

СТРАТИФИКАЦИЯ ХРОНОЛОГИЧЕСКОГО СРЕЗА ЛЕКСИКИ (ПЕРВАЯ ЧЕТВЕРТЬ XXI ВЕКА)

С. А. Ермаков

Воронежский государственный лесотехнический университет имени Г. Ф. Морозова

STRATIFICATION OF THE CHRONOLOGICAL SECTION OF VOCABULARY (FIRST QUARTER OF THE XXI CENTURY)

S. A. Ermakov

Voronezh State University of Forestry and Technologies named after G. F. Morozov

Аннотация: в предложенной статье подробно описан эксперимент по системной стратификации хронологического среза лексики первой четверти XXI в. Цель эксперимента – выявить ключевые аспекты лексической системы языка и определить, какие слова составляют ядро лексики, а какие относятся к периферии. Анализ ядерной лексики, особенно в динамике через хронологические срезы, позволяет лучше понять и изучить процессы, происходящие в лексической системе языка. Выявление изменений, происходящих в ядре лексики (количественных и качественных), имеет первостепенное значение для исследования, так как именно эти изменения наиболее важны для системы языка. Параметрический анализ лексики (ПАЛ) является ценным инструментом для определения системного ядра лексики. ПАЛ включает четыре основных параметра: функциональный, синтагматический, деривационный и парадигматический. Только учет всех четырех параметров дает полное представление о положении слова в иерархии лексико-семантической системы языка, что позволяет исследователям точнее оценивать значимость и влияние отдельных лексических единиц. Статья также содержит детальный обзор параметрического анализа лексики – от подготовки словника и его конкорданса до выделения МинТеМ-лексики, вычисления веса каждого слова по указанным параметрам и анализа результатов исследования. Кроме того, статья описывает оригинальные подходы и решения, использованные в ходе работы, которые могут быть полезны для дальнейших исследований в области исторической лексикологии и изучения лексико-семантической системы современного русского языка. Полагаем, что данная статья внесет вклад в современную лексикологию и может служить основой для будущих исследований в области лексико-семантической системы языка.

Ключевые слова: историческая лексикология, параметрический анализ лексики, ядро, периферия, словоформа.

Abstract: the proposed article describes in detail an experiment on the systematic stratification of the chronological section of the vocabulary of the first quarter of the XXI century. The purpose of the experiment is to identify key aspects of the lexical system of the language and determine which words form the core of the vocabulary and which belong to the periphery. The analysis of core vocabulary, especially in dynamics through chronological sections, makes it possible to better understand and study the processes occurring in the lexical and semantic system of the language. The identification of changes occurring in the core of vocabulary (quantitative and qualitative) is of paramount importance for research, since these changes that are most important for the language system. Parametric Vocabulary Analysis (PAL) is a valuable tool for determining the system core of vocabulary. PAL includes four main parameters: functional, syntagmatic, derivational and paradigmatic. Only taking into account all four parameters gives a complete picture of the position of a word in the hierarchy of the lexical and semantic system of the language, which allows researchers to more accurately assess the significance and influence of individual lexical units. The article also contains a detailed overview of the parametric analysis



of vocabulary – from the preparation of a dictionary and its concordance to the allocation of MinTeM-vocabulary, calculating the weight of each word according to the specified parameters and analyzing the results of the study. In addition, the article describes the original approaches and solutions used in the course of the work, which may be useful for further research in the field of historical lexicology and the study of the lexical and semantic system of the modern Russian language. We believe that this article will contribute to modern lexicology and can serve as a basis for future research in the field of lexical language system.

Key words: *historical lexicology, parametric analysis of vocabulary, core, periphery, word form.*

Введение

Для понимания устройства и функционирования лексико-семантической системы современного русского языка необходимо исследовать и описать динамику параметрического ядра русской лексики в различных хронологических срезах.

В данной статье рассматривается эксперимент по лексико-семантической стратификации лексики русского литературного языка первой четверти XXI в.

Инструментом системной стратификации лексики служит ПАЛ – Параметрический Анализ Лексики, активно разрабатываемый Воронежской лингвистической школой.

Новизна исследования заключается в том, что стратификация лексики проводится на примере русской повествовательной (нарративной) прозы.

Объектом исследования является словник лексики русского языка первой четверти XXI в., состоящий из 13 918 слов.

Словник

Хронологический срез лексики (далее – словник) был составлен путем объединения отрывков из произведений современных русских писателей. При формировании словника мы руководствовались только двумя критериями – год создания литературного произведения и повествовательный стиль текста. В словник вошли повествовательные отрывки с различной тематикой из следующих произведений:

– обучение средневекового мальчика у деда-знахаря (1266 слов) – роман-житие «Лавр» (2012) [1, с. 9–12], автор Е. Г. Водолазкин;

– рассказ о жизни прадеда-лагерника (1339 слов) – исторический роман «Обитель» (2014) [2, с. 6–8], автор Захар Прилепин;

– история Большого каменного моста (1452 слова) – роман «Каменный мост» (2009) [3, с. 29–35], автор А. М. Терехов;

– ограбление инкассаторов участником войны в Афганистане (1438 слов) – роман «Ненастье» (2015) [4, с. 139–145], автор А. В. Иванов;

– первый бой главного персонажа – майора Жилина (1481 слово) – роман «Асан» (2008) [5, с. 76–79], автор В. С. Маканин;

– события вокруг смерти маленькой девочки (1328 слов) – роман «Будьте как дети» (2008) [6, с. 9–15], автор В. А. Шаров;

– детство главного персонажа во времена «вождя народов» в закрытом военном городке (1387 слов) – роман «Всё поправимо: Хроники частной жизни» (2006) [7, с. 11–17], автор А. А. Кабаков;

– описание жизни старшего брата до переезда в деревню (1413 слов) – роман «Елтышевы» (2009) [8, с. 44–51], автор Р. В. Сенчин;

– возвращение из тюрьмы бывшего студента-филолога (1364 слова) – реалистический роман «Текст» (2017) [9, с. 13–18], автор Д. А. Глуховский;

– студенческая жизнь советской девушки (1450 слов) из романа «Петровы в гриппе и вокруг него» (2016) [10, с. 380–386], автор А. Б. Сальников.

Конкорданс и лемматизация

Следующим этапом стало создание электронного конкорданса исследуемого словника. Конкорданс был сформирован с использованием программ текста PrProcess и Sentences, затем загружен в электронные таблицы Microsoft Excel. После обработки с помощью макроса конкорданс словника представляет собой упорядоченный список всех словоформ словника. Для каждой словоформы указаны контексты, в которых она используется в тексте.

Затем была проведена лемматизация – преобразование всех слов словника в их базовые формы: существительные приводятся к именительному падежу единственного числа, прилагательные – к именительному падежу единственного числа мужского рода, а глаголы, причастия и деепричастия – к инфинитиву.

Таким образом, предложение «Я сел на землю и сидел» [5, с. 78] в конкордансе словника выглядит следующим образом, например:

Словоформа	Лемма	Слово1	Слово2	Слово3	Слово4	Слово5	Слово6
Я	Я	V	сел	на	землю	и	сидел.
сел	СЕСТЬ	Я	V	на	землю	и	сидел.
на	НА	Я	сел	V	землю	и	сидел.
землю	ЗЕМЛЯ	Я	сел	на	V	и	сидел.
и	И	Я	сел	на	землю	V	сидел.
сидел.	СИДЕТЬ	Я	сел	на	землю	и	V

Лемматизировать словоформы можно как с помощью компьютерной программы текста Add.3_2018, так и в ручном режиме.

МинТеМ (Минимально-Текстуально Маркированная) лексика

Выделение МинТеМ (Минимально-Текстуально Маркированной) лексики осуществлялось с помощью «Программы выделения тематически маркированной лексики – ProTeMaL» (авторы: И. В. Попова, И. Е. Воронина, А. А. Кретов).

Для анализа использовался Индекс Текстуальной Маркированности слова (ИнТеМ), который рассчитывается как разность между частотным весом словоформы (Ч-весом) и ее весом по длине (Д-весом). Значение ИнТеМа варьируется от +1 до –1. Слова с ИнТеМ, близким к +1, отражают специфику текста, тогда как слова с ИнТеМ, приближающимся к –1, относятся к периферии текста и словаря. Слова с ИнТеМ, близким к 0, составляют ядро языка и являются наименее маркированными (рис. 1).

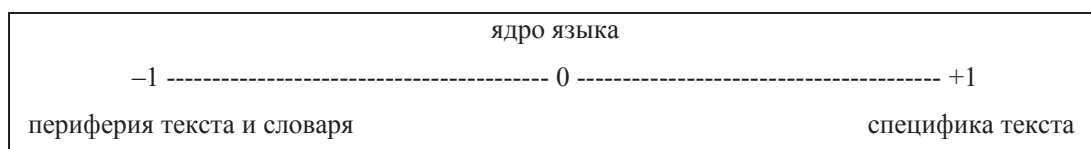


Рис. 1. МинТеМ-лексика

Граница, отделяющая МинТеМ-лексику от периферии, была определена средним значением ИнТеМа по модулю, которое в нашем случае составило 0,59992.

В исследовании рассматриваются только значимые части речи, такие как существительные, прилагательные и глаголы. В связи с этим из МинТеМ-лексики были исключены:

– служебные слова, наречия, местоимения, числительные, имена собственные, например: И, ТРИ, КОГДА, АЛЕКСАНДРОВСКИЙ, МАРИНА и т. д.;

– текстуально маркированная лексика, например: ГЛАВЛИКЕРВОДКА, ПЕСЬЕГОЛОВЕЦ, МОЛИТВОСЛОВИЦА и т. д.;

– все леммы с Частотой 1 (и Ч-весом 0).

В результате проведенной работы была отобрана МинТеМ-лексика, состоящая из 1066 слов, среди которых 520 существительных, 362 глагола, 182 прилагательных и 2 слова (ЗНАКОМЫЙ и СВЯТОЙ), употребляющихся как в значении существительного, так и прилагательного.

Дальнейшая работа в исследовании проводилась исключительно с использованием МинТеМ-лексики.

Функциональный анализ

В своем исследовании мы акцентируем внимание на объективной составляющей. Объективный показатель – это то, на что автор текста не может повлиять – длина словоформ, используемых в тексте. Функциональный вес каждого слова рассчитывался из средней длины словоформы.

Для корректного вычисления длины словоформы необходимо предварительно провести практическую транскрипцию всех словоформ слов, входящих в МинТеМ-лексику. С учетом особенностей общих

правил произношения (литературной нормы) необходимо последовательно заменить указанные в табл. 1 буквосочетания на соответствующие звуки. Например, слово МУРАВЬИ после транскрипции будет звучать как МУРАВЬИ, слово ЕЛКА преобразуется в ЙОЛКА, ЗВАТЬ станет ЗВАТ, а ЮНОША изменится на ЙУНОША.

Т а б л и ц а 1

Транскрипция звуков

Буквы	ье	ьс	ьи	ью	ья	#е	#с	#ю	#я	ь
Звуки	йэ	йо	йи	йу	йа	йэ	йо	йу	йа	Ø

– начало слова, Ø – пробел.

Разделив суммарную длину словоформы (число длины всех словоформ конкретного слова) на суммарную частоту словоформы (число частоты всех словоформ этого слова) и округлив результат до целого числа, мы рассчитали среднюю длину словоформы (среднее число длины всех словоформ конкретного слова). Например, слово ГОЛОС встречается в словнике 4 раза:

– «И говорил скрипучим древесным *голосом*» [1, с. 10];

– «Прадед быстро, в двух-трех фразах, хриплым и обширным своим *голосом* набрасывал какую-то картинку из прошлого» [2, с. 6–8];

– «Как плакал, был этот чужой *голос!*...» [5, с. 78];

– «В гостиной старательно пел *голосом* Джо Дассена пластиночный проигрыватель и качались парочки...» [10, с. 385].

Как видно из примеров, слово ГОЛОС представлено тремя словоформами ГОЛОСОМ (21 звук) и одной словоформой ГОЛОС (5 звуков). Их суммарная

длина составляет 26 звуков. Разделив эту сумму на количество словоформ (4) и округлив полученный результат (6,5) до целого числа, мы получили среднюю длину словоформы ГОЛОС 7 звуков.

Динамика распределения слов МинТеМ-лексики по средней длине словоформы представлена следующим графиком (рис. 2).

График демонстрирует интересную закономерность: подавляющее количество слов МинТеМ-лексики (701) имеют среднюю длину словоформы от 5 до 8 звуков, что свидетельствует о том, что эта длина является оптимальной для восприятия и использования в языке.

По формуле, предложенной В. Т. Титовым [11, с. 25]: $ВЕС = (\Sigma r - Pr) / \Sigma r$ (где Σr – сумма единиц всех рангов, Pr – сумма единиц от первого до данного), был рассчитан функциональный вес всех слов МинТеМ-лексики. Динамика функционального веса слов МинТеМ-лексики может быть представлена следующим графиком (рис. 3).

Как видно из графика, сначала представлены короткие слова, обладающие наибольшим весом, а затем более длинные. Самые длинные слова имеют вес 0, что свидетельствует о том, что чем короче слово, тем больший функциональный вес оно имеет и тем важнее оно для языка и его системы.

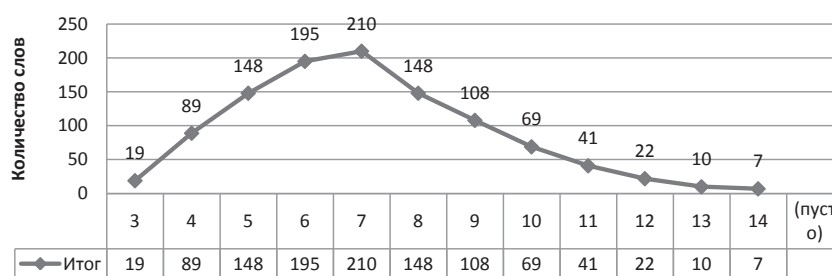


Рис. 2. Распределение МинТеМ-лексики по средней длине словоформы

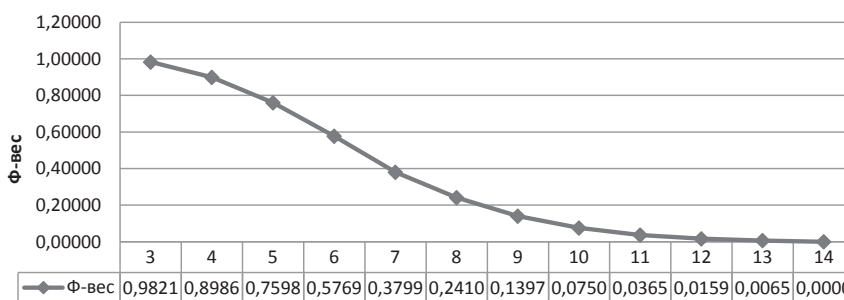


Рис. 3. Динамика функционального веса слов МинТеМ-лексики

Синтагматический анализ

Синтагматический вес каждого слова был рассчитан как среднее количество разных слов (не словоформ!), приходящихся на одну синтаксическую связь слова. Например, слово ВОЗДУХ встречается в словнике 4 раза:

1. «Он хотел лишь вдохнуть его лазурный **воздух** и, если получится, наконец увидеть бабушку» [1, с. 10].

ВДОХНУТЬ что? – **воздух**; **воздух** какой? – ЛАЗУРНЫЙ;

воздух чей? – ЕГО.

Предложение 1: С-связей 3, Слов 3.

2. «Тихим блаженством было ощущать все это – усилия крепких ног и плеч, объем груди, прохладу

воздуха, тяжесть тележки, ясность своего сознания» [4, с. 139].

ПРОХЛАДУ чего? – **воздуха**.

Предложение 2: С-связей 1, Слов 1.

3. «...грязь, тугая желтая глина, тяжелой каймой облеплявшая Мишкины галоши, на отцовы сапоги не цеплялась, он шел, будто по **воздуху**» [7, с. 12].

ШЕЛ по чему? – по **воздуху**.

Предложение 3: С-связей 1, Слов 1.

4. «В Лобне **воздух** был прозрачней, морозней; с неба тут сыпало холодной солью, секло щеки» [9, с. 13].

воздух какой? – ПРОЗРАЧНЕЕ, МОРОЗНЕЕ; **воздух** что делал? – БЫЛ.

Предложение 4: С-связей 2, Слов 3.

Итого разных слов: ВДОХНУТЬ, ЛАЗУРНЫЙ, ЕГО, ПРОХЛАДА, ИДТИ, ПРОЗРАЧНЫЙ, МОРОЗНЫЙ, БЫТЬ – всего 8 слов.

Итого синтаксических связей: *что?*, *какой?* *чей?* *чего?* *по чему?* *какой?* *что делал?* – всего 7 синтаксических связей.

Следовательно, у существительного ВОЗДУХ среднее количество слов на одну синтаксическую связь составляет 8/7 слова, т. е. С-показатель равен 1,14285 (1,14 округленно).

Аналогичным образом, были получены С-показатели для всех слов МинТеМ-лексики. Динамика распределения МинТеМ-лексики по С-показателям представлена следующим графиком (рис. 4).

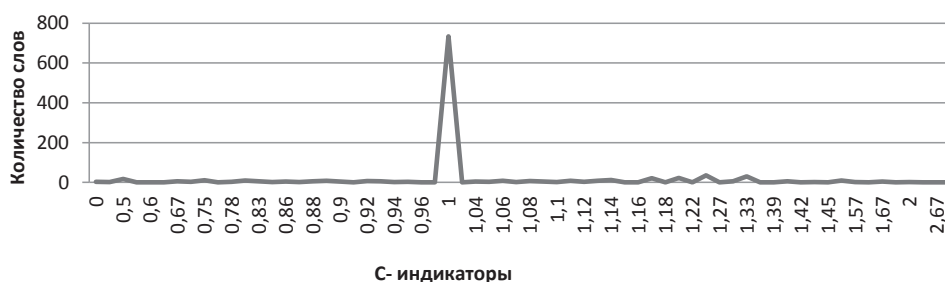


Рис. 4. Распределения МинТеМ-лексики по С-показателям

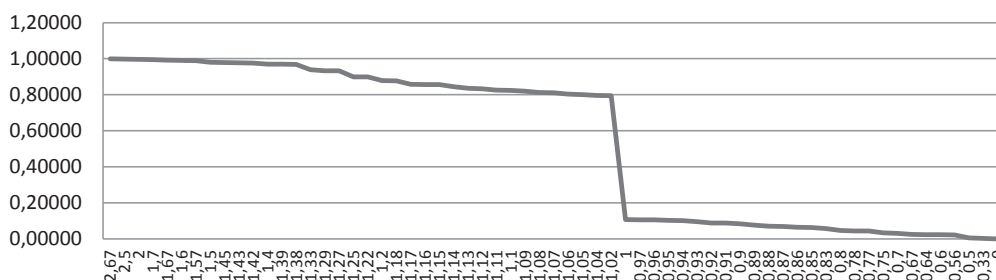


Рис. 5. Динамика синтагматического веса

Деривационный анализ

Установление деривационных связей проводилось по морфемному составу слова, который был взят из словаря БРУМС (Большой русский морфемный словарь) [12; 13].

Деривационная продуктивность (Д-показатель) каждого слова рассчитывался из количества слов, производных от него. Производным считается слово, которое отличается от производящего не менее чем на одну терминальную морфему – приставку или суффикс.

Для каждого слова надо установить, в какое количество слов входит его корень – от 1 до максимума и является ли оно производящим или производным. Например:

ГЛАЗ (глаз_ь) – корневое слово, но корень представлен один раз. Его Д-показатель 1.

График показывает, что большинство слов МинТеМ-лексики имеет значение С-показателя равное 1 (единице), т. е. для большинства слов характерно одно слово, приходящееся на синтаксическую связь.

По формуле В. Т. Титова был рассчитан синтагматический вес всех слов МинТеМ-лексики. Динамика синтагматического веса слов МинТеМ-лексики может быть представлена следующим графиком (рис. 5).

На графике видно, что слова с С-показателями (среднее количество слов, приходящихся на одну синтаксическую связь), превышающими единицу, имеют значительно больший синтагматический вес.

ПЕСНЯ (п>=сн'_а) – слово с суффиксом, но корень представлен один раз. Его Д-показатель также 1.

ПОКУПАТЬ (по-куп=А=тi_#) – слово с приставкой и суффиксом, но корень представлен один раз. Его Д-показатель тоже 1.

Следовательно, все слова, у которых корень слова представлен только один раз в составе МинТеМ-лексики, имеют Д-показатель 1.

У слов ВЫМОЛИТЬ (вы-мол=и=тi_#) и МОЛИТВА (мол=И=тв_а) корень МОЛ= представлен 2 раза, но они имеют равное количество терминальных морфем и от каждого из них ничего не образовано. Д-показатель каждого из однокоренных слов 1.

Пара слов УСПЕТЬ (у-сп>=0=тi_#) и УСПЕВАТЬ (у-сп<=vА=тi_#) является парадигмой степеней глагольного действия (с нулевым степенным суффиксом у глагола УСПЕТЬ) и является словоизменением, а

не словообразованием (не лексическая деривация). Следовательно, Д-показатель этих слов равен 1.

Другое дело пары ВЕК (в>к_ь) > ВЕЧНЫЙ (в>к=льн_ый), ВИСЕТЬ (вис=>т i_ь) > ЗАВИСЕТЬ (завИс=<=тi_#), ИГРАТЬ (иг=р=А=тi_#) > ВЫИГРАТЬ (выИ-иг=р=а=тi_#), КРАЙ (крАй_ь) > ОКРАИНА (о крАй=ин_а). Первое слово в каждой паре является производящим, а второе производным, так как отличается от производящего слова не менее чем на один аффикс. У производящих слов Д-показатель 2, а у производных Д-показатель 1.

Более сложная парадигма: УЧИТЬ (ук=И=тi_#) > УЧИТЬСЯ (ук=И=тi_# сын), УЧИТЕЛЬ (ук=И=тел_ь) > УЧИЛИЩЕ (ук=И=дл=искй_е). Глагол УЧИТЬ является производящим для всех слов в парадигме, поэтому его Д-показатель 4. Слова УЧИТЬСЯ и УЧИТЕЛЬ имеют одинаковое количество аффиксов и могут быть признаны производящими по отношению к слову УЧИЛИЩЕ, поэтому их Д-показатели 2. Только слово УЧИЛИЩЕ не имеет производных в этой парадигме, поэтому его Д-показатель 1.

Также в работе мы реконструируем производящие (даже если они и не представлены в тексте самостоятельными словами) и учитываем их в подсчетах, поскольку они представлены в составе производных.

Например: в парадигме НОСИТЬ (нос=И=тi_#) > ОТНОСИТЬСЯ (от-нос=И=тi_# сын) > ОТНОШЕНИЕ (от-нос=и=Е=н=ий_е) можно восстановить глагол **ОТНОСИТЬ** (от-нос=И=тi_#) и считать его присутствующим в деривационном ряде: НОСИТЬ (нос=И=тi_#) > **ОТНОСИТЬ** (от-нос=И=тi_#) > ОТНОСИТЬСЯ (от-нос=И=тi_# сын) > ОТНОШЕНИЕ (от-нос=и=Е=н=ий_е). В этом случае Д-показатели следующие: у слова НОСИТЬ – 4, ОТНОСИТЬСЯ – 2 и ОТНОШЕНИЕ – 1.

Аналогичным образом, были получены Д-показатели для всех слов МинТеМ-лексики.

По формуле В. Т. Титова был рассчитан деривационный вес всех слов МинТеМ-лексики. Размерность деривационного графа МинТеМ-лексики представлена следующим графиком (рис. 6).

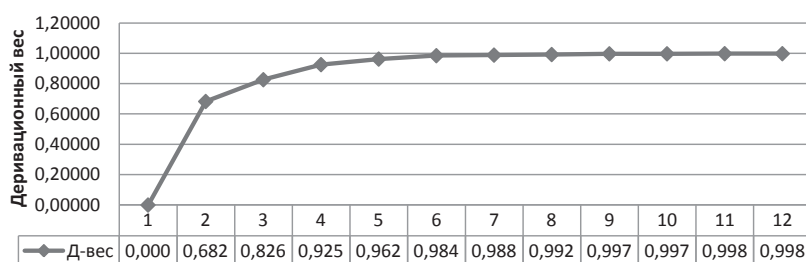


Рис. 6. Размерность деривационного графа

График показывает закономерность, что слова с Д-показателями (количество производных от слова), превышающими единицу, имеют значительно больший деривационный вес.

Парадигматический анализ

При определении парадигматического веса применялся более широкий подход, чем просто синонимический ряд, а именно: эквонимы (когипонимы с общим гиперонимом). Принимая во внимание многозначность слов, определение П-показателей проводилось с учетом отдельных значений слов (семем). П-показатели вычислялись по интегральному принципу, путем сложения количества эквонимов всех семем одного слова, образующих различные парадигмы.

Для выявления семантических множеств в работе использовался семантический метаязык (см.: [14; 15]) и лексикографические источники, такие как толковые словари.

Метаязык предназначен для семантической классификации, т. е. для обобщения, а не для различения

лексических значений. Также при определении П-показателей применялся более широкий подход, чем просто синонимический ряд, а именно: эквонимы (когипонимы с общим гиперонимом).

Например: значения БРЕЗЕНТОВЫЙ, ЛЕНТА, МАРЛЯ, ПОЛОТНО, САТИНОВЫЙ, ТРЯПОЧКА объединены в единую семантическую парадигму гиперонимом «ТКАНЬ» и имеют в ней равный П-показатель – 6.

При описании значений слов с помощью метаязыка используются 9 классов семантических констант: «Стихия», «Натурфакт», «Растения», «Животные», «Человек», «Артефакт», «Ментефакт», «Социум», «Сверхъестественное» и «Прочее» и 18 классов семантических переменных.

Распределение значений слов МинТеМ-лексики по классам семантических констант представлено следующим графиком (рис. 7).

Как видно из графика, количество значений слов (1204 семем) превосходит число слов МинТеМ-лексики (1066).

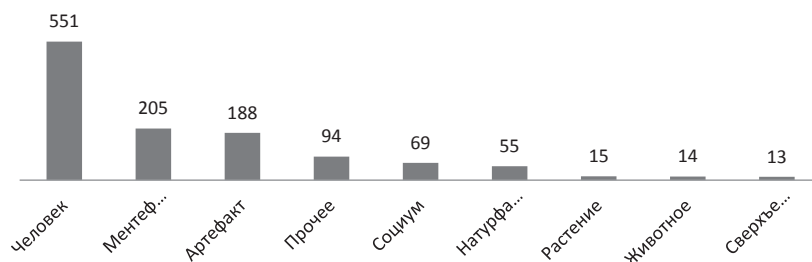


Рис. 7. Распределение МинТеМ-лексики по константам

После определения принадлежности значений слов к классу констант работа проводилась в рамках каждого класса. Для определения П-показателей необходимо выявить семантические множества, основанные на тождестве как можно большего числа метаслов. Например:

ГЕРБ: социум – знак

ЗВЕЗДА-1: социум – знак

ПОГОН: социум – знак

Как видно, значения ГЕРБ, ЗВЕЗДА семема-2 и ПОГОН объединены семантической константой «социум» и переменной «знак». Для верификации парадигматических связей между этими значениями необходимо использовать лексикографические источники.

ГЕРБ – «отличительный **знак** государства, города, сословия, рода и т. п., изображаемый на флагах, монетах, печатях и т. п.» [16, с. 306], например: «*А вот упаковка с валютой – президенты, **гербы**, орлы, дворцы какие-то*» [4, с. 143].

ЗВЕЗДА-1 – «геометрическая фигура с остроко-
нечными выступами, равномерно расположенными по окружности, фигура с лучами, исходящими от центра || **Знак** отличия, орден, имеющий такую форму» [16, с. 600], например: «*Еще и солнце!.. Оно так и играло на **звезде** моего левого плеча, зазывая снайперов...*» [5, с. 76].

ПОГОН – «наплечные **знаки** различия в виде нашивной или накладной полосы на военной и другой форменной одежде» [17, с. 168], например: «*...влажной тряпочкой он проходил по майорским звездам на серебряных инженерских **погонах**...*» [7, с. 13].

Очевидно, что во всех словарных определениях присутствует слово «ЗНАК», которое является для значений ГЕРБ, ЗВЕЗДА-1 и ПОГОН гиперонимом и объединяет их в единую парадигму. Все члены семантической парадигмы имеют равные П-показатели. Таким образом, значения ГЕРБ, ЗВЕЗДА-1 и ПОГОН имеют П-показатель 3.

В процессе описания слова ЗВЕЗДА было выявлено второе значение ЗВЕЗДА-2: ментефакт – форма – вид. Это значение имеет П-показатель 7. Сложив П-показатели обеих семем (ЗВЕЗДА-1 и ЗВЕЗДА-2), мы получили П-показатель семантемы ЗВЕЗДА – 10.

Аналогичным образом были получены П-показатели для всех слов МинТеМ-лексики.

По формуле В. Т. Титова был рассчитан парадигматический вес всех слов МинТеМ-лексики. Динамика парадигматического веса слов МинТеМ-лексики может быть представлена следующим графиком (рис. 8).



Рис. 8. Динамика парадигматического веса

График демонстрирует закономерность: чем выше у слова П-показатель, тем большим парадигматическим весом оно обладает.

Результаты

Суммировав функциональный, синтагматический, деривационный и парадигматический веса, мы

получили интегральный вес каждого слова МинТеМ-лексики. Округление интегрального веса до целого числа позволило стратифицировать МинТеМ-лексику по ее системному весу. Были выделены: малое ядро (8 слов), среднее ядро (85 слов), большое ядро (356 слов), ядро словника (503 слова) и периферия (114 слов).

Результаты лексико-семантической стратификации (малое и среднее ядра) МинТеМ-лексики хронологического среза лексики первой четверти XXI в. представлены в табл. 2.

Т а б л и ц а 2

Малое и среднее ядра МинТеМ-лексики по интегральному весу

Лемма	Ф-вес	С-вес	Д-вес	П-вес	И-вес	Окр.	Ядра
1	2	3	4	5	6	7	8
ИДТИ	0,98218	0,83302	0,99719	0,99906	3,81145	4	Малое ядро
МАТЬ	0,89869	0,96904	0,92589	0,92402	3,71764	4	
ДАЧА	0,89869	0,93902	0,92589	0,90150	3,66510	4	
ЖИТЬ	0,89869	0,79550	0,96248	0,90150	3,55817	4	
БАБА	0,89869	0,87805	0,82645	0,94465	3,54784	4	
ДОМ	0,89869	0,81051	0,92589	0,90150	3,53659	4	
ГОРА	0,89869	0,80300	0,82645	0,99531	3,52345	4	
ОТЕЦ	0,89869	0,85647	0,82645	0,92402	3,50563	4	
ЖЕНА	0,89869	0,82552	0,82645	0,92402	3,47468	3	Среднее ядро
ЖИЗНЬ	0,89869	0,80300	0,82645	0,84709	3,37523	3	
ХОДИТЬ	0,75985	0,80019	0,99812	0,81051	3,36867	3	
ДОЧЬ	0,89869	0,83583	0,68293	0,92402	3,34147	3	
ВЕК	0,98218	0,82552	0,68293	0,84709	3,33772	3	
ДАТЬ	0,89869	0,85741	0,96248	0,60788	3,32646	3	
ИМЕТЬ	0,89869	0,93246	0,68293	0,78143	3,29551	3	
ПЕТЬ	0,98218	0,89962	0,68293	0,71201	3,27674	3	
РЕКА	0,89869	0,93340	0,68293	0,75985	3,27487	3	
РОДИТЬ	0,57692	0,89962	0,92589	0,84709	3,24952	3	
КРИК	0,75985	0,93902	0,82645	0,71201	3,23733	3	
ПИТЬ	0,98218	0,85741	0,82645	0,56098	3,22702	3	
СЕМЬЯ	0,75985	0,83583	0,68293	0,92402	3,20263	3	
ВОДА	0,89869	0,84428	0,68293	0,75985	3,18575	3	
СТУПАТЬ	0,57692	0,82552	0,92589	0,81051	3,13884	3	
ЦЕНТР	0,57692	0,84428	0,82645	0,88180	3,12945	3	
ЗВАТЬ	0,75985	0,81238	0,98874	0,56098	3,12195	3	
БРОСИТЬ	0,57692	0,84428	0,68293	0,99625	3,10038	3	
БОРОДА	0,57692	0,85741	0,68293	0,96998	3,08724	3	
МЕСТО	0,75985	0,79550	0,68293	0,83771	3,07599	3	
ВИДЕТЬ	0,75985	0,81051	0,82645	0,65666	3,05347	3	
СКАЗАТЬ	0,57692	0,79550	0,96248	0,71201	3,04691	3	
ВЕЗТИ	0,75985	0,89962	0,82645	0,56098	3,04690	3	
ПАРА	0,89869	0,83583	0,68293	0,60788	3,02533	3	
ОКАЗАТЬСЯ	0,24109	0,93340	0,96248	0,88180	3,01877	3	
ЧЕЛОВЕК	0,57692	0,79456	0,68293	0,94465	2,99906	3	
ТЕЧЬ	0,57692	0,96998	0,68293	0,75985	2,98968	3	
ОСТАВИТЬ	0,37992	0,85741	0,96248	0,78143	2,98124	3	
ШКОЛА	0,75985	0,85741	0,68293	0,65666	2,95685	3	
ДЕВОЧКА	0,37992	0,93902	0,68293	0,94465	2,94652	3	
ЛИЦО	0,89869	0,10694	0,96248	0,96998	2,93809	3	
СТАТЬ	0,89869	0,07598	0,96248	0,99812	2,93527	3	
ТЯНУТЬ	0,57692	0,80300	0,68293	0,84146	2,90431	3	
ЗАНЯТЬСЯ	0,37992	0,93902	0,92589	0,65666	2,90149	3	
БЫТЬ	0,89869	0,07036	0,96248	0,96717	2,89870	3	
БЛЕСК	0,57692	0,87805	0,82645	0,60788	2,88930	3	
ДВИНУТЬ	0,57692	0,87805	0,82645	0,60788	2,88930	3	
ДУХ	0,98218	0,10694	0,82645	0,96998	2,88555	3	

Окончание табл. 2

1	2	3	4	5	6	7	8
КУХНЯ	0,75985	0,87805	0,68293	0,56098	2,88181	3	
ДЕЛАТЬ	0,57692	0,80300	0,82645	0,65666	2,86303	3	
НАЙТИ	0,75985	0,80300	0,82645	0,43152	2,82082	3	
САД	0,89869	0,06379	0,96248	0,88180	2,80676	3	
ЗУБ	0,89869	0,93340	0,00000	0,96998	2,80207	3	
СВОД	0,57692	0,87805	0,82645	0,48405	2,76547	3	
СЫН	0,98218	0,85741	0,00000	0,92402	2,76361	3	
МАМА	0,89869	0,93902	0,00000	0,92402	2,76173	3	
МУЖ	0,89869	0,10694	0,82645	0,92402	2,75610	3	
КРЫША	0,75985	0,10694	0,98499	0,90150	2,75328	3	
РОДИТЕЛИ	0,24109	0,89869	0,68293	0,92402	2,74673	3	
ПОТ	0,89869	0,85741	0,00000	0,96998	2,72608	3	
СТЕПЬ	0,89869	0,93902	0,00000	0,88180	2,71951	3	
ОСТАТЬСЯ	0,24109	0,81895	0,68293	0,96904	2,71201	3	
ПУЛЯ	0,89869	0,93902	0,68293	0,18480	2,70544	3	
ШЕЯ	0,89869	0,83583	0,00000	0,96998	2,70450	3	
ДЕД	0,98218	0,10694	0,68293	0,92402	2,69607	3	
ЖИВОЙ	0,89869	0,10694	0,82645	0,84709	2,67917	3	
НОЧЬ	0,89869	0,10694	0,82645	0,84709	2,67917	3	
ЧИТАТЬ	0,57692	0,98874	0,82645	0,28236	2,67447	3	
РУКА	0,89869	0,10694	0,68293	0,96998	2,65854	3	
ПОЛ	0,89869	0,85741	0,00000	0,90150	2,65760	3	
КЛАСТЬ	0,75985	0,08818	0,98499	0,81051	2,64353	3	
ТЯГА	0,89869	0,10694	0,92589	0,71201	2,64353	3	
УЗНАТЬ	0,57692	0,85741	0,92589	0,28236	2,64258	3	
ВСТАТЬ	0,75985	0,10694	0,82645	0,94465	2,63789	3	
ИМЯ	0,75985	0,84428	0,92589	0,09944	2,62946	3	
ДОСКА	0,75985	0,85741	0,82645	0,18480	2,62851	3	
ВЫЙТИ	0,75985	0,10694	0,82645	0,92402	2,61726	3	
ВЫЗВАТЬ	0,57692	0,84428	0,82645	0,36210	2,60975	3	
ЛЕС	0,89869	0,82552	0,00000	0,88180	2,60601	3	
УЙТИ	0,89869	0,07036	0,82645	0,81051	2,60601	3	
РАСТИ	0,57692	0,82552	0,82645	0,36210	2,59099	3	
ДЕЛО	0,89869	0,10225	0,92589	0,65666	2,58349	3	
ВЫХОД	0,57692	0,10694	0,99812	0,90150	2,58348	3	
ПАРЕНЬ	0,75985	0,87805	0,00000	0,94465	2,58255	3	
РУЧКА	0,75985	0,82364	0,00000	0,99719	2,58068	3	
ВОРОТА	0,57692	0,96998	0,92589	0,09944	2,57223	3	
БАШНЯ	0,75985	0,89962	0,00000	0,90150	2,56097	3	
ДОЖДЬ	0,89869	0,89962	0,00000	0,75985	2,55816	3	
ДЕВИЦА	0,57692	0,10694	0,92589	0,94465	2,55440	3	
БРАТ	0,75985	0,85741	0,00000	0,92402	2,54128	3	
СНЯТЬ	0,89869	0,10694	0,68293	0,84709	2,53565	3	
РОДНОЙ	0,57692	0,10694	0,92589	0,92402	2,53377	3	
КРУЖКА	0,57692	0,84428	0,92589	0,18480	2,53189	3	
ГЛИНА	0,75985	0,87805	0,00000	0,88180	2,51970	3	
ДРУГ	0,75985	0,87711	0,68293	0,18480	2,50469	3	

Выводы

Результаты проведенного эксперимента в рамках исследования лексической системы русского языка позволяют сделать следующие выводы:

– успешная лексико-семантическая стратификация лексики русского языка первой четверти XXI в. подтверждает правильность выбора методики и инструментов, использованных в исследовании (выделены: малое ядро, среднее ядро, большое ядро, ядро словника и периферия);

– получена ценная информация о месте (весе) каждого слова словника в системе языка;

– информация и результаты, полученные в ходе проведения эксперимента, могут быть полезны в дальнейших изысканиях, направленных на исследование лексико-семантической системы русского языка.

Думается, что дальнейшая работа должна быть направлена на проведение подобных экспериментов на словниках лексики первой четверти XX в., затем первой четверти XIX и т. д. Анализ и сопоставление данных по нескольким хронологическим срезам позволит выявить общие тенденции в динамике развития лексико-семантической системы русского языка.

ЛИТЕРАТУРА

1. Водлазкин Е. Г. Лавр. М. : АСТ, 2012. 448 с.
2. Прилепин З. Обитель. М. : Редакция Елены Шубиной: АСТ, 2014. 745 с.
3. Терехов А. М. Каменный мост. М. : АСТ: Редакция Елены Шубиной, 2021. 832 с.
4. Иванов А. В. Ненастье. М. : АСТ: Редакция Елены Шубиной, 2019. 640 с.
5. Маканин В. С. Асан. М. : Эксмо, 2017. 218 с.
6. Шаров В. А. Будьте как дети. М. : АСТ: Редакция Елены Шубиной, 2017. 448 с.
7. Кабаков А. А. Все поправимо : Хроники частной жизни. М. : АСТ: Редакция Елены Шубиной, 2023. 603 с.
8. Сенчин Р. В. Елтышевы. М. : Эксмо, 2009. 320 с.
9. Глуховский Д. А. Текст. М. : АСТ, 2017. 320 с.
10. Сальников А. Б. Петровы в гриппе и вокруг него. М. : АСТ, 2017. 411 с.
11. Титов В. Т. Принципы квантитативной лексикологии (на примере романских языков) : автореф. дис. ... д-ра филол. наук. Тверь, 2005. 36 с.
12. Кретов А. А. Большой русский морфемный словарь в обучении иностранцев // Русское слово в мировой культуре : материалы X Конгресса Междунар. ассоциации преподавателей русского языка и литературы. Санкт-Петербург, 30 июня – 5 июля 2003 г. СПб. : Политехника, 2003. С. 273–283.
13. Кретов А. А. БРУМС как электронный словарь и база данных // От языковых машинных фондов к лингвистическим корпусам : памяти В. М. Андриященко : тез. Всерос. конф. Москва, 28–29 сентября 2018 г. / Институт русского языка имени В. В. Виноградова. М.,

2018. С. 38–42. URL: <http://lcl.srcc.msu.ru/library/abstracts.pdf>

14. Кретов А. А., Подтележникова Е. Н. К проблеме создания метаязыка семантического описания лексики // Вестник Воронеж. гос. ун-та. Сер.: Лингвистика и межкультурная коммуникация. 2014. № 4. С. 5–10.

15. Кретов А. А., Подтележникова Е. Н. Опыт метаязыкового описания семантических констант русской лексики // Вестник Воронеж. гос. ун-та. Сер.: Лингвистика и межкультурная коммуникация. 2016. № 4. С. 5–10.

16. Словарь русского языка : в 4 т. / АН СССР, Ин-т рус. яз. ; под ред. А. П. Евгеньевой. 3-е изд. стереотип. М. : Русский язык, 1985–1988. Т. 1. А–Й. 1985. 696 с.

17. Словарь русского языка : в 4 т. / АН СССР, Ин-т рус. яз. ; под ред. А. П. Евгеньевой. 3-е изд., стереотип. М. : Русский язык, 1985–1988. Т. 3. П–Р. 1987. 752 с.

REFERENCES

1. Vodolazkin E. G. *Lavr* [Laurus]. Moscow: AST, 2012. 448 p.
2. Prilepin Z. *Obitel'* [The Monastery]. Moscow: Editorial office of Elena Shubina: AST, 2014. 745 p.
3. Terekhov A. M. *Kamennyj most* [The Stone Bridge]. Moscow: Editorial office of Elena Shubina: AST, 2021. 832 p.
4. Ivanov A. V. *Nenast'e* [Nenastye]. Moscow: Editorial office of Elena Shubina: AST, 2019. 640 p.
5. Makanin V. S. *Asan* [Asan]. Moscow: Eksmo, 2017. 218 p.
6. Sharov V. A. *Bud'te kak deti* [Be As Children]. Moscow: Editorial office of Elena Shubina: AST, 2017. 448 p.
7. Kabakov A. A. *Vsyo popravimo: Hroniki chastnoj zhizni* [Everything is fixable: Chronicles of private life]. Moscow: Editorial office of Elena Shubina: AST, 2023. 603 p.
8. Senchin R. V. *Elytshevy* [The Elytshevs]. Moscow: Eksmo, 2009. 320 p.
9. Gluhovskij D. A. *Tekst* [Text]. Moscow: AST, 2017. 320 p.
10. Salnikov A. B. *Petrovy v grippe i vokrug nego* [Petrov's Flu]. Moscow: AST Publishing House, 2017. 411 p.
11. Titov V. T. *Principy kvantitativnoj leksikologii (na primere romanskix yazykov)* [Principles of quantitative lexicology (on the example of Romance languages)]: abstract. diss. ... Doctor of Philology. Tver, 2005. 36 p.
12. Kretov A. A. *Bol'shoj russkij morfemnyj slovar' v obuchenii inostrancev* [Big Russian morphemic dictionary in the teaching of foreigners]. In: *Russkoe slovo v mirovoj kul'ture: Materialy X Kongressa Mezhdunarodnoj assotsicii prepodavatelej russkogo yazyka i literatury*. St. Petersburg: Polytechnic, 2003. Pp. 273–283.
13. Kretov A. A. *BRUMS kak elektronnyj slovar' i baza dannyh* [BRUMS as an electronic dictionary and database]. In: *Tezisy vserossijskoj konferencii «Ot yazykovykh mashinnyh fondov k lingvisticheskim korpusam: pamyati V. M. An-*

dryushchenko». The Vinogradov Institute of the Russian Language. Moscow, 2018. Pp. 38–42. Available at: <http://lcl.srcc.msu.ru/library/abstracts.pdf>

14. Kretov A. A., Podtelezhnikova E. N. *K probleme sozdaniya metazyka semanticheskogo opisaniya leksiki* [On the problem of creating a metalanguage of semantic description of vocabulary]. In: *Vestnik Voronezh. gos. un-ta. Seriya: Lingvistika i mezhkulturnaja kommunikatsija*. 2014. No. 4. Pp. 5–10.

15. Kretov A. A., Podtelezhnikova E. N. *Opyt metazykovogo opisaniya semanticheskikh konstant russkoj leksiki* [Experience of metalanguage description of semantic

constants of Russian vocabulary]. In: *Vestnik Voronezh. gos. un-ta. Seriya: Lingvistika i mezhkulturnaja kommunikatsija*. 2016. No. 4. Pp. 5–10.

16. *Slovar' russkogo yazyka: v 4 t.* [Dictionary of the Russian language: in 4 volumes]. Edited by A. P. Evgenieva. 3rd ed. stereotype. Moscow: Russian language, Vol. I. A–Й. 1985–1988. 696 p.

17. *Slovar' russkogo yazyka: v 4 t.* [Dictionary of the Russian language: in 4 volumes]. Edited by A. P. Evgenieva. 3rd ed. stereotype. Moscow: Russian language, 1987. Vol. III. П–Р. 1987. 752 p.

Воронежский государственный лесотехнический университет имени Г. Ф. Морозова

Ермаков С. А., кандидат филологических наук, доцент кафедры иностранных языков
E-mail: sergey-ermakov.1811@mail.ru

Поступила в редакцию 19 февраля 2025 г.

Принята к публикации 26 марта 2025 г.

Для цитирования:

Ермаков С. А. Стратификация хронологического среза лексики (первая четверть XXI века) // Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация. 2025. № 2. С. 131–141. DOI: <https://doi.org/10.17308/lic/1680-5755/2025/2/131-141>

Voronezh State University of Forestry and Technologies named after G. F. Morozov

Ermakov S. A., Candidate of Philology, Associate Professor of the Foreign Languages Department
E-mail: sergey-ermakov.1811@mail.ru

Received: 19 February 2025

Accepted: 26 March 2025

For citation:

Ermakov S. A. Stratification of the chronological section of vocabulary (first quarter of the XXI century). *Proceedings of Voronezh State University. Series: Linguistics and Intercultural Communication*. 2025. No. 2. Pp. 131–141. DOI: <https://doi.org/10.17308/lic/1680-5755/2025/2/131-141>