



OPEN ACCESS

[Научный диалог = Nauchnyi dialog = Nauchnyy dialog, 14(10), 2025]

[ISSN 2225-756X, eISSN 2227-1295]

Check for updates

ИНФОРМАЦИЯ ДЛЯ ЦИТИРОВАНИЯ:

Козловский Д. В. Автоматизированный анализ эвиденциальности в медиадискурсе : опыт применения нейросетевых моделей / Д. В. Козловский // Научный диалог. — 2025. — Т. 14. — № 10. — С. 103—122. — DOI: 10.24224/2227-1295-2025-14-10-103-122.

Kozlovsky, D. V. (2025). Automated Analysis of Evidentiality in Media Discourse: Experience with Neural Network Models. *Nauchnyi dialog*, 14 (10): 103-122. DOI: 10.24224/2227-1295-2025-14-10-103-122. (In Russ.).



Web of Science™

Scopus®

Q1



DOAJ

ERIH PLUS



Перечень рецензируемых изданий ВАК при Минобрнауки РФ

Автоматизированный анализ эвиденциальности в медиадискурсе : опыт применения нейросетевых моделей

Козловский Дмитрий Валентинович
orcid.org/0000-0002-1413-8882
доктор филологических наук, доцент,
профессор кафедры языкоznания и
переводоведения
kimo3006@gmail.com

Московский городской
педагогический университет
(Москва, Россия)

Automated Analysis of Evidentiality in Russian Media Discourse: Experience with Neural Network Models

Dmitry V. Kozlovsky
orcid.org/0000-0002-1413-8882
Doctor of Philology, Associate Professor,
Professor of Department of Linguistics
and Translation Studies
kimo3006@gmail.com

Moscow City University
(Moscow, Russia)



ОРИГИНАЛЬНЫЕ СТАТЬИ

Аннотация:

Рассматривается проблема анализа функционирования категории эвиденциальности в медиадискурсе с опорой на методы машинного обучения. Актуальность исследования обусловлена необходимостью интерпретации имплицитных и контекстуально обусловленных модусных смыслов, недостаточно описываемых с помощью традиционных методов. Целью работы является разработка методики автоматизированного анализа эвиденциальности с учетом ее взаимодействия с другими модусными категориями. Исследован корпус англоязычных медиатекстов. Предлагается авторская методика, основанная на алгоритмах векторного представления слов, комбинаторного моделирования и построения рекуррентной нейронной сети. В результате выделены и классифицированы эвиденциальные операторы, установлены устойчивые модели их сцепления с маркерами категорий авторизации, аппроксимации, восприятия, модальности, отрицания, оценочности, персональности, персуазивности, экспрессивности, эмотивности и темпоральности. Представлены типовые траектории полимодального развертывания эвиденциальных высказываний. Уделяется внимание лингвосинергетическому подходу, в рамках которого модусные смыслы рассматриваются как элементы многоуровневой самоорганизующейся семантической системы. Предложенная методика демонстрирует высокую эффективность нейросетевых методов в задачах автоматизированного дискурсивного анализа и лингвистической экспертизы.

Ключевые слова:

эвиденциальность; модусные категории; нейросетевые методы; *BiLSTM*; *Word2Vec*; *Language Modeling*.

ORIGINAL ARTICLES

Abstract:

This article addresses the challenge of analyzing evidentiality within media discourse by leveraging contemporary machine learning methods. The study's relevance stems from the need to interpret implicit and context-dependent modal meanings that are inadequately captured using traditional approaches. The aim is to develop a methodology for automated analysis of evidentiality while accounting for its interaction with other modal categories. A corpus of English-language media texts serves as the empirical basis for this research. An original method based on word vector representation algorithms, combinatorial modeling, and recurrent neural networks is proposed. As a result, evidential operators have been identified and classified, along with their stable patterns of cohesion with markers representing authorization, approximation, perception, modality, negation, evaluation, personalization, persuasiveness, expressivity, emotionality, and temporality. Typical trajectories of polymodal unfolding of evidential statements are presented. Special attention is given to the linguo-synergetic approach, which considers modal meanings as elements of a multilevel self-organizing semantic system. This methodology demonstrates high efficiency of neural network techniques in tasks related to automated discourse analysis and linguistic expertise.

Key words:

evidentiality; modal categories; neural network methods; BiLSTM; Word2Vec; language modeling.



УДК 811.111'221+811.93

DOI: 10.24224/2227-1295-2025-14-10-103-122

Научная специальность ВАК

5.9.6. Языки народов зарубежных стран

5.9.8. Теоретическая, прикладная и

сравнительно-сопоставительная лингвистика

5.9.9. Медиакоммуникации и журналистика

Автоматизированный анализ эвиденциальности в медиадискурсе : опыт применения нейросетевых моделей

© Козловский Д. В., 2025

1. Введение = Introduction

Современная лингвистика характеризуется стремлением к интеграции с высокотехнологичными методами анализа, в том числе к использованию инструментов машинного обучения [Suleimanova, 2020]. В контексте лингвосинергетической парадигмы особое внимание уделяется изучению модусных категорий, формирующих субъективную составляющую дискурса. Одной из таких категорий является эвиденциальность, отражающая отношение говорящего к источнику и способу получения информации [Дзюбенко, 2024, с. 3965].

Научная проблема, лежащая в основе настоящего исследования, заключается в выявлении и описании способов актуализации категории эвиденциальности в медиадискурсе, а также в моделировании ее взаимодействия с другими модусными категориями. Немаловажно, что в рамках медиадискурса, характеризующегося активным применением стратегий дистанцирования и деавторизации, эвиденциальность выполняет важные коммуникативные и суггестивные функции, способствует снижению ответственности говорящего за достоверность информации и усилинию доверия со стороны адресата [Басовец, 2024, с. 91; Кобринова, 2021, с. 54].

Эвиденциальность как модусная категория активно изучается в лингвистике на стыке прагматики, семантики и дискурсивного анализа. Она отражает отношение субъекта речи к способу получения информации, а также к степени достоверности и проверяемости передаваемых сведений [Иванова и др., 2024, с. 815; Хазиева, 2017, с. 188]. В зависимости от характера источника и степени вовлеченности говорящего исследователи выделяют три основных типа эвиденциальности: прямую непосредственную (основанную на личном восприятии информации), косвенную непосредственную (представляющую логические выводы говорящего) и косвенную опосредованную (актуализирующую сведения, полученные из внешнего источника) [Сусов, 2006; Плунгян, 2012].



Актуализация эвиденциальности в языке осуществляется посредством лингвистических маркеров, к которым относятся вводные конструкции, глаголы, наречия, прилагательные, существительные и устойчивые словосочетания. Данные языковые средства объединяются в десять семантических групп эвиденциальных операторов: выражения мнения, косвенного указания, мыслительной деятельности, прогнозирования, сомнения, уверенности в знании, усиления значимости высказывания, умственного восприятия, чувственного восприятия и явного указания [Козловский, 2017, с. 31].

В силу того, что один и тот же эвиденциальный оператор может выражать различные модусные смыслы в зависимости от контекста, традиционные методы анализа (сравнительно-исторический, описательный, сопоставительный, структурный) часто оказываются недостаточными для комплексной интерпретации полимодальной и контекстуально обусловленной эвиденциальной семантики (под *полимодальной семантикой* в рамках данного исследования понимается совокупность модусных смыслов различных категорий (эвиденциальности, модальности, оценочности, экспрессивности и др.), актуализирующихся в пределах одного высказывания). Проблема усугубляется тем, что эвиденциальные операторы, как правило, не функционируют изолированно: они взаимодействуют с маркерами, других модусных категорий и формируют устойчивые смысловые модели, актуализируемые в дискурсивном пространстве. Обращение к лингвосинергетическому подходу позволяет обеспечить целостное представление модусной организации эвиденциального высказывания (ЭВ). Лингвосинергетика трактует эвиденциальность не как автономную единицу, а как компонент многоуровневой самоорганизующейся семантической системы, в рамках которой модусные смыслы соотносятся, взаимно усиливаются и образуют сложные синергетические модели. Применение данной методологии открывает возможность выявления скрытых закономерностей категориального взаимодействия и описания типовых моделей сцепления в медиадискурсе. Методологическая опора на базовые синергетические принципы — взаимодополнение, взаимопроникновение и самоорганизацию — позволяет раскрыть структурные и семантические механизмы функционирования эвиденциальности в медиадискурсе. Принцип взаимодополнения реализуется с помощью включения в структуру эвиденциального высказывания дополнительных модусных показателей; принцип взаимопроникновения — через дополнение эвиденциальной семантики смыслами других модусных категорий; принцип самоорганизации — посредством формирования устойчивых категориальных синергетических моделей сцепления [Черкунова и др., 2022; Kramchenko, 2023]. Синергетическое сцепление понимается как устойчивое сочетание модус-



ных смыслов, возникающее в дискурсивной структуре высказывания на основе принципов самоорганизации и категориальной коактуализации.

Особенностью эвиденциальности (*Ev*) является ее способность пересекаться с другими модусными категориями: авторизацией (*A*), аппроксимацией (*Aprx*), восприятием (*Ap*), модальностью (*M*), отрицанием (*D*), оценочностью (*Evl*), персональностью (*Ps*), персуазивностью (*P*), экспрессивностью (*Ex*), эмотивностью (*Emt*) и темпоральностью (*T*) [Козловский, 2023]. Эти пересечения особенно ярко проявляются в медиадискурсе, где говорящий стремится представить информацию как объективную, сознательно дистанцируясь от ее источника или размывая границы собственной авторской позиции [Кобриня, 2021]. В пространстве массмедиийного дискурса эвиденциальная семантика может быть представлена в двух формах: эксплицитной, выраженной языковыми средствами, и имплицитной, реализуемой через контекст, интонацию или синтаксическую структуру. Наибольшую сложность при интерпретации представляет имплицитная реализация эвиденциальной семантики, лишенная явных языковых маркеров. Однако как в эксплицитных, так и в имплицитных случаях актуализации модусных смыслов эффективным инструментом анализа выступают методы, основанные на обучении нейронных сетей, поскольку они способны учитывать не только ярко выраженные лексико-синтаксические индикаторы, но и скрытые контекстуальные зависимости, определяющие смысловую структуру высказывания [Гафаров и др., 2018; Burnashev et al., 2023].

Искусственные нейронные сети (ИНС) представляют собой вычислительные модели, вдохновленные принципами работы человеческого мозга. Они состоят из взаимосвязанных узлов (нейронов), способных воспринимать входные данные, обрабатывать их, извлекать закономерности и адаптироваться на основе обратной связи. В отличие от традиционных алгоритмов, ИНС не требуют явного программирования правил, а обучаются на примерах, что особенно важно при анализе естественного языка, обладающего высокой степенью вариативности и контекстуальной зависимости [Гафаров и др., 2018, с. 6].

В лингвистике нейросетевые методы находят применение в широком спектре задач от морфологического и синтаксического анализа до машинного перевода, генерации текста и выявления дискурсивных паттернов. Наиболее востребованными при анализе модусных категорий являются следующие архитектуры и технологии:

— *Word2Vec* — метод векторного представления слов, позволяющий кодировать лексемы в виде числовых векторов, обученных на контексте. Такая репрезентация позволяет выявлять семантическую близость между словами, что актуально для анализа модусных маркеров [Naymushin, 2021, с. 47].



— *Language Modeling* — языковое моделирование, основанное на вероятностном предсказании следующего элемента (слова, словосочетания, конструкции) на основе предыдущего контекста. Этот подход особенно важен при изучении имплицитной реализации эвиденциальной семантики, где модусные смыслы зачастую имеют размытый и контекстуально обусловленный характер [Burnashev et al., 2023].

— *BiLSTM* (*Bidirectional Long Short-Term Memory*) — тип рекуррентной нейронной сети, обрабатывающий входную информацию в двух направлениях (слева направо и справа налево), что позволяет учитывать контекст до и после слова. *BiLSTM*-модели хорошо подходят для задач категоризации модусных операторов в ЭВ, поскольку способны запоминать длинные зависимости и интерпретировать семантику целых высказываний [Инь Суузюнь и др., 2023, с. 55; Кузнецов, 2023, с. 20].

Применение ИНС в лингвистических исследованиях позволяет не только автоматизировать идентификацию языковых единиц, но и учитывать сложные сочетания модусных смыслов, а также моделировать синергетическое сцепление категорий в дискурсивном пространстве [Петрова, 2022]. Нейросети открывают принципиально новые возможности для анализа функционирования эвиденциальности: они позволяют работать с неявно выраженным (имплицитным) смыслами, распознавать контекстно обусловленные модусные сдвиги, а также группировать языковые единицы по семантической схожести, выявляя устойчивые модели сочетания модусных категорий. Под *модусным сдвигом* в рамках данного исследования понимается изменение интерпретации модусной семантики в зависимости от контекста, не сопровождающееся изменением языковой формы используемого маркера.

Научная новизна работы заключается в том, что впервые для анализа функционирования эвиденциальности в медиадискурсе применяются методы, основанные на обучении нейронных сетей. Предлагаемая методика сочетает векторное представление лексических единиц (*Word2Vec*), комбинаторное моделирование сочетаний полимодальных показателей, классификацию на основе *BiLSTM*-модели и языковое прогнозирование (*Language Modeling*), направленное на выявление устойчивых паттернов взаимодействия эвиденциальности с другими модусными категориями и прогнозирование траекторий развертывания эвиденциальных высказываний. Такой подход позволяет интерпретировать имплицитные модусные смыслы, учитывать контекстуальную полисемию и моделировать полимодальную семантику медиадискурса.

Актуальность работы обусловлена необходимостью создания инструментария, позволяющего автоматизированно обрабатывать большие объ-



емы медиатекстов с целью выявления скрытых модусных смыслов, возникающих в результате взаимодействия эвиденциальности с другими модусными категориями. Значимость работы также обусловлена обращением к лингвосинергетической парадигме, в рамках которой модусные смыслы трактуются как элементы самоорганизующейся семантической системы. Такая трактовка открывает возможность системного описания модусной организации дискурса и способствует развитию технологий автоматизированной лингвистической экспертизы и анализа речевых стратегий в медиапространстве.

Целью настоящего исследования является разработка и апробация методики автоматизированного анализа эвиденциальности в медиадискурсе с использованием нейросетевых моделей. Задачи исследования включают: (1) выделение и классификацию эвиденциальных операторов; (2) моделирование их взаимодействия с другими модусными категориями; (3) выявление устойчивых паттернов развертывания эвиденциальных высказываний (ЭВ).

2. Материал, методы, обзор = Material, Methods, Review

В последние годы активно развиваются нейросетевые методы, применяемые к изучению модусных категорий, в частности категории эвиденциальности, прежде всего в задачах моделирования и оценки степени неопределенности при формулировании выводов и принятии решений. Одним из перспективных направлений является глубокое эвиденциальное обучение (*Evidential Deep Learning*), в рамках которого нейросеть формирует предсказания на основе «доказательств» — параметров, отражающих степень уверенности модели в полученном результате [Deep evidential learning, 2023].

Наряду с этим широкое распространение получила интеграция нейросетевых архитектур с теорией Демпстера-Шафера, что позволяет аккумулировать противоречивую информацию, а также учитывать несоответствия между различными источниками сведений. Эвиденциальные модели эффективно применяются в задачах классификации, кластеризации и обнаружения объектов вне обучающего распределения (*out-of-distribution*), что особенно важно при обработке нестабильных или контекстно неоднозначных входных данных [Tong et al., 2021].

Наконец, внимание исследователей привлекают мультимодальные интерпретируемые модели, способные учитывать вклад различных каналов данных и обеспечивать прозрачность работы эвиденциальных моделей при принятии решений. Несмотря на преимущественное использование таких подходов в прикладных областях науки, например, в компьютерном зрении, медицинской диагностике и молекулярном моделировании, они



обладают значительным потенциалом для лингвистических исследований [Shao et al., 2024].

Методологическая основа настоящего исследования сочетает лингвистический анализ модусных категорий с нейросетевыми методами обработки естественного языка. Применяемые процедуры направлены на выявление, классификацию и интерпретацию эвиденциальных операторов в медиадискурсе, а также на анализ их сочетаемости с другими модусными маркерами.

Материалом исследования послужили размеченные высказывания из новостных текстов, отобранных из British National Corpus [BNC, 2025] (около 15 миллионов слов), включающих информационные, аналитические и публицистические материалы, отражающие различные типы медиадискурса. Отбор осуществлялся на основе наличия клауз с эвиденциальной семантикой — эвиденциальных высказываний (ЭВ), актуализирующихся как в эксплицитной, так и в имплицитной форме.

Для обработки материала использовались следующие нейросетевые методы:

- *Word2Vec* — для векторизации лексических единиц и определения семантической близости между модусными маркерами;
- *Language Modeling* — для предсказания модусного развертывания эвиденциальных высказываний (ЭВ) и выявления контекстуальных сдвигов;
- *BiLSTM* — для анализа двунаправленного контекста и автоматической классификации модусных показателей в синтаксических структурах.

Обработка корпуса осуществлялась с использованием библиотек Python (gensim, NLTK, PyTorch), а также с применением собственных скриптов для анализа взаимодействия модусных признаков.

Разработанная методика анализа эвиденциальности включает четыре этапа: (1) выделение эвиденциальных операторов; (2) моделирование их сочетаний с маркерами других модусных категорий; (3) контекстуальную классификацию модусных показателей; (4) выявление случаев взаимодополнения и взаимопроникновения полимодальной семантики в рамках ЭВ. Ниже последовательно представлены результаты каждого из этапов, дополненные пояснениями и примерами из корпуса [BNC, 2025].

3. Результаты и обсуждение = Results and Discussion

3.1. Автоматическое выделение эвиденциальных операторов

На первом этапе с помощью модели *Word2Vec* был сформирован семантический вектор для базовых эвиденциальных операторов. С целью повышения валидности исследования была взята максимальная выборка



по существующим классификациям эвиденциальных операторов, полученных из работ предшественников (работы А. Айхенвальд [Aiknenvald, 2004; Aiknenvald, 2021], О. А. Кобриной [Кобрина, 2005], Н. А. Козинцевой [Козинцева, 1994] и др.). На основании смысловой близости в векторном пространстве были автоматически выявлены лексемы, обладающие схожей модусной семантикой. Это позволило расширить корпус эвиденциальных операторов, включив в него как общепринятые показатели, так и менее очевидные полимодальные маркеры. Немаловажно, что при классификации операторов мы обращались к типологии эвиденциальности, предложенной А. А. Сусовым и В. А. Плунгяном [Сусов, 2006; Плунгян, 2012]. Результатом данного этапа стало расширенное множество валидных эвиденциальных операторов, пригодных для последующей автоматической классификации и анализа. Примеры эвиденциальных операторов, а также ЭВ с семантикой прямой непосредственной, косвенной непосредственной и косвенной опосредованной эвиденциальности представлены в таблице (табл. 1).

Таблица 1

Эвиденциальные операторы и примеры их употребления в корпусе

Оператор	Тип эвиденциальности	Пример
I saw (Я видел) / I heard (Я слышал)	Прямая непосредственная	I saw him enter the building through the back door (Я видел, как он вошёл в здание через черный ход) (здесь и далее перевод наш. — Д. К.).
Apparently (Очевидно)	Косвенная непосредственная	Apparently, the system failed during the test (Очевидно, система дала сбой во время тестирования).
According to (По словам)	Косвенная опосредованная	According to the minister, the reform will start in June (По словам министра, реформа начнется в июне).

Представленные в таблице данные (табл. 1) подтверждают важное для медиадискурса различие между типами эвиденциальности, проявляющееся не только в способе маркировки источника информации, но и в том, как говорящий (субъект коммуникации) обозначает свое отношение к содержанию сообщения. Эвиденциальные операторы явного указания, такие как *see* (видеть) или *hear* (слышать), выполняют функцию эксплицитного свидетельства: они апеллируют к личному восприятию событий говоря-



щим, тем самым повышают достоверность транслируемых сведений. Операторы косвенного указания, например, *apparently* (*очевидно*), *according to* (*по словам*), напротив, создают дистанцию между автором и утверждением, снижают степень ответственности говорящего за достоверность содержания сведений и актуализируют стратегию деавторизации. Такого рода механизмы активно используются в массмедиийных текстах (в частности, в аналитических заметках и редакторских колонках) и могут привлекаться как средство целенаправленного воздействия на восприятие информации аудиторией (читателем). Таким образом, тип эвиденциальности выступает не только в качестве семантического параметра, но и как элемент прагматической структуры ЭВ.

3.2. Моделирование сочетаний эвиденциальных операторов с маркерами других модусных категорий

На втором этапе с целью выявления сочетаний полимодальных показателей применялся метод комбинаторного анализа. В программу помешались индексы модусных категорий: авторизации (*A*), аппроксимации (*Aprx*), восприятия (*Ap*), модальности (*M*), отрицания (*D*), оценочности (*Evl*), персональности (*Ps*), персуазивности (*P*), эвиденциальности (*Ev*), экспрессивности (*Ex*), эмотивности (*Emt*) и темпоральности (*T*), а также корпус соответствующих маркеров.

Разработанный алгоритм позволил определить все возможные катетериальные сочетания по принципу полного перебора комбинаций, а также соотнести каждый маркер с конкретной модусной категорией. В результате анализа было выделено 128 возможных синергетических моделей сцепления, различающихся как по количеству элементов (от 1 до 8), так и по набору входящих в них категорий. Каждая модель сопоставлялась с текстовыми примерами из корпуса, в которых в пределах одного и того же ЭВ актуализировались сразу несколько модусных смыслов.

Немаловажно, что процесс моделирования осуществлялся с учетом основополагающих лингвосинергетических принципов взаимодополнения, взаимопроникновения и самоорганизации, а также двух уровней эвиденциальной семантики — базовой ($Evs = Ev + A + Ap + Ps + T$) и добавочной ($Evs + Aprx, Evs + D, Evs + Emt, Evs + Ex, Evs + Evl, Evs + M, Evs + P$) [Козловский, 2023]. Отметим, что на данном этапе нейросетевые методы не действовали, поскольку комбинаторный анализ выполнялся с использованием логических и статистических процедур. Тем не менее полученные модели межкатетериального сцепления послужили основой для последующей нейросетевой классификации и сопоставления модусных признаков. Наиболее показательные случаи взаимодействия эвиденциальности с другими модусными категориями представлены далее (табл. 2).



Таблица 2

Примеры синергетических моделей сцепления эвиденциальности и других модусных категорий

Сочетание модусных категорий	Пример из корпуса
Evs + Aprx + Evl	It seems that the estimates were roughly calculated, which is troubling (Похоже, оценки были рассчитаны приблизительно, что вызывает беспокойство).
Evs + D + P	They definitely didn't inform the audience to influence public opinion (Они определенно не проинформировали аудиторию, чтобы повлиять на общественное мнение).
Evs + M + Ex	Apparently, this must have been a horrible mistake! (Очевидно, это была ужасная ошибка!)
Evs + Emt + Evl	Reportedly, he was furious — completely outraged by the results (Как сообщается, он был в ярости — абсолютно возмущен результатами).

Анализ представленных данных (табл. 2) позволяет сделать вывод о том, что эвиденциальные операторы не функционируют изолированно, образуют устойчивые модели сцепления с маркерами других модусных категорий и формируют полимодальную семантику медиадискурса. Например, модель *Evs + D + P* демонстрирует актуализацию эвиденциальности в контексте отрицания и суггестивного воздействия на адресата (читателя), что типично для дискурса опровержений, политической полемики и разоблачительных публикаций в СМИ [Карданова-Бирюкова, 2013]. Модели *Evs + M + Ex* или *Evs + Emt + Evl* показывают, каким образом семантика модальности и экспрессивности усиливает манипулятивный потенциал ЭВ. Подобные модели позволяют говорить о существовании типовых паттернов взаимодействия модусных категорий, отражающих не только семантическую направленность, но и модусную организацию медиатекстов. С позиции лингвосинергетики модели категориального сцепления представляют собой самоорганизующиеся смысловые блоки, характерные для определенных типов и жанров медиадискурса. Таким образом, выявленные в ходе комбинаторного анализа модусные сочетания могут служить отправными точками для автоматизированной идентификации и классификации речевых стратегий автора (субъекта коммуникации) в рамках медиадискурса.

3.3. Контекстуальная классификация модусных показателей (*BiLSTM*)

Поскольку один и тот же показатель может выполнять различные модусные функции в зависимости от контекста, была обучена модель *BiLSTM*



с *attention*-механизмом, способная классифицировать добавочную модусную функцию эвиденциального оператора в том или ином ЭВ. Обучение проводилось на размеченной выборке из 1000 ЭВ с ручной аннотацией показателей. Результаты эксперимента подтвердили, что модусная функция эвиденциального оператора определяется контекстом: одни и те же маркеры могут актуализировать различные смыслы в зависимости от дискурсивной среды. В таблице 3 представлены типовые случаи интерпретации эвиденциальных операторов в зависимости от контекста, выявленные в ходе автоматической классификации с использованием *BiLSTM*-модели.

Таблица 3

Классификация маркеров по модусной функции (результаты *BiLSTM*)

Предложение	Вывод модели	Уверенность
<i>They say she opposed the decision</i> (Говорят, что она выступила против этого решения).	Эвиденциальность	0.93
<i>It might be the correct approach</i> (Возможно, это правильный подход).	Модальность	0.85
<i>Allegedly, the government ignored the warning signs</i> (Предположительно, правительство проигнорировало предупреждения).	Оценочность	0.89
<i>We almost completed the task</i> (Мы почти завершили задание).	Аппроксимация	0.87
<i>I don't think it happened that way</i> (Я не думаю, что все было именно так).	Отрицание	0.90
<i>She was absolutely furious about the result</i> (Она была абсолютно в ярости из-за результата).	Эмотивность	0.88
<i>This is clearly intended to provoke a reaction</i> (Это явно направлено на то, чтобы спровоцировать реакцию).	Экспрессивность	0.86
<i>It must be true, they always report facts accurately</i> (Должно быть, это правда, они всегда точно передают факты).	Перsuазивность	0.84

Результаты автоматической классификации маркеров с использованием *BiLSTM*-модели демонстрируют высокую чувствительность модусных показателей к контексту. Один и тот же маркер (*they say* (говорят), *allegedly* (предположительно), *might* (возможно)) может актуализировать различные модусные смыслы (от эвиденциальности до модальности, оценочности или аппроксимации) в зависимости от языкового окружения. Это подтверждает гипотезу о полифункциональности и семантической вариа-



тивности эвиденциальных операторов в медиадискурсе. Немаловажно, что нейросетевая модель способна учитывать подобные сдвиги и с высокой точностью соотносить маркеры с соответствующими модусными смыслами, что делает ее эффективным инструментом, применимым для аннотирования и семантического парсинга. Кроме того, выявленная полифункциональность маркеров обуславливает необходимость отказа от их жесткой типологизации в пользу контекстно-ориентированного подхода, при котором определяющим становится не сама форма, а ее модусная презентация в дискурсивном пространстве. Тем самым анализ эвиденциальности сближается с моделями динамической семантики, а использование обучаемых ИНС получает теоретическое и методологическое обоснование в контексте интерпретации полимодальных смыслов.

3.4. Прогнозирование модусного развития ЭВ (*Language Modeling*)

На завершающем этапе исследования применялся метод языкового моделирования (*Language Modeling*), направленный на прогнозирование и интерпретацию вероятных сценариев развертывания ЭВ в медиадискурсе. Целью этапа стало выявление устойчивых паттернов формирования полимодальной семантики на основании базовой эвиденциальной семантики (*Evs*).

Реализация данного этапа потребовала предварительного рассмотрения структурной организации ЭВ. С позиций лингвосинергетики любое ЭВ включает две функционально взаимосвязанные фазы: профазу и эпифазу [Борбелько, 2011].

Профаза представляет собой семантический каркас ЭВ, формирующий его интенциональную и тематическую направленность. Она включает указание на автора или субъекта коммуникации, а также эвиденциальный оператор. На этом этапе возможно взаимопроникновение модусной семантики — когда один и тот же показатель одновременно выражает эвиденциальность и смыслы других модусных категорий.

Эпифаза раскрывает основное содержание ЭВ и включает передаваемый материал. Она развивает и конкретизирует ранее заложенную эвиденциальную семантику, обеспечивая ее расширение через взаимодополнение — за счет включения в структуру ЭВ дополнительных модусных маркеров, репрезентирующих смыслы других модусных категорий.

Таким образом, при анализе модусной динамики категории эвиденциальности с использованием метода языкового моделирования (*Language Modeling*) учитывается двухфазная структура высказывания, в рамках которой прослеживается взаимодействие (взаимопроникновение, взаимодополнение) эвиденциальности с другими модусными категориями. На основании результатов языкового моделирования были выделены типовые



траектории развертывания полимодальной эвиденциальной семантики (табл. 4).

Таблица 4

Специфика вероятного развертывания ЭВ (*Language Modeling*)

Оператор	Пример продолжения	Тип сцепления	Добавочная семантика
<i>Maybe</i> (<i>Может быть</i>)	<i>you are right</i> (<i>ты прав</i>)	Взаимопроникновение	<i>M</i>
<i>Bill feels</i> (<i>Билл считает</i>)	<i>this is likely to continue</i> (<i>что это, скорее всего, продолжится</i>)	Взаимодополнение	<i>M</i>
<i>Evidently</i> (<i>Очевидно</i>)	<i>we need to be careful</i> (<i>нам нужно быть осторожными</i>)	Взаимопроникновение	<i>P</i>
<i>According to experts</i> , (<i>Согласно экспертам</i>)	<i>red colour urges you to stay alert</i> (<i>красный цвет заставляет вас быть начеку</i>)	Взаимодополнение	<i>P</i>
<i>It is impossible that</i> (<i>Невозможно, чтобы</i>)	<i>each person will always choose the good</i> (<i>каждый человек всегда выбирал добро</i>)	Взаимопроникновение	<i>D</i>
<i>They claim</i> (<i>Они утверждают, что</i>)	<i>it never happened</i> (<i>этого никогда не было</i>)	Взаимодополнение	<i>D</i>
<i>Reportedly</i> , (<i>Как сообщается</i>)	<i>these measures are more than effective</i> (<i>эти меры более чем эффективны</i>)	Взаимодополнение	<i>Evl</i>
<i>It's awful</i> (<i>Ужасно, что</i>)	<i>that they lost their home</i> (<i>они потеряли свой дом</i>)	Взаимопроникновение	<i>Evl</i>
<i>It is striking</i> (<i>Поразительно</i>)	<i>how scientific studies help us</i> (<i>как научные исследования помогают нам</i>)	Взаимопроникновение	<i>Ex</i>
<i>I saw</i> (<i>Я видел</i>)	<i>him walk out like a storm</i> (<i>как он вышел, будто буря</i>)	Взаимодополнение	<i>Ex</i>
<i>She is almost certain</i> (<i>Она почти наверняка</i>)	<i>to make it to the second round</i> (<i>выйдет во второй тур</i>)	Взаимопроникновение	<i>Aprx</i>
<i>I believe</i> (<i>Я полагаю</i>)	<i>the figures were roughly estimated</i> (<i>данные были рассчитаны приблизительно</i>)	Взаимодополнение	<i>Aprx</i>
<i>We understood</i> (<i>Мы поняли, что</i>)	<i>he was absolutely furious</i> (<i>он был абсолютно яростны</i>)	Взаимодополнение	<i>Emt</i>
<i>It's a shame</i> (<i>Жаль, что</i>)	<i>to let all this beauty go to waste</i> (<i>вся эта красота пропадёт впустую</i>)	Взаимопроникновение	<i>Emt</i>



Представленные результаты языкового моделирования (табл. 4) позволяют выделить устойчивые траектории развертывания ЭВ, что существенно расширяет понимание их внутренней структуры. Использование профазы и эпифазы как фазового каркаса ЭВ подтверждает, что эвиденциальность в медиадискурсе не ограничивается функцией указания на источник информации. Эвиденциальность инициирует смысловое развертывание высказывания, в котором принимают участие другие модусные категории. Такие операторы, как *maybe* (может быть), *evidently* (очевидно), *according to* (согласно) и др., не только задают информационную рамку, но и прогнозируют характер модусного развертывания ЭВ, характеризующегося добавочной семантикой — оценочности, персуазивности, экспрессивности и т. д. Данное обстоятельство представляет интерес не только для лингвистической теории, но и для практики автоматического прогнозирования смысла в системах генерации текста. Синергетические сцепления в структуре ЭВ представляют собой не случайные, а закономерные конфигурации модусных смыслов, поддающиеся формализации и моделированию. Тем самым подтверждается эффективность использования нейросетевых подходов не только для классификации, но и для предиктивной интерпретации полимодальной семантики.

4. Заключение = Conclusions

Проведенное исследование подтвердило высокую эффективность применения методов, основанных на обучении нейронных сетей, в ходе анализа модусной категории эвиденциальности в медиадискурсе. В отличие от традиционных подходов нейросетевые алгоритмы позволяют учитывать контекстуальную обусловленность значений, фиксировать имплицитные смысловые сдвиги, а также автоматически классифицировать модусные показатели с высокой степенью точности.

Разработанная комплексная методика анализа актуализации эвиденциальности в медиадискурсе включает векторизацию лексических единиц (*Word2Vec*), комбинаторное моделирование модусных сочетаний, *BiLSTM*-классификации и языковое прогнозирование (*Language Modeling*). В ходе исследования были выявлены типовые маркеры эвиденциальности, установлены их контекстуальные функции, а также описаны устойчивые синергетические модели сцепления эвиденциальности с другими модусными категориями, такими как «авторизация», «аппроксимация», «восприятие», «модальность», «отрицание», «оценочность», «персональность», «персуазивность», «экспрессивность», «эмотивность» и «темпоральность».

Особую значимость полученные результаты приобретают в свете лингвосинергетической парадигмы, согласно которой модусная структу-



ра ЭВ рассматривается как многослойная система, включающая процессы взаимодополнения, взаимопроникновения и самоорганизации модусных смыслов различных категорий. Использование нейросетевых методов исследования позволяет не только количественно описать такое взаимодействие, но и выявить скрытые закономерности межкатегориального синергетического сцепления, тем самым расширяя инструментарий лингвосинергетического подхода к дискурсивному анализу.

Перспективы дальнейших исследований эвиденциальности с применением методов, основанных на обучении нейронных сетей, заключаются в решении следующих задач:

- увеличение объема материала исследования за счет мультиязычных и мультимодальных источников;
- адаптация предложенных моделей к различным жанрово-стилевым типам дискурса;
- интеграция нейросетевых методов в практику автоматизированной лингвистической экспертизы и анализа манипулятивных стратегий в медиадискурсе.

Таким образом, применение технологий искусственного интеллекта в лингвистике позволяет по-новому взглянуть на такие сложные модусные категории, как эвиденциальность, и открывает возможности для более точного анализа полимодальной семантики в медиадискурсе.

Автор заявляет об отсутствии конфликта интересов.	The author declares no conflicts of interests.
---	--

Источники и принятые сокращения

BNC — British National Corpus [Electronic resource]. — Access mode : <http://www.natcorp.ox.ac.uk/> (accessed 23.07.2025).

Литература

1. Басовец И. М. Лучевидная субъектная модель текстовой деавторизации в медиажанрах (на материале английского и белорусского языков) / И. М. Басовец // Известия Смоленского государственного университета. — 2024. — № 1 (65). — С. 90—102. — DOI: 10.35785/2072-9464-2024-65-1-90-102.
2. Борбелько В. Г. Принципы формирования дискурса : от психолингвистики к лингвосинергетике / В. Г. Борбелько. — Москва : Книжный дом «ЛиброКом», 2011. — 288 с. — ISBN 978-5-397-01802-9.
3. Гафаров Ф. М. Искусственные нейронные сети и приложения / Ф. М. Гафаров, А. Ф. Галимьянов. — Казань : Изд-во Казан. ун-та, 2018. — 121 с.
4. Дзюбенко А. И. Прагматика художественного вымысла : дискурсивные маркеры эвиденциальности / А. И. Дзюбенко // Филологические науки. Вопросы теории и практики. — 2024. — Т. 17. — № 11. — С. 3964—3970. — DOI: 10.30853/phil20240558.



5. Иванова С. В. Эвиденциальность и эпистемическая модальность : сопряжение на уровне семантики слова / С. В. Иванова, Г. Ш. Хакимова // Вестник Санкт-Петербургского университета. Язык и литература. — 2024. — Т. 21. — № 4. — С. 802—824. — DOI: 10.21638/spbu09.2024.403.
6. Инь Суузюнь. Метод применения нейронных сетей BERT-BILSTM-Attention для определения эмоционального отношения автора к тексту / Инь Суузюнь, Г. И. Афанасьев, А. П. Калистратов // Современная наука : актуальные проблемы теории и практики. Серия : Естественные и технические науки. — 2023. — № 7—2. — С. 55—58. — DOI: 10.37882/2023-2982.2023.7-2.13.
7. Карданова-Бирюкова К. С. Реализация принципа рекурсивности речевой деятельности в политическом дискурсе (на основе анализа стенограммы телепередачи «К барьеру!») / К. С. Карданова-Бирюкова // Вестник МГПУ. Серия : Филология. Теория языка. Языковое образование. — 2013. — № 2 (12). — С. 77—86.
8. Кобриной О. А. Категория эвиденциальности : ее статус и формы выражения в разных языках / О. А. Кобриной // Вопросы когнитивной лингвистики. — 2005. — № 1 (2). — С. 86—98.
9. Кобриной О. А. Модусная составляющая лингво-коммуникативной категории / О. А. Кобриной // Актуальные проблемы филологии и лингводидактики : Сборник материалов Второй всероссийской конференции с международным участием, Нижний Новгород, 21—22 мая 2021 года. — Нижний Новгород : Нижегородский государственный лингвистический университет им. Н. А. Добролюбова, 2021. — С. 50—56.
10. Козинцева Н. А. Категория эвиденциальности (проблемы типологического анализа) / Н. А. Козинцева // Вопросы языкоznания. — 1994. — № 3. — С. 92—104.
11. Козловский Д. В. Категоричная и некатегоричная эвиденциальность в англоязычном дискурсивном пространстве / Д. В. Козловский // Филология и культура. — 2017. — № 1 (47). — С. 30—36.
12. Козловский Д. В. Реализация категории эвиденциальность в дискурсивном пространстве массмедиа : лингвосинергетический аспект / Д. В. Козловский. — Саратов : Амирит, 2023. — 184 с. — ISBN 978-5-00207-175-3.
13. Кузнецов Р. С. Прогнозирование биржевых котировок Amazon Inc. С использованием BILSTM-Attention нейронной сети / Р. С. Кузнецов // Экономика и бизнес: теория и практика. — 2023. — № 10—2 (104). — С. 19—23. — DOI: 10.24412/2411-0450-2023-10-2-19-23.
14. Петрова И. М. Современные цифровые технологии в лингвистических исследованиях / И. М. Петрова, А. М. Иванова, В. В. Никитина. — Москва : Языки народов мира, 2022. — 259 с. — ISBN 978-5-6048046-8-1.
15. Плунгян В. А. «Что такое грамматические категории?» [Электронный ресурс] / В. А. Плунгян // ПостНаукаТВ. — 2012. — Режим доступа : <http://www.youtube.com/watch?v=HGSVYt8UpYc> (дата обращения 25.12.2024).
16. Сусов И. П. История языкоznания / И. П. Сусов. — Москва : ACT : Восток-Запад, 2006. — 295 с. — ISBN 5-17-038557-9.
17. Хазиева Р. Р. Информационные технологии использования эвиденциальных маркеров в политическом дискурсе СМИ / Р. Р. Хазиева // Один пояс — один путь. Лингвистика взаимодействия : Материалы Международной научной конференции, Екатеринбург, 16—21 октября 2017 года / Ответственные редакторы А. П. Чудинов, Сунь Юйхуа. — Екатеринбург : Уральский государственный педагогический университет, 2017. — С. 188—190. — ISBN 978-5-7186-0945-5.



18. Черкунова М. В. Функциональный синергизм текста как объект функциональной лингвосинергетики / М. В. Черкунова, Е. В. Пономаренко, А. А. Харьковская // Вестник Самарского университета. История, педагогика, филология. — 2022. — Т. 28. — № 4. — С. 106—114. — DOI: 10.18287/2542-0445-2022-28-4-106-114.
19. Aikhenvald A. Y. *Evidentiality* / A. Y. Aikhenvald. — Oxford : Oxford University Press, 2004. — 481 p.
20. Aikhenvald A. Y. The grammaticalization of evidentiality / A. Y. Aikhenvald // The Oxford handbook of grammaticalization. — Oxford : Oxford University Press, 2021. — Pp. 605—606. — ISBN 978-0-19-958678-3.
21. Burnashev R. F. The role of neural networks in linguistic research / R. F. Burnashev, A. S. Alamova // Science & Education. — 2023. — Vol. 4. — № 3. — Pp. 558—569.
22. Deep evidential learning in diffusion convolutional recurrent neural network / Z. Feng, K. Qi, B. Shi, H. Mei, Q. Zheng, H. Wei // Electronic Research Archive. — 2023. — Vol. 31. — Issue 4. — Pp. 2252—2264. — DOI: 10.3934/era.2023115.
23. Khramchenko D. S. The power of synergy in discourse : exploring persuasive language in English mass media / D. S. Khramchenko // Indonesian Journal of Applied Linguistics. — 2023. — Vol. 13. — № 2. — Pp. 368—379. — DOI: 10.17509/ijal.v13i2.63068.
24. Naymushin M. Word2vec semantic model and human language processing / M. Naymushin // Речевые технологии. — 2021. — № 1—2. — Pp. 47—60. — DOI: 10.58633/23058129_2021_1-2_47.
25. Shao Z. Dual-level Deep Evidential Fusion : Integrating multimodal information for enhanced reliable decision-making in deep learning / Z. Shao, W. Dou, Y. Pan // Information Fusion. — 2024. — Volume 103. — Pp. 102—113. — DOI: 10.1016/j.inffus.2023.102113.
26. Suleimanova O. A. The Learning and Educational Potential of Digital Tools in Humanities and Social Science / O. A. Suleimanova, I. A. Guseinova, A. A. Vodyanitskaya // Society. Integration. Education : Proceedings of the International Scientific Conference. — 2020. — Vol. 4. — Pp. 657—669. — DOI: 10.17770/SIE2020VOL4.4851.
27. Tong Z. An evidential classifier based on Dempster-Shafer theory and deep learning / Z. Tong, P. Xu, T. Denoeux // Neurocomputing. — 2021. — Vol. 450. — Pp. 275—293. — DOI: 10.1016/j.neucom.2021.03.066.

Статья поступила в редакцию 30.07.2025,
одобрена после рецензирования 29.10.2025,
подготовлена к публикации 15.12.2025.

Material resources

BNC — British National Corpus. Available at: <http://www.natcorp.ox.ac.uk/> (accessed 23.07.2025).

References

- Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford: Oxford University Press. 481 p.
- Aikhenvald, A. Y. (2021). The grammaticalization of evidentiality. In: *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press. 605—606. ISBN 978-0-19-958678-3.
- Basovets, I. M. (2024). The Xiphoid subject model of textual deauthorization in media genres (based on the material of English and Belarusian languages). *Proceedings of the Smolensk State University*, 1 (65): 90—102. DOI: 10.35785/2072-9464-2024-65-1-90-102. (In Russ.).



- Borbotko, V. G. (2011). *Principles of discourse formation: from psycholinguistics to linguosynergetics*. Moscow: Librocom Book House. 288 p. ISBN 978-5-397-01802-9. (In Russ.).
- Burnashev, R. F., Alamova, A. S. (2023). The role of neural networks in linguistic research. *Science & Education*, 4 (3): 558—569.
- Cherkunova, M. V., Ponomaryovskaya, E. V. (2022). Functional synergism of the text as an object of functional linguosynergetics. *Bulletin of Samara University. History, pedagogy, philology*, 28 (4): 106—114. DOI: 10.18287/2542-0445-2022-28-4-106-114. (In Russ.).
- Dzyubenko, A. I. (2024). Pragmatics of artistic fiction: discursive markers of evidentiality. *Philological Sciences. Questions of theory and practice*, 17 (1): 3964—3970. DOI: 10.30853/phii20240558. (In Russ.).
- Feng, Z. (2023). Deep evidential learning in diffusion convolutional recurrent neural network. *Electronic Research Archive*, 31 (4): 2252—2264. DOI: 10.3934/era.2023115.
- Gafarov, F. M., Galimyanov, A. F. (2018). *Artificial neural networks and applications*. Kazan: Kazan Publishing House. University. 121 p. (In Russ.).
- Ivanova, S. V., Khakimova, G. Sh. (2024). Evidentiality and epistemic modality: conjugation at the level of word semantics. *Bulletin of St. Petersburg University. Language and literature*, 21 (4): 802—824. DOI: 10.21638/spbu09.2024.403. (In Russ.).
- Kardanova-Biryukova, K. S. (2013). Implementation of the principle of recursiveness of speech activity in political discourse (based on the analysis of the transcript of the TV show “To the barrier!”). *Bulletin of the Moscow State Pedagogical University. Series: Philology. Theory of language. Language education*, 2 (12): 77—86. (In Russ.).
- Khazieva, R. R. (2017). Information technologies for the use of evidential markers in the political discourse of the media. In: *One Belt—One Way. Linguistics of interaction: Proceedings of the International Scientific Conference, Yekaterinburg, October 16—21, 2017*. Yekaterinburg: Ural State Pedagogical University. 188—190. ISBN 978-5-7186-0945-5. (In Russ.).
- Khramchenko, D. S. (2023). The power of synergy in discourse: exploring persuasive language in English mass media. *Indonesian Journal of Applied Linguistics*, 13 (2): 368—379. DOI: 10.17509/ijal.v13i2.63068.
- Kobrina, O. A. (2005). The category of evidentiality: its status and forms of expression in different languages. *Questions of cognitive linguistics*, 1 (2): 86—98. (In Russ.).
- Kobrina, O. A. (2021). The modus component of the linguistic and communicative category. In: *Actual problems of philology and linguodidactics: Collection of materials of the Second All-Russian Conference with international participation, Nizhny Novgorod, May 21—22, 2021*. Nizhny Novgorod: Nizhny Novgorod State Linguistic University named after N. A. Dobrolyubov. 50—56. (In Russ.).
- Kozintseva, N. A. (1994). Category of evidentiality (problems of typological analysis). *Questions of linguistics*, 3: 92-104. (In Russ.).
- Kozlovsky, D. V. (2017). Categorical and uncategorical evidentiality in the English-speaking discursive space. *Philology and Culture*, 1 (47): 30—36. (In Russ.).
- Kozlovsky, D. V. (2023). *Realization of the evidentiality category in the discursive space of mass media: linguistic and synergetic aspect*. Saratov: Amrit Limited Liability Company. 184 p. ISBN 978-5-00207-175-3. (In Russ.).
- Kuznetsov, R. S. (2023). Forecasting stock quotes of Amazon Inc. Using BiLSTM-Attention neural network. *Economics and Business: theory and practice*, 10—2 (104): 19—23. DOI: 10.24412/2411-0450-2023-10-2-19-23. (In Russ.).



- Naymushin, M. (2021). Word2vec semantic model and human language processing. *Speech technology*, 1—2: 47—60. DOI: 10.58633/2305 8129_2021_1-2_47.
- Petrova, I. M., Ivanova, A. M., Nikitina, V. V. (2022). *Modern digital technologies in linguistic research*. Moscow: Limited Liability Company “Languages of the Peoples of the World”. 259 p. ISBN 978-5-6048046-8-1. (In Russ.).
- Plungyan, V. A. (2012). “What are grammatical categories?”. *Post-science*. Available at: <http://www.youtube.com/watch?v=HGSVYi8UpYc> (accessed 25.12.2024). (In Russ.).
- Shao, Z., Dou, W., Pan, Y. (2024). Dual-level Deep Evidential Fusion: Integrating multimodal information for enhanced reliable decision-making in deep learning. *Information Fusion*, 103: 102—113. DOI: 10.1016/j.inffus.2023.102113.
- Suleimanova, O. A., Guseinova, I. A., Vodyanitskaya, A. A. (2020). The Learning and Educational Potential of Digital Tools in Humanities and Social Science. *Society. Integration. Education: Proceedings of the International Scientific Conference*, 4: 657—669. DOI: 10.17770/SIE2020VOL4.4851.
- Susov, I. P. (2006). *History of linguistics*. Moscow: AST: East-West. 295 p. ISBN 5-17-038557-9. (In Russ.).
- Tong, Z., Xu, P., Denoeux, T. (2021). An evidential classifier based on Dempster-Shafer theory and deep learning. *Neurocomputing*, 450: 275—293. DOI: 10.1016/j.neucom.2021.03.066.
- Yin, S., Afanasyev, G. I., Kalistratov, A. P. (2023). Method of applying neural networks BERT-BILSTM-Attention to determine the emotional attitude of the author to the text. *Modern science: actual problems of theory and practice. Series: Natural and Technical Sciences*, 7—2: 55—58. DOI: 10.37882/2223-2982.2023.7-2.13. (In Russ.).

*The article was submitted 30.07.2025;
approved after reviewing 29.10.2025;
accepted for publication 15.12.2025.*