

# Ментальные действия и предметы в пространстве научного дискурса\*

Д. А. Девяткин, Ю. М. Кузнецова

Федеральное государственное учреждение Федеральный исследовательский центр «Информатика и управление» Российской академии наук, г. Москва, Россия

**Аннотация.** В работе исследуется возможность представления научного дискурса как ментального пространства. Его объектами являются ментальные действия и ментальные предметы, под которыми предложено понимать оформляемое с помощью дефинитивных конструкций содержание, выражающее основные идеи научного текста. На материале более чем 7000 первичных научных публикаций показано, что состав и выраженность таких связей специфичны для текстов различной предметной принадлежности, что свидетельствует о чувствительности данного метода к особенностям интеллектуальных средств исследовательской и коммуникативной деятельности в той или иной дисциплине. Отображение публикаций в виде содержащихся в них констелляций ментальных действий и ментальных предметов может рассматриваться как одно из средств оптимизации интеллектуального анализа научного текста.

**Ключевые слова:** интеллектуальный анализ текста, научный дискурс, научный текст, ментальные действия, ментальные предметы.

DOI 10.14357/20718594200106

## Введение

Ранее [1, 2] была показана возможность формализованного представления научного дискурса как пространства ментальных действий, под которыми предлагается понимать действия, направленные на сознательное изменение (конструирование новых или реконструирование уже имеющихся) идеальных объектов. Ментальные действия осуществляются с помощью набора интеллектуальных операций, отвечающих как внутренним условиям (требованиям задачи и намерениям автора), так и внешним (научной традиции). К настоящему времени создан метод идентификации ментальных действий с опорой на текстовые средства их выражения – речевые шаблоны (формальные

описания множеств разноуровневых лингвистических признаков, характерных для реализации определенных видов ментальных операций в тексте) [3]. Привлечение пространственной метафоры для моделирования научного дискурса подразумевает использование понятий метрики и топологии, благодаря чему выявляемым в нем предметам может быть приписано определенное пространственное положение, а статистические связи между ними трактуются в категориях близости-удаленности. В терминах, принятых в настоящем исследовании, пространственная метафора позволяет отражать содержание научных дисциплин в виде многомерных констелляций ментальных действий, состав, конфигурация и сила связей в которых характеризуют общее и специфическое в процессе научного познания.

Работа выполнена при частичной поддержке РФФИ, грант № 17-29-07049.

✉ Кузнецова Юлия Михайловна. E-mail: kuzjum@yandex.ru

## 1. Выявление ментальных действий в научных текстах

Описанная в работе [1] процедура позволяет выявлять в научных текстах ментальные действия и связи между ними. Объектами классификации в этой процедуре являются клаузы. Для построения признаковых описаний клауз использовался метод реляционно-ситуационного анализа [4]. В качестве признаков предикатного слова использовался идентификатор его словарной статьи и векторное представление, полученное с помощью модели Fasttext [5]. В качестве признаков аргументов использовались идентификаторы их ролей и векторные представления, также построенные с помощью Fasttext.

Из-за небольшого объема и несбалансированности обучающего корпуса выявление реализаций ментальных действий в текстах выполнялось в два шага. На первом выявлялись гиперкатегории ментальных действий, для этого использовалась нейронная сеть со сверточным и рекуррентным слоями GRU (Gated Recurrent Unit) [6]. Затем строилась композиция из нескольких таких сетей, причем каждая обучалась на отдельном, отобранном случайным образом подмножестве признаков. Результат работы этой композиции определялся путем голосования (soft-voting). На втором шаге с помощью сверточной нейронной сети производилось уточнение категорий ментальных действий.

Связи между текстовыми категориями вычисляются как статистические показатели совместной встречаемости соответствующих им текстовых категорий и имеют две основные характеристики: вес, отражающий частотность данной связи в текстах, и направленность связи, определяющая следование «если в тексте обнаруживается категория *A*, то с вероятностью *p* будет обнаружена и категория *B*». Характеристика направленности связи является «пространственной» сама по себе, а вес связи в пространственной интерпретации трактуется как мера близости: близкие точки в пространстве связаны сильнее, далекие – слабее.

С применением данной процедуры было показано, что в составе и конфигурации констелляций ментальных действий, восстанавливаемых для различных дисциплин, выделяются инвариантные и вариативные компоненты. В инвариантных отражаются фундаментальные закономерности научного познания и общие

нормы научной коммуникации. Вариативные компоненты ментальных схем отражают предметную и жанровую специфику анализируемых текстов. Была продемонстрирована также чувствительность примененной процедуры к качеству научных текстов, то есть уровню их соответствия содержательным и формальным требованиям к научной публикации определенного типа и предметной принадлежности, при критическом снижении которого ментальные схемы претерпевают структурную деформацию [1, 7].

## 2. Понятие ментального предмета и его представленность в научном тексте

Описанное представление о ментальном пространстве науки как системе познавательных действий может быть дополнено посредством введения новых компонентов – ментальных предметов – как того, на что направлены ментальные действия в процессе исследовательской деятельности и в ходе научной коммуникации. В ткани научного текста ментальные предметы должны иметь собственную представленность, что технически означает существование специальных средств для их текстового оформления, позволяющего идентифицировать содержание как ментальный предмет.

Известно, что особый статус предмету придает его выделенность по отношению к тому содержанию, которое его окружает, к нему не относится и является для него фоном. Для описания закономерностей выбора и организации содержания, которое предстает как предмет на фоне, в когнитивной лингвистике применяются те же объяснительные принципы (отношение фигуры и фона, фокусировка внимания и его распределение), что и в психологии восприятия, а конкретные механизмы, обеспечивающие действие этих принципов на языковом материале, исследуются разными авторами под такими названиями, как фокусирование, высвечивание, профилирование, перспективизация, конструирование и т.д. [8, 9].

Прагматически выдвигание содержания в позицию ментального предмета может рассматриваться как один из приемов управления вниманием читателя, направленных на создание оптимальных условий понимания текста. С помощью специальных средств автор осу-

ществляет смысловое или логическое усиление части текста, или интенсификацию [10, с. 62], благодаря чему речевая единица приобретает качества особой выделенности, подчеркнутости, актуализированности [11, с. 141]. Близкими являются понятия эмфатичности (языковое указание на степень эмоциональной маркированности смыслового содержания высказывания и интенциональности текста [12, с. 35]) и экспрессивности (усиление выразительности высказывания с целью воздействия на восприятие адресата). Под экспрессивностью научного дискурса предлагается понимать выраженность акцентуализации, конкретизации мыслей, логического подчеркивания, выделения, усиления аргументированной мысли автора с целью активизации внимания читателя [13].

В исследовании смыслового акцентирования Г. И. Климовской [14] показано, как с помощью расстановки смысловых ударений автор достигает такого распределения семантических элементов текста на «смысловый авансцене» или на «заднем плане», которое отражает его представление о степени их важности и актуальности, и позволяет организовать «режиссуру» восприятия текста читателем, его наиболее однозначное и эффективное продвижение по всем смысловым уровням текста. В качестве свойственного научной речи способа акцентирования, обеспечивающего «приподнимание» предмета над фоновым содержанием, Г. И. Климовская называет определение. Для нас это означает, что именно определению, или дефиниции, может принадлежать в научном тексте роль средства «фигурирования», то есть оформления ментальных предметов. Другие способы акцентирования, такие, как восклицательные конструкции, риторические вопросы, вопросно-ответные комплексы, диалогические единства, лексико-синтаксические и лексические повторы, характерны, прежде всего, для устной научной и научно-популярной речи [15], но имеют малую представленность в текстах первичных научных публикаций. Поэтому для решения задач формализации научного дискурса в его исходной форме более перспективным представляется привлечение в качестве маркера выдвижения содержания в позицию ментального предмета дефиниций, являющихся одним из видов ментальных действий и выражающихся в тексте с помощью шаблонов, которые, как и шаблоны других

ментальных действий, также могут автоматически распознаваться.

### 3. Дефиниция как текстовая форма ментального предмета

По сравнению с другими ментальными действиями дефиниция обладает существенной спецификой. Имея строго фиксированный функционал, связанный с соответствующим компонентом научного познания, остальные ментальные действия, как правило, локализируются в определенных частях научного текста, в то время как действие дефинирования полифункционально, а выражающие его высказывания распределяются в тексте в зависимости от ряда обстоятельств гносеологической и коммуникативной природы.

Среди выполняемых дефиницией функций наиболее значимыми для научной речи являются: раскрытие темы (определение основных понятий исследования, характеристика изучаемых объектов, их идентификация, репрезентация знаний об описываемых объектах), дифференциация понятий, аргументация (выражение авторского мнения посредством отбора упомянутых в дефиниции признаков понятия), представление предшествующего знания, введение нового знания (представление результатов исследования, формулирование выводов), связь старого и нового знания и др. [16]. Множественности функций дефиниции в научном тексте соответствует разнообразие лексических, грамматических и синтаксических средств их выражения, формирующих дискурсивные шаблоны дефинитивных конструкций; так, для русского языка насчитывается не менее сорока вариантов схем построения дефинитивных высказываний [17].

В силу своей полифункциональности дефиниция обладает свойством свободно сочетаться с другими ментальными действиями:

- «Описание комплекса наиболее важных дифференциальных признаков нового объекта»: *Органоизвестковый сапропель оз. Анжелы является источником кальция (кратность накопления -20,2), стронция (6,16), бора (4,60), молибдена (4,60), азота (4,40) и титана (2,91).*
- «Определение места нового объекта в системе известных явлений»: *Другие исторические науки изучают результаты деятельности*

людей и «общественное сознание», а биографика – непосредственно человека в истории, во всех проявлениях его существования, в том числе его психический мир.

- «Описание методики эксперимента»: В экспериментах использовался 96% раствор медицинского спирта, который, как и водные растворы, относится к жидкостям с быстрой релаксацией зарядов.

- «Подтверждение или отвержение исходной гипотезы на основе интерпретации опытных данных»: Таким образом, полученные данные дают основания для подтверждения гипотезы о зависимости выраженности нейрокогнитивных нарушений от возраста дебюта заболевания: поздняя манифестация расстройств психического уровня соотносится с более мягким нейрокогнитивным дефицитом...

Такая удобосочетаемость дефиниции в предельном случае могла бы привести к тому, что в научном дискурсе упоминание каждого из ментальных действий сопровождалось бы указанием на его предмет. Отвечая требованиям точности и однозначности научной речи, такое положение, однако, нарушало бы универсальный принцип «лингвистической экономии» [18], в общем виде выражающийся в тенденции к минимизации речевых усилий, а для научной речи – в соблюдении правила «минимального толкования», то есть экспликации только необходимых основополагающих признаков и упоминания о таких при знаках, которые без ущерба для достижения коммуникативной цели могут пребывать в статусе фоновых [19]. Поэтому в реальности акцентированные компоненты научного текста, содержащие информацию о ментальных предметах, распределяются в тексте не равномерно, а сопровождают описания таких ментальных действий, которые имеют принципиальное значение для реализации замысла автора. Именно благодаря этому, происходит «расслоение» текста, когда акцентированная «фигура», то есть ментальный объект, «приподнимается» над «фоном». Образующие же «фон» исходные понятия (указывает М. П. Котурова) относятся к принятому автором и читателям знанию, они не определяются и не развиваются в данном тексте [20].

Таким образом, выдвигание с помощью действия дефинирования какого-либо содержания в позицию ментального предмета может выступать в качестве маркера значимости,

которое это содержание имеет для реализации намерений автора по изменению представлений читателя об исследуемой реальности. Именно совокупность обнаруживаемых в научном тексте ментальных предметов можно рассматривать в качестве наиболее корректной формы компрессии его содержания [21].

В зону действия «приподнимающего» над фоном акцентирования входят и те ментальные действия, с которыми дефиниция сочетается, и которые, тем самым, получают определенные смысловые предпочтения по сравнению с действиями, также упоминаемыми в тексте, но не сопровождающимися дефинитивными конструкциями. Можно предположить, что на этапе текстопорождения разделение содержания на ментальные предметы и связанные с ними ментальные действия, с одной стороны, и фоновое неопределяемое содержание, с другой, происходит под влиянием тех особенностей распределения внимания исследователя между различными составляющими процесса научного познания, которые задаются логикой конкретного исследования и развиваемого автором подхода, а также существующими в дисциплине общими представлениями об иерархии значимости рассматриваемых вопросов. Проведенный нами анализ подтверждает существование специфики сочетаемости дефиниции с другими ментальными действиями в научных публикациях различной дисциплинарной принадлежности.

#### 4. Исследование связей ментальных предметов и действий

С помощью описанной выше процедуры на корпусе из более чем 7000 статей на русском языке, находящихся в открытом доступе на сайтах научных рецензируемых журналов различной тематики и дисциплинарной принадлежности (Табл. 1), выявлялось 31 ментальное действие, относящееся к гиперкатегориям: 1. «Описание нового для науки явления», 2. «Классификация: представление результатов классификации данных опыта», 3. «Сообщение об эмпирической закономерности причинно-следственного типа», 4. «Критический анализ наличного теоретического знания», 5. «Определение принципов разрабатываемой теории», 7. «Проверка теории экспериментальным методом», а также ментальное действие «Дефиниция» (здесь сохранена нумерация, используемая в базовой процедуре [1]).

Табл. 1. Связи категории «Дефиниция» с другими категориями для научных текстов различной предметной принадлежности

Предметная принадлежность публикаций	Текстовые категории											Средний вес для текстов по дисциплине
	4.1.	4.2.	4.3.	4.4.	1.1.	3.1.	7.1.	7.2.	7.3.	7.4.	7.5.	
1. Математика и теоретическая физика	0,13			0,11	0,14	0,15		0,11	0,18	0,16	0,11	0,14
2. Природопользование						0,19		0,12	0,12	0,19	0,12	0,15
3. Техника, материаловедение								0,24	0,15	0,11	0,12	0,16
4. Медицина								0,19	0,19	0,14	0,12	0,16
5. Физика, химия, геология								0,18	0,15	0,13	0,17	0,16
6. Сельское хозяйство и пищевая промышленность								0,25	0,12	0,12		0,16
7. Здравоохранение: эпидемиология		0,17						0,17	0,33	0,17	0,17	0,20
8. Медицина: обзоры и тактики лечения				0,12					0,14	0,14		0,13
9. Биология, физиология				0,16				0,21				0,19
10. Финансы	0,16										0,12	0,14
11. Прикладная математика и теоретическая информатика	0,14	0,12						0,11	0,11		0,14	0,12
12. Информатика	0,12	0,11						0,15	0,11		0,15	0,13
13. Право	0,23	0,13								0,12	0,11	0,15
14. Политология	0,23	0,15								0,13	0,11	0,16
15. История	0,17	0,15								0,15	0,28	0,19
16. Социология: эмпирические исследования	0,24	0,24								0,12	0,12	0,18
17. Экономика	0,17	0,10				0,11	0,10					0,12
18. История искусств	0,16	0,15								0,33		0,21
19. Искусствоведение	0,21	0,24								0,17		0,21
20. Психология: эмпирические исследования	0,31	0,11						0,12				0,18
21. Культурология, этнография	0,27	0,24		0,14						0,16		0,20
22. Здравоохранение и демография	0,19	0,14		0,14	0,14							0,15
23. Социология	0,36	0,23										0,30
24. Лингвистика, филология	0,30	0,25										0,28
25. Педагогика, система образования	0,27	0,19										0,23
26. Философия	0,45	0,23	0,10									0,26

Предметная принадлежность публикаций	Текстовые категории											Средний вес для текстов по дисциплине
	4.1.	4.2.	4.3.	4.4.	1.1.	3.1.	7.1.	7.2.	7.3.	7.4.	7.5.	
27. Психология: обзоры и теория	0,33	0,26	0,10									0,23
28. Международные отношения	0,16	0,26	0,16									0,19
<b>Средний вес связей для категории</b>	0,23	0,18	0,12	0,13	0,14	0,15	0,10	0,17	0,16	0,16	0,14	
<b>Общее количество значимых связей категории</b>	20	19	3	5	2	2	1	11	10	15	13	

**Примечание.** Содержание упоминаемых категорий: 4.1. Анализ и интерпретация наиболее влиятельных теорий, формулирование их понятийных систем; 4.2. Анализ линии развития научного знания как средство обоснования актуальности изучаемой проблематики; 4.3. Определение перспективных направлений разработки проблемы; 4.4. Ознакомление с наиболее важными результатами новых исследований и др.; 1.1. Систематизированное описание нового объекта; 3.1. Фиксация данных опыта в разных исследуемых условиях; 7.1. Формулирование и конкретизация проверяемой гипотезы, определение сути проверяемого предложения; 7.2. Описание методики эксперимента; 7.3. Описание эксперимента, его условий, факторов, которые могут оказать влияние на его результаты, и самих результатов; 7.4. Анализ опытных данных; 7.5. Интерпретация опытных данных в форме ответов на вопросы.

Для извлечения дефиниций из текстов использовался метод, предложенный в работе [22]. В этом методе для выявления наличия в тексте дефиниции анализируется лексико-синтаксический и семантический контекст словоупотреблений. Контекст употребления дефиниций был выявлен экспертами в области лингвистики и психологии, а затем сформирован соответствующий набор шаблонов.

Поиск фрагментов текста, соответствующих шаблону, проводится за два шага. На первом шаге в списке лемм всех словоупотреблений текста, ищутся «ключевые слова» шаблона. В каждом шаблоне предусмотрена хотя бы одна проверка по лемме. На втором шаге проверяются все остальные условия шаблона. После нахождения клаузы, содержащей шаблон, выделяется собственно дефиниция. Для этого с каждым шаблоном связано правило, которое задает, какую часть клаузы необходимо выделить в качестве дефиниции, учитывая, что она может состоять из нескольких слов. Затем с помощью системы эвристических правил проводится фильтрация результатов.

Поскольку в используемой процедуре ментальные действия операционализируются как выявляемые в тексте категории, при описании полученных результатов эти названия будут использоваться как синонимы.

По данным Табл. 1 можно видеть, что из всех выявленных 31 категории, выражающих ментальные действия, статистически значимые связи с дефиницией определяются только для 11 категорий, относящихся к четырем гиперкатегориям: 1, 3, 4 и 7; гиперкатегории 2. «Классификация» и 5. «Определение принципов разрабатываемой теории» здесь не представлены совсем, что свидетельствует о статистически незначимом количестве случаев акцентирования относящихся к ним ментальных действий. Напротив, чаще других сочетаются с дефиницией ментальные действия, выражаемые категориями: 4.1. *Анализ и интерпретация наиболее влиятельных теорий, формулирование их понятийных систем* и 4.2. *Анализ линии развития научного знания как средство обоснования актуальности изучаемой проблематики* (20 и 19 случаев из 28 возможных). Эти же категории сильнее других связаны с дефиницией (средние веса 0,23 и 0,18). Можно видеть, что использование дефиниций, то есть описание ментальных предметов, в рамках проведения концептуального анализа не связано с дисциплинарной принадлежностью исследуемых текстов, но отражает общенаучный канон.

Наиболее полно представлены вошедшие в таблицу категории в текстах по математике и теоретической физике (№ 1 в Табл. 1): для них

выявлено 8 значимых связей из 11 возможных, что отражает высокие требования к определенности и точности изложения, характерные для данных наук. По сравнению с этой, в других группах публикаций картина связей имеет вырожденный характер, особенно в текстах по таким дисциплинам, как 9. *Биология, физиология*, 10. *Финансы*, 23. *Социология*, 24. *Лингвистика, филология*, 25. *Педагогика, система образования* – всего по две значимые связи. Для всех анализировавшихся текстов из 31 категории, определяемой с помощью применявшейся процедуры, на одну дисциплину в среднем пришлось менее чем 4 (точнее, 3,6) связи, что свидетельствует о выраженной избирательности в акцентировании составляющих научного познания.

В целом по признаку сочетания с дефиницией можно выделить три группы текстов, в которых ментальные предметы упоминаются в связи с ментальными действиями «эмпирической» направленности 4.1–4.4 (строки 2–6 Табл. 1), «теоретической» направленности (сочетание с ментальными действиями 1.1, 3.1 и 7.1–7.5, строки 23–28) и смешанные (строки 7–22).

Дисциплинарная специфика проявляется, во-первых, в составе констелляций ментальных действий, включающих в себя дефиницию. Например, только в текстах по экономике обнаруживается статистически значимая связь с категорией 7.1. *Формулирование и конкретизация проверяемой гипотезы, определение сути проверяемого предложения*, в то время как для других анализировавшихся текстов не характерно участие дефинитивных конструкций в упоминании об этом ментальном действии.

О значимости для дисциплины того или иного ментального действия можно судить также по весу его связи с дефиницией. Закономерным образом, максимального значения (0,45) он достигает в текстах философской тематики (№ 26 в Табл. 1) для «теоретической» категории 4.1. *Анализ и интерпретация наиболее влиятельных теорий, формулирование их понятийных систем*.

Самые высокие средние веса связей между дефиницией и другими ментальными действиями (последняя колонка Табл. 1) определяются для дисциплин: 23. *Социология* (средний вес 0,30, это максимальное значение из всех полученных), 24. *Лингвистика, филология* (0,28), 25. *Педагогика, система образования* (0,23), 26. *Философия* (0,26) и 27. *Психология: обзоры и теория* (0,23). При этом в текстах перечисленных групп все

обнаруженные связи с дефиницией приходятся на «теоретические» ментальные действия 4.1, 4.2 и 4.3, что можно интерпретировать как относительно более высокую значимость для этих дисциплин средств исследования именно концептуального характера.

В «эмпирических» и «смешанных» дисциплинах акцентирование в целом менее представлено, то есть упоминание ментальных действий чаще происходит без использования дефинитивных конструкций. Минимальное значение среднего веса (0,12, что почти в три раза меньше, чем максимальное 0,30) приходится на тексты из групп: 11. *Прикладная математика и теоретическая информатика* и 17. *Экономика*. Другие группы текстов из верхней и средней части таблицы имеют средний вес около 0,20 и ниже. Для нас это значит, что в эмпирически более ориентированных дисциплинах не считается неправильным относить значительную часть излагаемого содержания к знанию фоновому, не обсуждаемому, предположительно одинаково известному автору и читателю.

Показательными являются различия между условно более «эмпирическими» и более «теоретическими» направлениями исследований в одной и той же предметной области, например: 16. *Социология: эмпирические исследования* и 23. *Социология* или 20. *Психология: эмпирические исследования* и 27. *Психология: обзоры и теория*. В «эмпирических» работах (№№ 16 и 20) с ментальными предметами сочетаются «эмпирические» ментальные действия (категории 7.4, 7.5 и 7.2). Для «теоретических» исследований по тем же дисциплинам (№№ 23 и 27) все значимые связи с дефиницией выявляются только для «теоретических» же ментальных действий (категории 4.1, 4.2, 4.3).

Из всех исследовавшихся текстов зависимость смыслового акцентирования на определенных ментальных действиях от эмпирической или концептуальной ориентации особенно выражена для работ с медицинской тематикой: 4. *Медицина*, 7. *Здравоохранение: эпидемиология*, 8. *Медицина: обзоры и тактики лечения* и 22. *Здравоохранение и демография*. В Табл. 1 тексты из двух групп с тематикой «Здравоохранение» оказываются в окружении работ разной дисциплинарной принадлежности. Для эпидемиологических исследований выявляется близость, помимо общемедицинских и обзорных статей, с работами по дисциплинам: 6. *Сельское*

хозяйство и пищепром и 9. Биология, физиология. Согласно определению из [23], эпидемиология имеет тесные связи с такими биологическими науками, как микробиология, иммунология, инфекционные болезни, паразитология и др., активно заимствуя их терминологию и методологию. В то же время тексты по медицинской демографии, также относящиеся к общей рубрике «Здравоохранение», оказываются ближе к работам с тематикой: 21. *Культурология, этнография* и 23. *Социология*. Исследуя взаимосвязь воспроизводства населения с медико-социальными факторами, медицинская демография оперирует такими понятиями, как демографические характеристики (пол, возраст, место жительства, образование, профессия, национальность и т.д.), а также явления социальной динамики (миграция, рождаемость, смертность и т.п.) [24]. Можно видеть, что в рассматриваемых случаях сходными по критерию сочетания ментальных предметов и ментальных действий оказываются дисциплины с пересекающимися предметными полями и способами исследования.

## Заключение

В научной литературе дефинитивные структуры применяются для выражения определенного содержания с полнотой, выделяющей его как ментальный предмет на фоне также представленного, но не обсуждаемого в тексте знания. Сочетание ментальных предметов и связанных с ними ментальных действий образует ментальное пространство научных публикаций. Проведенное исследование показывает, что на дискурсивном уровне такое сочетание отражает существующие в дисциплине приоритеты исследовательской парадигмы и нормы научной коммуникации. Для выявления специфики на текстовом уровне с использованием разработанной процедуры планируются сравнительные исследования ментальной структуры научных публикаций, принадлежащих разным авторам и авторским коллективам.

Рассматривая ментальные предметы как узловые элементы смысловой структуры научной публикации, мы полагаем, что методика извлечения констелляций дефиниции с ментальными действиями, позволяющая фиксировать наиболее существенные моменты содержания и прагматики, расширяет возможности интеллек-

туального анализа научного текста в плане его корректной формализации.

## Литература

1. Девяткин Д.А., Кузнецова Ю.М., Чудова Н.В. Методы автоматического выявления ментальных действий в текстах научных публикаций. Часть I // Искусственный интеллект и принятие решений. 2018. № 2. С. 36–46.
2. Крылов С.А., Девяткин Д.А., Кузнецова Ю.М., Чудова Н.В., Швец А.В. Опыт построения когнитивной модели научного текста с помощью методов искусственного интеллекта // Когнитивные штудии: символичный, модульный и нейросетевой подходы. Материалы VII междунар. междисциплинар. конф. Вып. 7 / Под ред. А.П. Лобанова, Н.П. Радчиковой. Минск: БПГУ. 2017. С. 214–221.
3. Девяткин Д.А., Каджая Л.А., Салимовский В.А. Жанры речи как объект компьютерного анализа (на материале научных текстов) // Жанры речи. 2019. № 2(22). С. 86–104.
4. Osipov G.S., Smirnov I.V., Tikhomirov I.A. Relational-situational method for text search and analysis and its applications // Sci. Tech. Inf. Proc. 2010 37(6). P. 432–437.
5. Bojanowski P. et al. Enriching word vectors with subword information // Transactions of the Association for Computational Linguistics. 2017. 5. P. 135–146.
6. Chung J., et al. Empirical evaluation of gated recurrent neural networks on sequence modeling // arXiv preprint arXiv:1412.3555. 2014.
7. Кузнецова Ю.М. Ментальная схема научной публикации как средство выявления псевдонаучных текстов // Лженаука в современном обществе: теоретико-методологические подходы и стратегии противодействия. Тезисы Международной научной конференции. М.: Факультет журналистики МГУ им. М.В. Ломоносова. 2015. С. 77–78.
8. Голубева Н.А. Когнитивные принципы, механизмы и операции в грамматике // Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация. 2010. № 1. С. 32–35.
9. Токарева М.В. Языковая концептуализация когнитивного выделенных компонентов в ситуации слухового восприятия // Вестник Иркутского государственного лингвистического университета. 2011. № 3(15). С. 97–103.
10. Арнольд И.В. Стилистика современного английского языка. М.: Просвещение. 1981. 295 с.
11. Энциклопедический словарь-справочник. Выразительные средства русского языка и речевые ошибки и недочеты / Под ред. А.П. Сковородникова. М.: Флинта. 2005. 480 с.
12. Жирова И.Г. Лингвистическая категория «эмфатичность» в антропоцентризме. Языковая личность Маргарет Тетчер в эмоционально-оценочном дискурсе. М.: Либроком. 2012. 256 с.
13. Скрипак И.А. Синтаксические средства экспрессивности в текстах научного дискурса // Известия Российского государственного педагогического университета им. А.И. Герцена. 2008. № 60. С. 252–256.
14. Климовская Г.И. Русская атрибутивная конструкция в стилистическом аспекте. Томск: Изд-во Томского университета. 1986. 228 с.

15. Нистратова С.Л. К вопросу о системе экспрессивных синтаксических средств в научной речи // Язык, сознание, коммуникация: Сб. статей / Отв. ред. В.В. Красных, А.И. Изотов. М.: МАКС Пресс. 2004. Вып. 28. С. 142–150.
16. Шилова Е.В. Терминологическая дефиниция как метатекст в русскоязычной и англоязычной научно-технической литературе. Автореф. дис. ... канд. филол. наук. Екатеринбург. 2005. 24 с.
17. Kuznetsova Y.M. Functions, types, means of expression, and opportunities for the identification of definitions in scientific texts // Sci. Tech. Inf. Proc. 2016.43 (5). P. 16-25.
18. Мартине А. Основы общей лингвистики // Новое в зарубежной лингвистике. Вып. 3. М.: Издательство иностранной литературы. 1963. С. 366–566.
19. Извольская И.В. Текстобразующая роль термина // Вопросы филологических наук. 2008. № 6. С. 56–58.
20. Котюрова М.П. Автореферат диссертации юриста: «Клиповое» мышление и речевые погрешности // Вестник Челябинского государственного университета. 2010. № 17(198). Филология. Искусствоведение. Вып. 44. С. 49–53.
21. Kuznetsova Y.M. Definition in the Semantic Structure of a Scientific Text // Sci. Tech. Inf. Proc. 2017. 44 (5). P. 309-313.
22. Шелманов А.О. Метод автоматического выделения многословных терминов из текстов научных публикаций // Тринадцатая национальная конференция по искусственному интеллекту с международным участием КИИ-2012. 2012. С. 268–274.
23. Большая Медицинская Энциклопедия. Под редакцией Петровского Б.В., 3-е издание. Т. 28. URL: <https://бмэ.орг/index.php/эпидемиология>.
24. Тищенко Е.М., Заборовский Г.И. Общественное здоровье и здравоохранение. Гродно: Гродненский государственный медицинский университет. 2004. 156 с.

## Mental Actions and Mental Objects in the Space of Science Discourse

D.A. Devyatkin, Y.M. Kuznetsova

Federal Research Center "Computer Science and Control" of Russian Academy of Sciences, Moscow, Russia

**Abstract.** The paper presents an approach to discourse of science as a mental space whose structure displays the logic and tools of the scientific cognition and communication. We propose that this space is composed of the mental actions have been described previously [1], and a mental objects being described here as a definitions express the main ideas of an academic text. Based on more than 7000 research publications from peer-reviewed journals, an efficiency of the procedures of relational-situational analysis and machine learning to reveal the structure of correlations between mental objects and mental actions is shown. It is also shown that in publications of various disciplines the correlations between mental objects and mental actions vary in composition and strength, and therefore this method can capture the specifics of the intellectual tools of research and communication in a particular discipline. A presenting of research publications content as a constellation of extracted mental actions and mental objects can be one of the way of the text mining enhancement.

**Keywords:** text mining, science discourse, research publication, mental actions, mental objects.

DOI 10.14357/20718594200106

## References

1. Devyatkin D.A., Kuznetsova Y.M., Chudova N.V. Metody avtomaticheskogo vyyavleniya mental'nyh dejstvij v tekstah nauchnyh publikatsij. Ch. I [Methods for mental operations detection in scientific publications. P. I] // Iskustvennyj intellekt i prinyatie reshenij [Artificial Intelligence and Decision Making], 2018, 2: 36–46.
2. Krylov S.A., Devyatkin D.A., Kuznetsova Y.M., Chudova N.V., Shvets A.V. Opyt postroeniya kognitivnoj modeli nauchnogo teksta s pomoshch'yu metodov iskusstvennogo intellekta [The attempt of constructing a cognitive model of a scientific text using artificial intelligence methods] // Kognitivnye shtudii: simvol'nyj, modul'nyj i nejrosetevoj podhody: materialy VII mezhdunar. mezhdisciplinar. konf. Vyp. 7 / Pod red. A.P. Lobanova, N.P. Radchikovoj [A.P. Lobanov, N.P. Radchikova (eds.) Cognitive studies: symbolic, modular and neural network approaches: Materials of the VII international interdisciplinary conference. Issue 7]. Minsk: Belarusian State Pedagogical University, 2017: 214–221.
3. Devyatkin D.A., Kadzhaya L.A., Salimovskiy V.A. Zhanry rechi kak ob'ekt komp'yuternogo analiza (na materiale nauchnyh tekstov) [Speech Genres as on Object of Computer Analysis (Based on Academic Texts)] // Zhanry rechi [Speech Genres], 2019, 2(22): 86–104.
4. Osipov, G.S., Smirnov, I.V., & Tikhomirov, I.A. Relational-situational method for text search and analysis and its applications // Sci. Tech. Inf. Proc., 2010, 37(6): 432–437.

5. Bojanowski, P. et al. Enriching word vectors with subword information // *Transactions of the Association for Computational Linguistics*. 2017, 5: 135–146.
6. Chung J., et al. Empirical evaluation of gated recurrent neural networks on sequence modeling // *arXiv preprint arXiv:1412.3555*. 2014.
7. Kuznetsova Y.M. Mental'naya skhema nauchnoj publikacii kak sredstvo vyavleniya psevdonauchnyh tekstov [The mental scheme of scientific publication as a means of identifying pseudoscientific texts] // *Lzhenauka v sovremennom obshchestve: teoretiko-metodologicheskie podhody i strategii protivodejstviya. Tezisy Mezhdunarodnoj nauchnoj konferencii [Pseudoscience in modern society: theoretical and methodological approaches and countermeasures. Abstracts of the International Scientific Conference.]*. Moscow: Faculty of Journalism, Lomonosov Moscow State University, 2015: 77–78.
8. Golubeva N.A. Kognitivnye principy, mekhanizmy i operacii v grammatike [Cognitive principles, mechanisms and operations in grammar] // *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Lingvistika i mezhkul'turnaya kommunikaciya [Bulletin of Voronezh State University. Series: Linguistics and Intercultural Communication]*, 2010, 1: 32–35.
9. Tokareva M.V. Yazykovaya konceptualizaciya kognitivnogo vydelennyh komponentov v situacii sluhovogo vospriyatiya [Linguistic conceptualization of cognitive distinguished components in a situation of auditory perception] // *Vestnik Irkutskogo gosudarstvennogo lingvisticheskogo universiteta [Bulletin of Irkutsk State Linguistic University]*, 2011, 3(15): 97–103.
10. Arnold I.V. Stilistika sovremennogo anglijskogo yazyka [Stylistics of modern English]. Moscow: Prosveshchenie, 1981. 295 p.
11. Enciklopedicheskij slovar'-spravochnik. Vyrazitel'nye sredstva russkogo yazyka i rechevye oshibki i nedochety / Pod red. A.P. Skovorodnikova [A.P. Skovorodnikov (ed.) Encyclopedic reference dictionary. Expressive means of the Russian language and speech errors and shortcomings]. Moscow: Flinta, 2005. 480 p.
12. Zhirova I.G. Lingvisticheskaya kategoriya «emfaticnost» v antropocentrizme. Yazykovaya lichnost' Margaret Tetcher v emocional'no-ocenocnom diskurse [The linguistic category "emphaticity" in anthropocentrism. The language personality of Margaret Thatcher in an emotionally-evaluative discourse]. Moscow: Librokom, 2012. 256 p.
13. Skripak I.A. Sintaksicheskie sredstva ekspressivnosti v tekstah nauchnogo diskursa [Syntactic means of expressiveness in the texts of scientific discourse] // *Izvestiya Rossijskogo gosudarstvennogo pedagogicheskogo universiteta im. A.I. Gercena [Bulletin of the Herzen State Pedagogical University of Russia]*, 2008, 60: 252–256.
14. Klimovskaya G.I. Russkaya atributivnaya konstrukciya v stilisticheskom aspekte [Russian attributive construction in a stylistic aspect]. Tomsk: Tomsk University Press, 1986. 228 p.
15. Nistratova S.L. K voprosu o sisteme ekspressivnyh sintaksicheskikh sredstv v nauchnoj rechi [On a system of expressive syntactic means in scientific speech] // *Yazyk, soznanie, kommunikaciya: Sb. statej. Vyp. 28. / Otv. red. V.V. Krasnyh, A.I. Izotov [V.V. Krasnyh, A.I. Izotov (eds.) Language, consciousness, communication]*. Moscow: MAKSPress, 2004, 28: 142–150.
16. Shilova E.V. Terminologicheskaya definiciya kak metatekst v russkoyazychnoj i angloyazychnoj nauchno-tekhnicheskoy literature [Terminological definition as a metatext in the Russian and English scientific and technical literature]. PhD Thesis. Ekaterinburg, 2005. 24 p.
17. Kuznetsova, Y.M. Functions, types, means of expression, and opportunities for the identification of definitions in scientific texts // *Sci. Tech. Inf. Proc.* 2016, 43 (5): 16–25.
18. Martine A. Osnovy obshchej lingvistiki [Fundamentals of General Linguistics] // *Novoe v zarubezhnoj lingvistike. Vyp. 3 [News of Foreign Linguistics. Is. 3]*. Moscow: Izdatel'stvo inostrannoj literatury, 1963: 366–566.
19. Izvol'skaya I.V. Tekstoobrazuyushchaya rol' termina [The text-forming role of the term] // *Voprosy filologicheskikh nauk [Questions of philological sciences]*, 2008, 6: 56–58.
20. Kotyurova M.P. Avtoreferat dissertacii yurista: «Klipovoe» myshlenie i rechevye pogreshnosti [Abstract of the dissertation of a lawyer: “Clip-thinking” and speech errors] // *Vestnik Chelyabinskogo gosudarstvennogo universiteta. Filologiya. Iskusstvovedenie. Vyp. 44 [Bulletin of the Chelyabinsk State University. Philology. Arts. Is. 44]*, 2010, 17(198): 49–53.
21. Kuznetsova, Y.M. Definition in the Semantic Structure of a Scientific Text // *Sci. Tech. Inf. Proc.* 2017, 44 (5): 309–313.
22. Shelmanov A.O. Metod avtomaticheskogo vydeleniya mnogoslownykh terminov iz tekstov nauchnyh publikacij [The method of automatic extraction of verbose terms from texts of scientific publications] // *13th Russian Conference on Artificial Intelligence RCAI-2012, 2012: 268–274*.
23. Bol'shaya Medicinskaya Enciklopediya (BME), pod redakciej Petrovskogo B.V., 3-e izdanie. T. 28. [B.V. Petrovsky (ed.) Big Medical Encyclopedia, 3rd edition]. URL: <https://bme.org/index.php/EPIDEMIOLOGIYA>
24. Tishchenko E.M., Zaborovskij G.I. Obshchestvennoe zdorov'e i zdavoohranenie [Public health and healthcare]. Grodno: Grodno State Medical University, 2004. 156 p.