

КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ИНДИКАТОРОВ ФЕНОМЕНА TRANSLATIONESE В ПЕРЕВОДАХ МЕДИЦИНСКИХ НАУЧНЫХ СТАТЕЙ

O. U. Серова

Аннотация

Введение. Феномен переводческого языка (translationese) представляет собой комплекс лингвистических характеристик, отличающих переводные тексты от оригинальных (непереводных) текстов на языке перевода. В переводе каждая из этих характеристик выражается рядом формальных индикаторов, наличие которых в тексте нередко связывают со снижением качества перевода и его неестественностью. Все это делает важным учет данных характеристик при переводе и его постредактировании, особенно если речь идет о переводе медицинских научных статей, где точность передачи информации критически значима для клинической практики и, как следствие, здоровья пациентов. Более того, в условиях активной интеграции международной медицинской практики в российскую клиническую практику качество перевода становится критически важным фактором, определяющим эффективность внедрения зарубежных клинических рекомендаций и протоколов лечения. В связи с этим целью исследования является установление индикаторов переводческого языка в англо-русских переводах медицинских научных статей.

Материалы и методы. Исследование проведено на материале сопоставимого корпуса из 80 текстов, включающего 40 переводных и 40 оригинальных медицинских научных статей. Все тексты отобраны из рецензируемых профильных журналов и сопоставимы по жанру, объему и тематике. Для анализа характеристик использовался специализированный инструмент — автоматический анализатор феномена translationese, позволяющий исследовать формальные индикаторы таких переводческих характеристик, как упрощение, нормализация, экспликация и интерференция.

Результаты исследования. Анализ индикаторов указанных характеристик позволил выявить наличие в переводах таких особенностей переводческого языка, как снижение лексической плотности, лексического разнообразия, уникальности используемой в тексте лексики, увеличение частоты использования модальных конструкций с инфинитивом, причастных оборотов, предложных конструкций, придаточных предложений, повышение частоты употребления дискурсивных маркеров противопоставления, введение новой информации и примеров, увеличение доли глаголов в страдательном залоге.

Заключение. На основе полученных результатов предлагаются предварительные рекомендации, касающиеся терминологической точности, контроля синтаксической интерференции, модальности и глагольного залога в переводе. Данные рекомендации могут быть полезными для переводчиков и постредакторов переводов медицинских научных статей.

Ключевые слова: научная статья, медицинский перевод, translationese, индикаторы translationese, упрощение, нормализация, экспликация, интерференция

A QUANTITATIVE ANALYSIS OF TRANSLATIONESE INDICATORS IN MEDICAL LITERATURE TRANSLATIONS

O. U. Serova

Abstract

Introduction. Translationese comprises a range of linguistic characteristics that distinguish translated texts from originally composed (non-translated) texts in a target language. These characteristics manifest through specific formal indicators frequently associated with reduced translation quality and stylistic

unnaturalness. Accounting for such features is particularly crucial in the translation and post-editing of medical scientific literature, where the accurate transmission of information is paramount for clinical practice and patient safety. Moreover, in the context of integrating international medical practice into Russian clinical practice, translation quality becomes a critical determinant for the effective implementation of foreign clinical guidelines and treatment protocols. This study aims to identify indicators of *translationese* in English-to-Russian translations of medical research articles.

Materials and Methods. The investigation utilized a comparable corpus of 80 texts, comprising 40 translated and 40 non-translated medical research papers. All texts were sourced from peer-reviewed specialized journals and are matched for genre, length, and thematic focus. The analysis was performed using a specialized automated tool designed to detect formal indicators associated with the *translationese* characteristics of simplification, normalization, explication, and interference.

Results. The analysis confirmed the presence of distinct *translationese* characteristics in the translated corpus. These include decreased lexical density, reduced lexical variety, and lower lexical uniqueness. The translations also exhibited increased frequency of modal constructions with infinitives, participial phrases, prepositional constructions, and subordinate clauses. Furthermore, translations were identified to have higher frequency of discourse markers for introducing new information, expressing contrast, and providing examples, alongside an increased proportion of passive voice verbs.

Conclusions. Based on these results, the study proposes preliminary recommendations for translators and post-editors focusing on enhancing terminological precision, mitigating syntactic interference, and carefully managing modality and verb voice.

Keywords: scientific literature, medical translation, *translationese*, *translationese* indicators, simplification, normalization, explication, interference

Введение

Феномен переводческого языка представляет собой комплекс специфических лингвистических характеристик, отличающих переводы от оригинальных (непереводных) текстов на языке перевода. В современной лингвистике данный феномен принято называть термином *translationese* (Gellerstam 1986).

На практике *translationese* проявляется в виде таких характеристик, как *упрощение*, *экспликация*, *нормализация* и *интерференция*, а их совокупность позволяет говорить о переводе как о специфическом «промежуточном языке» (Toury 1979), отличном как от исходного языка (ИЯ), так и от языка перевода (ПЯ). При этом под *упрощением* принято понимать снижение сложности переводного текста на морфологическом, лексическом и синтаксическом уровнях с целью облегчения его восприятия читателем (Baker 1993; Volansky et al. 2013). *Нормализация* выражается в более частом употреблении в переводе стандартизованных конструкций ПЯ, в избегании лексических повторов и коллокаций,

нехарактерных для ПЯ (Volansky et al. 2013). *Экспликация* состоит в повышении объяснятельного потенциала текста перевода в сравнении с текстом оригинала посредством более частого использования языковых средств, повышающих когезию и когерентность текста (Blum-Kulka 1986; Volansky et al. 2013). Под *интерференцией* подразумевается перенос морфо-синтаксических черт ИЯ в текст перевода (Toury 1979; Volansky et al. 2013).

Феномен *translationese* связывают со снижением качества переводов, выражаящимся в неестественности и стилистической неуклюжести текста и, как следствие, в искажении смысла оригинала (Baker 1993; Blum-Kulka 1986). Исследование М. А. Куниловской (Kunilovskaya 2023) показало корреляцию между степенью выраженности индикаторов *translationese* и ухудшением качества перевода (при холистической оценке качества на уровне всего документа).

Критическую значимость проблема качества приобретает при переводе медицинских научных статей, поскольку неточности и искажения в медицинском переводе способны напрямую влиять на клиническую практику

и, как следствие, на здоровье пациентов. Актуальность проблемы возрастает и в контексте интеграции мировой науки в российскую медицинскую практику, что требует корректной передачи содержания зарубежных исследований, клинических рекомендаций и протоколов лечения, публикуемых в виде переводов научных статей в русскоязычных медицинских изданиях.

Тем не менее, несмотря на значительное количество исследований, посвященных изучению *translationese* на материале различных языковых пар и текстовых жанров (см., например: (Беляева и др. 2023; Краснопеева 2015; Серова 2025; Koppel, Ordan 2011; Kunilovskaya, Corpas Pastor 2021; Rabinovich 2015; Volansky et al. 2013)), особенности его проявления в медицинском научном дискурсе остаются недостаточно изученными. При этом медицинские научные статьи в силу своей жанровой специфики и pragматической направленности могут демонстрировать особые паттерны переводческого языка, выявление которых представляет научный интерес. В связи с этим целью настоящей работы является исследование индикаторов феномена *translationese* в англо-русских переводах медицинских научных статей.

Материалы и методы исследования

Материалом исследования послужил сопоставимый корпус из 80 текстов, включивший в себя подкорпус английских ручных переводов статей и подкорпус непереводных русскоязычных статей. Все статьи были отобраны из рецензируемых научных журналов по медицине (Портал научных журналов ELPUB 2025). Из текстов предварительно были удалены ссылки на литературу, заголовки, аннотации, ключевые слова и иллюстративные материалы. Для обеспечения сопоставимости подкорпусов применялись следующие критерии отбора:

- 1) принадлежность к научному функциональному стилю;
- 2) соответствие жанру «научная статья»;
- 3) тематическая однородность — область «медицина»;

4) сопоставимость по объему — по 40 статей общим объемом в 140 000 словоформ для каждого подкорпуса.

Подкорпусы были проанализированы с помощью анализатора феномена *translationese* (Анализатор феномена *translationese* 2025) на наличие формальных индикаторов каждой из характеристик. Рассмотрим эти индикаторы более подробно.

Для анализа *упрощения* анализировались следующие индикаторы (Volansky et al. 2013).

Лексическое разнообразие (lexical variety) текста рассчитывалось как:

$$TTR = \frac{V}{N} \times 100 \text{ (базовая версия),}$$

Log TTR = $\frac{\log V}{\log N} \times 100$ (нивелирует зависимость показателя от длины текста),

Modified TTR = $\frac{100 \times \log N}{1 - \frac{V_1}{V}}$ (учитывает слова с единичной встречаемостью (*hapax legomena*)),

где V — количество уникальных слов, N — общее количество словоформ, V_1 — количество *hapax legomena*.

Лексическая плотность (lexical density) текста вычислялась как отношение знаменательных частей речи к общему количеству слов.

Средняя длина слов (mean word length) — в символах и слогах.

Средняя длина предложений (mean sentence length) — в символах и токенах.

Средний ранг слов (mean word rank) вычислялся на основе ранжирования слов текста по частоте их встречаемости в частотном словаре (Ляшевская, Шаров 2009) по формуле (Volansky et al. 2013):

$MWR = \frac{\sum_{i=1}^n R_i}{n}$, где R_i — ранг i -го слова, n — общее количество словарных токенов в тексте. Ранг слова равен его позиции в частотном списке. В анализе используются первые 6000 слов списка, максимальный ранг слов — 5000. Отсутствующим в списке словам присваивается ранг 6000.

Если гипотеза об *упрощении* верна, в переводе ожидается снижение показателей перечисленных выше индикаторов.

Перейдем к индикаторам *нормализации*, они включили (Volansky et al. 2013) следующие:

лексическая повторяемость (repetition) — отношение абсолютной частоты знаменательных слов, встречающихся более одного раза к общему количеству слов. Считается, что переводчики стремятся избегать лексических повторов;

поточечная взаимная информация (point-wise mutual information, PMI) биграмм слов, выражаяющая силу ассоциативной связи между двумя соседствующими словами (Church, Hanks 1990):

$PMI(w_1, w_2) = \log\left(\frac{P(w_1 w_2)}{P(w_1)P(w_2)}\right)$, где $P(w_1 w_2)$ — вероятность совместного появления слов, $P(w_1)$ и $P(w_2)$ — вероятности их появления по отдельности. Показатель рассчитывается как доля биграмм с $PMI > 0$. Ожидается, что в переводах доля таких биграмм выше.

Для характеристики параметра *экспликация* рассчитывались следующие индикаторы (Volansky et al. 2013):

эксплицитное называние (explicit naming), рассчитываемое как соотношение числа личных/притяжательных местоимений и общему числу именованных сущностей. Предполагается, что в переводе показатель выше, так как местоимения обеспечивают более точную идентификацию именованных сущностей;

средняя длина именованных сущностей (mean named entities length) в токенах. В переводе ожидается увеличение данного показателя, поскольку полные версии имен повышают вероятность идентификации референта;

дискурсивные маркеры (discourse markers) научного текста (Koppel, Ordan 2011; Kunilovskaya, Corpas Pastor 2021; Volansky et al. 2013). Для их анализа подсчитываются нормализованные частоты двух групп маркеров (Когут 2016):

1) маркеры когезии, связывающие между собой предложения и их части (маркеры введения новой информации, примеров, пояснения, противопоставления);

2) маркеры когеренции, связывающие разные фрагменты текста в единое целое (маркеры введения в тему, порядка следования информации, внутритекстовой референции). Ожидается увеличение частоты использования маркеров обеих групп в переводе.

Наконец, индикаторы *интерференции* включают (Kunilovskaya, Corpas Pastor 2021; Volansky et al. 2013):

частеречные n-граммы (POS n-grams), а именно нормализованные частоты частеречных би- и триграмм. Для тегирования использовались библиотеки ruTagger2 и Natasha (Библиотека для обработки... 2023; Морфологический анализатор... 2020);

позиционная частота токенов (positional token frequency) — нормализованные частоты токенов на первой, второй, третьей с конца и предпоследней позициях в предложении.

Дополнительно анализировались показатели *глаголов в страдательном залоге* (passive verbs ratio), а именно отношение их числа к общему числу глаголов в тексте (Volansky et al. 2013).

Везде при вычислении нормализованных частот показателей производилось деление абсолютного числа вхождений подсчитывающей единицы на общее количество соответствующих единиц в тексте, полученные значения выражались в процентах.

При подсчете средних показателей индикаторов для подкорпуса использовалась формула взвешенного арифметического среднего — метод усреднения, учитывающий объем (вес) текстов в токенах:

$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$, где \bar{x} — взвешенное среднее арифметическое, x_i — значение показателя для i -го текста, w_i — вес i -го текста, n — количество текстов в корпусе.

Результаты исследования

Рассмотрим результаты анализа характеристики *упрощения* (табл. 1). Здесь и далее в таблицах непереводные тексты обозначаются как НТ, переводные — как ПТ.

Средние показатели индикаторов упрощения

Table 1

Average values of simplification indicators in translated vs. non-translated texts

Показатель	НТ	ПТ	Разница
Lexical density	61,7 %	59,33 %	2,37 %
TTR	27,33 %	24,02 %	3,31 %
Log TTR	84 %	82,3 %	1,7 %
Modified TTR	1731	1625	106
Mean word length			
(1) characters	6,93	6,76	0,17
(2) syllables	3	2,9	0,1
Mean sentence length			
(1) tokens	26,3	29,04	2,74
(2) characters	154,26	167,75	13,49
Mean word rank	2525,55	2225,77	299,78

Показатели всех индикаторов, кроме *средней длины предложений*, подтверждают действие характеристики *упрощение* в переводе. Снижение лексической плотности говорит о снижении доли знаменательных и увеличении доли служебных частей речи в переводах.

Все три версии коэффициента лексической вариативности оказались существенно ниже для переводных текстов. Причем наиболее примечательно снижение модифицированной версии коэффициента (*Modified TTR*), которое говорит о снижении частоты использования лексических единиц, встречающихся в тексте оригинала единожды. Вероятно, в целях облегчения понимания текста читателем переводчики или опускают такие лексемы вообще, или отдают предпочтение уже использованным лексемам.

Особого внимания в контексте перевода медицинских статей требует анализ среднего ранга слов. Для медицинских текстов в целом характерно использование специализированной терминологии, что закономерно обусловливает высокий показатель среднего ранга слов. Тем не менее полученные результаты демонстрируют существенное снижение этого показателя в переводном подкорпусе, что свидетельствует о тенденции

к использованию более распространенных, общеупотребительных лексических единиц вместо более редких и специфичных вариантов. Данное наблюдение особенно критично для медицинской литературы — медицинская терминология представляет собой сложно организованную систему, где каждый термин имеет устоявшиеся формы выражения в ПЯ. Замена специализированных (и, соответственно, более редких) терминов на более общеупотребительные варианты может приводить к снижению информационной насыщенности и «размытию» смысла текста. Например, в клинической практике общепринятыми являются термины *гипотиреоз* или *миопия*, в то время как их более общеупотребительные эквиваленты *снижение функции щитовидной железы* и *близорукость* не обладают необходимой терминологической точностью, хотя и могут на практике заменять закрепленные в медицинской литературе термины.

Более низкий показатель средней длины предложений в переводе противоречит ожидаемому проявлению характеристики *упрощение*, подразумевающей разбиение сложных предложений на более простые и короткие. Вероятно, такой результат объясняется влиянием *экспликации*, при которой пере-

водчики стремятся сделать имплицитно заложенную информацию текста оригинала более явной, что проявляется в добавлении уточнений, пояснений, требующих введения дополнительных связующих и структурирующих текст элементов (союзов, союзных слов, дискурсивных маркеров). Учитывая, что точность передачи информации в медицинских текстах особенно важна, повышенный уровень экспликации является вполне ожидаемым явлением.

Сопоставление показателей индикаторов нормализации не подтвердило ее действия в переводе. *Лексическая повторяемость* была даже несколько выше в переводном подкорпусе (43,4% в непереводных текстах против 43,83% в переводах) что, однако, может свидетельствовать о стремлении переводчиков к повышению связности текста — повторное употребление лексических единиц способствует созданию более прочных связей между фрагментами текста. Доля биграмм с РМІ > 0 оказалась практически идентичной для непереводных (97%) и переводных (96%) текстов.

Перейдем к результатам анализа экспликации (табл. 2).

Более высокий показатель *эксплицитного называния* для непереводных текстов, вопреки ожиданиям, демонстрирует, что в переводах соотношение личных/притяжательных местоимений к именованным сущно-

стям оказалось ниже. Анализ именованных сущностей в корпусе показал, что в медицинских статьях они используются преимущественно для именования исследований и клинических испытаний, международных организаций, профессиональных ассоциаций, стран, номенов (названия лекарственных препаратов, заболеваний, синдромов), а также исследователей, результаты работы которых упоминаются в тексте или к авторитету которых обращается автор. Снижение показателя *эксплицитного называния* может быть обусловлено особой осторожностью переводчиков — использование местоимений вместо именованных сущностей может приводить к референциальной неоднозначности. В связи с этим более выгодной стратегией будет повторить сущность, нежели создавать ситуацию, в которой читатель может неверно интерпретировать, к чему именно отсылает местоимение.

Показатель *средней длины именованных сущностей* оказался практически идентичным между подкорпусами.

Рассмотрим показатели дискурсивных маркеров, выявивших наибольшие отличия. Нормализованные частоты маркеров *введения новой информации* (в дополнение к, включая, кроме того, более того), *противопоставления* (однако, с одной/другой стороны, тем не менее) и *введения примеров* (например, такие как, взять хотя бы) оказались выше

Средние показатели индикаторов экспликации

Average values of explication indicators in translated vs. non-translated texts

Показатель	НТ	ПТ	Разница
Explicit naming	114,3 %	100 %	14,3 %
Mean named entity length	1,62	1,6	0,02
Маркеры введения новой информации	0,6%	0,7 %	0,1 %
Маркеры противопоставления	0,664 %	0,7 %	0,036 %
Маркеры порядка следования информации	0,43 %	0,37 %	0,06 %
Маркеры введения примеров	0,08 %	0,12 %	0,04 %

для переводов, что свидетельствует об экспликации. В то же время увеличение маркеров противопоставления может говорить и об интерференции английского языка, где предложения нередко начинаются с вводных конструкций, выражающих противоречие со сказанным ранее (*however, nevertheless, on the contrary*). Также анализ показал снижение частоты использования маркеров порядка следования информации (*во-первых, во-вторых, прежде всего, перейдем к*), что может снижать связность текста как единого целого.

Наконец, рассмотрим показатели индикаторов интерференции. Частеречные *n*-граммы, продемонстрировавшие наибольшие различия, приведены в таблице 3. В анализе использовались обозначения тегов библио-

теки ruymorphy2 (Частеречные теги библиотеки ruymorphy2 2020), а для более детального анализа были введены специальные теги*.

Увеличение частоты комбинаций модальных слов с инфинитивом (типа *можно объяснить, необходимо отнести, нельзя исключать*) может быть результатом интерференции английского языка, для которого

* Дополнительные теги для частеречного анализа:
MDL_WRD — модальное слово (*нужно, можно, нельзя, необходимо, следует* и т. д.),

CMP_MOD — наречия для образования аналитических форм ср. степени (*более, менее*),

PRN_CONJ_W — местоимения-союзные слова (*что, который, чей, какой* и т. д.),

DMSTR_PRN — указательное местоимение (*этот, том, такой* и т. д.).

Таблица 3

Нормализованные частоты частеречных *n*-грамм

Table 3

Normalized frequencies of part-of-speech *n*-grams distinguishing translated and non-translated texts

	N-грамма	НТ (%)	ПТ (%)	Разница
Интерференция	MDL_WRD, INFN	0,12	0,19	0,07
	COMMA, PRTF NOUN, COMMA, PRTF	0,67 0,60	0,85 0,78	0,18 0,18
	CMP_MOD, ADJF	0,12	0,19	0,07
	PREP, NOUN, NOUN NOUN, PREP, NOUN PREP, NOUN, PREP NOUN, NOUN, PREP	2,30 3,34 0,90 1,59	2,50 3,51 1,07 1,70	0,20 0,17 0,17 0,11
	COMMA, SCONJ COMMA, PRN_CONJ_W SCONJ, NOUN PRN_CONJ_W, VERB SCONJ, PREP NOUN, COMMA, PRN_CONJ_W NOUN, COMMA, SCONJ COMMA, PRN_CONJ_W, VERB VERB, COMMA, SCONJ COMMA, SCONJ, NOUN	0,62 0,13 0,34 0,10 0,20 0,12 0,20 0,06 0,1 0,24	0,9 0,30 0,43 0,19 0,27 0,27 0,30 0,14 0,16 0,30	0,28 0,17 0,09 0,09 0,07 0,15 0,10 0,08 0,06 0,06
	DMSTR_PRN, NOUN PREP, DMSTR_PRN	0,51 0,38	0,68 0,51	0,17 0,13

характерно более аккуратное выражение эпистемической возможности или необходимости (Хомутова 2010), в то время как русскоязычный научный стиль более утвердителен — ему в большей мере свойственны констатирующие высказывания без избыточной модальности. При этом избыточное использование модальных конструкций может создавать впечатление недостаточной уверенности автора в своих высказываниях, что не такично для русского научного стиля.

В переводе значительно чаще встречаются конструкции с полным причастием (*PRTF*) после запятой, что является случаем *интерференции* — в английском языке причастия с зависимыми словами (типа *included in the meta-analysis*) не могут разрывать определяемую ими именную группу, т. е. стоять в пропозиции, и обязательно размещаются после определяемого существительного (например, *The quality of the studies included in the meta-analysis was...*, но не *The quality of the included in the meta-analysis studies was...*). На русский язык это можно перевести двумя способами — как *Качество исследований, включенных в мета-анализ,...* и, что более компактно, как *Качество включенных в мета-анализ исследований....* Тем не менее, следуя за структурой оригинала, переводчики чаще сохраняют ее в переводе, даже в случае, когда это может быть излишним.

Более частое употребление комбинации наречия, используемого для формирования аналитической формы сравнительной степени, с прилагательным является следствием *интерференции* английских аналитических сравнительных конструкций с многосложными прилагательными — в английском широко используются конструкции типа *more effective, more sensitive*, что в переводе чаще передается как *более эффективный, более чувствительный*, в то время как в некоторых контекстах можно было бы предпочесть варианты *эффективнее, чувствительнее*.

Также в переводе значительно чаще используются разнообразные комбинации частей речи с предлогами. Это является пря-

мым следствием *интерференции* английского языка, в котором предлоги выполняют ключевую структурообразующую функцию из-за отсутствия развитой падежной системы, коей обладает русский язык. Переводчики часто сохраняют предложные конструкции оригинала даже в случае, когда в русском языке можно найти более компактные синтетические альтернативы. Вследствие этого конструкции типа *changes in parameters, difficulties in detection, system for data output, relationship between effects* переводятся как *изменения в параметрах, трудности в выявлении, система для вывода данных, связь между эффектами*, хотя приемлемыми (и иногда предпочтительными) вариантами являются *изменения параметров, трудности выявления, система вывода данных, связь эффектов*.

Повышение частоты использования конструкций с подчинительными союзами и союзовыми словами-местоимениями свидетельствует об *экспликации* — переводчики вводят придаточные клаузы, раскрывающие имплицитно заложенную в исходном тексте информацию и выражающие причинно-следственные, временные и другие логические отношения более четко. Другое свидетельство *экспликации* — повышение частоты использования комбинаций с указательными местоимениями — с целью облегчения идентификации имен существительных, к которым они отсылают.

Рассмотрим позиционную частоту токенов в предложениях. В таблице 4 приведены данные для токенов с частотой не менее 0,5%.

В непереводном подкорпусе на первой позиции чаще используются предлоги. Это связано с тем, что в русском языке предложные конструкции на начальной позиции — типичный способ выражения обстоятельственных значений (места, времени, причины и т. д.), особенно когда обстоятельство является известной информацией (темой высказывания), что к тому же допустимо благодаря относительно свободному порядку слов русского языка. В английском языке, напротив, наблюдается синтаксическая

Таблица 4

Нормализованные частоты токенов на разных позициях в предложении (по убыванию)

Table 4

Normalized token frequencies at different sentence positions (presented in descending order)

Непереводные тексты (%)			
Первая	Вторая	Третья с конца	Предпоследняя
Предлоги (28,74)	Им. сущ. (6,92)	Предлоги (5,08)	Пунктуация (12,6)
В (12,07), При (4,58), Для (1,92), У (1,88), На (1,46), По (1,39), Кроме (1,2), С (0,93), Среди (0,85), К (0,79), Согласно (0,65), После (0,52), Несмотря (0,5)	образом (1,82), пациентов (0,83), данные (0,76), данным (0,72), исследования (0,65), исследование (0,62), связи (0,51), больных (0,51), результаты (0,5)	в (2,14), с (1,32), на (0,6), при (0,52), до (0,5)) (12,6)
Местоимения (3,79)	Пунктуация (5,2)	Союзы (4,12)	Им. сущ. (4,955)
Таким (1,83), Это (0,94), Эти (0,5), Все (0,52)	, (5,2)	и (3,52), или (0,6)	заболевания (0,96), больных (0,67), пациентов (0,64), беременности (0,56), лет (0,56), лечения (0,545), исследования (0,52), пищевода (0,5)
Союзы (3,24)	Предлоги (3,31)	Пунктуация (3,33)	
Однако (1,87), Также (0,71), Если (0,66)	в (1,25), с (1,1), на (0,96)	, (2,56),) (0,77)	
Наречия (0,93)		Им. числит. (2,34)	
Так (0,93)		2 (0,65), 3 (0,59), 0,005 (0,54), 0,001(0,56)	
Им. сущ. (0,8)	Местоимения (3,04)	Частицы (1,86)	
Результаты (0,8)	того (1,28), этом (1,21), то (0,55)	не (1,86)	
Глаголы (0,56)	Союзы (0,7)	Им. сущ. (1)	
Следует (0,56)	и (0,7)	мм (1)	
	Им. прил. (0,67)		
	настоящее (0,67)		
Переводы (%)			
Первая	Вторая	Третья с конца	Предпоследняя
Предлоги (25,3)	Им. сущ. (7,11)	Предлоги (5,62)	Пунктуация (13,04)
В (11,97), При (2,6), Кроме (2,34), У (2,15), Для (1,78), По (1,34), На (1,1), С (0,77), После (0,67), Несмотря (0,58)	образом (1,2), больных (1), данные (0,94), исследование (0,9), исследования (0,87), пациентов (0,86), исследовании (0,72), результаты (0,62)	в (2,3), с (1,23), на (0,96), к (0,6), без (0,53)) (13,04)

Табл. 4. Продолжение

Непереводные тексты (%)			
Первая	Вторая	Третья с конца	Предпоследняя
Местоимения (7,25)	Пунктуация (5,6)	Союзы (4,01)	Им. сущ. (7,82)
Это (1,48), Тем (1,2), Таким (1,13), Эти (1,12), Все (0,98), Мы (0,84), Этот (0,5)	, (5,6)	и (3,2), или (0,81)	терапии (1,37), исследования (1), пациентов (0,94), лечения (0,86), плацебо (0,83), заболевания (0,7), ИБС (0,54), ЛПНП (0,54), жизни (0,52), исследований (0,52)
Союзы (4,41)	Местоимения (3,85)	Им. числит. (2,33)	
Однако (2,06), Поэтому (0,87), Если (0,78), Хотя (0,7)	того (3,05), этом (0,8)	2 (0,65), 3 (0,63), 1(0,55), 0,001 (0,5)	
Им. сущ. (3,19)	Предлоги (3,52)	Им. сущ. (2,28)	
Результаты (0,84), Исследование (0,73), Анализ (0,57), Данные (0,53), Лечение (0,52)	в (1,6), с (1), на (0,92)	лечения (0,62), ХС (0,6), мм (0,53), таблица (0,53)	
Наречия (1,71)			
Как (0,9), Более (0,81)			
Вводные слова (0,53)	Частицы (1,53)	Пунктуация (1,5)	
Например (0,53)	не (1,53)	, (1,5)	
	Союзы (0,92)	Частицы (1,3)	
	и (0,92)	не (1,3)	

ригидность с преобладанием структуры SVO, а обстоятельственные группы преимущественно локализуются в конце предложения даже в случае, если эта информация является известной. Если же предложение начинается с предложной группы, оно обычно маркировано и требует либо инверсии, либо обособления. В результате этого различия переводчики могут воспроизводить характерный для английского порядок слов, что также подтверждается более частым использованием в переводе предлогов с существительными в конце предложения (третья с конца и предпоследняя позиции).

Также на первых двух позициях в переводе чаще используются токены, в сочетании составляющие вводные конструкции *таким образом и кроме того*, а также вводное слово *например* на первой позиции. Дополнительный анализ вводных слов на первой позиции (n-грамма *S START, PARENTH, COMMA*)

также подтвердил это — в переводе их частота составляет 0,08 % против 0,05 % для непереводных текстов. Этот результат согласуется с приведенным ранее анализом дискурсивных маркеров.

Более частое употребление местоимений на первой позиции в переводе может быть результатом *экспликации* — они делают референциальные связи в тексте более явными. В то же время гораздо более часто употребление как местоимений, так и существительных на первой позиции может говорить об *интерференции*, так как английский язык требует явного указания подлежащего, преимущественно выражаемого местоимением или существительным. Увеличение частоты употребления местоимения *Мы* связано с существующей в англоязычных научных статьях тенденции к активному стилю изложения, в том числе от первого лица (например, *We investigated..., We found..., We conducted...*).

Рост частоты союзов *Однако, Поэтому, Хотя* на первой позиции является следствием *интерференции*, т. к. в английском в начале предложения нередко используются слова типа *However, Therefore, Thus, Although*.

Наконец, анализ глаголов в форме страдательного залога показал, что их доля в переводе составляет 36,6% — это превышает аналогичный показатель 33,8% для непереводных текстов на 2,8%. Более того, анализ кратких причастий (PRTS) показал, что их доля в переводе выше (1,04% против 0,96%); также существенную разницу продемонстрировала n-грамма, состоящая из вспомогательного глагола и краткого причастия (AUX, PRTS) — 0,44% для переводных текстов против 0,26% для непереводных. Достаточно высокая доля страдательных конструкций в переводах с английского языка, казалось бы, противоречит ожиданиям, так как в современных англоязычных журналах существует тенденция к более активному стилю изложения, что нередко прописывается в требованиях к оформлению статей (Plotnick 2025). Тем не менее использование страдательного залога остается широко распространенной практикой в англоязычных медицинских публикациях (Millar et al. 2013), особенно если речь идет о таких разделах статьи, как «Методы» и «Результаты исследования». Таким образом, увеличение доли глаголов в страдательном залоге может быть вызвано как *интерференцией* английского языка, так и стремлением переводчиков к более нормализованному — безличному стилю изложения, в большей мерециальному научному стилю, в котором страдательные формы позволяют сместить фокус с автора на объект исследования, тем самым подчеркивая нейтральность и беспристрастность повествования.

Заключение

Проведенный анализ позволил выявить ряд черт, свойственных переводческому языку в англо-русском переводе медицинских научных статей.

Упрощение. Переводам в значительной мере свойственно снижение лексического разнообразия, лексической плотности и среднего ранга слов. Последний показатель особенно важен для медицинских текстов, т. к. если при переводе предпочтение отдается более частотной и общеупотребительной лексике, это может приводить к снижению терминологической точности текста.

Нормализация. Анализ данной характеристики не выявил ее действия в переводе — лексическая повторяемость в переводах оказалась несколько выше, а процент биграмм с $PMI > 0$ был практически идентичным для обоих подкорпусов.

Экспликация. Анализ этой характеристики показал неоднозначные результаты. С одной стороны, показатель *эксплицитного называния* оказался на 14,3% ниже в переводных текстах, что, вероятно, отражает осознанную стратегию переводчиков избегать референциальной неоднозначности при работе с медицинскими текстами — предпочтение отдается не местоимениям, а повторному использованию именованных сущностей, что оправдано необходимостью максимальной точности текста при упоминании исследований, лекарственных средств и т. д. С другой стороны, повышенная частота маркеров противопоставления, введения новой информации и примеров подтвердила тенденцию к *экспликации логических связей* между фрагментами текста. Особенно показателен анализ частеречных n-грамм, выявивший значительное увеличение частоты использования комбинаций с подчинительными союзами и союзными словами — эти конструкции используются для введения придаточных предложений, выражающих логические связи и отношения между фрагментами текста более явно.

Интерференция проявилась наиболее ярко. Частеречный анализ n-грамм продемонстрировал перенос некоторых черт, свойственных английскому языку: более частое использование модальных конструкций с инфинитивом, причастных оборотов, аналитических форм сравнительной степени при-

лагательных и предложных конструкций. Анализ позиционной частоты токенов подтвердил *синтаксическую интерференцию* — в переводах значительно чаще прослеживается характерный для английского языка порядок слов, когда в силу требований английской грамматики на первой позиции обязательно стоит подлежащее, выраженное существительным или местоимением.

Также в переводах доля глаголов в страдательном залоге превышает на 2,8% аналогичный показатель в непереводных текстах. Это явление можно интерпретировать двояко: либо как результат интерференции английского научного стиля, характерного для медицинских статей, либо как проявление *нормализации* — сознательной адаптации активных конструкций оригинала к нормам русскоязычного научного дискурса, где страдательный залог традиционно используется для придачи тексту объективности и отстраненности.

На основе выявленных особенностей можно предварительно сформулировать некоторые рекомендации, которые важно учитывать при переводе/постредактировании медицинских статей:

1. Контроль *терминологической точности*: с целью обеспечения точности интерпретации текста необходимо следить за использованием специализированной терминологии — быть осторожным при замене терминов на общеупотребительные варианты, даже если они кажутся более простыми и понятными для читателя.
2. Контроль *синтаксической интерференции*: там, где это представляется возможным, следует отдавать предпочтение синтетическим формам сравнительной степени прилагательных, заменять избыточные пред-

ложные конструкции беспредложными, чаще использовать безличные конструкции, избегать чрезмерного употребления местоимений, учитывать различия в актуальном членении предложения в русском и английском языках, обращая особое внимание на предложные обстоятельственные конструкции, которые, будучи уже известной для читателя информацией (темой), в русском языке обычно ставятся на первую позицию в предложении.

3. Контроль *модальности*: следует ограничивать избыточное использование модальных конструкций, свойственное английскому научному стилю, если они не несут значимой смысловой нагрузки.
4. Контроль *экспликации*: необходимо ограничить использование придаточных предложений для экспликации логических связей, чтобы избежать перегруженности текста. Чаще использовать маркеры порядка следования информации для улучшения связности текста.
5. Контроль *глагольного залога*: следует использовать страдательный залог избирательно, с учетом конвенций русскоязычного научного стиля, не калькируя механически англоязычные страдательные конструкции.

Необходимо подчеркнуть, что эти рекомендации несут предварительный характер. Перспективным направлением дальнейших исследований представляется установление связи между выявленными индикаторами *translationese* и экспертной оценкой качества переводов, чтобы определить, какие из индикаторов действительно негативно влияют на восприятие качества перевода, а какие — нет.

ИСТОЧНИКИ

- Анализатор феномена translationese*. (2025) [Электронный ресурс]. URL: <https://pypi.org/project/translationese-analyzer/> (дата обращения 15.04.2025).
- Библиотека для обработки естественного русского языка Natasha*. (2023) [Электронный ресурс]. URL: <https://pypi.org/project/natasha/> (дата обращения 15.04.2025).
- Морфологический анализатор русского языка pymorphy2*. (2020) [Электронный ресурс]. URL: <https://pymorphy2.readthedocs.io/en/stable/> (дата обращения 15.04.2025).

Портал научных журналов ELPUB. (2025) [Электронный ресурс]. URL: <https://elpub.ru/elpub-journals> (дата обращения 15.04.2025).

Частеречные теги библиотеки *rutmorphy2* (2020) Документация по грамматическим тегам. [Электронный ресурс]. URL: <https://rutmorphy2.readthedocs.io/en/stable/user/grammemes.html> (дата обращения 15.04.2025).

СЛОВАРИ

Ляшевская, О. Н., Шаров, С. А. (2009) *Новый частотный словарь русской лексики*. М.: Азбуковник. [Электронный ресурс]. URL: dict.ruslang.ru/freq.php (дата обращения 15.04.25).

СПИСОК ЛИТЕРАТУРЫ

Беляева, Л. Н., Камшилова, О. Н., Щубина, Н. Л. (2023) *Научная статья в технологическом пространстве машинного перевода: правила и процедуры редактирования*. СПб.: Книжный дом, 90 с.

Когут, С. В. (2016) Дискурсивные маркеры в письменном научном дискурсе. *Сибирский филологический журнал*, № 2, с. 157–163.

Краснопеева, Е. С. (2015) *Лексические особенности русскоязычного переводного дискурса: корпусное сравнительно-сопоставительное исследование на материале современной художественной прозы*. Диссертация на соискание степени кандидата филологических наук. Челябинск, Челябинский государственный университет, 211 с.

Серова, О. У. (2025) Автоматизированное исследование индикаторов переводческого языка на материале переводов научных статей. *Terra Linguistica*, т. 16, № 1, с. 61–81. <https://doi.org/10.18721/JHSS.16105>

Хомутова, Т. Н. (2010) Научный текст: интегральный анализ лексики. *Язык и культура*, № 4 (12), с. 77–98.

Baker, M. (1993) Corpus linguistics and translation studies: Implications and applications. In: *Text and technology: In honour of John Sinclair*. Amsterdam: John Benjamins Publ., pp. 233–252. <https://doi.org/10.1075/z.64.15bak>

Blum-Kulka, S. (1986) Shifts of cohesion and coherence in translation. In: J. House, S. Blum-Kulka (eds.). *Interlingual and intercultural communication: Discourse and cognition in translation and second language acquisition studies*. Tübingen: Narr Publ., pp. 17–35.

Church, K. W., Hanks, P. (1990) Word association norms, mutual information, and lexicography. *Computational Linguistics*, vol. 16, no. 1, pp. 22–29.

Gellerstam, M. (1986) Translationese in Swedish novels translated from English. In: L. Wollin, H. Lindquist (eds.). *Translation studies in Scandinavia: Poceedings from the Scandinavian symposium on translation theory (SSOTT) II*. Lund: CWK Gleerup Publ., pp. 88–95.

Koppel, M., Ordan, N. (2011) Translationese and its dialects. In: *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*. Portland: Association for Computational Linguistics Publ., pp. 1318–1326.

Kunilovskaya, M. (2023) *Translationese indicators for human translation quality estimation (based on English-to-Russian translation of mass-media texts)*. Extended abstract of the PhD dissertation (Science and Engineering). Wolverhampton, University of Wolverhampton. [Online]. Available at: <http://hdl.handle.net/2436/625250> (accessed 15.04.2025).

Kunilovskaya, M., Corpas Pastor, G. (2021) Translationese and register variation in English-to-Russian professional translation. In: *New Perspectives on Corpus Translation Studies*. Singapore: Springer Publ., pp. 133–180. https://doi.org/10.1007/978-981-16-4918-9_6

Millar, N., Budgell, B., Fuller, K. (2013) “Use the active voice whenever possible”: The impact of style guidelines in medical journals. *Applied Linguistics*, vol. 34, no. 4, pp. 393–414. <https://doi.org/10.1093/applin/ams059>

Plotnick, J. (2025) *How to use active voice in the sciences*. [Online]. Available at: <https://www.uc.utoronto.ca/active-voice-in-science> (accessed 16.04.2025).

Toury, G. (1979) Interlanguage and its manifestations in translation. *Meta: Journal des Traducteurs*, vol. 24, no. 2, pp. 223–231. <https://doi.org/10.7202/004502ar>

Volansky, V., Ordan, N., Wintner, Sh. (2013) On the features of translationese. *Digital Scholarship in the Humanities*, vol. 30, no. 1, pp. 98–118. <https://doi.org/10.1093/llc/fqt031>

SOURCES

Analizator fenomena translationese [Automated translationese analyzer]. (2025) [Online]. Available at: <https://pypi.org/project/translationese-analyzer/> (accessed 15.04.2025). (In English)

Biblioteka dlya obrabotki estestvennogo russkogo yazyka Natasha [Library for natural Russian language processing Natasha]. (2023) [Online]. Available at: <https://pypi.org/project/natasha/> (accessed 15.04.2025). (In English)

Chasterechnye tegi biblioteki pymorphy2 [Part-of-speech tags of pymorphy2 library]. (2020) Dokumentatsiya po grammaticheskim tegam — Documentation on grammatical tags. [Online]. Available at: <https://pymorphy2.readthedocs.io/en/stable/user/grammemes.html> (accessed 15.04.2025). (In Russian)

Morfologicheskij analizator russkogo yazyka pymorphy2 [Russian language morphological analyzer pymorphy2]. (2020) [Online]. Available at: <https://pymorphy2.readthedocs.io/en/stable/> (accessed 15.04.2025). (In Russian)

Portal nauchnykh izdanij ELPUB [Catalog of scientific journals ELPUB]. (2025) [Online]. Available at: <https://elpub.ru/elpub-journals> (accessed 15.04.2025). (In Russian).

DICTIONARIES

Lyashevskaya, O. N., Sharov, S. A. (2009) *Novyj chastotnyj slovar' russkoj leksiki [New frequency dictionary of modern Russian]*. Moscow: Azbukovnik Publ. [Online]. Available at: dict.ruslang.ru/freq.php (accessed 15.04.25). (In Russian).

REFERENCES

Baker, M. (1993) Corpus linguistics and translation studies: Implications and applications. In: *Text and technology: In honour of John Sinclair*. Amsterdam: John Benjamins Publ., pp. 233–252. <https://doi.org/10.1075/z.64.15bak> (In English)

Belyaeva, L. N., Kamshilova, O. N., Shubina, N. L. (2023) *Nauchnaya stat'ya v tekhnologicheskem prostранстве машинного перевода: правила и процедуры редактирования* [Scientific article in the technological space of machine translation: Editing rules and procedures]. Saint Petersburg: Knizhnyj dom Publ., 90 p. (In Russian)

Blum-Kulka, S. (1986) Shifts of cohesion and coherence in translation. In: J. House, S. Blum-Kulka (eds.). *Interlingual and intercultural communication: Discourse and cognition in translation and second language acquisition studies*. Tübingen: Narr Publ., pp. 17–35. (In English)

Church, K. W., Hanks, P. (1990) Word association norms, mutual information, and lexicography. *Computational Linguistics*, vol. 16, no. 1, pp. 22–29. (In English)

Gellerstam, M. (1986) Translationese in Swedish novels translated from English. In: L. Wollin, H. Lindquist (eds.). *Translation studies in Scandinavia: Poceedings from the Scandinavian symposium on translation theory (SSOTT) II*. Lund: CWK Gleerup Publ., pp. 88–95. (In English)

Khomutova, T. N. (2010) Nauchnyj tekst: integral'nyj analiz leksiki [Research text: integral analysis of lexis]. *Yazyk i kul'tura — Language and Culture*, no. 4 (12), pp. 77–98. (In Russian)

Kogut, S. V. (2016) Diskursivnye markery v pis'mennom nauchnom diskurse [Discourse markers in written scientific discourse]. *Sibirskij filologicheskij zhurnal — Siberian Journal of Philology*, no. 2, pp. 157–163. (In Russian)

Koppel, M., Ordan, N. (2011) Translationese and its dialects. In: *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*. Portland: Association for Computational Linguistics Publ., pp. 1318–1326. (In English)

Krasnopoeva, E. S. (2015) *Leksicheskie osobennosti russkoyazychnogo perevodnogo diskursa: korpusnoe srovnitel'no-sopostavitel'noe issledovanie na materiale sovremennoj khudozhestvennoj prozy* [Lexical features of Russian-language translated discourse: A Corpus-based comparative-contrastive study based on contemporary fiction]. PhD dissertation (Philology). Chelyabinsk, Chelyabinsk State University, 211 p. (In Russian)

Kunilovskaya, M. (2023) *Translationese indicators for human translation quality estimation (based on English-to-Russian translation of mass-media texts). Extended abstract of the PhD dissertation (Science and Engineering)*. Wolverhampton, University of Wolverhampton. [Online]. Available at: <http://hdl.handle.net/2436/625250> (accessed 15.04.2025). (In English)

- Kunilovskaya, M., Corpas Pastor, G. (2021) Translationese and register variation in English-to-Russian professional translation. In: *New Perspectives on Corpus Translation Studies*. Singapore: Springer Publ., pp. 133–180. https://doi.org/10.1007/978-981-16-4918-9_6 (In English)
- Millar, N., Budgell, B., Fuller, K. (2013) “Use the active voice whenever possible”: The impact of style guidelines in medical journals. *Applied Linguistics*, vol. 34, no. 4, pp. 393–414. <https://doi.org/10.1093/applin/ams059> (In English)
- Plotnick, J. (2025) *How to use active voice in the sciences*. [Online]. Available at: <https://www.uc.utoronto.ca/active-voice-in-science> (accessed 16.04.2025). (In English)
- Serova, O. U. (2025) Avtomatizirovannoe issledovanie indikatorov perevodcheskogo yazyka na materiale perevodov nauchnykh statej [An automated study of translationese indicators in scientific paper translations]. *Terra Linguistica*, vol. 16, no. 1, pp. 61–81. <https://doi.org/10.18721/JHSS.16105> (In Russian)
- Toury, G. (1979) Interlanguage and its manifestations in translation. *Meta: Journal des Traducteurs — Meta: Translators' Journal*, vol. 24, no. 2, pp. 223–231. <https://doi.org/10.7202/004502ar> (In English)
- Volansky, V., Ordan, N., Wintner, Sh. (2013) On the features of translationese. *Digital Scholarship in the Humanities*, vol. 30, no. 1, pp. 98–118. <https://doi.org/10.1093/lrc/fqt031> (In English)

СВЕДЕНИЯ ОБ АВТОРЕ

СЕРОВА Олеся Усмановна — *Olesya U. Serova*

Российский государственный педагогический университет им. А. И. Герцена, Санкт-Петербург, Россия.

Herzen State Pedagogical University of Russia, Saint Petersburg, Russia.

ORCID: [0009-0009-4902-1738](https://orcid.org/0009-0009-4902-1738), e-mail: serovaolesyau@gmail.com

Аспирант кафедры образовательных технологий в филологии.

Научный руководитель: Беляева Лариса Николаевна, доктор филологических наук, профессор, профессор кафедры образовательных технологий в филологии РГПУ им. А. И. Герцена.

Поступила в редакцию: 19 июня 2025.

Прошла рецензирование: 30 июля 2025.

Принята к печати: 30 сентября 2025.