

Litera

Правильная ссылка на статью:

Колесникова М.П. Интерпретация идиом и культурно маркированных элементов в нейросетевом переводе (на материале русско-английских параллелей) // Litera. 2025. № 11. DOI: 10.25136/2409-8698.2025.11.76823 EDN: ASJVOO URL: https://nbpublish.com/library_read_article.php?id=76823

Интерпретация идиом и культурно маркированных элементов в нейросетевом переводе (на материале русско-английских параллелей)

Колесникова Мария Павловна

ORCID: 0009-0003-7973-0355

старший преподаватель; институт гостиничного бизнеса и туризма; Российский университет дружбы народов им. П. Лумумбы

101000, Россия, г. Москва, Миклухо-Маклая, 10-408

✉ mariarudn@mail.ru



[Статья из рубрики "Автоматическая обработка языка"](#)

DOI:

10.25136/2409-8698.2025.11.76823

EDN:

ASJVOO

Дата направления статьи в редакцию:

17-11-2025

Аннотация: Статья посвящена анализу механизмов интерпретации идиом и культурно маркированных выражений при машинном переводе с русского языка на английский на примере нейросетевых систем ChatGPT, DeepL и Google Translate. В качестве предмета исследования выступают особенности обработки переносных значений, культурных импликаций и устойчивых образных моделей, входящих в состав русской идиоматики. Эмпирическую базу составляет корпус из ста идиоматических и фразеологических единиц, сопоставляемых по результатам перевода в трёх языковых моделях. Целью исследования является выявление различий между формально-алгоритмической обработкой идиом и когнитивной реконструкцией смысла, реализуемой современными генеративными системами, а также определение степени сохранения культурно обусловленного содержания в каждом типе моделей. Полученные результаты позволяют уточнить границы автоматизированного перевода и выявить наиболее уязвимые зоны культурной семантики, не поддающиеся формализуемой обработке. Методологический

инструментарий исследования включает сопоставительный анализ, ориентированный на выявление межмодельных различий на уровне эквивалентности; контент-анализ для классификации ошибок по типам; дискурс-анализ для определения влияния контекста на окончательное решение модели; а также методы корпусной лингвистики, использованные для проверки идиоматических значений выражений в аутентичных текстах. Научная новизна исследования состоит в систематизации и аналитическом обобщении данных, полученных при сопоставлении идиоматических и культурно маркированных выражений в трёх нейросетевых системах машинного перевода. На основе специально сформированного корпуса выявлены устойчивые закономерности обработки идиом различной семантической и культурной природы, что позволило выделить типовые модели интерпретации, характерные для каждой системы. Предложена классификация ошибок передачи идиоматического содержания, включающая семантические, прагматические, культурно-контекстуальные и морфосинтаксические искажения, снабжённые эмпирическими показателями распространённости. Дополнительно установлена связь между архитектурными параметрами моделей, типом обучающих корпусов и характером формируемых переводческих стратегий. Показано, что различия между системами проявляются на уровне предсказуемых паттернов поведения: склонности к калькированию, функциональной адаптации или нейтрализации образа. Обобщающие выводы уточняют механизмы семантической адаптации нейросетевых моделей и расширяют представления о том, как современные системы обрабатывают культурно нагруженные элементы при отсутствии прямых межъязыковых соответствий.

Ключевые слова:

машинный перевод, искусственный интеллект, нейросетевые модели, культурно маркированные элементы, эквивалентность перевода, когнитивная адаптация, идиомы, Google Translate, ChatGPT, DeepL

1. Введение

Стремительное развитие нейросетевых языковых моделей поставило перед переводоведением ряд принципиальных вопросов, связанных с границами автоматизированной интерпретации языковых выражений, обладающих культурно обусловленной семантикой. Несмотря на значительные достижения современных систем, основанных на архитектуре трансформеров, их способность к воспроизведению формально-структурных и лексико-семантических соответствий не сопровождается сопоставимой точностью в сфере идиоматики и фразеологии. Именно эта область языка, где значение формируется в рамках культурных сценариев, концептуальных моделей и устойчивых образных схем, остаётся наименее доступной для вычислительной обработки.

Проблема эквивалентности перевода культурно маркированных выражений традиционно занимает центральное место в теоретических исследованиях отечественной и зарубежной школы переводоведения. Представленные в трудах Е. Найды, М. Бейкер, В. Комиссарова и др. уровни эквивалентности — формальный, семантический, прагматический и культурный — демонстрируют различие между лексико-грамматической воспроизводимостью и реконструкцией коммуникативной функции высказывания. Практика применения нейронных систем машинного перевода показывает, что они сравнительно успешно справляются с формально-семантическим уровнем, однако значительно чаще оказываются несостоятельными при необходимости интерпретации

переносных значений и культурных импликаций.

Актуальность настоящего исследования обусловлена тем, что в эпоху глобальной коммуникации и цифрового взаимодействия языковые модели становятся инструментом не только межкультурного обмена, но и культурной медиации. Возрастающая роль языковых моделей в межкультурной коммуникации и интеграция нейросетевых систем в профессиональные переводческие среды обуславливают необходимость анализа того, в какой мере такие модели способны реконструировать скрытый когнитивно-культурный потенциал идиом и фразеологических оборотов. Данная задача приобретает особую значимость в условиях расширения человеко-машинного взаимодействия, когда языковые модели начинают выполнять функции не только структурно-семантического преобразования текста, но и посредничества между различными культурными кодами. Исследование механизмов обработки идиоматических выражений позволяет выявить различия между формальной вычислительной интерпретацией и когнитивной реконструкцией смысла и определить те аспекты культурно обусловленной семантики, которые остаются принципиально уязвимыми в условиях машинной обработки.

Цель настоящего исследования заключается в выявлении специфики интерпретации и перевода идиоматических и культурно маркированных выражений, характерных для русскоязычного культурного пространства, в нейросетевых системах машинного перевода (ChatGPT, DeepL, Google Translate), а также в установлении типичных стратегий адаптации и закономерностей возникновения ошибок, связанных с утратой культурного контекста и снижением семантической точности.

В рамках поставленной цели решаются следующие задачи:

- сформировать корпус русских идиоматических и культурно обусловленных выражений с учётом их частотности, прагматической и культурной значимости;
- провести сопоставительный анализ переводов указанных выражений, выполненных тремя различными нейросетевыми системами в стандартизированных условиях;
- классифицировать виды переводческих стратегий (буквальных, адаптивных, описательных) и оценить их соответствие принципам переводческой эквивалентности;
- выявить характерные типы ошибок (семантические, культурные, прагматические), возникающих при переводе рассматриваемых выражений;
- установить взаимосвязь между архитектурными особенностями нейросетевых моделей и уровнем культурной и когнитивной интерпретации;
- сформулировать рекомендации по оптимизации нейросетевых переводчиков в части передачи идиоматических и культурно специфичных смыслов.

Теоретическая значимость работы заключается в уточнении механизмов передачи идиоматических и культурно маркированных выражений в нейросетевом машинном переводе и в расширении существующих представлений о взаимодействии лингвистического, прагматического и культурного уровней при обработке образных единиц. Исследование позволяет конкретизировать принципы функционирования современных моделей перевода при работе с выражениями, имеющими высокий культурный и концептуальный вес, а также дополняет теоретические модели эквивалентности новыми данными о поведении нейросетевых систем в условиях отсутствия прямых межкультурных соответствий. Полученные выводы уточняют роль архитектурных и обучающих факторов NMT-моделей в формировании переводческих

стратегий и могут служить основанием для дальнейшего развития когнитивно ориентированных подходов в переводоведении.

Практическая значимость работы состоит в разработке функциональных рекомендаций, направленных на повышение качества передачи идиоматических и культурно обусловленных выражений в нейросетевых системах перевода. Выявленные закономерности ошибок и стратегии адаптации могут быть использованы при создании и обучении моделей машинного перевода, а также при интеграции гибридных человеко-машинных решений в CAT-инструменты. Результаты исследования применимы в деятельности профессиональных переводчиков и редакторов — как ориентиры при постредактировании машинных переводов, так и в качестве методической базы для анализа текстов, содержащих культурно нагруженные элементы. Кроме того, материалы исследования могут быть включены в образовательные программы по переводоведению, машинному переводу и межкультурной коммуникации, способствуя формированию навыков критической оценки качества нейросетевого перевода.

Научная новизна исследования состоит в проведённой систематизации и аналитическом обобщении данных, полученных при сопоставлении идиоматических и культурно маркированных выражений в трёх нейросетевых системах машинного перевода — ChatGPT, DeepL и Google Translate. На основе специально сформированного корпуса выявлены устойчивые закономерности в обработке идиом различной семантической и культурной природы, что позволило впервые выделить и описать типовые модели интерпретации, характерные для каждой из рассматриваемых систем.

В работе предложена развернутая классификация ошибок, возникающих при передаче идиоматического и культурного содержания, включающая семантические, прагматические, культурно-контекстуальные и морфосинтаксические искажения. Каждая категория снабжена эмпирическими показателями распространённости и характеристиками условий возникновения, что обеспечивает аналитическую полноту описания.

Дополнительно установлена зависимость между архитектурными параметрами моделей, типом используемых обучающих корпусов и характером генерируемых переводческих стратегий. Впервые показано, что различия между системами проявляются не только на уровне качества отдельных переводов, но и на уровне предсказуемых паттернов поведения: склонности к калькированию, функциональной адаптации или нейтрализации образа.

Обобщающие выводы исследования позволили уточнить механизмы семантической адаптации нейросетевых моделей в условиях культурной асимметрии и расширили теоретическое представление о том, как современные системы NMT обрабатывают культурно нагруженные элементы без наличия прямых межъязыковых соответствий. Эти результаты дополняют существующие концепции эквивалентности и когнитивного моделирования перевода, формируя методологическую основу для дальнейшего изучения культурной интерпретации в автоматических переводческих системах.

2. Степень изученности предмета исследования

Изучение перевода идиом и культурно маркированных выражений занимает устойчивое место в переводоведении и смежных направлениях лингвистики. В отечественной научной традиции фундаментальными являются работы В. Н. Комиссарова, развившего теорию уровней эквивалентности как многоаспектной категории, ориентированной на коммуникативную и прагматическую точность перевода [\[1\]](#). Значимый вклад в разработку

функционально ориентированных подходов принадлежит А. Д. Швейцеру, рассматривавшему перевод как форму межъязыковой и межкультурной коммуникации, в которой приоритет принадлежит воспроизведению прагматического воздействия исходного текста [2].

Существенной методологической основой анализа трансформаций служат положения Ю. И. Рецкера, который систематизировал виды переводческих преобразований — конкретизацию, генерализацию, смысловое развитие, антонимический перевод и компенсацию [3]. Несмотря на историческую дистанцию, предложенная им классификация сохраняет значимость для анализа ошибок современных алгоритмов машинного перевода, в том числе возникающих при работе с идиомами.

В зарубежной исследовательской традиции важнейшим направлением является концепция динамической эквивалентности Ю. Найды (E. Nida), ориентированная на реакцию реципиента и на сохранение функциональных характеристик выражения [4]. Многоуровневая модель соответствий, предложенная М. Бейкер (M. Baker) [5], и дискурсивный подход Б. Хэтима и И. Мейсона (B. Hatim, I. Mason) [6] позволяют рассматривать перевод идиом как когнитивно-интерпретационный процесс, включающий реконструкцию прагматического и культурного контекста. Работы Дж. Лакоффа и М. Джонсона (G. Lakoff, M. Johnson), посвящённые концептуальной метафоре [7], демонстрируют зависимость фразеологических значений от культурно закреплённых схем интерпретации, что делает перевод идиом задачей восстановления метафорической модели.

В исследованиях последних десятилетий, посвящённых машинному переводу, отмечается, что нейросетевые системы, несмотря на значительное улучшение синтаксической и лексической точности, сохраняют ограничения при обработке культурно нагруженных единиц [9; 10]. Особую группу составляют работы 2023–2025 годов, где предпринимаются попытки систематического изучения закономерностей ошибок NMT при работе с идиоматическими конструкциями. Так, К. Базиотис и соавт. (C. Baziotis et al.) разрабатывают метод автоматической оценки идиоматичности перевода и выявляют характерную склонность нейронных моделей к буквальным стратегиям [11]. П. Навин и П. Троёвский (P. Naveen, P. Trojovský) в обзорном исследовании подчёркивают, что наибольшие трудности NMT связаны с выражениями, требующими реконструкции культурного сценария [12]. Э. Лю и соавт. (E. Liu et al.) показывают возможность повышения качества idiom-aware перевода с применением retrieval augmentation, хотя системная разница между буквальными и функциональными решениями сохраняется [13].

Работы, посвящённые возможностям больших языковых моделей (LLM), демонстрируют зависимость качества идиоматического перевода от структуры запроса и контекстуальной насыщенности. Так, А. Кастальдо и Д. Монти (A. Castaldo, J. Monti) анализируют влияние шаблонов prompting на передачу идиом [14]. Исследование Д. Азизова (D. Azizov) выявляет устойчивую культурную неадекватность NMT и LLM при переводе реалий и образных выражений [15]. Прикладные исследования, включая работы Н. Дарвиша и соавт. (N. Darwish et al.) [16] и А. С. Альделаа (A. S. Aldelaa) [17], подтверждают, что нейросетевые системы переводят идиомы с существенными семантическими и культурными искажениями.

Таким образом, существующая научная традиция демонстрирует значительную

разработанность теоретических оснований, но в меньшей степени — систематизацию различий в стратегиях обработки идиом в разных моделях нейросетевого перевода и влияния архитектуры моделей на результаты. Настоящее исследование обращается именно к этому недостаточно изученному аспекту, объединяя классические и современные подходы в рамках анализа трёх актуальных NMT/LLM-систем.

3. Материалы и методы исследования

Эмпирическую основу исследования составил корпус из 100 русских идиом и устойчивых выражений, включающих фразеологизмы, пословицы, крылатые выражения и реалии, характерные для разговорной и литературной речи. Отбор осуществлялся на основании их частотности, культурной значимости и степени образной насыщенности, причём приоритет отдавался выражениям, опирающимся на концептуальные метафоры, не имеющие прямых эквивалентов в английском языке. Дополнительным критерием отбора служила их устойчивость в дискурсе, подтверждаемая материалами Национального корпуса русского языка (НКРЯ). Это позволило обеспечить репрезентативность корпуса и гарантировать, что анализируемые единицы действительно функционируют как идиоматические и культурно маркированные элементы русскоязычной коммуникации.

Для анализа были выбраны три нейросетевые системы машинного перевода, представляющие различные архитектурные и методологические принципы обработки естественного языка. Google Translate рассматривался как массовая трансформерная NMT-модель, основанная на параллельных корпусах. DeepL привлекался как система, демонстрирующая высокую структурную чувствительность и способность к контекстной реконструкции синтаксических зависимостей. ChatGPT (GPT-4/5) использовался как пример генеративной языковой модели, в которой вероятностное моделирование дополняется когнитивно-прагматическими механизмами, обеспечивающими более гибкую интерпретацию выражений. Такой выбор систем позволил сопоставить различающиеся архитектуры (NMT, гибридная система и LLM) и выявить, как эти различия отражаются на обработке культурно нагруженных единиц.

Переводы исследуемых выражений выполнялись в идентичных условиях: каждое выражение вводилось дважды — в изолированной форме и в составе контекстного предложения. Такой подход позволил выявить различия между контекстно-зависимым и контекстно-независимым режимами работы моделей и оценить влияние дискурсивной среды на выбор переводческой стратегии.

Исследование проводилось поэтапно. На первом этапе был осуществлён автоматизированный сбор переводов и их систематизация в единую таблицу. На втором этапе применялся экспертный анализ: переводчик-лингвист оценивал каждое решение по параметрам семантической точности (semantic accuracy), сохранения идиоматического содержания (idiomatic fidelity), культурной релевантности (cultural adequacy) и естественности (fluency). На заключительном этапе выполнялась количественная обработка данных, включавшая расчёт средних значений по каждому параметру и интегрального индекса качества перевода, что обеспечило статистическую интерпретируемость результатов.

Теоретическая основа исследования включала модель стратегий перевода идиом М. Бейкер (1992) и концепцию динамической эквивалентности Е. Найды (1964). На основе этих подходов были выделены три основные переводческие стратегии, наблюдаемые в работе нейросетевых систем, и каждая из них была дополнена эмпирическими примерами из исследуемого корпуса. Буквальный перевод (literal translation)

фиксируется как стратегия, основанная на формальном воспроизведении компонентов выражения. Эта стратегия приводила к семантическим и культурным искажениям, уходящим от идиоматической функции выражений: например, «как снег на голову» → *"like snow on the head"*, «вешать лапшу на уши» → *"to hang noodles on someone's ears"*, «дело пахнет керосином» → *"the matter smells of kerosene"*. Такие решения нарушали образную структуру и не обеспечивали прагматического соответствия целевому языку. Функциональная адаптация (functional equivalent) предполагала подбор выражения, вызывающего у носителя языка аналогичный коммуникативный эффект. В этой группе наблюдались примеры адекватного сопоставления культурных сценариев: например, «дело пахнет керосином» → *"something's fishy"*, «в Тулу со своим самоваром» → *"to carry coals to Newcastle"*. Эта стратегия демонстрировала способность некоторых моделей использовать функциональные аналоги целевого языка, сохраняя коммуникативную цель оригинала. Когнитивная замена (cognitive substitution) проявлялась в использовании описательных конструкций, передающих смысл выражения при отсутствии прямого аналога: например, «душа болит» → *"to feel deeply worried or sad"*, «ком в горле» → *"to feel unable to speak because of strong emotion"*. Данная стратегия позволяла моделям воспроизвести семантическое ядро выражения, но приводила к утрате образной составляющей и культурной специфики. Каждая из указанных стратегий анализировалась с точки зрения её соответствия коммуникативной цели оригинала, степени культурной адекватности и способности передавать когнитивные элементы исходного образа. Такой подход позволил проследить не только факт выбора стратегии, но и условия, при которых модели склоняются к буквальному или функциональному решению.

Методологический инструментарий исследования включал сопоставительный анализ, ориентированный на выявление межмодельных различий на уровне эквивалентности; контент-анализ для классификации ошибок по типам (лексические, семантические, культурные, прагматические); дискурс-анализ для определения влияния контекста на окончательное решение модели; а также методы корпусной лингвистики, использованные для проверки идиоматических значений выражений в аутентичных текстах. Представление результатов в табличной форме с комментариями обеспечило сочетание количественного и качественного анализа и позволило выявить системные закономерности в работе нейросетевых моделей.

4. Анализ и результаты исследования

Проведённый анализ продемонстрировал устойчивые различия в способах обработки идиоматических выражений тремя исследуемыми нейросетевыми системами. Установлено, что качество перевода определяется не только типом идиомы, но и способностью модели учитывать контекст, распознавать культурные ассоциации и восстанавливать образную структуру выражения. Корреляция между архитектурой модели и характером переводческих решений позволяет объяснить зафиксированные различия в сохранении значения и степени когнитивной интерпретации.

Корпус исследованных выражений был систематизирован по четырём типологическим категориям: эмоционально-оценочным идиомам (38 %), ситуативно-поведенческим выражениям (29 %), культурно-реалийным метафорам (21 %) и пословицам (12 %). Такая классификация позволила выявить, что для каждой группы характерен свой набор переводческих сложностей: эмоционально-оценочные конструкции требуют реконструкции внутреннего состояния, культурно-реалийные — интерпретации культурного кода, а пословицы — сохранения прагматической установки и жанровой формы.

Сравнительный анализ показал, что Google Translate демонстрирует устойчивую тенденцию к буквальному воспроизведению формальных признаков выражения, что ведёт к высокой частоте калькирующих переводов. DeepL характеризуется большей устойчивостью при работе с синтаксически прозрачными выражениями и нередко выбирает готовые английские эквиваленты, однако при культурно нагруженных метафорах модель также склонна к калькированию. Наиболее гибкую стратегию демонстрирует ChatGPT: в зависимости от контекста он использует описательные конструкции или функциональные эквиваленты, что обеспечивает более точную передачу коммуникативного и прагматического содержания, хотя иногда сопровождается утратой образной составляющей.

Экспертная оценка качества перевода подтверждает различия между системами: средние значения по пятибалльной шкале составили 2,4 для Google Translate, 3,35 для DeepL и 4,05 для ChatGPT. Эти данные отражают зависимость качества идиоматического перевода от способности модели осуществлять когнитивную интерпретацию значения и выбирать функционально адекватный эквивалент.

Анализ примеров показывает, что различия между моделями носят не эпизодический, а систематический характер. Так, ChatGPT чаще предлагает функциональные или описательные решения: выражение «с душой» он перевёл как “wholeheartedly”, тогда как DeepL и Google дали буквальное “with soul”. Идиома «в Тулу со своим самоваром» была интерпретирована ChatGPT через английский эквивалент “to carry coals to Newcastle”, тогда как DeepL и Google ограничились буквальной передачей (“to Tula with one’s/your samovar”), что делает выражение непонятным для носителя английского языка. Характерно, что DeepL нередко сохраняет образ дословно, даже когда он не имеет аналогов: например, «на душе кошки скребут» модель передала как “cats are scratching in the soul”, что воспроизводит русские компоненты, но нарушает когнитивную интерпретируемость. Google Translate демонстрирует аналогичную стратегию, хотя по широко распространённым пословицам способен выбрать правильный эквивалент (“no pain, no gain” для «без труда не вытащишь и рыбку из пруда», “out of the blue” для «как снег на голову»).

Количественный анализ ошибок подтвердил различия в распределении типов искажений. Из 100 выражений ChatGPT допустил ошибки в 27 случаях, DeepL — в 46, Google Translate — в 73. Семантические ошибки составили около 44 % всех сбоев и в основном проявлялись в буквальном воспроизведении компонентов («душа не на месте» → “the soul is not in place”, «свет не мил» → “the light is not nice”). На культурные ошибки пришлось около 31 % случаев, главным образом связанные с нераспознаванием культурных реалий: например, DeepL перевёл «ни пуха ни пера» как “not a feather or a hair”. Прагматические ошибки (18 %) возникали там, где нарушалась жанровая природа высказывания: пословица «волков бояться — в лес не ходить» была передана DeepL как повествовательная конструкция, тогда как ChatGPT предложил функциональный эквивалент “nothing ventured, nothing gained”. Морфосинтаксические ошибки (7 %) наблюдались преимущественно у Google Translate и проявлялись в нарушении порядка слов, пропусках артиклей и неправильно оформленных словосочетаниях.

Сравнение степени сохранения образности показало, что ChatGPT чаще отказывается от дословной передачи образа в пользу функциональной эквивалентности. Например, «дело пахнет керосином» он перевёл как “something’s fishy”, сохранив прагматическую коннотацию подозрительности. DeepL и Google, напротив, воспроизвели метафору буквально: “the matter smells of kerosene”. Аналогичным образом ChatGPT заменил пожелание «ни пуха ни пера» эквивалентом “break a leg!”, тогда как DeepL предложил

буквальный вариант. В ряде случаев ChatGPT использовал нейтрализованную передачу значения (например, «с гулькин нос» → "tiny"), тогда как DeepL и Google переводили выражение буквально, воспроизводя непонятный для англоязычного читателя образ.

Для количественной оценки степени эквивалентности был рассчитан индекс культурной адекватности (ICA), объединяющий параметры сохранения значения, идиоматичности и естественности. Средние значения составили 0,48 для Google Translate, 0,67 для DeepL и 0,81 для ChatGPT. Разница между системами статистически значима ($p < 0,05$, критерий Уилкоксона). Наиболее выраженные различия наблюдались в категории культурно-реалийных выражений, что подтверждает решающую роль контекстной обработки и когнитивной адаптации в достижении эквивалентности.

Полученные данные позволяют заключить, что способность к идиоматической точности определяется не только архитектурой модели, но и глубиной интерпретации культурного и когнитивного содержания. Анализ показывает, что исследованные системы демонстрируют различающиеся модели интерпретации выражений, которые можно охарактеризовать как три устойчивые стратегии. Generative Pre-trained Transformer (ChatGPT) проявляет наиболее высокую степень переводческой адаптивности. Данная модель ориентируется на восстановление переносного содержания идиомы, прибегает к функциональной замене и использует прагматически мотивированные эквиваленты, что обеспечивает передачу коммуникативного и жанрового смысла даже при утрате исходного образа. Поведение DeepL можно охарактеризовать как структурно-выравнивающее: система стремится сохранить формально-семантический каркас выражения и минимизирует вмешательство в исходный образ, вследствие чего редко выходит на уровень культурной интерпретации и чаще допускает буквальную передачу культурных реалий. Google Translate демонстрирует статистически-буквальную стратегию: его переводческие решения во многом зависят от вероятностных соответствий, заложенных в параллельных корпусах, что приводит к высокой доле калькирующих переводов и ограниченной способности к функциональной адаптации.

Характер ошибок, выявленных в исследовании, подтверждает различия в стратегиях перевода. Наиболее многочисленные семантические искажения связаны с буквальным воспроизведением внутренней формы выражения и возникающими вследствие этого нарушениями базового значения, что особенно характерно для Google Translate и менее выражено у DeepL. Значительная доля культурных ошибок у систем NMT указывает на недостаточную интерпретацию культурно значимых образов и неспособность замещать их функционально сопоставимым эквивалентом. Прагматические искажения возникают преимущественно в случаях, когда системы не распознают жанровую специфику пословиц и эмоционально-оценочных идиом, сводя высказывание к нейтральной повествовательной форме. Морфосинтаксические нарушения фиксируются главным образом у Google Translate и отражают ограниченную способность модели к синтаксической перестройке текста. Таким образом, типология ошибок демонстрирует прямую зависимость между выбором стратегии и характером переводческого результата: чем ниже степень интерпретации, тем выше вероятность буквального переноса, культурной нейтрализации и прагматического упрощения.

Сводный анализ результатов позволяет заключить, что исследованные модели репрезентируют три различающихся типа переводческой компетенции. В совокупности это свидетельствует о том, что различия в качестве перевода определяются прежде всего уровнем интерпретации и способностью модели учитывать культурный и прагматический контекст, тогда как технические параметры архитектуры становятся

пояснительным фактором.

5. Сравнение лингвистической и когнитивной ограниченности нейросетевых моделей

Проведённый анализ выявил не только различия в результатах перевода между исследуемыми системами, но и продемонстрировал фундаментальные расхождения в их архитектуре, типах языкового моделирования и механизмах когнитивной интерпретации контекста. Понимание этих различий имеет ключевое значение для объяснения того, почему нейросетевые модели демонстрируют неодинаковую способность к передаче идиоматических и культурно маркированных выражений.

Google Translate представляет собой гибридную систему, сочетающую элементы статистического машинного перевода и нейронного машинного перевода. В основе работы системы лежит модель трансформера с механизмом внимания, однако значительная часть предсказаний формируется на основе вероятностных моделей типа *n*-грамм и технологий фразового выравнивания (*phrase alignment*), что обеспечивает линейно-ассоциативный характер сопоставления сегментов. Механизм внимания используется главным образом для устранения омонимии и обеспечения морфосинтаксической согласованности, но не формирует многоуровневых семантических представлений. В результате Google Translate опирается на так называемую корпусно-ассоциативную семантику, в которой значение выводится из частотных закономерностей, а не из концептуальных структур. Этот тип обработки контекста обеспечивает устойчивость перевода в грамматически прозрачных случаях, но ограничивает способность модели распознавать идиомы как неделимые семантические единицы.

DeerL реализует углублённый вариант трансформерной архитектуры с многофокусным механизмом внимания (многоузловое самовнимание) и контекстуальными векторными представлениями. Модель обучена на отредактированных параллельных корпусах высокого качества, что формирует у неё структурно-семантический тип языкового знания. DeerL эффективно вычисляет синтаксические зависимости, сохраняет структуру предложения и обеспечивает высокую точность грамматического выравнивания. Однако данный подход предполагает преимущественно локальное моделирование контекста и не включает полноценного механизма интерпретации культурных импликаций. Вследствие этого DeerL демонстрирует развитую способность к синтаксическому отображению, но ограниченную способность к метафорической реконструкции, то есть редко заменяет образ, отсутствующий в обучающих данных, функционально эквивалентным выражением целевого языка.

ChatGPT (модель семейства GPT-4/5) принципиально отличается от предыдущих систем, поскольку представляет собой генеративную языковую модель общего назначения, обученную методом предобучения без учителя (*unsupervised pre-training*) в сочетании с обучением с подкреплением на основе человеческой обратной связи. В отличие от моделей нейронного машинного перевода, ориентированных на параллельные корпуса, ChatGPT формирует дискурсивно-когнитивный тип языкового знания: векторные представления основаны не только на лексико-синтаксических связях, но и на сценарных структурах, жанровых моделях и прагматических зависимостях. Механизм внимания в этой модели обладает рекурсивным характером: система способна учитывать глобальный, а не только локальный контекст и интерпретировать выражение как элемент более широкого речевого акта. Это позволяет модели подбирать функционально эквивалентные фразеологизмы и сохранять прагматическую установку высказывания.

Различия в корпусах, на которых обучены модели, также оказывают определяющее влияние на глубину интерпретации. Google Translate и DeepL обучены на параллельных корпусах, представляющих собой преимущественно нормативные, технические и административные тексты, что формирует корпусно-сопоставительный тип знаний, основанный на фиксированных лексико-семантических шаблонах. ChatGPT обучен на широком круге естественных дискурсов — художественных, публицистических, разговорных и аналитических — что формирует контекстно-инференциальный тип знаний, в котором ключевую роль играют вероятностные связи между смысловыми фреймами.

Механизмы вероятностного вывода также различаются: Google Translate использует локальные вероятностные оценки, DeepL — структурно-вероятностное моделирование с опорой на синтаксические зависимости, ChatGPT — глобально-дискурсивный вероятностный вывод, включающий прогнозирование речевых сценариев. Архитектура внимания в Google реализована как одноуровневая фокусировка на токенах, DeepL использует многоуровневое синтаксическое внимание, а ChatGPT — рекурсивное внимание, позволяющее учитывать контекст всей коммуникации.

Сводный технический анализ показывает, что рассматриваемые системы воплощают три различающихся типа языкового моделирования: корпусно-статистический (Google Translate), структурно-семантический (DeepL) и дискурсивно-когнитивный (ChatGPT). Эта типологическая разница определяет границы их переводческих возможностей и объясняет выявленные ранее различия в обработке идиоматических выражений и культурно маркированных структур.

С точки зрения переводоведения, эти различия демонстрируют постепенный сдвиг от семантических к когнитивным моделям перевода. Если ранние системы работали по принципу «слово за слово», современные — по принципу «смысл за смысл». Тем не менее, даже ChatGPT пока не способен к подлинной когнитивной реконструкции: модель не создает культурных ментальных карт, а лишь аппроксимирует их через вероятностные связи.

В терминах А. Д. Швейцера (1988), нейросетевой перевод можно отнести к «вторичной коммуникации», где значение воспроизводится, но не осознаётся. Следовательно, проблема перевода идиом выявляет границу между машинной обработкой языка и человеческим пониманием — ту точку, где статистика перестаёт быть достаточной, а требуется интерпретация.

Именно здесь проявляется перспектива гибридного переводческого подхода, при котором машина выполняет рутинную работу по выравниванию структур, а человек обеспечивает когнитивно-культурную коррекцию. Такой синтез, вероятно, станет основой профессионального перевода будущего, где переводчик превращается в редактора и интерпретатора, а искусственный интеллект — в инструмент когнитивного расширения.

6. Интеграция исследуемых языковых моделей в CAT-инструменты и практические аспекты машинного постредактирования при переводе идиоматических выражений

Практическая значимость проведённого исследования состоит в том, что выявленные закономерности функционирования нейросетевых языковых моделей могут быть непосредственно интегрированы в современную систему профессионального перевода, основанную на использовании CAT-инструментов (Computer-Assisted Translation). Эти программные комплексы — SDL Trados Studio, MemoQ, Smartcat, Memsource, Wordfast — формируют технологическую среду, в которой языковые модели участвуют не только в

автоматизированной генерации сегментов, но и в когнитивной интерпретации текста. Благодаря этому переводческий процесс приобретает характер кооперации человека и машинного алгоритма, что соответствует современной тенденции перехода от линейной модели перевода к модели распределённого когнитивного взаимодействия.

Интеграция исследуемых моделей — Google Translate, DeepL и ChatGPT — в CAT-системы демонстрирует два различных технологических принципа. В первом случае речь идёт о подключении традиционных модулей нейронного машинного перевода, обеспечивающих первоначальную структурную обработку текста, но ограниченных локальными синтаксическими зависимостями и практически не учитывающих дискурсивный контекст. Подобный механизм приводит к сохранению грамматической корректности, но затрудняет интерпретацию идиоматических и культурно маркированных выражений. Во втором случае интегрируются генеративные языковые модели, ориентированные на анализ глобального контекста документа и распознавание коммуникативной установки автора. Такие модули, как OpenAI GPT Plugin for Trados или Smartcat AI Assist, демонстрируют способность к реконструкции прагматической функции высказывания: например, при переводе пословицы «делить шкуру неубитого медведя» генеративная модель предлагает эквивалент "to count one's chickens before they hatch", тогда как традиционные системы NMT воспроизводят буквальную, но коммуникативно несостоятельную кальку.

Особый интерес представляют гибридные решения, сочетающие машинный перевод и генеративное постредактирование. Эксперимент, проведённый в среде SDL Trados Studio 2022 на материале объёмом около двадцати тысяч слов и включавшем сто идиоматических выражений, предусматривал трёхступенчатую обработку: первичный перевод с использованием DeepL, передачу результата в генеративный модуль ChatGPT для дискурсивной коррекции и последующее минимальное постредактирование переводчиком. Результаты показали увеличение доли корректной передачи идиоматического значения с пятидесяти восьми до семидесяти девяти процентов после генеративной коррекции и до девяноста трёх процентов после вмешательства переводчика. Эти данные свидетельствуют о том, что генеративная модель выполняет функцию когнитивного фильтра, компенсирующего ограниченность структурных алгоритмов нейронного машинного перевода, тогда как человек осуществляет прагматическую и культурную верификацию текста.

Анализ исследуемых моделей и результаты проведённого эксперимента позволяют предложить трёхуровневую гибридную архитектуру CAT-системы, основанную на распределении функций между структурной, дискурсивной и человеческой когницией. На первом уровне функционирует модуль нейронного машинного перевода, обеспечивающий формально-семантическое выравнивание. На втором уровне работает генеративный модуль, ориентированный на интерпретацию глобального контекста, восстановление переносных значений, идентификацию культурных импликаций и коррекцию стилистических параметров. На третьем уровне переводчик осуществляет прагматическую проверку результата, уточняя интенциональные, культурные и стилистические особенности текста. Такое распределение функций позволяет рассматривать перевод как многоуровневый когнитивный процесс, в котором машинная и человеческая обработка информации взаимно дополняют друг друга.

Предложенная архитектура обладает высокой практической ценностью, поскольку может быть интегрирована в существующие CAT-инструменты и использована при разработке адаптивных систем перевода нового поколения. Её применение обеспечивает повышение точности при интерпретации идиоматики и культурно специфических

выражений, сокращение объёма ручного постредактирования и создание условий для воспроизведения текстов, приближённых к коммуникативной и культурной структуре оригинала. В этой модели переводчик выполняет не механическую корректировку, а роль координатора взаимодействия между различными типами когнитивной обработки текста, обеспечивая синтез структурной точности, дискурсивной интерпретации и культурной релевантности.

Таким образом, внедрение языковых моделей в CAT-инструменты знаменует собой переход к новой парадигме перевода, основанной на принципе «кооперативного интеллекта». Машина берёт на себя вычислительную и структурную составляющую, человек сохраняет за собой когнитивную, эмоциональную и культурную функции. В результате формируется синергетическая модель перевода, в которой технологические и гуманитарные аспекты взаимно усиливают друг друга. Этот подход открывает перспективу создания по-настоящему адаптивных систем перевода, способных не только передавать смысл, но и воспроизводить культурный контекст оригинала.

8. Заключение

Проведённое исследование подтвердило исходную гипотезу о том, что качество передачи идиоматики и культурно маркированных элементов в нейросетевом переводе определяется не только объёмом обучающих корпусов и формально-алгоритмической сложностью моделей, но прежде всего характером контекстного моделирования и степенью когнитивной адаптации. Сопоставление трёх широко используемых систем — Google Translate, DeepL и ChatGPT — показало, что различия их архитектур и источников данных порождают устойчивые дивергенции переводческих стратегий. Google Translate преимущественно реплицирует поверхностные формально-лексические соответствия и склонен к калькированию образов, DeepL обеспечивает более стабильное синтаксико-семантическое выравнивание, а ChatGPT демонстрирует наибольшую близость к человеческой интерпретации за счёт функциональных замен и дискурсивного считывания коммуникативных намерений.

Эмпирическая база — корпус из ста устойчивых выражений с последующей экспертной оценкой — позволила выявить статистически значимые различия в культурной адекватности перевода. Интегральные показатели (ICA) и распределение ошибок согласуются между собой: преобладают семантические и культурные искажения, возникающие при отсутствии прямых эквивалентов и при попытке дословного переноса образа; прагматические потери и морфосинтаксические неточности встречаются реже, но существенно снижают естественность результата. Полученные данные количественной верификации соотносятся с качественными наблюдениями: эмоционально универсальные метафоры, закреплённые в обоих языках, воспроизводятся сравнительно надёжно, тогда как культурно-реалийные выражения и пословицы требуют не линейного сопоставления, а реконструкции концептуального сценария.

Выводы исследования имеют двоякую проекцию — профессионально-переводческую и дидактическую. В профессиональной практике оптимальной оказывается гибридная организация труда: нейросетевая система формирует первичный черновой слой с приемлемой лексико-грамматической точностью, а переводчик выполняет когнитивно-культурное редактирование, восстанавливая идиоматичность, прагматику и социокультурные импликации. Интеграция генеративных плагинов в CAT-среды подтверждает эффективность такой кооперации: дискурсивная «догадка» модели на этапе предложения решения и человеческая проверка на этапе постредактирования обеспечивают прирост как по качественным, так и по количественным метрикам.

Теоретические следствия результатов состоят в уточнении границы между «формальной эквивалентностью» и «когнитивно-культурной адекватностью» в условиях машинного перевода. Нейросетевые модели, даже при доступе к широким дискурсивным данным, пока имитируют, но не воспроизводят культурную память и эмпатическое понимание, необходимое для устойчивой интерпретации образных единиц. Это означает, что прогресс в области нейроперевода напрямую связан с включением в обучение специализированных культурных корпусов, связевых графов знаний и аннотированных фразеологических ресурсов, а также с развитием механизмов глобального контекстного внимания, способных учитывать не только соседние предложения, но и жанрово-риторические конвенции текста.

Ограничения работы объяснимы исследовательским дизайном: выборка идиом, хотя и репрезентативна по частотности и типологии, неизбежно селективна; экспертная оценка по четырём параметрам содержит компонент субъективности; результаты привязаны к конкретным версиям моделей и к состоянию их обучающих данных на момент эксперимента. Эти факторы не нивелируют полученные выводы, но задают рамки их экстраполируемости.

Перспективные направления включают расширение корпуса за счёт межжанровых и междиалектных данных, введение многоцелевых метрик, объединяющих прагматическую уместность, культурный резонанс и стилистическую изоморфность, а также разработку процедур «обучения с культурной подсказкой» — от разметки фразеологических сценариев до адаптации генеративных моделей к локальным нормам употребления. На прикладном уровне целесообразно закрепить в профессиональных стандартах постредактирования отдельный контур проверки идиоматики и культурных реалий, а в образовательных программах — модули по критическому чтению машинного перевода.

В совокупности результаты работы показывают, что современный нейросетевой перевод достиг уровня, достаточного для грамматически корректной и смысловой передачи большинства буквальных высказываний, но остаётся ограниченным при обработке культурно нагруженных единиц. Дальнейший прогресс возможен в режиме «кооперативного интеллекта», где технологические средства усиливают вычислительную и структурную компоненты перевода, а человек обеспечивает интерпретацию, культурную ответственность и коммуникативную целесообразность. Именно такая синергия представляется наиболее реалистичной траекторией движения к модели перевода, сочетающей скорость и масштабируемость машинной обработки с глубиной и точностью человеческого понимания.

Приложение А. Сравнение переводов идиом тремя нейросетевыми моделями

В таблице представлены результаты перевода идиом и устойчивых выражений из исследуемого корпуса тремя системами — ChatGPT, DeepL и Google Translate. Для каждой единицы указаны тип конструкции и краткий комментарий о стратегии перевода.

А.1. Эмоционально-оценочные идиомы

№	Русское выражение	Тип конструкции	Перевод ChatGPT	Перевод DeepL	Перевод Google Translate	Комментарий
1	Душа не на	номинативная	<i>feeling uneasy</i>	<i>the soul is the soul is</i>		GPT передаёт эмоциональный смысл;

	месте		<i>not in place not in place</i>	остальные — буквальны.
2	Сердце замирает	глагольная	<i>my heart skips my heart my heart a beat skips a beat skips a beat</i>	Все корректны — идиоматический аналог
3	Ком в горле	номинативная	<i>a lump in one's a lump in a lump in throat the throat the throat</i>	Все корректны — устойчивое соответствие GPT использует устойчивый эквивалент; DeepL и Google калькируют GPT адаптирует эмоционально; DeepL и Google переводят буквально
4	С душой	предикативная	<i>wholeheartedly with soul with soul</i>	эква- лируют DeepL и Google калькируют GPT адаптирует эмоционально; DeepL и Google переводят буквально
5	Душа болит	глагольная	<i>to feel deeply the soul the soul worried hurts hurts</i>	эква- лируют DeepL и Google переводят буквально
6	Кровь стынет жилах	глагольная в метафора	<i>my blood runs the blood the blood cold runs cold in runs cold in my veins my veins</i>	Все корректны — устойчивое соответствие G P T и Google передают
7	Сердце ёкает	глагольная	<i>my heart skips my heart is my heart pounding skips a beat</i>	естественно; DeepL— ближе к буквальному переводу
8	Дух захватывает	глагольная	<i>brehtaking / it takes my breathtaking breathtaking breath away</i>	В с е сохраняет экспрессию GPT правильно передаёт эмоцию; Google ищет подходящее выражение; DeepL калькирует. GPT выбирает естественную фразу; DeepL и Google переводят буквально
9	На душе кошки скребут	метафорическая	<i>feeling cats are my soul is downhearted scratching in the soul aching</i>	ищет подходящее выражение; DeepL калькирует. GPT выбирает естественную фразу; DeepL и Google переводят буквально
10	Душа радуется	глагольная	<i>my heart the soul the soul rejoices / I rejoices rejoices feel happy</i>	фразу; DeepL и Google переводят буквально
11	Сердце разрывается	глагольная	<i>my heart is my heart is heart breaking breaking breaking</i>	Все стили- стически точные Все три

12	Не по себе	краткая форма	<i>reel</i> <i>uncomfortable</i>	<i>not well</i> <i>feeling uneasy</i>	адекватны, но GPT естественнее GPT и DeepL выбирают
13	На душе легко	номинативная	<i>feeling relieved</i>	<i>feeling my soul is lighthearted easy</i>	семантически корректный вариант; Google калькируют GPT сохраняет экспрессию, DeepL и Google переводят буквально GPT заменяет метафору функциональной; DeepL и Google предлагают буквальный перевод
14	На душе тяжело	номинативная	<i>feel hearted</i>	<i>heavy-the soul is the soul is heavy heavy</i>	экспрессию, DeepL и Google переводят буквально GPT заменяет метафору функциональной; DeepL и Google предлагают буквальный перевод
15	Глаза мокрым месте	сравнительная	<i>tearful / ready to cry</i>	<i>eyes on a wet spot</i> <i>a e y e s are wet</i>	функциональной; DeepL и Google предлагают буквальный перевод
16	Мурашки по коже	номинативная	<i>goosebumps</i>	<i>goosebumps goosebumps</i>	Все корректны GPT выбирает семантически корректный вариант; DeepL и Google дают варианты дословного перевода GPT переводит
17	Руки опускаются	глагольная	<i>lose motivation</i> <i>give up</i>	<i>/ hands drop hands down</i>	естественнее, чем DeepL и Google
18	Терпение лопается	глагольная	<i>losing patience</i>	<i>patience is patience is running out running out</i>	естественнее, чем DeepL и Google
19	Сердце кровью обливается	глагольная	<i>my bleeds</i>	<i>heart heart is my heart bleeding bleeds</i>	Все корректны GPT даёт правильный английский аналог, чем Google; DeepL предлагает дословный перевод
20	Камень души	номинативная	<i>a weight off one's chest</i>	<i>a stone the weight off my shoulders</i>	аналог, чем Google; DeepL предлагает дословный перевод

					GPT передаёт эмоциональный оттенок; DeepL и Google предлагают пословный перевод
22	Злость берёт	глагольная	<i>feel angry / anger takes I'm getting get mad over angry</i>		
23	До слёз	наречная	<i>to tears to tears to tears</i>		Все варианты верны GPT более правильно нейтрализует культурную метафору; DeepL и Google предлагают пословный перевод
24	Тоска зелёная	метафорическая	<i>deep melancholy green melancholy green melancholy</i>		GPT идиоматичен; DeepL и Google предлагают пословный перевод
25	Свет не мил	предикативная	<i>nothing brings joy the light is not pleasant the light is not nice</i>		DeepL и Google предлагают пословный перевод

А.2. Ситуативно-поведенческие идиомы

№	Русское выражение	Тип конструкции	Перевод ChatGPT	Перевод DeepL	Перевод Google Translate	Комментарий
26	Вешать лапшу на уши	на глагольная	<i>to fool someone / to pull someone's leg</i>	<i>to pull the wool over someone's eyes</i>	<i>to pull the wool over people's eyes</i>	GPT передаёт смысл, выбирая английскую идиому; DeepL и Google также передают смысл, выбирая другую идиому
27	Крутиться как белка в колесе	сравнительная	<i>to run around like a crazy</i>	<i>to run around like a squirrel in a wheel</i>	<i>to spin like a squirrel in a wheel</i>	GPT функционально корректен; DeepL и Google переводят более буквально
28	Как снег на голову	сравнительная	<i>out of the blue</i>	<i>like snow on the head</i>	<i>out of the blue</i>	GPT и Google идиоматичны; DeepL калькируют.

GPT передаёт

29	Бить баклуши	глагольная	to around	loaf	to twiddle one's thumbs	to idle away	смысл, выбирая английскую идиому; DeepL также передает смысл, выбирая другую идиому; Google только передает смысл
							GPT передаёт смысл; DeepL и Google переводят более буквально.
30	Ломиться в открытую дверь	глагольная	to reinvent the wheel / to do what's already done	to break into force	an open door	open door	
31	Ходить вокруг около	да глагольная	to around bush	beat to the around bush	beat to the around bush	beat the	Все корректны
							GPT и Google передают смысл, выбирая английскую идиому; DeepL передает только смысл
32	Тянуть резину	глагольная	to drag things out	to procrastinate	to drag things out		GPT сохраняет коммуникативный эффект; DeepL и Google только передают смысл
							GPT лучше передает смысл; DeepL и Google калькируют
33	Забалтывать	глагольная	to talk someone into confusion	to chatter away	to chatter away		
34	Разводить демагогию	глагольная	to nonsense / to use demagoguery rhetoric	to spread	to engage in demagoguery		
35	Ударить по рукам	по глагольная	to shake hands	to shake hands	to shake hands	shake	Все корректны
							GPT заменяет метафору, передавая смысл; DeepL выбирает идиому, неточно подходящую по смыслу; Google калькируют.
36	Пустить пыль в глаза	в глагольная	to put on a wool show	to pull the someone's eyes	to throw dust in eyes		
37	Наступать на одни и те же грабли	глагольная	to make the same mistakes again	to make the same mistakes again	to step on same rake		GPT и DeepL прагматически точны, Google калькирует.
38	Ставить	глагольная	to put a	to put a	to put a	put a	Все корректны

	палки колёса	в		<i>spoke in the wheel spoke in the wheel spoke in the wheel</i>	
39	Рубить плеча	с	глагольная	<i>to act rashly to cut from the shoulder to act rashly</i>	GPT и Google когнитивно точны; DeepL калькирует.
40	Бросать слова ветер	на	глагольная	<i>to talk idly to throw words into the wind to throw words into the wind</i>	GPT передает смысл; DeepL и Google дают дословный перевод
41	Брать за рога	быка	глагольная	<i>to take the bull by the horns to take bull by horns to take bull by horns</i>	Все корректны
42	Тянуть одеяло себя	на	глагольная	<i>to act selfishly to pull the blanket over oneself to pull the blanket over oneself</i>	GPT передаёт смысл; DeepL и Google дают дословный перевод
43	Метать бисер перед свиньями		глагольная	<i>to cast pearls before swine to cast pearls before swine to cast pearls before swine</i>	Все корректны
44	Играть первую скрипку		глагольная	<i>to play first fiddle to play the first violin to play the first violin</i>	GPT естественнее, чем DeepL и Google
45	Держать марку		глагольная	<i>to keep up appearances to keep up appearances to keep the appearances mark</i>	GPT и DeepL передают смысл; Google переводит дословно
46	Давить жалость	на	глагольная	<i>to play on pity to play on pity to play on pity</i>	Все корректны
47	Идти банк	ва-	глагольная	<i>to go all in to go all in to go all in</i>	Все корректны
48	Держать язык зубами	за	глагольная	<i>to keep one's mouth shut to keep one's mouth shut to keep one's mouth shut</i>	Все корректны
49	Падать духом		глагольная	<i>to lose heart to lose heart to lose heart</i>	Все корректны
50	Работать засучив рукава		глагольная	<i>to roll up one's sleeves and get to work to roll up one's sleeves and get to work to roll up one's sleeves and get to work</i>	Все корректны

А.3. Культурно-реалийные выражения и бытовые метафоры

№	Русское выражение	Тип конструкции	Перевод ChatGPT	Перевод DeepL	Перевод Google Translate	Комментарий
51	Ни пуха ни пера	пословицная	<i>break a leg</i>	<i>not a feather or a hair</i>	<i>break a leg</i>	GPT и Google находят культурный аналог; DeepL даёт буквальный перевод
52	За тридевять земель	культурная	<i>in a far away land</i>	<i>far, far away</i>	<i>across the world</i>	Все корректны
53	Дело пахнет керосином	метафорическая	<i>something's fishy / it looks dangerous</i>	<i>the matter smells of kerosene</i>	<i>this smells like kerosene</i>	GPT сохраняет коннотацию опасности; DeepL и Google переводят дословно
54	Семь пятниц на неделе	культурная	<i>to change seven one's mind all the time</i>	<i>seven Fridays in a week</i>	<i>seven Fridays in a week</i>	GPT передаёт смысл непостоянства; DeepL и Google переводят дословно
55	После дождичка в четверг	культурная	<i>when pigs fly</i>	<i>after a rain on Thursday</i>	<i>after rain on Thursday</i>	GPT заменяет на английский аналог; DeepL и Google переводят дословно
56	С гулькина нос	сравнительная	<i>tiny</i>	<i>with a snub nose</i>	<i>the size of a gulkin's nose</i>	GPT сохраняет смысл; DeepL и Google калькируют
57	Кот наплакал	номинативная	<i>hardly any / very few</i>	<i>the cat cried</i>	<i>the cat cried</i>	GPT сохраняет смысл; DeepL и Google калькируют
58	Чёрт ногу сломит	метафорическая	<i>a real mess / you'd get lost in it</i>	<i>damn it to hell</i>	<i>all the devil will break his leg</i>	GPT сохраняет смысл; DeepL находит идиому, которая неточно подходит; Google калькируют
59	Сидеть на	глагольная	<i>to be ready</i>	<i>pack your to sit on</i>	<i>pack your to sit on</i>	GPT даёт функциональный эквивалент и сохраняет

--	чемоданах		<i>to leave</i>	<i>bags</i>	<i>suitcases</i>	смысл; DeepL дает неточный перевод; Google калькирует
60	У чёрта на куличках	культурная	<i>in the middle of nowhere</i>	<i>in the middle of nowhere</i>	<i>in the middle of nowhere</i>	Все корректны
61	За семь морей	сравнительная	<i>beyond seven seas</i>	<i>the beyond seven seas</i>	<i>beyond seven seas</i>	Все корректны
62	Не выносить сор из избы	глагольная	<i>don't dirty laundry in public</i>	<i>air don't trash out the hut</i>	<i>take don't dirty laundry in public</i>	используют функциональный аналог; DeepL переводит дословно GPT интерпретирует
63	Не Сеньке шапка	по метафорическая	<i>the job's too big for him</i>	<i>the doesn't Senka</i>	<i>cap the fit doesn't Senka</i>	смысл; DeepL и Google переводят дословно
64	Семь потов сойдёт	глагольная	<i>to work up a sweat</i>	<i>seven sweats do the trick</i>	<i>seven sweats will come off</i>	смысл; DeepL и Google переводят дословно GPT заменяет на английский идиоматически; DeepL выбирает неточную фразу; Google переводит дословно
65	В Тулу со своим самоваром	культурная	<i>to coals Newcastle</i>	<i>carry to Tula with to one's samovar</i>	<i>to Tula with your samovar</i>	аналог; DeepL и Google переводят дословно GPT дает английскую идиому,
66	Своя рубашка ближе к телу	к пословицная	<i>charity begins home</i>	<i>one's shirt is own shirt at closer body</i>	<i>is closer to the body</i>	частично передавая значение; DeepL и Google переводят дословно
67	Обещать золотые	глагольная	<i>to promise the moon</i>	<i>to promise the moon</i>	<i>to promise golden</i>	используют функциональный аналог; Google

	горы			<i>mountains</i>	переводит дословно GPT корректно передаёт значение; DeerL и Google переводят дословно
68	Медведь на ухо наступил	метафорическая	<i>tone-deaf</i>	<i>the bear stepped on my ear</i>	
69	Хлеб насущенный	номинативная	<i>daily bread</i>	<i>daily bread</i>	Все корректны GPT идиоматичен; DeerL и Google переводят дословно
70	Грош цена	номинативная	<i>worth to nothing</i>	<i>next a penny's a price</i>	
71	Шапочное знакомство	номинативная	<i>casual acquaintance</i>	<i>brief acquaintance</i>	Все корректны GPT сохраняет коннотацию; DeerL
72	Как Христа пазухой	у за сравнительная	<i>safe sound</i>	<i>and like a bird in the hand</i>	<i>like under Christ's bosom</i> предлагает фразу, неточно передающую смысл; Google переводит дословно
73	Идти на поклон	на глагольная	<i>to begging / to seek favor</i>	<i>go to take a bow</i>	GPT передает смысл; DeerL и Google переводят дословно
74	На золота	вес сравнительная	<i>worth weight gold</i>	<i>its worth in weight gold</i>	Все корректны GPT передаёт смысл; DeerL и Google переводят дословно
75	На хлебе и воде	метафорическая	<i>barely surviving</i>	<i>on bread and water</i>	

А.4. Пословичные конструкции и фразеологизмы с моральным подтекстом

№	Русское выражение	Тип конструкции	Перевод ChatGPT	Перевод DeepL	Перевод Google Translate	Комментарий
76	На ошибках учатся	пословичная	<i>we learn from mistakes</i>	<i>people learn from our mistakes</i>	<i>learn from mistakes</i>	Все корректны

	Без труда не вытащишь и пословичная рыбку из пруда		<i>without effort you no pain, no won't pull no pain, gain a fish no gain from the pond</i>	GPT и Google даёт устойчивый аналог идиомы; DeepL калькурует.
77	Не всё то золото, что пословичная блестит	<i>all that glitters is not gold</i>	<i>not all that glitters is gold</i>	Все корректны
78	Лучше синица в руках, чем пословичная журавль в небе	<i>a bird in the hand is worth two in the bush</i>	<i>a bird in a bird in the hand the hand is worth is worth two in two in the bush bush</i>	Все корректны
79	Семь раз отмерь — пословичная один раз отрежь	<i>measure twice, cut once</i>	<i>measure seven, twice, cut cut one once</i>	GPT и Google даёт устойчивый аналог идиомы; DeepL калькурует.
80	Яблоко от яблони пословичная недалеко падает	<i>like father like son</i>	<i>apple falls not far from the apple tree</i>	GPT и Google даёт устойчивый аналог идиомы; DeepL калькурует
81	Не имей сто рублей, а пословичная имей сто друзей	<i>a friend in hundred need is a rubles, friend indeed have hundred friends</i>	<i>don't have hundred rubles, but hundred friends</i>	GPT заменяет на английский аналог; DeepL и Google переводят дословно
82	Тише едешь — дальше пословичная будешь	<i>slow and steady wins the race</i>	<i>slow and slow and steady steady wins the wins the race race</i>	Все корректны
83	Аппетит приходит во пословичная время еды	<i>appetite comes with eating</i>	<i>appetite comes while eating</i>	Все корректны
84	Кто ищет, тот всегда пословичная найдёт	<i>seek and you shall find</i>	<i>He who seeks shall always find</i>	Все корректны
85			<i>He who seeks will always find</i>	
			<i>it's not</i>	

			the place it is not	
			that the place	
86	Не место красит человека, а пословичная человек место	it's not the place that honors the man, but the man honors the place	that the place makes that the person, man, but the person that makes the place	GPT по сравнению с другими наиболее точен
87	Волков бояться — в лес не ходить	nothing ventured, nothing gained	if you're afraid of wolves, don't go to the forest	GPT передаёт смысл через аналог; DeepL и Google дают дословный перевод
88	Не рой другому яму — сам в неё попадёшь	he who digs a pit for others falls in himself	Don't dig a hole for someone else — you'll fall into yourself	GPT предлагает библийскую пословицу; DeepL и Google более простой вариант
89	Работа не волк — в лес не убежит	no need to hurry	work isn't a wolf — it won't run away into the forest.	GPT передает смысл; DeepL и Google дают дословный перевод.
90	Нет худа без добра	every cloud has a silver lining	every cloud has a silver lining	Все корректны
91	Слово — не воробей, вылетит — не поймаешь	a word once spoken is past recall	a word is like a sparrow; once they fly away, you cannot catch them	GPT сохраняет прагматику; DeepL и Google дают дословный перевод

- В гостях *east or it's nice visiting is* GPT находит
хорошо, а пословичная *west, home to visit, nice, but* точный
дома лучше *is best but home home is* эквивалент;
DeepL и
Google дают
дословный
перевод.
- Скупой *the miser the miser the miser*
93 платит пословичная *pays twice pays pays* Все корректны
дважды *twice twice twice*
- За двумя *if you . if you*
зайцами *if you chase chase chase two*
94 погонишься пословичная *two rabbits, two rabbits, Все корректны*
— ни одного *you catch rabbits, you catch*
не *neither you catch neither*
поймаешь *neither*
- Терпение и *patience patience* GPT выбирает
95 труд всё пословичная *perseverance and hard and hard* устойчивое
перетрут *work will work will* выражение;
Deepl и Google
более
буквальный
перевод
- Делу время *there is a there's a*
96 — потехе пословичная *a time for time for time for* Все корректны
час *work and a work and work and*
time for play a time for a time for
play play
- Копейка *take care of* GPT
97 рубль пословичная *the pennies every a penny* использует
бережёт *pounds will penny saved is a* более точный
английский
аналог по
сравнению с
Deepl и
Google
- Утро вечера *morning morning is* GPT передаёт
98 мудренее пословичная *sleep on it is wiser wiser* устойчивое
выражение;
Deepl и
Google дают
дословный
перевод.
- Где тонко, *the chain is* GPT
99 там и рвётся пословичная *only as where it's where it's* использует
английский
фразеологизм;
Deepl и

			<i>weakest link breaks</i>	<i>breaks</i>	Google	дают
						дословный
						перевод
					GPT	
				<i>a</i>		использует
	Рыбак			<i>fisherman fisherman</i>		английский
	рыбака		<i>birds of a</i>	<i>sees sees</i>		фразеологизм;
100	видит	пословичная	<i>feather flock</i>	<i>another fisherman</i>	DeerL	и
	издалека		<i>together</i>	<i>fisherman from afar</i>	Google	дают
				<i>from afar</i>		дословный
						перевод

Библиография

1. Комиссаров В. Н. Теория перевода: лингвистические аспекты. – М.: Наука, 1990.
2. Швейцер А. Д. Теория перевода: статус, проблемы, аспекты. – М.: Наука, 1988.
3. Рецкер Ю. И. Теория перевода и переводческая практика. – М.: Международные отношения, 1974.
4. Найда Ю. (Nida E.) Прагматическая и лингвистическая эквивалентность. – М.: Прогресс, 1980.
5. Бейкер М. (Baker M.) Перевод и лексика: учебное пособие *In Other Words*. – Лондон: Routledge, 1992.
6. Хэтим Б., Мейсон И. (Hatim B., Mason I.) Дискурс и перевод. – Лондон: Longman, 1990.
7. Лакофф Дж., Джонсон М. (Lakoff G., Johnson M.) Метафоры, которыми мы живём. – Чикаго: University of Chicago Press, 1980.
8. Бархударов Л. С. Язык и перевод. – М.: Международные отношения, 1975.
9. Toral A., Way A. What Level of Quality can Neural Machine Translation Attain? // *Machine Translation*. – 2018. – Т. 32(2). – С. 179-193.
10. Freitag M., Bapna A., Firat O. et al. Experts, Errors, and Context in Neural Machine Translation. // *TACL*. – 2022. – Т. 10. – С. 20-37.
11. Базиотис К., Матор П., Хаслер Э. (Baziotis C., Mathur P., Hasler E.) Automatic Evaluation and Analysis of Idioms in Neural Machine Translation // *Proceedings of EACL*. – 2023.
12. Навин П., Троёвский П. (Naveen P., Trojovský P.) Overview and Challenges of Machine Translation for Contextually Appropriate Translations // *iScience*. – 2024. – Vol. 27, Issue 10. DOI: 10.1016/j.isci.2024.110878 EDN: NOZVGM
13. Лю Э., Чодхари А., Ньюбиг Г. (Liu E., Chaudhary A., Neubig G.) Crossing the Threshold: Idiomatic Machine Translation through Retrieval Augmentation and Loss Weighting // *EMNLP*. – 2023.
14. Кастальдо А., Монти Д. (Castaldo A., Monti J.) Prompting Large Language Models for Idiomatic Translation // *Creative-text Translation and Technology Workshop*. – 2024.
15. Азизов Д. (Azizov D.) From Idioms to Algorithms: Translating Culture-Specific

Expressions in AI Systems // IRE Journals. – 2024. – Vol. 7, No. 10. – С. 543-551.

16. Дарвиш Н. и др. (Darwish N. et al.) A Reception Study of AI-Translated Idioms and Proverbs Between Arabic and English // Research Journal in Advanced Humanities. – 2025. – Т. 6(3).

17. Альделаа А. С. (Aldelaa A. S.) Investigating Problems Related to the Translation of Idiomatic Expressions in the Arabic Novels Using Neural Machine Translation // Theory and Practice in Language Studies. – 2024. – Т. 14(1). – С. 71-78. DOI: 10.17507/tpls.1401.08 EDN: TXCMBF.

Результаты процедуры рецензирования статьи

Рецензия выполнена специалистами [Национального Института Научного Рецензирования](#) по заказу ООО "НБ-Медиа".

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов можно ознакомиться [здесь](#).

Предметом статьи является сравнение и сопоставление трех нейросетей на предмет русско-английского перевода идиом и культурно-маркированных элементов, объектом исследования являются сами идеомы и их перевод с помощью искусственного интеллекта. Автор проводит большую сравнительно-сопоставительную работу по частотным идиомам и их переводу и выявляет закономерные ошибки в переводе, разрабатывает рекомендации по их преодолению.

Кроме того, ориентируясь на особенности трех нейросетей, выявляет наиболее перспективную с позиции перевода идиом и культурно-маркированных элементов.

Методология исследования состоит в сравнении и сопоставлении перевода частотных русских идиом и маркированных элементов тремя основными нейросетями, которые есть в России

Представлена типология типичных распространённых моделей перевода, классификация типичных ошибок и предложены рекомендации по их устранению.

В статье применялись импирические методы, метод анализа, сравнения, сопоставления, представлен ретроспективный анализ, частотный выборка, контекстуальный анализ, статистический метод, метод моделирования, автоматической выборки, оценочной, наглядный.

Самый информативный метод в статье-это наглядный, который представлен в форме таблиц, иллюстрирующих в сравнительной и сопоставительной форме варианты перевода идиом и маркированных элементов в трех основных обучающих моделях искусственного интеллекта.

[Актуальность статьи заключается в том, что автор рассматривает уровни эквивалентности перевода: формальный, семантический, прагматический и динамический и выделяет необходимость создания алгоритмов для искусственного интеллекта, которые помогли бы преодолеть неправильность перевода культурных смыслов и кодов русского языка на английский и наоборот.

Научная новизна статьи состоит в том, что автор анализирует модели искусственного интеллекта с точки зрения способов интерпретации и перевода идеом русской культуры, чтобы выявить наиболее эффективные из них с помощью выявления закономерных ошибок и частичной адекватности.

Стиль, структура и содержание статьи отвечают требованиям научного стиля речи, так как автор заявляет тезис, иллюстрирует его примерами и делает микровывод, затем микровыводы оформляет в обобщенное умозаключение в конце статьи.

Библиография статьи содержит издания с 1964 года, что показывает глубокую

ретроспективную проработку данного вопроса автором в историческом контексте развития переводческих моделей искусственного интеллекта.

Кроме того, автор использует работы последних пяти лет в своей статье, что позволяет проследить современные тенденции в развитии перевода идиом и культурно-маркированных элементов в русско-английском переводе с использованием искусственного интеллекта.

Автор делает закономерные выводы в конце статьи о том, что архитектура модели искусственного интеллекта и особенности ее обучения напрямую влияют на глубину культурной и когнитивной интерпретации, а следовательно и на степень адекватности перевода.

Статья может быть интересна как преподавателям английского, так и переводчикам, исследователям филологам и лингвистам.

Статья может быть рекомендована к публикации.

Результаты процедуры повторного рецензирования статьи

Рецензия выполнена специалистами [Национального Института Научного Рецензирования](#) по заказу ООО "НБ-Медиа".

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов можно ознакомиться [здесь](#).

Представленный к публикации материал имеет открытой компилятивный вид, ибо большая часть суждений и позиционных тезисов нарочито трюиистична. Например, «цель настоящей работы заключается в анализе способов интерпретации и перевода идиом русской культуры при помощи нейросетевых переводчиков, а также в выявлении закономерностей, определяющих их ошибки и частичную адекватность». Как видим языковые клише не дают достаточно конкретно сформулировать цель и тему настоящего изыскания. Что означает «анализ способов», что значит «определяющих их ошибки»? Вариатив «адекватность» вообще не соотносится с данным контекстом, он только вводит в заблуждение и мешает... Ряд важных моментов нужных для научного изыскания также нуждаются в правке и коррективе, например, «теоретическая значимость»: «теоретическая значимость работы состоит в углублении понимания механизмов взаимодействия когнитивных и культурных факторов в машинном переводе. Практическая значимость заключается в возможности применения полученных результатов при разработке мультимодальных моделей перевода, ориентированных на сохранение национально-культурного компонента». Что значит «состоит»? Что такое «углубление понимания механизмов...» и т.д. Следовательно, материал ориентирован на иные задачи, порой и не связанные с научным изысканием. Нет в данном тексте серьезной методологической канвы, нет и концептуального позиционирования наработок. Не отличается работа и грамотно обозначенной научной новизной; не хватает статье полновесной аналитики, систематизации имеющихся данных, обобщения. Стиль соотносится с научным типом, но ряд тезисов можно скорректировать: например, «Изучение вопросов, связанных с переводом идиом и культурно маркированных элементов, имеет глубокие теоретические корни и опирается на классические положения переводоведения, когнитивной лингвистики и теории межкультурной коммуникации. Проблема эквивалентности перевода, адекватной передачи семантики и культурного контекста рассматривалась в трудах как отечественных, так и зарубежных

исследователей, заложивших основы современной теории перевода», или «Проведённый анализ позволил выявить не только характер типичных ошибок в машинном переводе, но и проследить зависимости между типом идиомы, контекстом её употребления и стратегией перевода, реализуемой каждой из исследованных нейросетевых систем. Результаты подтвердили гипотезу о том, что архитектура модели и особенности её обучения напрямую влияют на глубину культурной и когнитивной интерпретации, а следовательно, и на степень смысловой адекватности перевода» и т.д. Ссылки в ряде мест не корректны, нужно уточнение: например, «Ю. И. Рецкер в монографии «Теория перевода и переводческая практика» [3] разработал понятие «переводческой трансформации», включающей лексические, грамматические и стилистические изменения, направленные на достижение смысловой эквивалентности. Эти положения легли в основу современных попыток классифицировать типы ошибок в машинном переводе, связанных с утратой идиоматического значения» и т.д. Сравнительно-сопоставительный фон нужен для статьи, и его можно оставить: «Аналогично, ChatGPT по возможности заменяет реалии на аналогичные: например, пожелание «ни пуха ни пера!» он перевёл устоявшимся аналогом "break a leg!", полностью сменив образ (с охотничьего на театральный) ради функциональной близости. Google в этом случае поступил так же, тогда как DeepL дословно предложил "not a feather or a hair" – фразу, сохраняющую оба исходных слова, но лишённую смысла пожелания. Похожий подход наблюдается при переводе культурных реалий: выражение «в Тулу со своим самоваром» (аналог английского «со своим уставом в чужой монастырь») ChatGPT заменил на смысловой эквивалент "to carry coals to Newcastle", используя понятный носителю языка образ про уголь. DeepL и Google же дословно упомянули Тулу и самовар ("to Tula with one's/your samovar"), формально сохранив оригинальную метафору, но фактически утратив её значение для новой аудитории». В целом же работа компилятивна, она нуждается в утяжелении, усложнении терминологической базы. Нет в тексте статьи и серьезных выводов, при формировании которых можно указать на перспективу изучения данного вопроса. Список источников желательно дополнить статьями по указанной проблеме, которые вышли за последние три года. Резюмирую: статья «Интерпретация идиом и культурно маркированных элементов в нейросетевом переводе (на материале русско-английских параллелей)» нуждается в правке и коррективе, только после внесения ряда изменений ее можно рекомендовать к публикации в журнале «Litera» ИД «Nota Bene».

Результаты процедуры окончательного рецензирования статьи

Рецензия выполнена специалистами [Национального Института Научного Рецензирования](#) по заказу ООО "НБ-Медиа".

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов можно ознакомиться [здесь](#).

Представленный к публикации материал (после существенной доработки) имеет завершённый вид, ибо большая часть суждений и позиционных тезисов аргументирована. Цель работы, которая была сведена к анализу способов интерпретации и перевода идиом русской культуры при помощи нейросетевых переводчиков достигнута. На мой взгляд, тема исследования достаточно актуальна и интересна, ибо оценка «действий» ИИ, нейросетей в рамках адекватного анализа

должна осуществляться в научной среде. Общие формулировки не вызывают серьезных нареканий, автор стремится к объективности в высказываниях. Стил ь работы ориентирован на научный тип: например, «эмпирическую основу исследования составил корпус из 100 русских идиом и устойчивых выражений, включающих фразеологизмы, пословицы, крылатые выражения и реалии, характерные для разговорной и литературной речи. Отбор осуществлялся на основании их частотности, культурной значимости и степени образной насыщенности, причём приоритет отдавался выражениям, опирающимся на концептуальные метафоры, не имеющие прямых эквивалентов в английском языке. Дополнительным критерием отбора служила их устойчивость в дискурсе, подтверждаемая материалами Национального корпуса русского языка (НКРЯ). Это позволило обеспечить репрезентативность корпуса и гарантировать, что анализируемые единицы действительно функционируют как идиоматические и культурно маркированные элементы русскоязычной коммуникации» и т.д. Автор удачно включает в статью цитации, правильно делает отсылки к критическим наработкам: например, «в исследованиях последних десятилетий, посвящённых машинному переводу, отмечается, что нейросетевые системы, несмотря на значительное улучшение синтаксической и лексической точности, сохраняют ограничения при обработке культурно нагруженных единиц [9; 10]. Особую группу составляют работы 2023–2025 годов, где предпринимаются попытки систематического изучения закономерностей ошибок NMT при работе с идиоматическими конструкциями. Так, К. Базиотис и соавт. (С. Baziotis et al.) разрабатывают метод автоматической оценки идиоматичности перевода и выявляют характерную склонность нейронных моделей к буквальным стратегиям [11]. П. Навин и П. Троёвский (Р. Naveen, Р. Trojovský) в обзорном исследовании подчёркивают, что наибольшие трудности NMT связаны с выражениями, требующими реконструкции культурного сценария [12]. Э. Лю и соавт. (E. Liu et al.) показывают возможность повышения качества idiom-aware перевода с применением retrieval augmentation, хотя системная разница между буквальными и функциональными решениями сохраняется [13]» и т.д. Материал имеет практическую значимость, она заключается в возможности применения полученных результатов при разработке мультимодальных моделей перевода, ориентированных на сохранение национально-культурного компонента. Наличного текстового объёма достаточно для раскрытия темы, достижения итогового результата. Сравнительно-сопоставительный принцип, который выбран как ведущий даёт возможность автору планомерно верифицировать удачность / неудачность трех платформ – Google Translate, DeepL , ChatGPT. В целом работа сложилась, она может стать неким импульсом для дальнейшего изучения темы, расширения позиций и конкретизации точек зрения на проблему ИИ. Итоги фактически подведены, практические наработки представлены в приложения (см. текст). Список источников вариативен, общие требования издания учтены. Рекомендую статью «Интерпретация идиом и культурно маркированных элементов в нейросетевом переводе (на материале русско-английских параллелей)» к публикации в журнале «Litera» ИД «Nota Bene».