.
- Grave, Edouard, Piotr Bojanowski, Prakhar Gupta, Armand Joulin & Tomas Mikolov. 2018. *Learning Word Vectors for 157 Languages*. *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*. <https://www.aclweb.org/anthology/L18-1550> (accessed 21 September 2022)

- Janda, Laura & Olga Lyashevskaya. 2011. Aspectual Pairs in the Russian National Corpus. *Scando-Slavica* 57 (2). 201–215.
- Janda, Laura, Anna Endresen, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nessel & Svetlana Sokolova. 2013. *Why Russian Aspectual Prefixes aren't Empty*. Bloomington: Slavica.
- Filip, Hana. 1999. *Aspect, Eventuality Types, and Nominal Reference*. London.
- Kutuzov, Andrey & Igor Andreev. 2015. Texts in, meaning out: Neural language models in semantic similarity task for Russian. *Proceedings of the Dialog 2015 Conference*, Moscow, Russia. 133–144.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva Aiden & Adrian Veres. 2011. Quantitative analysis of culture using millions of digitized books. *Science* 331 (6014). 176–182.
- Mikolov, Tomas, Kai Chen, Greg Corrado & Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.
- Pennington, Jeffrey, Richard Socher & Christopher D. Manning. 2014. Glove: Global vectors for word representation. *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 1532–1543.
- Plank, Frans. 1991. Inflection and derivation. *EUROTYP working papers* 10.
- Richey, Sean & Benjamin J. Taylor. 2020. Google Books Ngrams and political science: Two validity tests for a novel data source *PS: Political Science & Politics* 53. 72–77. <https://doi.org/10.1017/S1049096519001318>.
- Ryzhova, Daria & Denis Paperno. 2022. Constructing typological questionnaire with distributional semantic models. In Ekaterina Rakhilina, Tatiana Reznikova & Daria Ryzhova (eds.), *The Typology of Physical Qualities*. Amsterdam: John Benjamins Publishing Company.
- Sahlgren, Magnus. 2008. The Distributional Hypothesis. From context to meaning. Distributional models of the lexicon in linguistics and cognitive science. [Special issue]. *Rivista di Linguistica* 20 (1). 33–53.
- van Schooneveld, Cornelis. 1978. *Semantic Transmutations: Prolegomena to a Calculus of Meaning. V. 1: The Cardinal Semantic Structure of Prepositions, Cases, and Paratactic Conjunctions in Contemporary Standard Russian*. Bloomington,
- Shavrina, Tatiana & Olga Shapovalova. 2017. To the methodology of corpus construction for machine learning: “Taiga” syntax tree corpus and parser. *Proceedings of the International Conference «Corpus Linguistics–2017»*. St. Petersburg. 78–84.
- Shavrina, Tatiana. 2021. Methods of computational linguistics in the evaluation of artificial intelligence systems. *Voprosy yazykoznaniiya* 6. 117–138.
- Singh, Rajendra & Alan Ford. 1980. Flexion, dérivation et Panini. *Amsterdam Studies in the Theory and History of Linguistic Science Amsterdam* 20. 323–332.
- Sivakumar, Soubraylu, Lakshmi Sarvani Videla, T. Rajesh Kumar, J. Nagaraj, Shilpa Itnal & D. Haritha. 2020. Review on Word2Vec Word Embedding Neural Net. *2020 International Conference on Smart Electronics and Communication (ICOSEC)*. IEEE. 282–290.
- Solovyev, Vavery, Bochkarev Vladimir & Akhtyamova Svetlana. 2019. Google Books Ngram: Problems of representativeness and data reliability. *Communications in Computer and Information Science* 1223. Springer, Cham. 147–162. https://doi.org/10.1007/978-3-030-51913-1_10
- Solovyev, Valery, Marina Solnyshkina & Danielle McNamara. 2022. Computational linguistics and discourse complexology: Paradigms and research methods. *Russian Journal of Linguistics* 26 (2). 275–316. <https://doi.org/10.22363/2687-0088-30161>
- Zhilin, Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov & Quoc V Le. 2019. Xlnet: Generalized autoregressive pretraining for language understanding. *arXiv preprint arXiv:1906.08237*.

Интернет-ресурсы

Database Exploring Emptiness, <http://emptyprefixes.uit.no/> (accessed 05 September 2022).

База данных суффиксальных аспектуальных пар [Baza dannyh suffiksalyh aspektual'nyh par], <http://ru-eval.ru/go/resources.html> (accessed 12 February 2020).

ПРИЛОЖЕНИЕ

Таблица S1. Величина сходства между словами в векторной семантике, модель *fasttext* /
Table S1. The value of the similarity between words in vector semantics, *fasttext* model

№	Естественный перфектив	Расстояние	Базовый имперфектив	Расстояние	Специализированный перфектив	Расстояние	Вторичный имперфектив
1	сделать	0,76	делать	0,51	переделать	0,76	переделывать
2	показаться	0,74	казаться	0,63	оказаться	0,63	оказываться
3	попросить	0,82	просить	0,59	выпросить	0,65	выпрашивать
4	поставить	0,81	ставить	0,41	приставить	0,82	приставлять
5	сохранить	0,59	хранить	0,39	охранить	0,64	охранять
6	спрятать	0,80	прятать	0,65	припрятать	0,64	припрятывать
7	наплевать	0,89	плевать	0,49	оплевать	0,71	оплевывать
8	вырастить	0,74	растить	0,45	отрастить	0,78	отращивать
9	сожрать	0,63	жрать	0,61	пожрать	0,47	пожирать
10	понюхать	0,71	нюхать	0,51	вынюхать	0,71	вынюхивать
11	пощупать	0,67	щупать	0,45	нащупать	0,76	нащупывать
12	сыграть	0,76	играть	0,48	выиграть	0,75	выигрывать
13	поверить	0,72	верить	0,40	проверить	0,75	проверять
14	позвонить	0,76	звонить	0,54	перезвонить	0,63	перезванивать
15	нарисовать	0,75	рисовать	0,55	срисовать	0,80	срисовывать
16	повлечь	0,64	влечь	0,51	привлечь	0,79	привлекать
17	сжечь	0,66	жечь	0,59	зажечь	0,77	зажигать
18	замерзнуть	0,75	мерзнуть	0,60	промерзнуть	0,76	промерзать
19	выкопать	0,61	копать	0,58	раскопать	0,71	раскапывать
20	украсть	0,69	красть	0,50	обокрасть	0,53	обкрадывать
Среднее		0,73		0,52		0,70	

Таблица S2. Величина сходства между словами в векторной семантике, модель *skipgram*, корпус НКРЯ и Википедия /

Table S2. The value of the similarity between words in vector semantics, *skipgram* model, RNC corpus and Wikipedia

№	Естественный перфектив	Расстояние	Базовый имперфектив	Расстояние	Специализированный перфектив	Расстояние	Вторичный имперфектив
1	сделать	0,77	делать	0,19	переделать	0,73	переделывать
2	показаться	0,69	казаться	0,38	оказаться	0,76	оказываться
3	попросить	0,79	просить	0,53	выпросить	0,62	выпрашивать
4	поставить	0,86	ставить	0,39	приставить	0,69	приставлять
5	сохранить	0,53	хранить	0,39	охранить	0,45	охранять
6	спрятать	0,80	прятать	0,72	припрятать	0,67	припрятывать
7	наплевать	0,65	плевать	0,51	оплевать	0,56	оплевывать

Окончание табл. S2							
№	Естественный перфектив	Расстояние	Базовый имперфектив	Расстояние	Специализированный перфектив	Расстояние	Вторичный имперфектив
8	вырастить	0,61	растить	0,27	отрастить	0,58	отращивать
9	сожрать	0,51	жрать	0,51	пожрать	0,54	пожирать
10	понюхать	0,69	нюхать	-	вынюхать	-	вынюхивать
11	пощупать	0,65	щупать	0,51	нащупать	0,70	нащупывать
12	сыграть	0,70	играть	0,27	выиграть	0,82	выигрывать
13	поверить	0,78	верить	0,33	проверить	0,82	проверять
14	позвонить	0,75	звонить	0,63	перезвонить	0,60	перезванивать
15	нарисовать	0,80	рисовать	0,53	срисовать	0,52	срисовывать
16	повлечь	0,69	влечь	0,34	привлечь	0,73	привлекать
17	сжечь	0,56	жечь	0,49	зажечь	0,76	зажигать
18	замерзнуть	0,57	мерзнуть	0,51	промерзнуть	0,62	промерзать
19	выкопать	0,68	копать	0,45	раскопать	0,63	раскапывать
20	украсть	0,44	красть	0,39	обокрасть	-	обкрадывать
Среднее		0,69		0,42		0,66	

Article history:

Received: 13 September 2021

Accepted: 04 September 2022

Bionotes:

Valery SOLOVYEV is Doctor Habil. of Physics and Mathematics, Professor, Chief Researcher of “Text Analytics” Research Lab at the Institute of Philology and Intercultural Communication of Kazan Federal University, Kazan, Russia. He is a member of the Presidium of the Interregional Association for Cognitive Research, author of four monographs and over 60 publications on text complexity and computational linguistics.

e-mail: maki.solovyev@mail.ru<https://orcid.org/0000-0003-4692-2564>

Vladimir BOCHKAREV is Research Fellow of the scientific laboratory “Linguistics and artificial intelligence” at Kazan Federal University, Kazan, Russia. His area of interest is quantitative linguistics.

e-mail: vbochkarev@mail.ru<https://orcid.org/0000-0001-8792-1491>

Venera BAYRASHEVA holds a PhD in Physics and Mathematics. She is Associate Professor at the Department of Theoretical Cybernetics at Kazan Federal University, Kazan, Russia. Her areas of interest are computational and theoretical linguistics.

e-mail: vbayrasheva@gmail.com<https://orcid.org/0000-0002-1729-034X>**Сведения об авторах:**

Валерий Дмитриевич СОЛОВЬЕВ – доктор физико-математических наук, профессор, главный научный сотрудник НИЛ «Текстовая аналитика» Института филологии и межкультурной коммуникации Казанского федерального университета, Казань,

Россия. Он является членом президиума Межрегиональной ассоциации когнитивных исследований, а также автором четырех монографий и более 60 публикаций по проблемам сложности текста и компьютерной лингвистики.

e-mail: maki.solovyev@mail.ru

<https://orcid.org/0000-0003-4692-2564>

Владимир Владимирович БОЧКАРЕВ – научный сотрудник научно-исследовательской лаборатории «Лингвистика и искусственный интеллект». Сфера его интересов – квантитативная лингвистика.

e-mail: vbochkarev@mail.ru

<https://orcid.org/0000-0001-8792-1491>

Венера Рустамовна БАЙРАШЕВА – кандидат физико-математических наук, доцент кафедры теоретической кибернетики Казанского федерального университета. Сфера ее научных интересов – компьютерная и теоретическая лингвистика.

e-mail: vbayrasheva@gmail.com

<https://orcid.org/0000-0002-1729-034X>