



4. Vladlenov, R., et al. (2023). Philology, Linguistics, and the Digital Age: An Interdisciplinary Approach. *Journal of Language and Cognitive Development*, 10(1), 87-102.
5. Yang, L., et al. (2023). Classification of Medical Image Notes for Image Labeling by Using MinBERT. *Journal of Medical Informatics*, 22(4), 456-467.
6. Далиева, М. Х., & Сатибалдиев, Э. К. (2023). ПОЛИСЕМИЯ ТЕРМИНОЛОГИИ КАК ОБЪЕКТ ИССЛЕДОВАНИЙ В СОВРЕМЕННОМ ЯЗЫКОЗНАНИИ. In *ББК 81.2 я43 Методика преподавания иностранных языков и РКИ: традиции и инновации: сборник научных трудов VIII Международной научно-методической онлайн-конференции, посвященной Году педагога и наставника в России и Году русского языка в странах СНГ (11 апреля 2023 г.)*—Курск: Изд-во КГМУ, 2023.—521 с. (p. 47).
7. Kamilovich, S. E. (2023). EXPLORING LINGUISTIC UNIVERSALS AND TYPOLOGICAL PATTERNS: AN ANALYSIS OF THE COGNITIVE AND CULTURAL FACTORS THAT SHAPE LANGUAGE STRUCTURES ACROSS DIVERSE LANGUAGES. *American Journal of Pedagogical and Educational Research*, 10, 129-132.
8. Сатибалдиев, Э. К. (2022). ЯЗЫКОВОЕ КОНТАКТИРОВАНИЕ: БИЛИНГВИЗМ, ПОЛИЛИНГВИЗМ, ИНТЕРФЕРЕНЦИЯ. In *ИНОСТРАННЫЙ ЯЗЫК В ПРОФЕССИОНАЛЬНОЙ СФЕРЕ: ПЕДАГОГИКА, ЛИНГВИСТИКА, МЕЖКУЛЬТУРНАЯ КОММУНИКАЦИЯ* (pp. 144-149).
9. Temirova, N. A. (2023). CONSECUTIVE INTERPRETATION AS A TYPE OF ORAL TRANSLATION. *Academia Science Repository*, 4(6), 197-204.
10. Satibaldieva, N. (2024). Polysemy of Terms in Computational Linguistics. *International Journal of Scientific Trends*, 3(1), 82-84

Применение методов анализа компьютерной лингвистики при составлении каталога многозначных терминов

Урманова Динара

УзГУМЯ, магистрант

Аннотация

Эта статья посвящена анализу и интерпретации полисемичных терминов в контексте компьютерной лингвистики. Рассматриваются методы использования корпусного анализа, семантических алгоритмов и инструментов контекстного анализа для эффективного различения и категоризации значений полисемичных терминов. Подчеркивается важность вычислительных инструментов в лингвистике и их роль в понимании эволюции языка. Исследования демонстрируют, как контекст влияет на значение слов, и показывают значение междисциплинарного подхода к изучению полисемии.

Ключевые слова: полисемия, корпусный анализ, семантические алгоритмы, контекстный анализ, вычислительные инструменты, эволюция языка, лексическая неоднозначность, междисциплинарный подход.

В постоянно развивающейся области компьютерной лингвистики одной из наиболее интересных задач является анализ и интерпретация полисемичных терминов — слов, имеющих несколько значений. Эта сложность имеет значительные последствия для обработки естественного языка (NLP),



семантического анализа, языкового образования и машинного перевода. В данной статье рассматривается применение методов компьютерного лингвистического анализа в сложном процессе каталогизации полисемичных терминов. Основное внимание уделяется использованию корпусного анализа, семантических алгоритмов и инструментов контекстного анализа для эффективного различения и категоризации различных значений полисемичных терминов. Этот синтез подчеркивает важность вычислительных инструментов в лингвистике и акцентирует динамическую связь между эволюцией языка и технологическими достижениями в понимании человеческого языка.

В статье "Exploring Linguistic Patterns: Unveiling Insights from Text Corpus and Analytical Methods" Далиева (2023) предлагает всестороннее исследование использования больших текстовых корпусов и методов компьютерной лингвистики для анализа языковых моделей. Их работа подчеркивает важную роль вычислительных инструментов в разборе сложных языковых структур, что необходимо для понимания полисемии. Этот подход соответствует взглядам Джурафски и Мартина (2018), которые подчеркивают значимость вычислительных методов в обработке речи и языка, акцентируя их способность расшифровывать сложные языковые модели и нюансы. Исследование Далиевой и Са'дуллаевой использует эти методы для раскрытия многогранной природы языка, особенно в контексте полисемичных терминов.

Кроме того, исследование созвучно с основными концепциями, представленными в работе МакЭнери и Харди (1996) по корпусной лингвистике. МакЭнери и Харди выступают за использование крупных корпусов как инструмента для наблюдения и анализа языковых явлений. Применяя такие методы корпусного анализа, исследование Далиевой и Са'дуллаевой проливает свет на модели и частоту использования полисемичных терминов в различных языковых контекстах. Их подход иллюстрирует, как количественные данные, полученные из корпусной лингвистики, могут информировать качественные



интерпретации использования языка, особенно в понимании того, как контекст влияет на значение полисемичных слов.

Продолжая это исследование, Далиева и Каримбаева (2023) в статье "Unveiling Lexical Ambiguity: Investigating Polysemy through Corpus Analysis" конкретно рассматривают проблемы, возникающие при работе с полисемичными терминами. Их исследование эффективно демонстрирует применение корпусного анализа для различения различных значений одного и того же слова с учетом контекста его использования. Этот методологический подход особенно актуален при учете выводов Хансена (2015) о будущем, нереальном и контрфактическом дискурсе новостей. Корпусный лингвистический анализ Хансена показывает, как контекст может радикально изменить интерпретацию языковых элементов, что является центральной концепцией для понимания полисемии.

Представленные исследования развивают эту концепцию, применяя ее специально к полисемичным терминам. Они используют передовые вычислительные техники для парсинга и интерпретации больших наборов данных, что позволяет глубже понять, как контекст влияет на значение слов. Их работа является свидетельством силы корпусного анализа в раскрытии тонких различий в использовании языка. Изучая разнообразные текстовые источники, они предоставляют более полное представление о том, как полисемичные термины функционируют в различных языковых средах, значительно способствуя нашему пониманию лексической неоднозначности и ее последствий для лингвистики и обработки естественного языка.

В области понимания и обработки языка полисемия представляет собой значительную задачу. Эта сложность возникает из-за того, что одно слово может иметь несколько значений, каждое из которых зависит от контекста, что делает точную интерпретацию языка сложной задачей. Для решения этой проблемы важна роль корпусной лингвистики. Анализируя большие текстовые корпуса, лингвисты могут эффективно идентифицировать и категоризировать



полисемичные термины. Исследования таких ученых, как Далиева и Са'дуллаева, подчеркивают важность таких методологий для расшифровки сложных языковых моделей. Для облегчения этого анализа используются различные вычислительные техники, включая токенизацию, часть речи и контекстный анализ, которые необходимы для обработки и понимания текстовых данных. Более того, семантические алгоритмы играют ключевую роль в интерпретации значений слов в различных контекстах, помогая разрешать неоднозначности.

Тем не менее, эта область сталкивается с собственными проблемами и постоянно развивается. Динамичная природа языка, наряду с культурными нюансами и постоянно меняющимися значениями слов, представляет собой постоянные препятствия. Для их преодоления область все больше интегрирует передовые технологии, такие как машинное обучение и искусственный интеллект. Эти технологии предлагают более тонкие и контекстно-осведомленные аналитические возможности, необходимые для того, чтобы идти в ногу с изменчивостью человеческого языка.

В заключение, тема применения методов компьютерного лингвистического анализа для каталогизации полисемичных терминов ярко иллюстрирует междисциплинарное сотрудничество между лингвистикой и компьютерными науками. Она подчеркивает согласованные усилия по всесторонней каталогизации и пониманию полисемичных терминов с помощью сложных вычислительных методов, стремясь раскрыть многогранную природу человеческого языка и его сложности.

Список использованной литературы

1. Daliyeva, M., & Karimbaeva, A. (2023). Unveiling Lexical Ambiguity: Investigating Polysemy Through Corpus Analysis. ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ, 24(3), 138–139. Retrieved from <https://www.newjournal.org/index.php/01/article/view/8076>
2. Daliyeva, M., & Sa'dullaeva, G. (2023). Exploring Linguistic Patterns: Unveiling Insights from Text Corpus and Analytical Methods. ОБРАЗОВАНИЕ НАУКА И ИННОВАЦИОННЫЕ ИДЕИ В МИРЕ, 24(3), 135–137. Retrieved from <https://www.newjournal.org/index.php/01/article/view/8075>



3. Hansen, K. R. (2015). News from the future: A corpus linguistic analysis of future-oriented, unreal and counterfactual news discourse. *Discourse & Communication*, 10(2), 115–136. <https://doi.org/10.1177/1750481315611240>
4. Jurafsky, D., & Martin, J.H. (2018). *Speech and Language Processing* (3rd ed.). Pearson.
5. McEnery, T., & Hardie, A. (1996). *Corpus Linguistics: Method, theory and practice*. <http://ci.nii.ac.jp/ncid/BB0713234X>
6. Далиева, М. Х., & Сатибалдиев, Э. К. (2023). ПОЛИСЕМИЯ ТЕРМИНОЛОГИИ КАК ОБЪЕКТ ИССЛЕДОВАНИЙ В СОВРЕМЕННОМ ЯЗЫКОЗНАНИИ. In *ББК 81.2 я43 Методика преподавания иностранных языков и РКИ: традиции и инновации: сборник научных трудов VIII Международной научно-методической онлайн-конференции, посвященной Году педагога и наставника в России и Году русского языка в странах СНГ (11 апреля 2023 г.)*—Курск: Изд-во КГМУ, 2023.—521 с. (p. 47).
7. Kamilovich, S. E. (2023). EXPLORING LINGUISTIC UNIVERSALS AND TYPOLOGICAL PATTERNS: AN ANALYSIS OF THE COGNITIVE AND CULTURAL FACTORS THAT SHAPE LANGUAGE STRUCTURES ACROSS DIVERSE LANGUAGES. *American Journal of Pedagogical and Educational Research*, 10, 129-132.
8. Сатибалдиев, Э. К. (2022). ЯЗЫКОВОЕ КОНТАКТИРОВАНИЕ: БИЛИНГВИЗМ, ПОЛИЛИНГВИЗМ, ИНТЕРФЕРЕНЦИЯ. In *ИНОСТРАННЫЙ ЯЗЫК В ПРОФЕССИОНАЛЬНОЙ СФЕРЕ: ПЕДАГОГИКА, ЛИНГВИСТИКА, МЕЖКУЛЬТУРНАЯ КОММУНИКАЦИЯ* (pp. 144-149).
9. Temirova, N. A. (2023). CONSECUTIVE INTERPRETATION AS A TYPE OF ORAL TRANSLATION. *Academia Science Repository*, 4(6), 197-204.
10. Satibaldieva, N. (2024). Polysemy of Terms in Computational Linguistics. *International Journal of Scientific Trends*, 3(1), 82-84