

Litera

Правильная ссылка на статью:

Лю М. Стандарты оценки качества перевода политического дискурса с применением больших языковых моделей // Litera. 2025. № 4. DOI: 10.25136/2409-8698.2025.4.73946 EDN: VVAHQM URL: https://nbpublish.com/library_read_article.php?id=73946

Стандарты оценки качества перевода политического дискурса с применением больших языковых моделей

Лю Мяо

ORCID: 0000-0003-2346-0600

кандидат филологических наук

доцент; факультет русского языка и литературы; Пекинский университет

100871, Китай, г. Пекин, ул. Ихюань, 5

✉ liumiaolm@pku.edu.cn



[Статья из рубрики "Перевод"](#)

DOI:

10.25136/2409-8698.2025.4.73946

EDN:

VVAHQM

Дата направления статьи в редакцию:

29-03-2025

Дата публикации:

05-04-2025

Аннотация: Предметом настоящего исследования является разработка комплексной системы стандартов оценки качества перевода политического дискурса с применением больших языковых моделей. Исследование сфокусировано на создании интегративного подхода к оценке перевода китайского политического дискурса на русский язык, что особенно актуально в контексте укрепления российско-китайского стратегического партнерства. Разработанная трехуровневая модель оценки учитывает специфику политического дискурса, характеризующегося высокой степенью терминологической насыщенности, концептуальной специфичности и идеологической маркированности. Модель базируется на параметрах информационной целостности, языковой точности и межкультурной удобочитаемости, обеспечивая комплексный анализ перевода на всех

уровнях: от лексического и синтаксического соответствия до передачи семантических отношений и культурно-прагматических аспектов. Особое внимание в рамках исследования уделяется применению технологических возможностей больших языковых моделей для автоматизированной диагностики переводческих проблем и повышения объективности оценки перевода политических текстов. Предлагаемая система учитывает требования «Базовых норм перевода китайского политического дискурса», адаптируя их к контексту российско-китайской межкультурной коммуникации. Методология исследования основана на системном анализе классических и современных моделей оценки перевода, с последующим функциональным моделированием интегративного подхода, объединяющего лингвистические теории, компьютерные методы и возможности больших языковых моделей для оценки переводов. Научная новизна исследования заключается в преодолении разрыва между теоретическими моделями оценки перевода и их практической реализацией за счет интеграции традиционных лингвистических подходов, компьютерных методов и требований отраслевых стандартов с технологическими возможностями больших языковых моделей. Впервые предложена детализированная структура оценки качества перевода политического дискурса с 10 параметрами второго уровня и 36 параметрами третьего уровня, что обеспечивает всесторонний анализ переводческой деятельности. Доказано, что применение больших языковых моделей значительно повышает эффективность и объективность оценки благодаря глубокому семантическому анализу и автоматизированной диагностике переводческих проблем. Разработанная система имеет высокую практическую ценность, будучи применимой как для анализа готовых переводов, так и для прогнозирования потенциальных трудностей при переводе политических текстов в контексте российско-китайской межкультурной коммуникации.

Ключевые слова:

стандарты оценки, политический дискурс, большие языковые модели, интегративный подход, информационная целостность, языковая точность, межкультурная удобочитаемость, автоматизированная диагностика перевода, российско-китайская коммуникация, переводческие параметры

Введение

В условиях глобальной трансформации мирового порядка и формирования полицентрической системы международных отношений эффективная межкультурная коммуникация становится фундаментальным фактором геополитического взаимодействия. Российско-китайское стратегическое партнерство, достигшее беспрецедентно высокого уровня в XXI веке, выступает ярким примером конструктивного диалога цивилизаций, основанного на взаимном уважении и понимании национальных интересов. Важнейшую роль в обеспечении такого диалога играет адекватный перевод политических текстов, которые, в силу высокой степени терминологической насыщенности, концептуальной специфики и идеологической маркированности, требуют особых подходов к оценке качества переводческой деятельности. Именно соответствие перевода политического дискурса его прагматическим функциям определяет эффективность международной коммуникации на высшем уровне. Однако обеспечение этого соответствия становится всё более сложной задачей, поскольку в условиях информационного общества и цифровой дипломатии традиционные методологические парадигмы оценки качества перевода демонстрируют определенную ограниченность. Эта ограниченность

обусловлена как возросшими объемами межъязыковой коммуникации, так и прогрессирующим усложнением самого политического дискурса, что актуализирует необходимость разработки специализированных критериев оценки.

Проблематика оценки качества перевода исторически развивалась в рамках двух направлений: лингвистического и вычислительного. В области разработки стандартов оценки качества перевода российская школа переводоведения, заложенная А. В. Фёдоровым [11], сформировала фундаментальные критерии эквивалентности, которые были развиты в работах Л. С. Бархударова [2] через призму многоаспектности переводческой деятельности. Значительный вклад в создание стандартов оценки внес В. Н. Комиссаров [3], предложивший многоуровневую систему критериев, включающую функциональную адекватность и социокультурную конвенциональность как ключевые параметры оценки политических текстов. Западная переводоведческая традиция внесла существенный вклад в стандартизацию оценки через работы К. Reiss [4], связавшей критерии оценки с функциональной типологией текстов, J. House [5], разработавшей модель на основе регистрационных параметров, и M. Williams [6], предложившего аргументативно-центрированный подход, релевантный для оценки перевода политических текстов с их сложной аргументативной структурой.

Параллельно развивались компьютерные методы автоматической оценки качества перевода, такие как метрики BLEU и NIST, позволяющие квантифицировать степень совпадения переводов с эталонными текстами, и методы предиктивного моделирования качества без использования эталонов (Quality Estimation) [7]. Также развивались подходы на основе ключевых языковых явлений, где важную роль сыграла работа Юй Шивэня [8], фокусирующаяся на оценке способности систем точно передавать определенные языковые структуры, и методы сравнения качества переводов, в которых традиционные балльные оценки уступили место ранжированию, показывающему более высокую согласованность между экспертами, что подтверждено исследованиями [9] и с 2010 года применяется на семинарах WMT.

Анализ этих методов показывает, что автоматические метрики обеспечивают эффективность и объективность оценки, в то время как лингвистические подходы лучше учитывают семантическую глубину и культурные аспекты перевода. Однако оба направления столкнулись с методологическими ограничениями, и лишь с появлением больших языковых моделей в конце 2022 года открылись принципиально новые возможности для создания интегрированных систем оценки качества перевода. Исследования в области стандартизации оценки качества с применением БЯМ развиваются в нескольких направлениях. Разработаны комплексные метрики, позволяющие объективизировать оценку по множеству параметров одновременно [10], [11], предложены критерии оценки культурно-прагматической адекватности политического перевода [12]. Сформулированы стандарты многоуровневой оценки с учетом структурно-семантических и функциональных особенностей текста [13], созданы адаптивные системы критериев для различных типов политического дискурса [14]. В методологическом плане переосмыслены фундаментальные критерии эквивалентности и адекватности в контексте возможностей БЯМ [15], а также предложены новые параметры оценки, учитывающие прагматический потенциал перевода в сфере межгосударственной коммуникации [16-18]. Несмотря на то, что за последние 2-3 года появился ряд исследований, посвященных применению больших языковых моделей в оценке качества

перевода, эти работы либо носят обзорный характер, либо сконцентрированы исключительно на технических аспектах, либо ограничиваются только лингвистическими и переводческими теориями. Крайне редко встречаются исследования, которые объединяют эти два подхода и при этом ориентированы на специфические текстовые области с разработкой конкретных систем оценки.

В настоящем исследовании предлагается интегративный методологический подход к разработке таких стандартов, синтезирующий достижения лингвистической теории перевода и компьютерных методов анализа с функциональными возможностями БЯМ. Такая методологическая конвергенция позволяет преодолеть лимитирующие факторы отдельных подходов и сформировать комплексную систему стандартов оценки качества перевода. Особое внимание уделяется политическому дискурсу, характеризующемуся высокой степенью культурной обусловленности, имплицитностью, наличием идеологически нагруженной лексики и риторических приемов, что делает его перевод особенно сложной задачей. Предлагаемая система оценки включает многоуровневые критерии, охватывающие как структурно-семантическую эквивалентность, так и прагматическую адекватность, с применением БЯМ для углубленного анализа контекстуальных и коннотативных аспектов перевода политических текстов.

Целью данного исследования является разработка научно обоснованной и технологически реализуемой системы стандартов оценки качества перевода китайского политического дискурса на русский язык. Для достижения поставленной цели сформулированы следующие исследовательские задачи:

- (1) Проанализировать существующие лингвистические и переводоведческие модели оценки качества перевода;
- (2) Рассмотреть компьютерные методы автоматической оценки переводов;
- (3) Обосновать принципы интеграции традиционных теорий и технологий больших языковых моделей;
- (4) Разработать комплексную модель оценки качества перевода, учитывающую информационную целостность, языковую точность и межкультурную читабельность.

Научная новизна исследования заключается в разработке интегративной модели стандартов оценки качества перевода политического дискурса, основанной на методологическом синтезе лингвистического и технологического подходов с имплементацией функциональных возможностей больших языковых моделей.

Теоретическая значимость работы определяется вкладом в развитие методологии оценки качества перевода и концептуализацией роли БЯМ в переводческом процессе. Практическая ценность обусловлена возможностью применения разработанных стандартов для оптимизации качества перевода китайского политического дискурса на русский язык, что способствует интенсификации российско-китайского стратегического взаимодействия и повышению эффективности межкультурной коммуникации в контексте современных геополитических трансформаций.

1. Оценка качества перевода с позиций лингвистики и переводоведения

Лингвистические и переводоведческие подходы к оценке качества перевода представляют собой фундаментальную основу для понимания многоаспектной природы перевода. В данном разделе рассматриваются ключевые теоретические модели, которые легли в основу предлагаемой интегративной системы оценки.

1.1 Модель оценки перевода Reiss

K. Reiss, ведущий представитель немецкой школы переводоведения 1970-х годов, предложила функциональный подход к оценке качества перевода в работе «Критика перевода: Возможности и ограничения»^[4]. Опираясь на модель Karl Bühler^[19], выделившего три функции языка, Reiss разработала типологию текстов: информативные, экспрессивные и оперативные, позже добавив аудиовизуальные тексты. Ключевой тезис Reiss: «Тип текста является первостепенным фактором, влияющим на выбор метода перевода»^[4, с. 17]. Система оценки Reiss включает внутриязыковые факторы (семантические, лексические, грамматические, стилистические) и внеязыковые (ситуация, тема, время, место, адресат, отправитель, эмоциональный фактор).

Хотя Reiss не разработала количественные методы оценки, её подход стал основой для последующих исследований. Как отмечает Чжан Мэйфан^[20], теория имеет ограничения: классификация языковых функций упрощена, а критерии типологии текстов дискуссионны.

1.2 Функционально-прагматическая модель House

J. House является одним из наиболее влиятельных теоретиков в современном переводоведении. Ее модель оценки качества перевода, разработанная в конце 1970-х годов, до сих пор широко обсуждается и применяется. Теория House представлена в двух ключевых работах: «Модель оценки качества перевода»^[21] и переработанная версия «Оценка качества перевода: пересмотренная модель»^[5]. Теоретическая база модели House основана на системно-функциональной лингвистике с использованием теории регистра Halliday. House рассматривает перевод как замещение исходного текста текстом на целевом языке при сохранении семантической и прагматической эквивалентности. Ее модель оценки включает анализ регистра (поле, тональность и модус), где поле относится к теме и содержанию текста, тональность – к отношениям между участниками коммуникации, а модус – к каналу и способу использования языка. В пересмотренной версии модели House вводит понятие жанра как социально признанного типа текста, имеющего существенное значение для оценки качества перевода. House подчеркивает важность функциональной эквивалентности, предлагая концепции «явного перевода» и «скрытого перевода». Явный перевод очевидно представляет собой переводной текст, тогда как скрытый перевод стремится создать функционально эквивалентный текст в культуре целевого языка. Ошибки в переводе House классифицирует как «явные ошибки», включающие очевидные языковые ошибки, и «скрытые ошибки», связанные с нарушением функциональной эквивалентности.

Несмотря на значительное влияние модели House, она подвергалась критике со стороны некоторых ученых. По мнению Юань Хуна^[22], хотя модель предлагает детальный анализ, ей не хватает количественных стандартов для объективной оценки. Тем не менее, теория House обеспечивает комплексную основу для оценки качества перевода, особенно благодаря вниманию к контекстуальным факторам и функциональной эквивалентности.

1.3 Оценочная модель Williams

Malcolm Williams является одним из важнейших теоретиков в области оценки качества перевода начала XXI века. В своей книге «"Оценка качества перевода: подход, ориентированный на аргументацию»^[6] он предложил инновационную модель оценки

качества перевода, объединяющую традиционные лингвистические методы с теорией аргументации. Williams утверждает, что все тексты обладают универсальной аргументационной структурой, что делает его подход применимым к различным типам текстов в переводческой практике. Модель Williams включает анализ аргументационной схемы (тезис, основания, обоснование, поддержка, квалификатор и опровержение), организационных отношений (проблема-решение, вывод-причина), пропозициональных функций и связующих элементов, типов аргументации, риторических фигур и нарративных стратегий. По мнению Williams, эти параметры охватывают всю информацию и цели текста на макро- и микроуровнях, обеспечивая комплексный подход к оценке перевода. Процесс оценки в модели Williams предполагает определение аргументационной схемы и структурных отношений в исходном тексте, анализ связности целевого текста, оценку ключевых сегментов, сравнительный анализ пропозициональных функций, типов аргументации и нарративных стратегий с последующей общей оценкой качества перевода. Williams разработал оценочные шкалы с различными параметрами и весовыми коэффициентами для разных типов текстов. Например, для аргументативных текстов больший вес придается аргументационной схеме и стилистическим аспектам, тогда как для статистических отчетов более важны пропозициональные функции и терминология.

Несмотря на системность и практическую ценность модели Williams, особенно для оценки сложных аргументативных текстов, она подвергается критике за сложность применения, требующую глубоких знаний лингвистики и теории аргументации, а также за недостаточную четкость некоторых количественных критериев оценки.

1.4 Концепция оценки перевода Newmark

Peter Newmark, известный британский теоретик перевода, внес значительный вклад в исследование оценки качества перевода. Хотя он не разработал формальную модель оценки, его идеи по этому вопросу широко представлены в работах «Подходы к переводу» [23] и «Учебник по переводу» [24]. Подход Newmark к оценке качества перевода основан на его понимании сущности перевода как искусства воссоздания информации и идей оригинального автора на другом языке. Newmark предлагает функциональную оценку, учитывающую тип текста (экспрессивный, информативный или воздействующий), и подчеркивает важность языковой точности, для чего разработал метод «компонентного анализа», предполагающий детальное исследование оригинала и перевода на уровне слов и предложений. Важное место в теории Newmark занимает учет культурных факторов и введенное им понятие «культурных слов». Он различает семантический и коммуникативный перевод, считая, что разные методы подходят для разных типов текстов и должны оцениваться соответственно. Newmark также разработал систематическую классификацию переводческих ошибок, включающую языковые, культурные и прагматические ошибки, что позволяет проводить объективную оценку качества перевода.

1.5 Вклад китайских ученых в теорию оценки перевода

Исследования оценки качества перевода в Китае начались несколько позже, чем на Западе, но отличаются значительными результатами. Китайские ученые не только познакомили научное сообщество с западными теориями, но и разработали собственные модели оценки качества перевода. Рассмотрим три наиболее влиятельные китайские модели.

1.5.1 Модель наилучшей аппроксимации Гу Чжэнкуня

Гу Чжэнкунь предложил модель «наилучшей аппроксимации» [25], основанную на многомерной взаимодополняющей теории стандартов перевода. Он разработал систему стандартов: «абсолютный стандарт (оригинал) – высший стандарт (наилучшая аппроксимация) – конкретные стандарты». В этой системе высший стандарт – «наилучшая аппроксимация» – определяется как «степень достоверности, с которой перевод имитирует содержание и форму оригинала». Пэн Чунъянь [26] модифицировала эту систему, подчеркнув многомерность, социальный характер и эволюционность стандартов перевода.

1.5.2 Модель функциональной лингвистики Сы Сяньчжу

Сы Сяньчжу разработал модель оценки на основе системно-функциональной лингвистики. По его мнению, «сущность перевода заключается в эквивалентности значений оригинала и перевода на семантическом, прагматическом и текстовом уровнях» [27]. Эта модель предполагает анализ переходности, наклонения и модальности в предложениях оригинала и перевода для выявления концептуальных и межличностных значений, а также отклонений перевода от оригинала. Затем проводится классификация и оценка этих отклонений для определения степени эквивалентности текстов.

1.5.3 Модель теории релевантности Хэ Саннина

Хэ Саннин применил теорию релевантности к оценке качества перевода. Он считает, что максимальная релевантность перевода заключается в стремлении к совпадению с оригиналом, что является теоретической основой оценки качества [28]. Хэ Саннин подчеркивает, что оценка качества перевода «связана с совпадением значений оригинала и перевода на всех уровнях – от предложения до текста, от стиля до воздействия на читателя, стремясь к максимальной релевантности контекстуальных эффектов». Его модель включает микроэлементы оценки: «межъязыковую релевантность», «текстовую релевантность» и «общую релевантность».

Модели оценки качества перевода, разработанные как западными, так и китайскими учеными, отражают эволюцию теории перевода и стремление к объективной оценке. Хотя лингвистические модели создают теоретическую основу, их практическое применение часто затруднено сложностью и низкой эффективностью. В этом контексте автоматизированные методы оценки предлагают решения, повышающие скорость и объективность процесса.

2. Компьютерные методы автоматической оценки качества перевода

Современные методы автоматической оценки качества перевода предлагают эффективные решения для преодоления ограничений лингвистических моделей. Основанные на компьютерных технологиях, они обеспечивают количественный, быстрый и объективный анализ переводов. Четыре основных подхода – оценка с эталоном, оценка без эталона, тестирование ключевых языковых элементов и сравнительное ранжирование – дополняют традиционные теоретические модели оценки.

2.1 Оценка с использованием эталонного перевода

2.1.1 Алгоритм BLEU

Алгоритм BLEU, предложенный Раринепи коллегами из IBM в 2002 году, оценивает качество перевода путем сравнения n-граммного совпадения между машинным переводом и одним или несколькими эталонными переводами [29]. Метод вычисляет

точность n -грамм, применяет штрафной коэффициент за краткость и взвешенное геометрическое среднее для точности n -грамм разной длины. Преимущества BLEU включают простоту вычисления и быстроту, а также объективную оценку от 0 до 1. Однако метод имеет ограничения: он не оценивает семантическую эквивалентность, фокусируясь лишь на лексическом совпадении, недостаточно чувствителен к структурным изменениям и требует нескольких эталонных переводов для надежной оценки.

2.1.2 Алгоритм NIST

Алгоритм NIST, разработанный Doddington в 2002 году [\[30\]](#), является усовершенствованной версией BLEU. Он также использует n -граммное соответствие, но вводит концепцию информационного веса. NIST вычисляет точность n -грамм с учетом информационного веса, использует арифметическое среднее вместо геометрического и оптимизирует штрафной коэффициент для длины предложения. Ключевое нововведение — присвоение более высоких весов редким n -граммам, что повышает точность и различительную способность оценки. Несмотря на эти усовершенствования, NIST сохраняет некоторые ограничения методов на основе n -грамм.

2.2 Оценка качества без эталонного перевода

Оценка качества без эталона (Quality Estimation, QE), предложенная Specia и коллегами в 2010 году [\[71\]](#), использует методы машинного обучения с учителем для прогнозирования качества перевода без эталонных текстов. Этот подход включает извлечение лингвистических и статистических признаков из исходного текста и машинного перевода, обучение модели на размеченных данных и прогнозирование оценки качества. Преимущества QE включают гибкость и практичность, возможность применения на уровне слов, предложений и документов. Однако метод требует качественных обучающих данных, тщательного выбора признаков и имеет ограничения в обобщающей способности.

2.3 Тестирование на основе ключевых языковых элементов

Методы оценки на основе ключевых языковых точек оценивают качество перевода через заранее определенные языковые элементы, охватывающие ключевую семантику и распространенные ошибки. Пионерская работа в этой области была проведена Юй Шивэнь из Пекинского университета в начале 1990-х годов [\[8\]](#). Его команда разработала комплексную систему оценки для китайско-английского машинного перевода, включающую тестовые спецификации и специальный язык описания качества перевода. Система классифицировала языковые тестовые точки на 9 категорий, охватывающих аспекты от лексики до перевода сложных предложений. Хотя этот метод обеспечивает всесторонний анализ, он требует значительных затрат на разработку и имеет ограниченную гибкость.

2.4 Сравнительное ранжирование по оценке качества перевода

В области сравнительной оценки произошел переход от числовых оценок к методам ранжирования качества. Исследования показывают [\[9\]](#); [\[31\]](#), что при определении относительного превосходства переводов достигается большее согласие между оценщиками, чем при абсолютной оценке (коэффициент Каппа для ранжирования 0,37-0,56 против 0,22-0,25 для оценок). С 2010 года ранжирование стало официальным методом человеческой оценки на конференциях по машинному переводу (WMT). Задача ранжирования формализуется как упорядоченная классификация и решается с помощью

алгоритмов машинного обучения. Хотя ранжирование не отражает степень качественных различий, оно обеспечивает относительно надежный и практичный подход к оценке машинного перевода.

Автоматические методы оценки перевода предлагают эффективные решения для задач, где традиционные лингвистические подходы сталкиваются с трудностями. Однако они ограничены в понимании глубокой семантики и культурных нюансов. Оптимальный подход к оценке качества перевода, особенно для такой специфической области как китайский политический дискурс, должен сочетать лингвистическую теорию с вычислительными методами, учитывая также отраслевые руководства, такие как «Базовые нормы перевода китайского политического дискурса на английский язык».

3. «Базовые нормы перевода китайского политического дискурса на английский язык» как основа оценки

При разработке критериев оценки перевода китайского политического дискурса необходимо учитывать жанровые особенности текста. Как отмечает Reiss, тип исходного текста является определяющим фактором в процессе перевода, влияя не только на стратегию переводчика, но и на весь процесс принятия решений при переводе [4].

«Базовые нормы перевода китайского политического дискурса на английский язык» определяют китайский политический дискурс как «особую форму выражения, сформированную партией и правительством в процессе государственного управления» (далее «Нормы») [32, с. 2]. Согласно этому документу, основной принцип перевода политического дискурса — баланс между верностью оригиналу и читабельностью, при приоритете верности в случае невозможности достижения обоих. Документ устанавливает четыре общих требования: глубокое понимание оригинального текста перед началом перевода; учет особенностей мышления иностранной аудитории при сохранении верности оригиналу; совершенствование перевода ключевых концепций с углублением их понимания; международное сотрудничество и учет мнений носителей языка перевода. Эти принципы и требования формируют важную основу для разработки критериев оценки качества перевода китайского политического дискурса.

4. Интеграция теорий оценки качества перевода на основе БЯМ

Проведенный анализ лингвистических теорий и компьютерных методов оценки качества перевода показывает необходимость комплексного подхода к оценке перевода китайского политического дискурса. Такой подход должен учитывать как теоретическую глубину традиционных лингвистических моделей, так и эффективность компьютерных методов, а также специфику политического дискурса. С появлением крупных языковых моделей появилась возможность преодолеть ограничения существующих методов оценки.

История развития моделей оценки качества перевода отражает эволюцию от чисто лингвистического подхода к междисциплинарному. Лингвистические модели, представленные работами Reiss [4], House [5] и Newmark [23], заложили теоретическую основу, подчеркивая важность типа текста, функциональной эквивалентности и культурных факторов. Параллельно развивались компьютерные методы оценки — от анализа ключевых языковых элементов [8] до автоматических систем оценки (BLEU, NIST) и новейших методов машинного обучения [7]. Каждый из этих подходов имеет свои преимущества и ограничения в контексте оценки качества перевода. На основе анализа существующих моделей и с учетом требований «Норм», для создания эффективной

системы оценки перевода политических текстов были выбраны следующие теоретические основы:

- (1) Теория типов текста [\[4\]](#) как основа для оценки соответствия перевода жанровым особенностям политического дискурса;
- (2) Анализ типов ошибок [\[23\]](#) для выявления лексических, синтаксических и семантических отклонений;
- (3) Метод ключевых элементов оценки [\[8\]](#) для проверки перевода на лексическом, синтаксическом и прагматическом уровнях;
- (4) Оценка перевода культурно-маркированной лексики [\[23\]](#), особенно важная для политических текстов с китайской спецификой;
- (5) Система взвешенных параметров оценки, адаптированная к особенностям политического дискурса.

Использование крупных языковых моделей позволяет интегрировать преимущества различных подходов, обеспечивая как глубину анализа, так и эффективность автоматической оценки качества перевода китайского политического дискурса.

4.1 Параметры оценки качества перевода на основе больших языковых моделей

В соответствии с пятью классическими теориями оценки качества перевода, рассмотренными ранее, и требованиями «Норм», мы разработали систему оценки качества перевода китайского политического дискурса, основанную на трех ключевых критериях. Несмотря на терминологические различия в работах разных исследователей (например, «стиль» у Reiss и «прагматика» у House), наш подход основывается на фундаментальном принципе, указанном в «Нормах»: «Основной принцип перевода китайского политического дискурса заключается в балансе между точностью и удобочитаемостью» [\[32, с. 4\]](#). Исходя из этого, мы определили три основных измерения для оценки качества перевода: информационная целостность, языковая точность и межкультурная удобочитаемость.

4.1.1 Информационная целостность

Информационная целостность представляет собой критерий, ориентированный на исходный текст, и оценивает степень передачи информации оригинала в переводе. Большие языковые модели могут эффективно анализировать семантические структуры исходного и переводного текстов, идентифицировать ключевые информационные точки и сравнивать соответствие между ними, что повышает эффективность оценки и позволяет выявлять мельчайшие расхождения, которые могут быть пропущены при ручной оценке.

4.1.2 Языковая точность

Языковая точность ориентирована на переводной текст и оценивает степень правильного использования языковых норм и выразительных средств целевого языка. Большие языковые модели способны быстро идентифицировать лексические и синтаксические ошибки, обнаруживать несоответствующие словосочетания или способы выражения и предлагать более аутентичные альтернативы.

4.1.3 Межкультурная удобочитаемость

Межкультурная удобочитаемость ориентирована на читателя целевого языка и оценивает, насколько перевод соответствует языковым привычкам и культурным ожиданиям целевой аудитории. Большие языковые модели могут анализировать языковой стиль перевода, идентифицировать выражения, способные вызвать культурное недопонимание, и предлагать альтернативы, более соответствующие культуре целевого языка.

4.2 Параметры оценки качества перевода на основе больших языковых моделей

Процесс определения параметров представляет собой комплексный процесс, учитывающий многие факторы. Наш подход сочетает теоретические основы классических теорий перевода, требования отраслевых стандартов и технические возможности больших языковых моделей. Следуя стратегии «ориентации на отраслевые стандарты, руководства по лингвистике и теории перевода, реализации с помощью компьютерных технологий», мы разработали детальные параметры для каждого из трех основных измерений.

4.2.1 Параметры информационной целостности

1) Степень целостности информации

- (1) Переведены ли ключевые политические термины или концепции
- (2) Опущены ли важные данные или статистическая информация
- (3) Переведены ли важные компоненты предложений или абзацев
- (4) Переведены ли примеры или цитаты из оригинала

2) Степень достоверности информации

- (1) Изменено ли значение политических терминов или профессиональной лексики
- (2) Правильно ли переведены числа, даты или единицы измерения
- (3) Правильно ли переведены многозначные слова
- (4) Правильно ли переведены идиомы или устойчивые выражения

3) Семантические отношения

- (1) Правильно ли выражены причинно-следственные связи
- (2) Правильно ли выражены отношения противопоставления
- (3) Правильно ли выражены параллельные отношения
- (4) Правильно ли выражены условные отношения

4.2.2. Параметры языковой точности

1) Лексический уровень

- (1) Правильность словоизменения (склонение существительных, прилагательных, числительных и т.д.)
- (2) Правильность спряжения глаголов

- (3) Правильность глагольного вида
 - (4) Правильность использования частей речи
 - 2) Синтаксический уровень
 - (1) Правильность порядка слов
 - (2) Согласование подлежащего и сказуемого
 - (3) Правильность структуры придаточных предложений
 - (4) Правильность предложных сочетаний
 - 3) Избыточное выражение
 - (1) Наличие лишних знаменательных слов
 - (2) Наличие лишних служебных слов
 - (3) Наличие ненужных повторений
 - 4) Пунктуация и орфография
 - (1) Правильность использования знаков препинания
 - (2) Правильность написания слов
 - (3) Правильность использования прописных и строчных букв
- #### 4.2.3 Параметры межкультурной удобочитаемости
- 1) Уровень стиля
 - (1) Сохранен ли стиль политических документов
 - (2) Сохранены ли тон и интонация оригинала
 - (3) Адекватна ли длина предложений (не слишком длинные или короткие)
 - (4) Уместен ли перевод риторических приемов
 - 2) Эмоциональный уровень
 - (1) Точно ли передана эмоциональная направленность оригинала
 - (2) Точно ли выражена эмоциональная интенсивность оригинала
 - (3) Точно ли выражена убедительность или воздействие оригинала
 - 3) Культурный уровень
 - (1) Правильно ли переведены культурные ключевые слова
 - (2) Учтены ли культурный фон и когнитивные привычки читателей целевого языка
 - (3) Правильно ли обработаны различия в политической культуре

Разработанная система оценки качества перевода китайского политического дискурса охватывает основные элементы, рассматриваемые в традиционных теориях перевода, такие как точность передачи информации и нормативность использования языка, а также

особо подчеркивает важность межкультурной коммуникации для удовлетворения специфики и сложности перевода политического дискурса. Технические преимущества БЯМ, включая семантическое понимание, грамматический анализ, обширную базу знаний и понимание контекста, значительно повышают точность и эффективность оценки по всем трем измерениям.

Заключение

Проведенное исследование представляет интеграцию лингвистических теорий и технологических возможностей больших языковых моделей в сфере оценки качества перевода политического дискурса. Разработанная трехуровневая модель, охватывающая параметры информационной целостности, языковой точности и межкультурной удобочитаемости, создает методологически обоснованную систему для объективной оценки переводческой деятельности.

Предложенная система с 10 параметрами второго уровня и 36 параметрами третьего уровня, дополненная весовыми коэффициентами и стабильной градацией качества, обеспечивает необходимую гибкость при оценке различных жанров политического дискурса. Интеграция БЯМ существенно повышает эффективность выявления структурных и стилистических несоответствий, а также способствует автоматизированной диагностике переводческих проблем.

Библиография

1. Федоров А. В. Введение в теорию перевода: (лингвистич. проблемы). – 2-е изд., перераб. – Москва: Изд. лит. на иностр. яз., 1958. – 374 с.
2. Бархударов Л. С. Язык и перевод. – М.: Международные отношения, 1975. – 235 с.
3. Комиссаров В. Н. Лингвистика перевода. – Москва: Международные отношения, 1980. – 167 с.
4. Райс К. Возможности и границы критики перевода. – Мюнхен: М. Хюбер, 1971. – 124 с.
5. House J. Translation Quality Assessment: A Model Revisited. – Gunter Narr Verlag, 1997. – 207 р.
6. Williams M. Translation quality assessment: An argumentation-centred approach. – University of Ottawa Press, 2004. – 209 р.
7. Specia L., Raj D., Turchi M. Machine Translation Evaluation versus Quality Estimation // Machine Translation. – 2010. – Vol. 24. – P. 39-50. DOI: 10.1007/s10590-010-9077-2 EDN: QHHKQW
8. Юй Шивэнь. О некоторых исследованиях в области вычислительной лингвистики // Применение языка и письменности. – 1993. – № 3. – С. 55-64.
9. Callison-Burch C., Fordyce C. S., Koehn P., Monz C., Schroeder J. (Meta-) Evaluation of Machine Translation // In Proceedings of the Second Workshop on Statistical Machine Translation. – 2007. – P. 136-158.
10. Kocmi T., Federmann C. Large language models are state-of-the-art evaluators of translation quality // arXiv preprint arXiv:2302.14520. – 2023.
11. Денисенко В. В., Чесников Л. С. Стратегии оптимизации и методы оценки для точной настройки больших языковых моделей // Международный журнал гуманитарных и естественных наук. – 2024. – № 4-1 (91). – С. 180-184. DOI: 10.24412/2500-1000-2024-4-1-180-184 EDN: MGIFYI
12. Лю М., Шао Ц., Се Г. Автоматизированный перевод политического дискурса: от больших языковых моделей к мультиагентной системе MAGIC-PTF // Litera. 2024. № 11. С. 28-46. DOI: 10.25136/2409-8698.2024.11.72197 EDN: GFRZMO URL: https://nbpublish.com/library_read_article.php?id=72197

13. ЧжАО Цзюньфэн, Ли Сян. Исследование построения и применения переводческих агентов на основе больших языковых моделей // Преподавание иностранных языков с помощью электронных технологий. – 2024. – № 5. – С. 22-28, 75, 108.
14. Ли Дэфэн, Ван Хуашу, Лю Шицзе. Национальные возможности переводческих технологий и большие языковые модели // Шанхайский перевод. – 2025. – № 2. – С. 18-24.
15. Lu Q., et al. Error analysis prompting enables human-like translation evaluation in large language models // arXiv preprint arXiv:2303.13809. – 2023.
16. Huang H., et al. Towards making the most of llm for translation quality estimation // CCF International Conference on Natural Language Processing and Chinese Computing. – Cham: Springer Nature Switzerland, 2023.
17. Zhang B., Haddow B., Birch A. Prompting large language model for machine translation: A case study // International Conference on Machine Learning. – PMLR, 2023.
18. ЧжАО Янь, Чжан Хуэй, Ян Ичэнь. Сравнительное исследование качества больших языковых моделей в переводе текстов - на примере перевода "Цветения" // Преподавание иностранных языков с помощью электронных технологий. – 2024. – № 4. – С. 60-66, 109.
19. Бюлер К. Теория языка: Репрезентативная функция языка. – Йена: Густав Фишер, 1934. – 434 с.
20. Чжан Мэйфэн. Функциональный подход к исследованию перевода. – Шанхай: Шанхайское издательство образования по иностранным языкам, 2005. – 233 с.
21. House J. A Model for Translation Quality Assessment. – TBL-Verlag Narr, 1977. – 344 p.
22. Юань Хун. Объяснительная сила типологии текста и функциональной лингвистики в оценке качества перевода // Журнал Хунаньского института гуманитарных наук и технологий. – 2007. – № 1. – С. 158-161.
23. Newmark P. Approaches to Translation. – Oxford: Pergamon Press, 1981. – 344 p.
24. Newmark P. A textbook of translation. – Vol. 66. – New York: Prentice Hall, 1988. – 402 p.
25. Гу Чжэнкунь. О многомерной взаимодополняющей теории стандартов перевода // Китайский перевод. – 1989. – № 1. – С. 16-20.
26. Пэн Чунъянь. Новая теория стандартов перевода - пересмотр системы стандартов перевода, установленной профессором Гу Чжэнкунем // Научный журнал Университета Сунь Ятсена. – 2004. – № 5. – С. 237-241.
27. Сы Сяньчжу. Исследование модели оценки качества перевода с точки зрения функциональной лингвистики // Преподавание иностранных языков. – 2004. – № 4. – С. 45-50.
28. Хэ Саньнин. Исследование моделей оценки качества перевода. – Пекин: Центральное издательство переводов, 2015. – 256 с.
29. Papineni K., Roukos S., Ward T., Zhu W. J. "BLEU: A Method for Automatic Evaluation of Machine Translation." // In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics. – 2002. – Р. 311-318.
30. Doddington G. "Automatic Evaluation of Machine Translation Quality Using N-gram Co-occurrence Statistics." // In Proceedings of the Second International Conference on Human Language Technology Research. – 2002. – Р. 138-145.
31. Duh K. "Ranking vs. Regression in Machine Translation Evaluation." // In Proceedings of the Third Workshop on Statistical Machine Translation. – 2008. – Р. 191-194.
32. Базовые нормы перевода китайского политического дискурса на английский язык / Под ред. редакционной группы. – Пекин: Издательство литературы на иностранных языках, 2023. – 234 с.

Результаты процедуры рецензирования статьи

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

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

В рецензируемой статье предметом исследования выступает оценка качества перевода политического дискурса с применением больших языковых моделей, актуальность изучения которой аргументируется, с одной стороны, важнейшей ролью «адекватного перевода политических текстов, которые, в силу высокой степени терминологической насыщенности, концептуальной специфики и идеологической маркированности, требуют особых подходов к оценке качества переводческой деятельности» в обеспечении конструктивного межкультурного диалога, с другой стороны, недостаточным количеством комплексных исследований интегративного методологического подхода к разработке стандартов оценки качества перевода политического дискурса, синтезирующего достижения лингвистической теории перевода и компьютерных методов анализа с функциональными возможностями больших языковых моделей: «Несмотря на то, за последние 2-3 года появился ряд исследований, посвященных применению больших языковых моделей в оценке качества перевода, эти работы либо носят обзорный характер, либо сконцентрированы исключительно на технических аспектах, либо ограничиваются только лингвистическими и переводческими теориями. Крайне редко встречаются исследования, которые объединяют эти два подхода и при этом ориентированы на специфические текстовые области с разработкой конкретных систем оценки».

Теоретической основой работы выступили труды таких российских и зарубежных исследователей, как Л. С. Бархударов, В. Н. Комиссаров, К. А. Бюлер, В. Федоров, К. Райс, В. В. Денисенко, Л. С. Чесников, Хэ Саньнин, Сы Сяньчжу, Пэн Чунъянь, Р. Newmark, J. House, Т. Kostki, C. Federmann и др., посвященные различным аспектам теории и практики перевода, современным переводческим технологиям, большим языковым моделям, моделям оценки качества перевода и т. д. Библиография статьи насчитывает 32 источника, представляется достаточной для обобщения и анализа теоретического аспекта изучаемой проблематики, соответствует специфике изучаемого предмета, содержательным требованиям и находит отражение на страницах рукописи. Все цитаты ученых сопровождаются авторскими комментариями.

Методология исследования определена поставленной целью («разработка научно обоснованной и технологически реализуемой системы стандартов оценки качества перевода китайского политического дискурса на русский язык») и задачами («проанализировать существующие лингвистические и переводоведческие модели оценки качества перевода; рассмотреть компьютерные методы автоматической оценки переводов; обосновать принципы интеграции традиционных теорий и технологий больших языковых моделей; разработать комплексную модель оценки качества перевода, учитывающую информационную целостность, языковую точность и межкультурную читабельность») и носит комплексный характер: применяются общенаучные методы анализа и синтеза, обобщения; описательный метод; дискурсивный и сравнительно-сопоставительный методы; социокультурный анализ и др. В ходе работы последовательно рассматриваются ключевые теоретические модели, которые легли в основу предлагаемой интегративной системы оценки (модель оценки перевода Reiss, функционально-прагматическая модель House, оценочная модель Williams, концепция оценки перевода Newmark), а также вклад китайских ученых в теорию оценки перевода; компьютерные методы автоматической оценки качества перевода (оценка с эталоном, оценка без эталона, тестирование ключевых языковых

элементов и сравнительное ранжирование); «базовые нормы перевода китайского политического дискурса на английский язык» как основа оценки; интеграция теорий оценки качества перевода на основе больших языковых моделей. Разработана трехуровневая модель оценки качества перевода, охватывающая параметры информационной целостности, языковой точности и межкультурной удобочитаемости, что создает методологически обоснованную систему для объективной оценки переводческой деятельности.

Полученные результаты, безусловно, имеют теоретическую значимость и определяются вкладом в развитие методологии оценки качества перевода и концептуализацией роли больших языковых моделей в переводческом процессе. Практическая ценность обусловлена возможностью применения разработанных стандартов для оптимизации качества перевода политического дискурса с китайского на русский язык, что «способствует интенсификации российско-китайского стратегического взаимодействия и повышению эффективности межкультурной коммуникации в контексте современных геополитических трансформаций».

Стиль изложения отвечает требованиям научного описания и характеризуется логичностью и доступностью. Содержание работы соответствует названию, логика исследования четкая. Статья имеет завершенный вид; она вполне самостоятельна, оригинальна, будет интересна и полезна широкому кругу лиц и может быть рекомендована к публикации в научном журнале «Litera».