



УДК 81`32+612.821

EDN HIKNYL

<https://www.doi.org/10.33910/2687-0223-2023-5-1-14-25>

## Применение психофизиологического инструментария для изучения понятия «механизм» в лингвистическом исследовании

К. А. Овчинникова<sup>✉1</sup>, Е. Г. Вергунов<sup>2</sup>

<sup>1</sup> Новосибирский национальный исследовательский государственный университет,  
630090, Россия, г. Новосибирск, ул. Пирогова, д. 1

<sup>2</sup> Научно-исследовательский институт нейронаук и медицины,  
630117, Россия, г. Новосибирск, ул. Тимакова, д. 4

### Сведения об авторах

Овчинникова Кристина  
Алексеевна,  
ORCID: 0000-0002-7400-0586,  
e-mail: [k.ovchinnikova2@ngsu.ru](mailto:k.ovchinnikova2@ngsu.ru)

Вергунов Евгений Геннадьевич,  
SPIN-код: 9940-3675,  
Scopus AuthorID: 57191523873,  
ResearcherID: N-7962-2014,  
ORCID: 0000-0002-8352-5368,  
e-mail: [vergounov@gmail.com](mailto:vergounov@gmail.com)

### Для цитирования:

Овчинникова, К. А.,  
Вергунов, Е. Г.  
(2023) Применение  
психофизиологического  
инструментария для изучения  
понятия «механизм»  
в лингвистическом исследовании.  
*Комплексные исследования  
детства*, т. 5, № 1, с. 14–25.  
<https://doi.org/10.33910/2687-0223-2023-5-1-14-25> EDN HIKNYL

Получена 26 декабря 2022;  
прошла рецензирование  
30 января 2023; принята 30 января  
2023.

**Финансирование:** Исследование  
не имело финансовой поддержки.

**Права:** © К. А. Овчинникова,  
Е. Г. Вергунов (2023).  
Опубликовано Российским  
государственным педагогическим  
университетом им. А. И. Герцена.  
Открытый доступ на условиях  
лицензии [CC BY-NC 4.0](https://creativecommons.org/licenses/by-nc/4.0/).

**Аннотация.** «Механизм» как понятие в науке давно рассматривается исследователями, в каждой предметной области это значение приобретает характерные для нее коннотации. Однако в лингвистике, например, понятие механизмов представляется нечетко, количество исследований этого вопроса немногочисленно, математические инструменты выявления механизмов не предоставляют возможностей, которые в междисциплинарных исследованиях дает современный анализ.

В работе показана эффективность инструмента 2B-PLS, хорошо зарекомендовавшего себя в междисциплинарных исследованиях (включая нейролингвистику), для задач математической лингвистики. 2B-PLS модель можно с помощью вопросов (откликов) «обучить» разворачивать предикторы нужным образом для получения максимально информативного ответа на ряд поставленных перед моделью вопросов. Для примера анализа встречаемости термина «механизм» взят корпус из 100 русскоязычных научных публикаций 2005 – 2020 гг. Статьи корпуса отнесены к коллекциям наук (гуманитарных, естественных, точных, общественных и технических). Общее количество вхождений термина «механизм» — 104, общее число словоупотреблений в корпусе — 370,8 тыс.

На основе корпуса научных текстов показано, что термин «механизм» относится к значимым элементам научного стиля статей (встречается в 37 статьях из 100). Частота упоминаний термина не показала значимой частной корреляции с количеством статей по годам.

2B-PLS модель показала, что работа механизмов лингвистики выражает общие системные закономерности и частную специфику наборов данных. С помощью 2B-PLS модели было показано, что более молодые статьи в гуманитарных и точных науках обсуждают аспекты полученных результатов с точки зрения их механизмов более подробно, чем в более старых работах технических и социальных наук (контрастные группы). При этом в периоды с редким выходом статей корпуса более молодые работы гуманитарных наук реже использовали термин «механизм» в одной публикации (по сравнению с контрастной группой — естественных, общественных и технических наук).

**Ключевые слова:** механизм, 2B-PLS, гуманитарные науки, естественные науки, точные науки, технические науки, общественные науки

# The use of psychophysiological tools to study the concept of “mechanism” in linguistic research

К. А. Овчинникова<sup>✉1</sup>, Е. Г. Вергунов<sup>2</sup>

<sup>1</sup> Novosibirsk National Research State University, 2 Pirogova Str., Novosibirsk 630090, Russia

<sup>2</sup> Scientific Research Institute of Neurosciences and Medicine, 4 Timakova Str., Novosibirsk 630117, Russia

## Authors

Kristina A. Ovchinnikova,  
ORCID: 0000-0002-7400-0586,  
e-mail: [k.ovchinnikova2@ng.nsu.ru](mailto:k.ovchinnikova2@ng.nsu.ru)

Eugeny G. Vergunov,  
SPIN-код: 9940-3675,  
Scopus AuthorID: 57191523873,  
ResearcherID: N-7962-2014,  
ORCID: 0000-0002-8352-5368,  
e-mail: [vergounov@gmail.com](mailto:vergounov@gmail.com)

## For citation:

Ovchinnikova, K. A.,  
Vergunov, E. G.  
(2023) The use of  
psychophysiological tools to study  
the concept of “mechanism”  
in linguistic research.  
*Comprehensive Child Studies*, vol. 5,  
no. 1, pp. 14–25. <https://doi.org/10.33910/2687-0223-2023-5-1-14-25> EDN HIKNYL

**Received** 26 December 2022;  
reviewed 30 January 2023;  
accepted 30 January 2023.

**Funding:** The study did not receive  
any external funding.

**Copyright:** © K. A. Ovchinnikova,  
E. G. Vergunov (2023). Published by  
Herzen State Pedagogical University  
of Russia. Open access under  
CC BY-NC License 4.0.

**Abstract.** The concept of “mechanism” has long been considered by researchers, and its meaning acquires various connotations depending of the particular research area. However, in linguistics the concept of mechanism is vague, the studies on the topic are scarce, and mathematical tools for identifying mechanisms do not provide the opportunities that modern analysis provides in interdisciplinary research.

The paper shows the effectiveness of 2B-PLS, which has proven itself in interdisciplinary research, for problems of mathematical linguistics. The 2B-PLS model can be trained, using questions, to deploy predictors in the right way to get the most informative answer to a number of questions posed to the model.

We took 100 Russian-language research publications dated 2005-2020 to analyse the occurrence of the term “mechanism”. The articles of the corpus belong to scientific collections: humanitarian, natural, exact, social and technical sciences. The total number of occurrences of the term “mechanism” is 104, while the total number of words usages in the corpus is 370.8 thousand. Based on the corpus, we showed that the term “mechanism” is a significant element of the style of research articles: it is used in 37 articles out of 100. The frequency of the term showed no significant correlation with the number of articles by year.

The 2B-PLS model showed that the work of linguistics mechanisms expresses general system regularities and particular specifics of datasets.

According to the 2B-PLS model, more recent articles in humanities and exact sciences discuss results in terms of their mechanisms in more detail than older works in the technical and social sciences (contrast groups). At the same time, in the periods with rare publication of the articles in the corpus, more recent articles in humanities used the term “mechanism” in one publication less often than the articles in natural, social and technical sciences (contrast groups).

**Keywords:** mechanisms, 2B-PLS, humanities, natural sciences, exact sciences, technical sciences, social sciences

## Введение

Понятие «механизм» рассматривается учеными с момента появления данного термина. В словаре Д. Н. Ушакова, помимо значения, связанного с устройством машины или прибора, приведено его переносное значение: «...внутреннее устройство, система функционирования чего-нибудь, аппарат какого-нибудь вида деятельности» (Ушаков 2004). В каждой области это понятие дополняется семами, характерными для нее, а исследователи продолжают изучать специфику механизмов для научных областей.

Так, например, в работе Ю. В. Горбунова данное понятие исследовано для области экономики (Горбунов 2018). Авторы предлагают понимать под механизмом в экономике

характеристики процесса, то есть способы, методы, нормы, средства, формы функционирования чего-либо или воздействия на что-либо, а не совокупность ресурсов или состояний объекта. А одним из вариантов представления механизмов в области программирования (Greenhalgh et al. 2015) является взаимодействие между программными ресурсами и то, как участники интерпретируют их и реагируют (или не реагируют) на них.

## Механизмы в науке

Концепция механизма (Machamer et al. 2000) анализируется с точки зрения сущностей и действий, организованных так, чтобы они производили регулярные изменения. Понятие механизма включает в себя объекты и действия, организованные так, чтобы они производили регулярные изменения от первоначальных

установок до выполнения условий завершения (прекращения). Описать механизмы явления — объяснить явление, его генезис как системы. Именно генезис системы определяет ее функционирование в конкретных условиях.

В статье В. Ф. Мартюшова говорится о том, что обществоведы не уделяли достаточного внимания анализу понятия *социальных механизмов* и более широкого, родового по отношению к нему понятия *механизм* (Мартюшов 2015). Речь идет о том, что понятие *механизм* связано с философскими категориями необходимости и случайности. Категория *механизм* в исследовании политических, социальных и других общественных процессов позволяет рассматривать эти процессы как *включающие в себя явления, факторы и зависимости, образующие целостный феномен*, анализ устройства которого дает возможность более полно и глубоко разобраться в изучаемых закономерностях.

Выше отмечено, что механизм — система функционирования чего-либо; *система* — порядок, обусловленный правильным, закономерным расположением частей в определенной связи (Ушаков 2004).

Система обладает рядом свойств (Попов, Крайнюченко 2012), делящихся на *свойства строения* и *свойства функционирования*. Совокупность элементов образует систему тогда, когда отношения между ними порождают *эмерджентность* (англ. emergent «возникающий»), т. е. наличие у системы свойств, не приписуемых ее компонентам по отдельности и не сводящихся к сумме свойств (Комлев 2006). Так, например, химическое соединение двух элементов, образующих формулу воды: водорода и кислорода. Вода как единое целое обладает более существенными качествами и свойствами, чем простое соединение этих двух элементов, а тем более каждый отдельно взятый элемент.

В лингвистике структура и регулярность грамматики языков и языковые изменения являются также *эмерджентным* явлением (Keller 1994).

К функциональному блоку свойств системы относится *гомеостаз* — саморегуляция, способность открытой системы сохранять постоянство внутреннего состояния посредством скоординированных реакций, направленных на поддержание динамического равновесия.

Другая группа системных свойств связана с представлениями о развитии. К этой группе следует отнести *гомеокинез*. *Гомеокинез* (аллостаз) — это «достижение стабильности через изменения», приспособление системы к изменениям, ее обучение или адаптация.

Гомеостаз и гомеокинез — примеры механизмов, отражающих общие особенности, присущие всем системам. Кроме указанных выше, каждая система обладает свойствами в соответствии с ее строением и функционированием, которые отражают ее общие особенности и частную специфику.

### *Механизмы в лингвистике*

В статье Г. М. Костюшкиной отмечается, что в лингвистике для каждого языкового явления существует свой особый *системообразующий механизм*, позволяющий выявить как внутрипарадигмальные связи и отношения в изучаемом явлении, так и межпарадигмальные, позволяющие рассматривать данное явление в общей систематике с другими элементами системы. Именно в этом ключе происходит построение систематики языковых явлений, единиц, категорий, т. е. систематики языка (Костюшкина 2012).

Системообразующий механизм в языке, речи и речевой деятельности в каждом конкретном случае имеет свою специфику, однако его универсальная черта во всех исследованиях — *служить целям формирования и организации системы, которая нам никогда не дана вся целиком в прямом наблюдении*. По этой причине она нуждается в моделировании, что позволяет показать ее внутрипарадигмальные связи и межпарадигмальные связи по отношению к другим системам и тем самым вписать данную систему в общую систему систем и определить ее внутреннюю и внешнюю значимость.

Ф. Соссюр в своих трудах говорил, что если бы *механизм* языка был полностью рационален, его можно было бы изучать как вещь в себе, но он представляет собой «лишь частичное исправление хаотичной по природе системы» (де Соссюр 1977). На представлениях, восходящих к Ф. Соссюру, о языке как о механизме основано математическое описание языка. Формальный аппарат для описания строения естественных и некоторых искусственных языков разрабатывает математическая лингвистика. *Теории способов описания синтаксической структуры* являются одним из разделов математической лингвистики и изучают способы описания текстов.

Описание синтаксической структуры предложения может быть проведено в следующих направлениях:

- 1) выделение составляющих группы слов, функционирующие как цельные синтаксические единицы (Лайонз 1978);
- 2) выделение зависимых слов (Теньер 1988).

Различия между данными направлениями представляются довольно существенными. Во-первых, в синтаксисе непосредственно составляющих группы слов не важно разделение на главные и зависимые слова, в отличие от синтаксиса Л. Теньера. Во-вторых, в синтаксисе зависимостей строго разделяются структурный порядок слов и линейный. «*Структурный порядок слов — это порядок, в котором устанавливаются синтаксические связи*», то есть это порядок от главного члена предложения к зависимому, поэтому он может не совпадать с линейным порядком. В синтаксисе непосредственно составляющих, в свою очередь, важен линейный порядок, поскольку составляющая должна быть непрерывной. В-третьих, составляющие представляют собой не только слова. В качестве минимальных составляющих некоторые дескриптивисты выделяли даже не слова, а морфемы.

Для отечественной традиции подход синтаксических зависимостей кажется очевидным и чуть ли не единственно возможным.

Приведем пример для предложения «*Лошади кушают овёс*». Непосредственными составляющими являются все предложение, каждое отдельное слово и словосочетание «*кушают овёс*».

Теперь рассмотрим построение дерева синтаксических зависимостей для того же предложения. Зависимости в предложении будут построены следующим образом: от «*кушают*» к «*лошади*» и к «*овёс*».

Другой раздел математической лингвистики — *теория формальных грамматик*, предложенная Н. Хомским (Хомский 2015). Эта теория изучает способы описания закономерностей, которые характеризуют всю совокупность правильных текстов того или иного языка. Эти закономерности описываются путем построения «формальной грамматики», представляющей собой абстрактный «механизм», позволяющий с помощью единообразной процедуры получать правильные тексты данного языка вместе с описаниями их структуры.

Наиболее широко используемый тип формальной грамматики — *порождающая грамматика Хомского*. Рассмотрим вариант построения такой грамматики для предложения «*Лошади кушают овёс*». Допустим, грамматика имеет следующие правила:

- 1)  $I \rightarrow S_{x,y,им} V_y;$
- 2)  $V_y \rightarrow V_y^t S_{x,y,вин};$
- 3)  $S_{м, ед. ч., вин} \rightarrow \text{овёс};$
- 4)  $S_{ж, мн. ч., им} \rightarrow \text{лошади};$
- 5)  $V_{мн.ч.}^t \rightarrow \text{кушают}.$

$V_y$  означает категорию «группа глагола в числе  $y$ »,  $V_y^t$  — «переходный глагол в числе  $y$ »,  $S_{x,y,z}$  — «существительное рода  $x$  в числе  $y$  и падеже  $z$ ».

Опишем дерево синтаксических зависимостей на основе заданных правил (рис. 1). Слева представлено дерево на русском языке, а справа — перевод на английский.

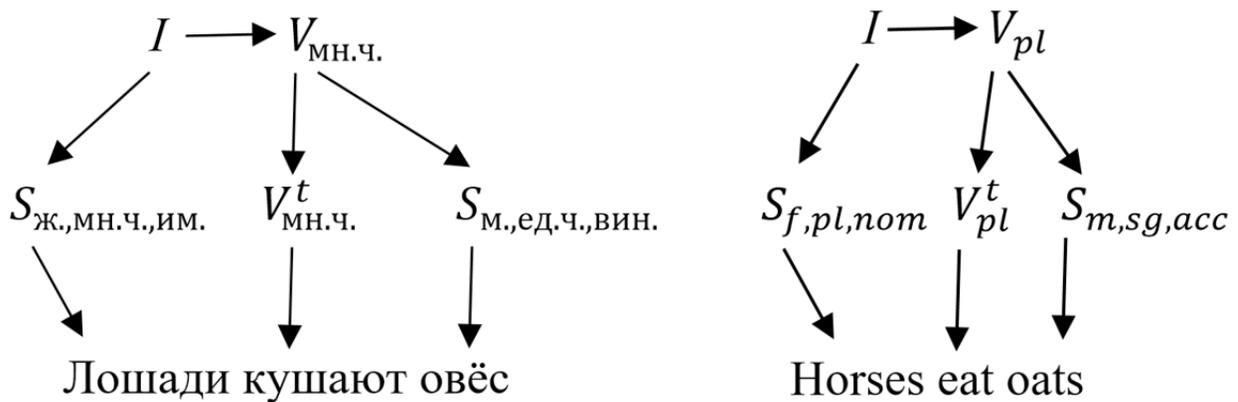


Рис. 1. Порождающая грамматика для предложения «*Лошади кушают овёс*» (англ. «*Horses eat oats*»). На схеме стрелки идут из левых частей применяемых правил к элементам соответствующих правых частей.  $V_y$  — группа глагола в числе  $y$ ,  $V_y^t$  — переходный глагол в числе  $y$ ,  $S_{x,y,z}$  — существительное рода  $x$  в числе  $y$  и падеже  $z$

Fig. 1. Generative grammar for the sentence “*Horses eat oats*”. In the diagram, the arrows go from the left parts of the applied rules to the elements of the corresponding right parts.  $V_y$  is a group of a verb in the number  $y$ ,  $V_y^t$  is a transitive verb in the number  $y$ ,  $S_{x,y,z}$  is a noun of gender  $x$  in the number  $y$  and case  $z$

Порождающие грамматики используются для описания не только естественных, но и искусственных языков, в особенности языков программирования.

### *Методы изучения механизмов в лингвистике*

*Контент-анализ* стал широко применяться в информационную эпоху (Дмитриев 2005). В работе И. Дмитриева описаны основные процедуры контент-анализа: выявление смысловых единиц контент-анализа, выделение единиц счета и процедура подсчета. Контент-анализ может использоваться для исследования большей части документальных источников, однако лучше всего он работает при относительно большом количестве однопорядковых данных.

*Контент-анализ* — анализ текста, основанный на выделении категорий содержания по эксплицитно заданным правилам и последующем подсчете представленности данных категорий в текстах, для которых учитывается один или более внешних факторов (Алмаев 2012).

Согласно мнению К. Криппендорп, контент-анализ наиболее успешен, когда он может разбить «лингвистически составленные факты» на четыре класса:

Attributions (атрибуция, приписывание свойств),

Social relationships (социальные взаимодействия),

Public behaviours (общественное поведение),

Institutional realities (институциональные реалии) (Krippendorp 2004).

*Латентное размещение Дирихле (LDA, Latent Dirichlet allocation)* — это неконтролируемый алгоритм, присваивающий каждому документу значение для каждой определенной темы (Blei et al. 2003). LDA рассматривает каждый документ как набор тем, а каждую тему — как набор слов. Оно перебирает общее количество тем и каждое слово; случайным образом присваивает каждое слово теме и оценивает, как часто это слово встречается в этой теме вместе с другими словами.

В статье Н. А. Алмаева и О. В. Мурашевой показана эффективность алгоритмов LDA для анализа дискуссий в социальных сетях. Темы одного и того же корпуса обсуждений были сгруппированы методами LDA и NMF/LSA. Один из выводов — применимость LDA к задаче тематического анализа (Алмаев, Мурашева 2022).

*Латентно-семантический анализ (LSA, Latent semantic analysis)* — это метод обработки информации на естественном языке, анализирующий взаимосвязь между библиотекой

документов и встречающимися в них терминами, а также выявляющий характерные факторы (тематики), присущие всем документам и терминам (Landauer et al. 1998).

Идея латентно-семантического анализа состоит в том, что если матрица  $A$  — это матрица термины-на-документы, то матрица  $\hat{A}$ , содержащая только  $k$  первых линейно независимых компонент  $A$ , отражает основную структуру различных зависимостей, присутствующих в исходной матрице («скрытых тем»). Структура зависимостей определяется весовыми функциями термов.

Таким образом, каждый терм и документ представляются при помощи векторов в общем пространстве размерности  $k$  (так называемом пространстве гипотез). Близость между любой комбинацией термов и/или документов легко вычисляется при помощи скалярного произведения векторов. Как правило, выбор  $k$  зависит от поставленной задачи и подбирается эмпирически.

*Неотрицательная матричная факторизация (NMF, Non-negative matrix factorization)* представляет собой группу алгоритмов в многомерном анализе и линейной алгебре, где матрица  $V$  разложена (обычно) на две матрицы  $W$  и  $H$  с тем свойством, что все три матрицы не имеют отрицательных элементов (Lee, Seung 1999). В статье Garla с помощью данного метода были проанализированы речи президента США на инаугурации и были выделены восемь тем, которые можно обозначить как законы, мир, лидерство, расходы, справедливость, доходы, идеалы, принятие решений (Garla 2021).

Термины «латентная тема» (LDA), «скрытая тема» (LSA) и «тема» (NMF) совпадают с нашим пониманием механизмов.

*PLS* — это метод получения проекций на латентные структуры (Projection to Latent Structure, или метод частичных наименьших квадратов, Partial Least Squares). Эффективным PLS-инструментом является **2B-PLS** (Two-Block PLS) (Rohlf, Corti 2000). **2B-PLS** показал высокую эффективность в междисциплинарных исследованиях в нейронауках (Savostyanov et al. 2022), генетике (Polunin et al. 2019), биологии (Ковалева и др. 2019), психофизиологии (Кривошецов и др. 2022; Nikolaeva et al. 2022), психологии (Vergunov 2022).

**2B-PLS** позволяет изучать имплицитные процессы различной природы с помощью анализа глубинных независимых латентных структур (механизмов), единых для обоих блоков многомерных показателей — матриц **B1** и **B2**. Для этого оба блока преобразуются в пары матриц:

$[B1] = [B1\text{-score}] \times [B1\text{-loadings}]^T$  и  $[B2] = [B2\text{-score}] \times [B2\text{-loadings}]^T$ , где

$[]$  — матрица,  $T$  — транспонирование, **loadings** — нагрузки, **score** — счета.

**2B-PLS** модель определяет систему пар единичных осей для обоих блоков, которые выражают максимальный шаблон ковариации для матриц **B1-score** и **B2-score**, которые будут проекциями **B1** и **B2** на искомые оси ортогональных латентных структур (Polunin et al. 2019). Это достигается выделением наиболее контрастных групп. Ряды данных центрируются, приводятся к единому масштабу, разворачиваются. Полученные латентные структуры описываются с помощью ортогональных матриц **B1-loadings** и **B2-loadings**, которые содержат коэффициенты перехода от новых латентных структур к исходным «явным структурам».

Число латентных структур равно минимальному числу переменных из двух исходных блоков. Соотношения для структур сырых данных в блоках остаются теми же самыми после любого числа (и порядка применения) операций центрирования, масштабирования и поворота. В **2B-PLS** допустимы ситуации с исходными данными, когда переменных больше, чем объектов, или ряды данных показывают тесную взаимную корреляцию (Rännar et al. 1994).

Если описываемая конкретная структура доли наблюдаемой дисперсии превышает 5% от общей дисперсии, то считается, что для данной структуры объема выборки достаточно.

В один блок (**B1**) можно поместить переменные-предикторы (данные), а в другой (**B2**) — отклики-признаки (состоят только из «0» и «1», это наши вопросы, на которые модель дает ответы). Блок **B2** позволяет управлять углом поворота блока **B1** — он должен быть таким, чтобы дать ответы на вопросы-признаки из **B2**. Это своеобразное обучение, настройка модели на конкретный ряд вопросов из **B2**. Данный подход отличается от известной PLS-регрессии тем, что вся информации из исходных рядов данных при построении **2B-PLS** модели собирается в первых латентных структурах.

Описанные особенности **2B-PLS** предоставляют больше возможностей, чем общеизвестные методы главных компонент (Principal Component Analysis, PCA) и главных координат (Principal Coordinates Analysis, PCO), латентное размещение Дирихле (Latent Dirichlet allocation, LDA), латентный семантический анализ (Latent semantic analyses, LSA или LSI — латентно-семантическое индексирование), неотрицательная матричная факторизация Lee-Seung (Nonnegative matrix factorization, NMF) или PLS-регрессия.

Целью данной работы является описание представлений о механизмах в науке, методов их изучения в лингвистике и примера **2B-PLS**-анализа частоты их употребления в научных текстах. Интерес в первую очередь представляет употребление этого термина в гуманитарных науках, к которым относится математическая лингвистика.

## Материалы и методы

### Характеристика корпуса научных статей

Под наукой понимается деятельность, направленная на выработку и систематизацию объективных знаний о действительности. Она может быть разделена на несколько областей.

Для проведения исследования был составлен корпус текстов, который включает в себя 100 русскоязычных научных публикаций за период с 2005 по 2020 г., взятых из открытого информационного источника (Научная электронная библиотека... 2022). Статьи относятся к пяти научным коллекциям, соответствующих гуманитарным, естественным, техническим, общественным и точным наукам. Общий объем корпуса составляет 370,8 тысяч словоупотреблений. Количество вхождений термина «механизм» (как термин в описанном выше понимании) составляет 104 (в 37 статьях).

Обработка корпуса состояла в следующих этапах:

- 1) предобработка корпуса: токенизация и удаление символов;
- 2) удаление стоп-слов и английских наименований;
- 3) подсчет вхождений каждого слова в корпус в нем самом и в каждой отдельной статье.

Таким образом, была получена таблица данных с вхождениями термина «механизм» для каждой статьи, представленной в корпусе.

## Методы анализа

В блоки **2B-PLS**-модели вошли переменные (табл. 1), которые представляют собой количество упоминаний термина «механизм» в статье, возраст статьи и общее число статей в базе, вышедших в данном году (3 переменные, блок № 1), а также ряды признаков наук, ряды признаков статей и годы их выхода (58 признаков, блок № 2). Соответственно, будет три латентные структуры. Все переменные блока 1 предварительно нормировались на размах.

Табл. 1. Блоки переменных для 2B-PLS-модели

Переменные	Блок
<b>&amp;count</b> — число упоминаний в статье термина механизм	№ 1
<b>&amp;all</b> — общее число статей в базе, вышедших в данном году	№ 1
<b>&amp;age</b> — возраст статьи	№ 1
<b>Tech, Soc, Nat, Exac, Hum</b> : 5 признаков технических, социальных, естественных, точных или гуманитарных наук соответственно	№ 2
<b>2005..2020</b> : 16 признаков года публикации	№ 2
<b>#1..#37</b> : 37 признаков публикации	№ 2

Table. 1. Blocks of variables used for 2B-PLS models

Variables	Block
<b>&amp;count</b> —number of mentions of the term “mechanism” in the article	No. 1
<b>&amp;all</b> —total number of articles in the database published in a given year	No. 1
<b>&amp;age</b> —article age	No. 1
<b>Tech, Soc, Nat, Exac, Hum</b> —five signs of technical, social, natural, exact or humanitarian sciences, respectively	No. 2
<b>2005..2020</b> —16 signs of the year of publication	No. 2
<b>#1..#37</b> —37 signs of publication (publication code)	No. 2

Таким образом, предикторы 2B-PLS-модели (блок № 1) должны дать ответы на вопросы о соотношении частоты упоминания термина «механизм» в статьях и науках, к которым эти статьи относятся. Обучение модели (выбор угла разворота блока данных) происходит за счет откликов (блок № 2), которые включают признаки года выхода статьи и статистической неоднородности ее данных (частоты термина «механизм»).

Включение в анализ откликов, представляющих собой коды отдельных статей, позволяет «оттянуть» из латентных структур, которые описывают общие особенности (обычно это первая структура) и частную специфику (несколько последующих структур), ту часть дисперсии, которая обусловлена индивидуальной неоднородностью данных и при изучении общих особенностей или частной специфики является «шумом». При этом наличие откликов-кодов для отдельных статей позволяет выполнить индивидуальный профайлинг таких статей, если, например, в задачи входило бы изучение статистически однородных кластеров среди данных.

Модель формировалась с помощью программного пакета JACOBI 4 (Polunin et al. 2019).

### Анализ частот употребления термина «механизм» в русскоязычных статьях различных наук

Доля статей с упоминанием термина «механизм» (37 из 100) значима на уровне  $p < 0,05$ . Таким образом, использование термина «механизм» является значимым элементом научного стиля статей.

Между показателями «Число статей с упоминанием термина “механизм”», «Общее число статей в корпусе», «Число упоминаний термина “механизм”» показана корреляция от средней до сильной. Однако если учесть показатель «Общее число статей в корпусе» как корреляту, то частная корреляция (+0,704) будет только между показателями «Число упоминаний термина “механизм”» и «Число статей с упоминанием термина “механизм”» (что логично). Если учесть показатель «Число статей с упоминанием термина “механизм”» как корреляту, то изучаемые переменные значимой частной корреляции не показывают.

Таким образом, можно сделать вывод, что частота упоминаний термина «механизм» в целом имеет иную природу, чем частота

выхода всех статей в корпусе по годам. Для описания этой природы была построена 2B-PLS-модель.

Согласно графику осыпи латентных структур, сформированных моделью (рис. 2), все латентные структуры благополучно преодолели порог в 5% описываемой дисперсии и показали значимые нагрузки с изучаемыми нами показателями. Структура № 1 находится до первого перегиба графика и описывает общие особен-

ности, структуры № 2 и № 3 описывают частную специфику.

Как следует из рисунка 3, первая пара контрастных групп — технические и точные науки, с одной стороны, социальные и естественные — с другой. В годы с меньшим числом публикаций в корпусе, термин «механизм» чаще употребляется в одной статье в более старых работах технических и точных наук (относительно контрастной группы). В годы с большим числом публикаций в корпусе термин «механизм» реже употребляется в одной статье

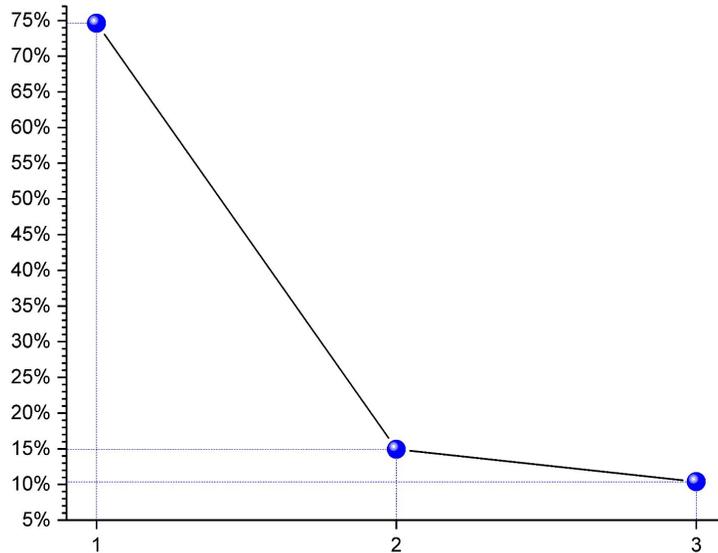


Рис. 2. График осыпи для латентных структур 2B-PLS-модели. Горизонтальная ось — латентные структуры, вертикальная — описываемая доля общей наблюдаемой дисперсии

Fig. 2. Scree plot for latent structures of the 2B-PLS model. The horizontal axis is the numbers of latent structures; the vertical axis is the described share of the total observed variance

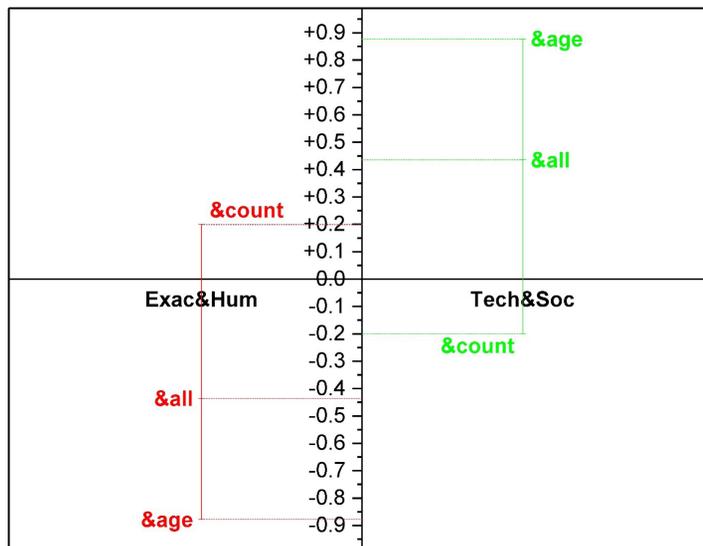


Рис. 3. Визуализация статистически значимых нагрузок для переменных (латентная структура № 1). Горизонтальная ось — контрастные группы коллекций статей из корпуса текстов, вертикальная — нагрузки (коэффициенты корреляции); обозначения на графике соответствуют таблице 1

Fig. 3. Visualization of statistically significant loadings for variables (latent structure No. 1). The horizontal axis—contrasting groups of collections of articles from the corpus of texts; the vertical axis—loadings (correlation coefficients); the symbols on the graph correspond to Table 1

в более молодых работах социальных и естественных наук (относительно контрастной группы).

Как следует из рисунка 4, вторая пара контрастных групп — естественные, технические и социальные науки, с одной стороны, гуманитарные науки — с другой. В годы с большим числом публикаций в корпусе термин «механизм» чаще употребляется в одной статье в более старых работах технических, социальных и естественных наук (относительно контрастной группы). В годы

с меньшим числом публикаций в корпусе термин «механизм» реже употребляется в одной статье в более молодых работах гуманитарных наук (относительно контрастной группы).

Как следует из рисунка 5, третья пара контрастных групп — гуманитарные и точные науки, с одной стороны, технические и социальные — с другой. В годы с меньшим числом публикаций в корпусе термин «механизм» чаще встречается в одной статье в более молодых работах

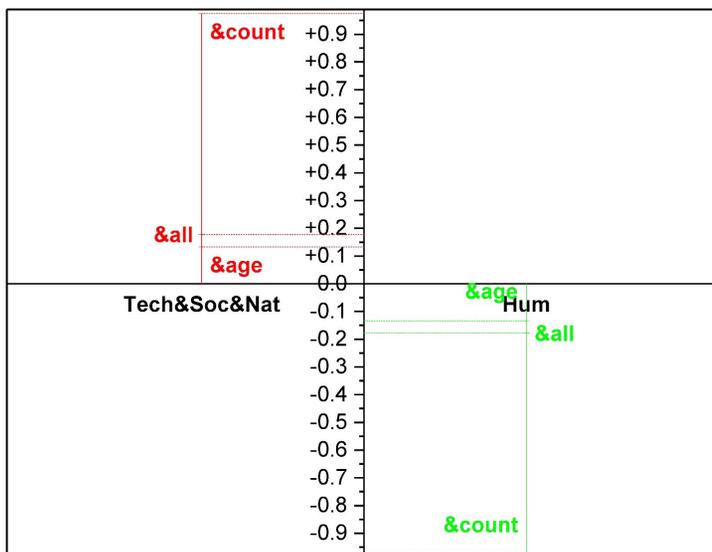


Рис. 4. Визуализация статистически значимых нагрузок для переменных (латентная структура № 2). Горизонтальная ось — контрастные группы коллекций статей из корпуса текстов, вертикальная — нагрузки (коэффициенты корреляции); обозначения на графике соответствуют таблице 1

Fig. 4. Visualization of statistically significant loadings for variables (latent structure No. 2). The horizontal axis—contrasting groups of collections of articles from the corpus of texts; the vertical axis—loadings (correlation coefficients); the symbols on the graph correspond to Table 1

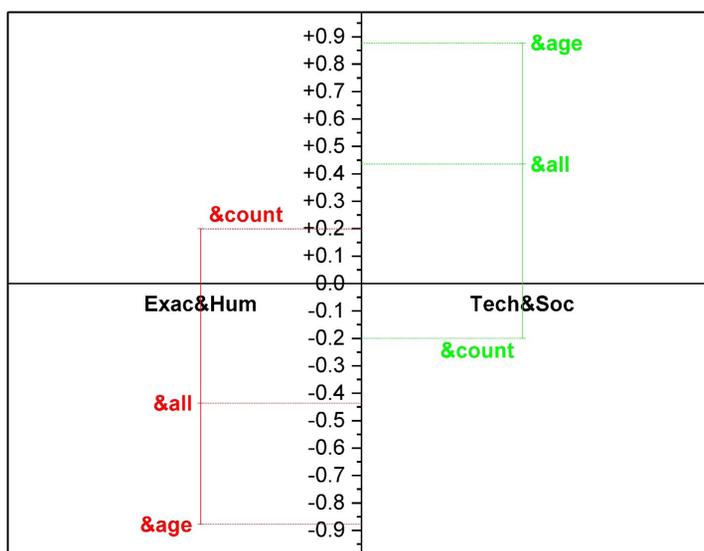


Рис. 5. Визуализация статистически значимых нагрузок для переменных (латентная структура № 3). Горизонтальная ось — контрастные группы коллекций статей из корпуса текстов, вертикальная — нагрузки (коэффициенты корреляции); обозначения на графике соответствуют таблице 1

Fig. 5. Visualization of statistically significant loadings for variables (latent structure No. 3). The horizontal axis—contrasting groups of collections of articles from the corpus of texts; the vertical axis—loadings (correlation coefficients); the symbols on the graph correspond to Table 1

точных и гуманитарных наук. В годы с бóльшим числом публикаций в корпусе термин «механизм» реже употребляется в одной статье в более старых работах технических и социальных наук.

Таким образом, можно сделать вывод, что более молодые статьи в гуманитарных и точных науках обсуждают аспекты полученных результатов с точки зрения их механизмов подробнее, чем в более старых работах технических и социальных наук (контрастные группы). При этом в периоды с редким выходом статей корпуса более молодые работы гуманитарных наук реже использовали термин «механизм» в одной публикации (по сравнению с контрастной группой — естественных, общественных и технических наук).

Можно предположить, что рост статей в корпусе в отдельные годы связан, например, с отчетностью по сравнительно недорогим грантам — и там важнее сам факт публикации статьи, чем ее качественное обсуждение.

## Выводы

При рассмотрении языка как системы действие механизмов лингвистики выражают его системные свойства: общие системные закономерности и частную специфику, которая связана с конкретными наборами данных.

**2B-PLS** позволяет обучить модель искать угол поворота предикторов, который обеспе-

чивает максимально информативный ответ для откликов (вопросов к модели).

Выявленные в ходе 2B-PLS-анализа латентные структуры обуславливают имплицитную связь между возрастом публикаций, отраслью наук и частотой упоминаний в статье термина «механизм». Можно предположить, что это связано с дискуссиями о механизмах в изучаемых науках, а в отдельные годы рост статей в корпусе связан, например, с отчетностью по сравнительно недорогим грантам — и там важнее сам факт публикации статьи, чем ее качественное обсуждение.

## Конфликт интересов

Авторы заявляют об отсутствии потенциального или явного конфликта интересов.

## Conflict of Interest

The authors declare that there is no conflict of interest, either existing or potential.

## Вклад авторов

Авторы внесли равный вклад в подготовку рукописи статьи.

## Author Contributions

The authors have made an equal contribution to the preparation of the manuscript of the article.

## Источники

Научная электронная библиотека «КиберЛенинка». (2022) *CYBERLENINKA*. [Электронный ресурс]. URL: <https://cyberleninka.ru/> (дата обращения 20.09.2022).

## Литература

- Алмаев, Н. А. (2012) *Применение контент-анализа в исследованиях личности*. М.: Институт психологии РАН, 167 с.
- Алмаев, Н. А., Мурашева, О. В. (2022) Тематический анализ дискуссий с применением метода латентного размещения Дирихле. *Институт психологии Российской академии наук. Социальная и экономическая психология*, т. 7, № 1 (25), с. 47–69. [https://doi.org/10.38098/ipran.sep\\_2022\\_25\\_1\\_03](https://doi.org/10.38098/ipran.sep_2022_25_1_03)
- Горбунов, Ю. В. (2018) О понятии «механизм» в экономических науках. *Экономика. Профессия. Бизнес*, т. 2, № 2, с. 17–21. <https://doi.org/10.14258/201819>
- Де Соссюр, Ф. (1977) *Труды по языкознанию*. М.: Прогресс, 695 с.
- Дмитриев, И. В. (2005) Контент-анализ: сущность, задачи, процедуры. *Пси-фактор*. [Электронный ресурс]. URL: <https://psyfactor.org/lib/k-a.htm> (дата обращения 20.10.2022).
- Ковалева, В. Ю., Поздняков, А. А., Литвинов, Ю. Н., Ефимов, В. М. (2019) Оценка сопряженности морфогенетических молекулярно-генетических модулей изменчивости серых полевков *Microtus S.L.* в градиентных условиях среды. *Экологическая генетика*, т. 17, № 2, с. 21–34. <https://doi.org/10.17816/ecogen17221-34>
- Комлев, Н. Г. (2006) *Словарь иностранных слов (более 4500 слов и выражений)*. М.: Эксмо, 669 с.
- Костюшкина, Г. М. (2012) В поисках системообразующего механизма в языке. *Вестник Иркутского государственного лингвистического университета*, № 2s (18), с. 128–133.
- Кривошеков, С. Г., Николаева, Е. И., Вергунов, Е. Г., Приходько, А. Ю. (2022) Многомерный анализ показателей тормозного и автономного контроля при ортостазе и в эмоциональных ситуациях. *Физиология человека*, т. 48, № 1, с. 26–37. <https://doi.org/10.31857/S0131164621060059>

- Лайонз, Д. (1978) *Введение в теоретическую лингвистику*. М.: Прогресс, 543 с.
- Мартюшов, В. Ф. (2015) Понятие «механизм» в контексте изучения социальных процессов. *Вестник Тверского государственного университета. Серия: Философия*, № 3, с. 94–103.
- Попов, В. П., Крайнюченко, И. В. (2012) *Теория и анализ систем*. Пятигорск: Изд-во ПГГТУ, 236 с.
- Теньер, Л. (1988) *Основы структурного синтаксиса*. М.: Прогресс, 656 с.
- Ушаков, Д. Н. (2004) *Большой толковый словарь русского языка*. М.: АСТ, 1268 с.
- Хомский, Н. (2015) *Создавая будущее: Оккупации, вторжения, имперское мышление и стабильность*. М.: Альпина нон-фикшн, 316 с.
- Blei, D. M., Ng, A. Y., Jordan, M. I. (2003) Latent Dirichlet allocation. *Journal of Machine Learning Research*, vol. 3, pp. 993–1022.
- Garla, A. (2021) NMF — A visual explainer and Python implementation. *Towards Data Science*, 18 March. [Online]. Available at: <https://towardsdatascience.com/nmf-a-visual-explainer-and-python-implementation-7ecdd73491f8> (accessed 18.10.22).
- Greenhalgh, T., Wong, G., Jagosh, J. et al. (2015) Protocol—the RAMESES II study: Developing guidance and reporting standards for realist evaluation. *BMJ Open*, vol. 5, article e008567. <http://doi.org/10.1136/bmjopen-2015-008567>
- Keller, R. (1994) *On language change: The invisible hand in language*. London; New York: Routledge Publ., 182 p.
- Krippendorp, K. (2004) *Content analysis: An introduction to its methodology*. Thousand Oaks: Sage Publ., 422 p.
- Landauer, T. K., Foltz, P. W., Laham, D. (1998) An introduction to latent semantic analysis. *Discourse Processes Journal*, vol. 25, no. 2-3, pp. 259–284. <https://doi.org/10.1080/01638539809545028>
- Lee, D., Seung, H. (1999) Learning the parts of objects by non-negative matrix factorization. *Nature*, vol. 401, no. 6755, pp. 788–791. <https://doi.org/10.1038/44565>
- Machamer, P., Darden, L., Craver, C. (2000) Thinking about mechanisms. *Philosophy of Science*, vol. 67, no. 1. [Online]. Available at: <https://doi.org/10.1086/392759> (accessed 18.10.22).
- Nikolaeva, E.I., Efimova, V. L., Vergunov, E. G. (2022) Integration of vestibular and auditory information in ontogenesis. *Children*, vol. 9, no. 3, article 401. <https://doi.org/10.3390/children9030401>
- Polunin, D, Shtager, I, Efimov, V. (2019) JACOBI4 software for multivariate analysis of biological data. *bioRxiv*. [Online]. Available at: <https://doi.org/10.1101/803684> (accessed 18.10.22).
- Rännar, S., Lindgren, F., Geladi, P., Wold, S. (1994) A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. *Journal of Chemometrics*, vol. 8, no. 2, pp. 111–125. <https://doi.org/10.1002/cem.1180080204>
- Rohlf, F. J., Corti, M. (2000) Use of two-block partial least-squares to study covariation in shape. *Systematic Biology*, vol. 49, no. 4, pp. 740–753. <https://doi.org/10.1080/106351500750049806>
- Savostyanov, A. N., Vergunov, E. G., Saprygin, A. E., Lebedkin, D. A. (2022) Validation of a face image assessment technology to study the dynamics of human functional states in the EEG resting-state paradigm. *Vavilov Journal of Genetics and Breeding*, vol. 26, no. 8, pp. 765–772. <https://doi.org/10.18699/VJGB-22-92>
- Vergunov, E. G. (2022) Coping space transformation at different levels of university training during the pandemic and the assessment of its integral indicators. *Kompleksnye issledovaniya detstva — Comprehensive Child Studies*, vol. 4, no. 2, pp. 115–123. <https://doi.org/10.33910/2687-0223-2022-4-2-115-123>

## Sources

- Nauch'naya elektronnyaya biblioteka "KiberLeninka" [Scientific electronic library "CyberLeninka"]. (2022) *CYBERLENINKA*. [Online]. Available at: <https://cyberleninka.ru/> (accessed 20.09.2022). (In Russian)

## References

- Almaev, N. A. (2012) *Primenenie kontent-analiza v issledovaniyakh lichnosti [The use of content analysis in personality research]*. Moscow: Institute of Psychology of Russian Academy of Sciences Publ., 167 p. (In Russian)
- Almaev, N. A., Murasheva, O. V (2022) Tematicheskij analiz diskussij s primeneniem metoda latentnogo razmeshcheniya Dirikhle [Thematic analysis of discussions using the latent Dirichlet allocation]. *Institut psikhologii Rossijskoj akademii nauk. Sotsial'naya i ekonomicheskaya psikhologiya — Institute of Psychology Russian Academy of Sciences. Social and Economic Psychology*, vol. 7, no. 1 (25), pp. 47–69. <https://doi.org/10.38098/ipran.sep.2022.25.1.03> (In Russian)
- Blei, D. M., Ng, A. Y., Jordan, M. I. (2003) Latent Dirichlet allocation. *Journal of Machine Learning Research*, vol. 3, pp. 993–1022. (In English)
- de Saussure, F. (1977) *Trudy po yazykoznaniyu [Works on linguistics]*. Moscow: Progress Publ., 695 p. (In Russian)
- Dmitriev, I. V. (2005) Kontent-analiz: sushchnost', zadachi, protsedury [Content analysis: Entity, tasks, procedure]. *Psi-faktor [Psyfactor]*. [Online]. Available at: <https://psyfactor.org/lib/k-a.htm> (accessed 20.10.2022). (In Russian)
- Garla, A. (2021) NMF—a visual explainer and Python implementation. *Towards Data Science*, 18 March. [Online]. Available at: <https://towardsdatascience.com/nmf-a-visual-explainer-and-python-implementation-7ecdd73491f8> (accessed 18.10.22). (In English)

- Chomsky, N. (2015) Making the future. Occupations, interventions, empire and resistance. Moscow: Alpina Non-Fiction Publ., 316 p (In Russian)
- Gorbunov, Yu. V. (2018) O ponyatii "mekhanizm" v ekonomicheskikh naukakh [Towards the mechanism conception in economic]. *Ekonomika. Professiya. Biznes — Economics. Profession. Business*, vol. 2, no. 2, pp. 17–21. <https://doi.org/10.14258/201819> (In Russian)
- Greenhalgh, T., Wong, G., Jagosh, J. et al. (2015) Protocol — the RAMESES II study: Developing guidance and reporting standards for realist evaluation. *BMJ Open*, vol. 5, article e008567. <http://doi.org/10.1136/bmjopen-2015-008567> (In English)
- Keller, R. (1994) *On language change: The invisible hand in language*. London; New York: Routledge Publ., 182 p. (In English)
- Khomsikij, N. (2015) *Sozdavaya budushchee: Okkupatsii, vtorzheniya, imperskoe myshlenie i stabil'nost'* [Creating the future: Occupations, invasions, imperial thinking and stability]. Moscow: Al'pina non-fikshn, 316 p. (In Russian)
- Komlev, N. G. (2006) *Slovar' inostrannykh slov (bole 4500 slov i vyrazhenij)* [Dictionary of foreign words (more than 4500 words and expressions)]. Moscow: Eksmo Publ., 669 p. (In Russian)
- Kostyushkina, G. M. (2012) V poiskakh sistemoobrazuyushchego mekhanizma v yazyke [Research of systemic mechanism in language]. *Vestnik Irkutskogo gosudarstvennogo lingvisticheskogo universiteta*, no. 2s (18), pp. 128–133. (In Russian)
- Kovaleva, V. Yu., Pozdnyakov, A. A., Litvinov, Yu. N., Efimov, V. M. (2019) Otsenka sopryazhennosti morfogeneticheskikh molekulyarno-geneticheskikh modulej izmenchivosti serykh polevok *Microtus S.L.* v gradientnykh usloviyakh sredi [Estimation of the conjugation between morphogenetic and molecular-genetic modules of gray voles *Microtus S.L.* variability along a climatic gradient conditions]. *Ekologicheskaya genetika — Ecological Genetics*, vol. 17, no. 2, pp. 21–34. <https://doi.org/10.17816/ecogen17221-34> (In Russian)
- Krippendorp, K. (2004) *Content analysis: An introduction to its methodology*. Thousand Oaks: Sage Publ., 422 p. (In English)
- Krivoshchekov, S. G., Nikolaeva, E. I., Vergunov, E. G., Prihod'ko, A. Yu. (2022) Mnogomernyj analiz pokazatelej tormoznogo i avtonomnogo kontrolya pri ortostaze i v emotsional'nykh situatsiyakh [Multivariate analysis of indicators of inhibitory and autonomic control in orthostasis and emotional situations]. *Fiziologiya cheloveka — Human Physiology*, vol. 48, no. 1, pp. 26–37. <https://doi.org/10.31857/S0131164621060059> (In Russian)
- Lajonz, D. (1978) *Vvedenie v teoreticheskuyu lingvistiku* [Introduction to Theoretical Linguistics]. Moscow: Progress Publ., 543 p. (in Russian)
- Landauer, T. K., Foltz, P. W., Laham, D. (1998) An introduction to latent semantic analysis. *Discourse Processes Journal*, vol. 25, no. 2-3, pp. 259–284. <https://doi.org/10.1080/01638539809545028> (In English)
- Lee, D., Seung, H. (1999) Learning the parts of objects by non-negative matrix factorization. *Nature*, vol. 401, no. 6755, pp. 788–791. <https://doi.org/10.1038/44565>. (In English)
- Machamer, P., Darden, L., Craver, C. (2000) Thinking about mechanisms. *Philosophy of Science*, vol. 67, no. 1. [Online]. Available at: <https://doi.org/10.1086/392759> (accessed 18.10.22). (In English)
- Martyushov, V. F. (2015) Ponyatie "mekhanizm" v kontekste izucheniya sotsial'nykh protsessov [The concept of "mechanism" in the context of the study of social processes]. *Vestnik Tverskogo gosudarstvennogo universiteta. Seriya: Filosofiya — Vestnik Tver State University. Series: Philosophy*, no. 3, pp. 94–103. (In Russian)
- Nikolaeva, E. I., Efimova, V. L., Vergunov, E. G. (2022) Integration of vestibular and auditory information in ontogenesis. *Children*, vol. 9, no. 3, article 401. <https://doi.org/10.3390/children9030401>. (In English)
- Polunin, D., Shtager, I., Efimov, V. (2019) JACOBI4 software for multivariate analysis of biological data. *bioRxiv*. [Online]. Available at: <https://doi.org/10.1101/803684> (accessed 18.10.22). (In English)
- Popov, V. P., Krajnyuchenko, I. V. (2012) *Teoriya i analiz sistem* [Theory and analysis of systems]. Pyatigorsk: Pyatigorskij gosudarstvennyj gumanitarno-tekhnologicheskij universitet Publ., 236 p. (In Russian)
- Rännar, S., Lindgren, F., Geladi, P., Wold, S. (1994) A PLS kernel algorithm for data sets with many variables and fewer objects. Part 1: Theory and algorithm. *Journal of Chemometrics*, vol. 8, no. 2, pp. 111–125. <https://doi.org/10.1002/cem.1180080204> (In English)
- Rohlf, F. J., Corti, M. (2000) Use of two-block partial least-squares to study covariation in shape. *Systematic Biology*, vol. 49, no. 4, pp. 740–753. <https://doi.org/10.1080/106351500750049806> (In English)
- Savostyanov, A. N., Vergunov, E. G., Saprygin, A. E., Lebedkin, D. A. (2022) Validation of a face image assessment technology to study the dynamics of human functional states in the EEG resting-state paradigm. *Vavilov Journal of Genetics and Breeding*, vol. 26, no. 8, pp. 765–772. <https://doi.org/10.18699/VJGB-22-92> (In English)
- Tesnière, L. (1988) *Osnovy strukturnogo sintaksisa* [Fundamentals of structural syntax]. Moscow: Progress Publ., 656 p. (In Russian)
- Ushakov, D. N. (2004) *Bol'shoj tolkovyj slovar' russkogo yazyka* [Big explanatory dictionary of the Russian language]. Moscow: AST Publ., 1268 p. (In Russian)
- Vergunov, E. G. (2022) Coping space transformation at different levels of university training during the pandemic and the assessment of its integral indicators. *Kompleksnye issledovaniya detstva — Comprehensive Child Studies*, vol. 4, no. 2, pp. 115–123. <https://doi.org/10.33910/2687-0223-2022-4-2-115-123> (In Russian)