



МЕЖДИСЦИПЛИНАРНЫЕ ИССЛЕДОВАНИЯ ЯЗЫКА/INTERDISCIPLINARY LANGUAGE STUDIES

DOI: <https://doi.org/10.60797/IRJ.2026.166.26> EDN: SMGNPW

ЦИФРОВИЗАЦИЯ И КОГНИТИВНАЯ ЛИНГВИСТИКА: КАК ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ МЕНЯЕТ НАШЕ ПОНИМАНИЕ ЯЗЫКОВЫХ ИССЛЕДОВАНИЙ

Научная статья

Анненкова А.В.^{1,*}¹ ORCID : 0000-0001-6870-8126;¹ Юго-Западный государственный университет, Курск, Российская Федерация

* Корреспондирующий автор (antonina-1984[at]yandex.ru)

Аннотация

В статье анализируется трансформация когнитивной лингвистики под влиянием цифровизации и искусственного интеллекта (ИИ). Традиционные методы дополняются машинным обучением, меняя природу знания и верификацию гипотез. Это порождает методологические парадоксы: закономерности ИИ требуют «вычислительной герменевтики», а проблема «чёрного ящика» ставит под вопрос когнитивную аргументированность моделей. Дополнительными вызовами становятся этические аспекты: предвзятость корпусов и деперсонализация данных. Для решения эпистемологических, этических и интерпретативных проблем предлагается гибридная методология, сочетающая вычислительный анализ с качественными исследованиями. Обосновывается сближение компьютерной и когнитивной лингвистики, требующее повышения цифровой грамотности и критического подхода к результатам, полученным при помощи ИИ.

Ключевые слова: когнитивная лингвистика, цифровизация, искусственный интеллект, большие языковые модели, машинное обучение, компьютерная лингвистика, междисциплинарность.

DIGITALISATION AND COGNITIVE LINGUISTICS: HOW ARTIFICIAL INTELLIGENCE IS CHANGING OUR UNDERSTANDING OF LANGUAGE RESEARCH

Research article

Annenkova A.V.^{1,*}¹ ORCID : 0000-0001-6870-8126;¹ Southwest State University, Kursk, Russian Federation

* Corresponding author (antonina-1984[at]yandex.ru)

Abstract

The article analyses the transformation of cognitive linguistics under the influence of digitalisation and artificial intelligence (AI). Traditional methods are being supplemented by machine learning, altering the nature of knowledge and the verification of hypotheses. This creates methodological paradoxes: the principles of AI require "computational hermeneutics", and the "black box" problem calls into question the cognitive soundness of models. Ethical aspects present further challenges: corpus bias and the depersonalisation of data. To tackle epistemological, ethical and interpretative problems, a hybrid methodology is suggested, combining computational analysis with qualitative research. The convergence of computational and cognitive linguistics is substantiated, requiring improved digital literacy and a critical approach to results obtained using AI.

Keywords: cognitive linguistics, digitalisation, artificial intelligence, large language models, machine learning, computational linguistics, interdisciplinarity.

Введение

Сегодня цифровизация охватывает все сферы человеческой деятельности, включая научное познание в рамках гуманитарной парадигмы, которые развиваются и эволюционируют, всё теснее вступая в симбиоз с естественными и математическими науками. Современная лингвистика переживает период глубокой трансформации, обусловленной стремительной цифровизацией научной среды. Если ещё несколько десятилетий назад когнитивная лингвистика преимущественно опиралась на применение качественных экспериментальных методов, таких как анализ метафор, концептуальных интеграций, психолингвистические эксперименты и корпусные наблюдения, то на сегодняшний день наблюдается сдвиг в сторону привлечения в исследовательскую практику таких инновационных методов как алгоритмы машинного обучения, нейросетевые архитектуры и автоматизированные системы анализа естественного языка. Искусственный интеллект (ИИ) в форме больших языковых моделей (БЯМ) не только трансформирует инструментарий лингвистической науки, но и видоизменяет природу лингвистического знания: теперь нужно переопределять то, что будет считаться данными, решать как их интерпретировать, какие вопросы задавать и как проверять гипотезы.

Основные результаты

Рассмотрим и проанализируем процесс трансформации методологии языковых исследований как отражения влияния цифровизации и внедрения систем ИИ в когнитивную лингвистику, а также коснемся вопроса возникновения новых научных практик и порождаемых ими теоретико-методологических дилемм.

Традиционно когнитивная лингвистика использовала язык для доступа к разуму: через анализ языковых структур исследователь реконструировал ментальные процессы говорящего. Такой подход требовал глубокого, часто



субъективного, но теоретически обоснованного анализа текстового и, соответственно, языкового материала [1]. Несколько лет назад корпусная лингвистика при помощи технических средств частично автоматизировала этот процесс, но всё ещё оставляла за исследователем ключевую роль в интерпретации коллокаций, частотных паттернов и семантических полей для верификации гипотез. С появлением БЯМ ситуация изменилась кардинально. Современные модели способны практически мгновенно не только извлекать миллионы примеров употребления искомого языковых паттернов, но и генерировать гипотезы о семантической структуре, выявлять скрытые аналогии, реконструировать исторические изменения значения или даже предлагать интерпретации контекстуальных метафор. Например, языковая модель для обработки естественного языка двунаправленной архитектуры BERT от корпорации Google позволяет визуализировать семантические пространства и отслеживать динамику значений в зависимости от контекста за несколько секунд, тогда как ранее для решения подобной задачи требовались месяцы ручной работы.

Тем не менее, такая автоматизация приводит к возникновению методологического парадокса: если машина и ИИ «видят» закономерности, недоступные человеческому восприятию, то каким образом исследователь может их грамотно верифицировать? Следовательно, появляется необходимость новой формы понимания и анализа текста — вычислительной герменевтики, сочетающей статистическую достоверность с классической теоретической интерпретацией, что требует от лингвиста не только владения традиционными методами, но и базовой компетенции в области data science и особенностях интерпретации результатов, полученных при помощи ИИ.

Цифровизация привела к возникновению ещё одного вызова для когнитивной лингвистики — проблемы «чёрного ящика», связанной с непрозрачностью работы БЯМ [2]. Архитектуры нейронных сетей языковых моделей типа GPT или Llama принимают решения на основе анализа огромного количества параметров, что делает невозможным проследивание логики формирования итогового ответа. Для когнитивной науки, где важна реконструкция ментальных механизмов, такая ситуация представляет серьёзную проблему: если нет возможности объяснить, почему и каким образом модель использует ту или иную языковую структуру или категоризацию, то можно ли считать её поведение «когнитивным»?

Эта проблема напрямую затрагивает методологию когнитивной лингвистики: допустимо ли использовать ИИ как инструмент генерации гипотез без понимания работы его внутренней логики? Некоторые исследователи (напр., [3]) настаивают на необходимости развития в гуманитарной парадигме феномена объяснимого ИИ. Другие (напр., [4]) предлагают рассматривать нейронные языковые модели как «симуляторы» когнитивных процессов, чья ценность не в точности воспроизведения, а в способности выявлять частотные тенденции, которые уже затем будут подвергаться теоретическому описанию человеком. Следовательно, цифровизация не нивелирует роль лингвиста, но смещает её от прямого анализа к критической интерпретации вычислительных результатов.

Важно понимать, что классическая когнитивная лингвистика исходила из того, что язык представляет собой продукт конкретного когнитивного субъекта, обладающего телесным опытом, культурной памятью и социальной идентичностью. Цифровые модели, напротив, обучаются на агрегированных, деперсонифицированных данных, лишённых контекста происхождения, что приводит к стиранию индивидуального и культурного многообразия в пользу доминирующих дискурсов, представленных в интернете [5]. С точки зрения исследователя подобная трансформация языковых данных инициирует понимание необходимости и обоснованности внедрения критического подхода к данным работы моделей ИИ: нельзя принимать выводы модели как универсальную истину [6]. Более того, использование ИИ в лингвистических исследованиях требует соблюдения этических норм — учёта когнитивных искажений в обучающих корпусах, прозрачности источников и ответственности за проведённую интерпретацию.

С теоретической точки зрения указанная дилемма поднимает вопрос о том, может ли лингвистическое исследование, опирающееся на данные, полученные от языковых моделей ИИ, сохранять когнитивную направленность, будучи «оторванным» от живого субъекта исследования [7]? Решением может стать внедрение гибридной методологии когнитивной лингвистики, комбинирующей технологии вычислительного анализа с качественными методами исследования (напр., интервью, экспериментами, наблюдением), что позволит вернуть человека в центр исследования.

Примером практической реализации гибридного подхода применительно к когнитивной лингвистике может служить, например, изучение концептуальных метафор. На первом этапе исследователями будут применяться нейросетевые алгоритмы категоризации и кластеризации к большому корпусу текстов для выявления наличия статистических связей между целевым доменом (например, «время») и областью-источником (например, «пространство»). Затем необходимо провести качественный контекстуальный анализ в целях верификации когнитивной зависимости указанных связей и исключения артефактов и статистически незначимых шумов. Завершающим этапом станет верификация через проведение психолингвистических экспериментов, результаты которых смогут подтвердить достоверность выявленной при помощи ИИ закономерности, которая действительно отражает ментальные процессы в сознании индивида, а не является неким артефактом в «цифровом» корпусе данных, случайно обнаруженным нейросетями ИИ. Аналогичным образом можно поступить и при работе с эмоциональными концептами, векторное пространство которых будет строиться с помощью нейросетевых моделей ИИ, а затем сопоставляться с данными проведённых ассоциативных экспериментов. Это позволит выявить наличие зон расхождения между пониманием значения индивидом и искусственным интеллектом, что может быть обусловлено предвзятостью обучающих данных интеллектуальных систем и систематическими ошибкам.

Следовательно, трансформация когнитивной лингвистики под влиянием цифровизации открывает не только новые методологические возможности, но и приводит к возникновению этических и эпистемологических проблем, связанных с природой данных ИИ, интерпретацией результатов и самим статусом языкового знания в новой цифровой эпохе. Осмыслить эти основания и изменения помогает компьютерная лингвистика, изначально направленная на работу с формализацией, автоматизацией и вычислительным моделированием языка и языковых моделей [8]. Компьютерная лингвистика предоставляет техническую базу для дальнейшей разработки, обучения и имплементации



языковых моделей ИИ, кроме того, она позволяет формировать концептуальный аппарат лингвистики, описывающий динамику взаимодействия пары «естественный язык — вычислительные системы». Таким образом, в современной языковой парадигме компьютерную лингвистику и когнитивную лингвистику можно охарактеризовать как взаимодополняющие направления, в общую сферу интересов которых входит структура, функционирование и обработка естественного языка, но которые отличаются разными методологическими подходами, установками и целями исследования. Несмотря на различия, в последнее время наблюдается всё большее сближение и интеграция этих дисциплин, когда данные когнитивных исследований внедряются для повышения интерпретируемости и семантической точности языковых моделей ИИ, а достижения компьютерной лингвистики создают предпосылки для эмпирической верификации когнитивных гипотез на больших корпусах языкового материала.

Компьютерная лингвистика в условиях цифровизации трансформируется от правило-ориентированных систем к нейросетевым архитектурам, формируя лингвистическое знание на стыке гуманитарных и точных наук. Ключевыми инновациями стали методы машинного обучения (МО) и глубокие нейронные сети (ГНС). МО — нацелено на создание алгоритмов, автоматически извлекающих знания из данных и в лингвистике оно выступает не только как инструмент обработки языка [9], но и как новое методологическое основание когнитивной лингвистики, позволяющее реконструировать когнитивные репрезентации на основе больших корпусов, обеспечивая тем самым эмпирическую верификацию теорий о связи языка и мышления. ГНС, имитирующие иерархическую организацию биологических нейронных систем, благодаря многослойной архитектуре эффективно решают сложные задачи обработки естественного языка [10] — от распознавания речи и машинного перевода до генерации текста, демонстрируя потенциал для моделирования как формальных, так и когнитивных аспектов языка.

Заключение

Таким образом, цифровизация и ИИ не просто дополняют арсенал когнитивной лингвистики, но и вызывают необходимость пересмотра её методологических основ. Языковые исследования становятся всё более междисциплинарными, требуя от исследователя сочетания лингвистической компетенции с цифровой грамотностью, при этом сохраняя ключевую задачу в понимании связей между языком и познанием.

Будущее лингвистических исследований, ориентированных на когнитивные процессы, видится в разработке методологии, которая использует ИИ как инструмент расширения аналитических возможностей, критически оценивает ограничения и предвзятости моделей, сохраняет культурно-историческую и теоретическую глубину, интегрирует количественные и качественные подходы. Только такой синтез позволит когнитивной лингвистике сохранить свою научную релевантность в эпоху цифровой трансформации.

Конфликт интересов

Не указан.

Conflict of Interest

None declared.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Чилингарян К.П. Корпусная лингвистика: теория vs методология / К.П. Чилингарян // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. — 2021. — № 1 (12). — С. 196–218. — DOI: 10.22363/2313-2299-2021-12-1-196-218
2. Bathaee Y. The Artificial Intelligence Black Box and the Failure of Intent and Causation / Y. Bathaee // Harvard Journal of Law & Technology. — 2018. — № 2 (31). — P. 890–934.
3. Райков А.Н. Субъектность объяснимого искусственного интеллекта / А.Н. Райков // Философские науки. — 2022. — № 1 (65). — С. 72–90. — DOI: 10.30727/0235-1188-2022-65-1-72-90
4. Bender E.M. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? / E.M. Bender, T. Gebru, A. McMillan-Major et al. // Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT'21). Association for Computing Machinery; — New York: ACM, 2021. — P. 610–623.
5. Tao Y. Cultural Bias and Cultural Alignment of Large Language Models / Y. Tao, O. Viberg, S.R. Baker et al. // PNAS Nexus. — 2024. — № 3 (9). — P. 346.
6. Ивашко К.С. Компьютерная лингвистика в психологии: ключ к пониманию языка и поведения человека / К.С. Ивашко, С.А. Изосимова, В.Н. Пигуз и др. // Язык и текст. — 2025. — № 2 (12). — С. 140–162. — DOI: 10.17759/langt.2025120213
7. Мишанкина Н.А. Специфика метафорического моделирования научного дискурса / Н.А. Мишанкина // Вопросы когнитивной лингвистики. — 2010. — № 1 (22). — С. 37–46.
8. Томарева И.Г. Взаимодействие ИИ и когнитивной лингвистики: актуальное положение и будущие направления / И.Г. Томарева // Международный научно-исследовательский журнал. — 2024. — № 6 (144). — DOI: 10.60797/IRJ.2024.144.86
9. Найденова К.А. Машинное обучение в задачах обработки естественного языка: обзор современного состояния исследований / К.А. Найденова, О.А. Невзорова // Ученые записки Казанского государственного университета. Серия: Физико-математические науки. — 2008. — № 4 (150). — С. 5–24.



10. Zhu D. Couplet Analysis of Linguistic Topology Using Deep Neural Networks in Cognitive Linguistics / D. Zhu, N, Wang, F. Yang // *Computational Intelligence and Neuroscience*. — 2022. — № 11. — P. 912–922.

Список литературы на английском языке / References in English

1. Chilingaryan K.P. Korpusnaya lingvistika: teoriya vs metodologiya [Corpus Linguistics: Theory Vs Methodology] / K.P. Chilingaryan // *PFUR Journal of Language Studies, Semiotics and Semantics*. — 2021. — № 1 (12). — P. 196–218. — DOI: 10.22363/2313-2299-2021-12-1-196-218 [in Russian]
2. Bathaee Y. The Artificial Intelligence Black Box and the Failure of Intent and Causation / Y. Bathaee // *Harvard Journal of Law & Technology*. — 2018. — № 2 (31). — P. 890–934.
3. Rajkov A.N Sub"ektnost' ob"yasnimogo iskusstvennogo intellekta [Subjectivity of Explainable Artificial Intelligence] / A.N Rajkov // *Russian Journal of Philosophical Sciences*. — 2022. — № 1 (65). — P. 72–90. — DOI: 10.30727/0235-1188-2022-65-1-72-90 [in Russian]
4. Bender E.M. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? / E.M. Bender, T. Gebru, A. McMillan-Major et al. // *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT'21)*. Association for Computing Machinery; — New York: ACM, 2021. — P. 610–623.
5. Tao Y. Cultural Bias and Cultural Alignment of Large Language Models / Y. Tao, O. Viberg, S.R. Baker et al. // *PNAS Nexus*. — 2024. — № 3 (9). — P. 346.
6. Ivashko K.S. Komp'yuternaya lingvistika v psixologii: klyuch k ponimaniyu yazy'ka i povedeniya cheloveka [Computational Linguistics in Psychology: a Key to Understanding Language and Human Behavior] / K.S. Ivashko, S.A. Izosimova, V.N. Piguz et al. // *Language and Text*. — 2025. — № 2 (12). — P. 140–162. — DOI: 10.17759/langt.2025120213 [in Russian]
7. Mishankina N.A. Specifika metaforicheskogo modelirovaniya nauchnogo diskursa [Specificity of Metaphorical Modelling of the Scientific Discourse] / N.A. Mishankina // *Issues of Cognitive Linguistics*. — 2010. — № 1 (22). — P. 37–46. [in Russian]
8. Tomareva I.G. Vzaimodejstvie II i kognitivnoj lingvistiki: aktual'noe polozhenie i budushhie napravleniya [Interaction Between AI and Cognitive Linguistics: Current Status and Future Directions] / I.G. Tomareva // *International Research Journal*. — 2024. — № 6 (144). — DOI: 10.60797/IRJ.2024.144.86 [in Russian]
9. Najdenova K.A. Mashinnoe obuchenie v zadachax obrabotki estestvennogo yazy'ka: obzor sovremennogo sostoyaniya issledovaniy [Machine Learning for Natural Language Processing: Contemporary State] / K.A. Najdenova, O.A. Nevzorova // *Proceedings of Kazan University. Physics and Mathematics Series*. — 2008. — № 4 (150). — P. 5–24. [in Russian]
10. Zhu D. Couplet Analysis of Linguistic Topology Using Deep Neural Networks in Cognitive Linguistics / D. Zhu, N, Wang, F. Yang // *Computational Intelligence and Neuroscience*. — 2022. — № 11. — P. 912–922.