

©Лагутина Н. С., Лагутина К. В., Бойчук Е. И., Воронцова И. А., Парамонов И. В., 2019

DOI: 10.18255/1818-1015-2019-3-420-440

УДК 004.912

Автоматизированный поиск средств ритмизации художественного текста для сравнительного анализа оригинала и перевода на материале английского и русского языков

Лагутина Н. С., Лагутина К. В., Бойчук Е. И., Воронцова И. А., Парамонов И. В.

Поступила в редакцию 12 августа 2019

После доработки 11 сентября 2019

Принята к публикации 13 сентября 2019

Аннотация. Анализ функциональной эквивалентности перевода, основанный на достижении ритмической эквивалентности, представляет собой чрезвычайно важную задачу современной лингвистики. При этом ритмическая составляющая является неотъемлемой частью функциональной эквивалентности, которая не может быть достигнута без передачи ритмических характеристик текста. Для анализа ритмических средств в оригинальном тексте и переводе художественного произведения авторами был разработан программный инструмент ProseRhythmDetector, позволяющий находить и визуализировать лексические и синтаксические средства в англоязычных и русскоязычных прозаических текстах: анафору, эпифору, симплоку, анадиплозис, эпаналепсис, редупликацию, эпистрофу, многосоюзию и апозиопезу. Целью данной работы является представление результатов апробации ProseRhythmDetector на двух произведениях английских авторов и их переводах на русский язык: Ш. Бронте «Городок» (Ch. Bronte “Villette”) и А. Мердок «Черный принц» (I. Murdoch “The Black Prince”). На основе результатов работы инструмента авторы сопоставили ритмические характеристики в оригинале текста и его переводе и сравнили как аспекты ритма, так и их контексты. Данный эксперимент позволил выявить особенности передачи стиля автора художественного произведения переводчиком, обнаружить и объяснить случаи несовпадения ритмических средств оригинала и перевода. Применение программного инструмента ProseRhythmDetector позволило существенно сократить объем работы экспертов-лингвистов за счет автоматизированного выявления лексических и синтаксических средств с достаточно высокой точностью (от 62% до 93%) для различных ритмических средств.

Ключевые слова: ритм текста, анализ ритма, обработка естественного языка, ритмические средства, автоматизация

Для цитирования: Лагутина Н. С., Лагутина К. В., Бойчук Е. И., Воронцова И. А., Парамонов И. В., "Автоматизированный поиск средств ритмизации художественного текста для сравнительного анализа оригинала и перевода на материале английского и русского языков", *Моделирование и анализ информационных систем*, **26:3** (2019), 420–440.

Об авторах:

Лагутина Надежда Станиславовна, orcid.org/0000-0002-6137-8643, канд. физ.-мат. наук, доцент, Ярославский государственный университет им. П.Г. Демидова, ул. Советская, 14, г. Ярославль, 150003 Россия, e-mail: lagutinans@rambler.ru

Лагутина Ксения Владимировна, orcid.org/0000-0002-1742-3240, аспирант, Ярославский государственный университет им. П.Г. Демидова, ул. Советская, 14, г. Ярославль, 150003 Россия, e-mail: lagutinakv@mail.ru

Бойчук Елена Игоревна, orcid.org/0000-0001-6600-2971, доктор фил. наук, доцент, Ярославский государственный педагогический университет им. К.Д. Ушинского,

ул. Республиканская, д. 108/1, г. Ярославль, 150000 Россия, e-mail: elena-boychouk@rambler.ru

Воронцова Инна Алексеевна, orcid.org/0000-0001-5897-9299, канд. фил. наук, доцент,
Ярославский государственный педагогический университет им. К.Д. Ушинского,
ул. Республиканская, д. 108/1, г. Ярославль, 150000 Россия, e-mail: arinna1@yandex.ru

Парамонов Илья Вячеславович, orcid.org/0000-0003-3984-8423, канд. физ.-мат. наук, доцент,
Ярославский государственный университет им. П.Г. Демидова,
ул. Советская, 14, г. Ярославль, 150003 Россия, e-mail: Пуа.Paramonov@fruct.org

Благодарности:

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00243.

Введение

В современной лингвистике существует задача определения и анализа ритма художественных произведений. Под ритмом обычно понимается регулярное повторение сходных и соизмеримых единиц естественного языка, выполняющих структурирующую, текстообразующую и экспрессивно-эмоциональную функции. Ритмические средства прозы и стихов имеют самостоятельные характеристики и не только являются феноменом организации и восприятия художественного текста, но и отражают особенности стиля автора произведения [1].

Работы в области автоматического анализа ритма в основном посвящены исследованию стихотворных текстов и основаны на определении чередования ударных и безударных слогов и простых количественных характеристиках, таких как длина предложений и частота встречаемости слов. Однако средства ритмики включают в себя большое количество аспектов на фонетическом, лексическом, морфологическом и синтаксическом уровнях. В связи с этим интересной и трудной задачей становится создание методов и инструментов определения ритма, позволяющих выразить его в конкретных параметрах. По таким параметрам можно сравнивать ритм текстов, созданных разными писателями, разных частей текста, ритмы оригинального текста и перевода [2]. Лингвисты также отмечают важность сохранения авторского ритма при переводе произведения на другой язык и частое искажение этого ритма переводчиками прозаических текстов [3].

Особенностью исследований, посвященных ритму в художественном переводе, являются чрезвычайная комплексность самого объекта исследования — ритма прозы, сложность и многоаспектность феномена художественного перевода, а также ограниченность математического инструментария и отсутствие эффективных автоматизированных систем оценки параметров ритма оригинального и переведенного текстов. Тем не менее, математические методы применяются для оценки средств ритмизации в исследованиях [4].

Учитывая все вышесказанное, авторы статьи поставили перед собой задачу построить систему автоматизированного определения комплекса ритмических средств художественных произведений и применить её к сравнительному анализу оригинального текста на английском языке и русскоязычного перевода. Важной особенностью разрабатываемого инструмента является визуализация результатов поиска средств ритмики конкретного произведения с целью проведения экспертного анализа отдельных аспектов.

1. Рассматриваемые ритмические средства

В данной работе ритмические характеристики текста определяются, исходя из употребления средств создания ритма, в основе которых находится повтор в определенной конфигурации, в определенной позиции, с определенным количеством повторяющихся элементов. На основе этого выделяются следующие ритмические средства:

1. Анафора — скрепление речевых отрезков (частей фразы, стихов) с помощью повтора слова или словосочетания в начальной позиции.
2. Эпифора — скрепление речевых отрезков (частей фразы, стихов) с помощью повтора слова или словосочетания в конечной позиции.
3. Симплога — фигура синтаксического параллелизма в смежных стихах или фразах, у которых одинаковые начало и конец при разной середине или наоборот, разные начало и конец при одинаковой середине.
4. Анадиплосис — риторическая фигура, в которой следующее предложение начинается теми же словами, которыми оканчивается предыдущее.
5. Эпаналепсис — фигура речи, состоящая в повторении одного и того же слова или словосочетания с небольшими вариациями.
6. Редупликация — фигура речи, заключающаяся в удвоении одного и того же слова или выражения.
7. Эпистрофа — стилистическая фигура, заключающаяся в повторении одного и того же слова или выражения в длинной фразе или периоде.
8. Многосоюзиe — стилистическая фигура, состоящая в намеренном увеличении количества союзов в предложении, обычно для связи однородных членов.
9. Апозиопеза — умолчание, неоконченность мысли, внезапная остановка в устной или письменной речи, чаще прерывание предложения и начало нового.

Выбор данных средств для анализа ритма, а именно для их автоматизированного поиска и количественной обработки обусловлен тем, что это наиболее частотные ритмические средства, употребляемые в прозаических текстах, и именно они выделяются в качестве ритмических средств на лексико-грамматическом уровне большинством лингвистов, проводящих исследования в области ритмизации текста.

2. Обзор смежных работ

Определение ритмических свойств текста чаще всего основывается на фонетических средствах ритмики, опирающихся на структурный анализ произведения на уровне слогов, его метроритмические параметры, определение рифмы. На этом подходе основаны доступные инструменты SPARSAR¹ для английского языка [6], Metricalizer²

¹<https://sparsar.wordpress.com>

²<https://www.poetron-zone.de/metricalizer.php>

для немецкого [7], Rhymes³ и RitmInMe⁴ для русского. Все эти инструменты работают с поэтическими текстами.

Аналогичный подход часто применяют и к прозаическому тексту. Фонетический ритм английских, румынских и французских текстов анализируется в цикле работ [8, 9], на его основе определяется стиль: стихи, художественная литература и политическая речь. Аналогичная работа проделана для русскоязычных текстов [10]. Бойчук Е. И. и соавторы [11] анализируют ритм французской художественной литературы с целью определения авторского стиля писателя. Авторы работы [12] преобразуют текст в набор сигналов, разбивая его на буквы, а затем исследуют ритм текста, применяя математические методы разложения сигналов на функции (Empirical Mode Decomposition). Такой подход анализирует текст на уровне слогов, букв и звуков, реже слов, и удобен тем, что позволяет легко использовать аппарат математической статистики для расчета частоты, периодичности, корреляции и т. п. Однако для углубленного исследования ритма прозы требуется работа на более высоком смысловом уровне словосочетаний (синтагм), предложений, абзацев.

Исследователи в области компьютерной лингвистики говорят о необходимости изучения различных фигур речи естественного языка и обращают внимание на сложность формализации определения таких фигур и построения методов их автоматического поиска [13]. Авторы этой работы описывают процесс создания онтологии и базы данных метафор, риторических вопросов, противопоставлений, а также фигур речи, основанных на повторении слов и фраз. Они обращают внимание на то, что взаимодействие лингвистов и специалистов в области информационных технологий дает возможность получить глубокое понимание человеческого языка с точки зрения его структуры и функционирования.

К сожалению, на сегодняшний момент существует недостаточно работ, посвященных автоматизации поиска и применения лексических и синтаксических аспектов ритмики текста. Авторы работы [14] предлагают метод обнаружения анадипозиса, анафоры, антиметаболы (повторения во второй фразе слов первой, но в другом порядке и значении) и эпифоры, основанный на применении регулярных выражений. Эксперимент проведен на корпусе из более чем 163 тысяч фрагментов текстов, однако рассчитана только частота встречаемости каждого из рассмотренных средств ритмики, а точность и полнота результатов не оцениваются.

Российские ученые [15] сравнивают шесть программных систем обнаружения анафоры. Для автоматизации процесса поиска в системах использовались методы машинного обучения, морфологические и синтаксические анализаторы, онтологии и правила. В ходе экспериментов с общим корпусом из 85 текстов лучшие показатели F-меры были достигнуты с использованием лингвистических алгоритмов, основанных на правилах и онтологиях.

Подробно описано исследование методов поиска хиазмов (крестообразное изменение последовательности элементов в двух параллельных рядах слов), эпифор и эпанафор (повтор начальных и конечных слов в соседних фразах или предложениях) в работе [16]. Авторы отмечают сложность определения и формализации параметров текста и его частей для качественного поиска фигур речи. Поверхностный анализ структуры текста даёт слишком много «ложных» результатов, так как на-

³<http://rifmovnik.ru>

⁴<http://www.ritminme.ru>

ходит фрагменты текста, содержащие повторяющиеся элементы, но не соответствующие смыслу ритмического средства, что особенно характерно для хиазма. Проявления ритмических средств могут быть очень редкими, например, одна фигура на несколько сотен страниц текста. Для отбора кандидатов авторы выбрали линейную модель, используя которую разработали детекторы, настраиваемые вручную. После того, как была накоплена достаточно большая обучающая выборка, систему обучали, используя логистическую регрессию. В результате значение F-меры оказалось в большинстве случаев более 60%. Авторы позиционируют свой инструмент как помощник лингвиста, дополняющий традиционный ручной анализ в условиях огромного количества литературных произведений.

Таким образом, построение и исследование алгоритмов автоматического анализа текстов на уровне смысловых конструкций, таких как ритмические средства, является важной и малоисследованной задачей компьютерной лингвистики. Решение этой задачи включает в себя разработку инструментов для лингвистов. Эти инструменты необходимы для проведения экспериментов с корпусами текстов большого объёма, поскольку их ручная обработка не представляется возможной. В качестве результатов исследований в этом случае выступают как параметры оценки произведений, так и материалы для использования в методах машинного обучения.

3. Алгоритмы поиска ритмических средств

Для того, чтобы проанализировать ритм прозаического произведения и сравнить его с ритмом перевода, был разработан комплекс алгоритмов, автоматически находящих в тексте ритмические средства, а именно лексические и синтаксические.

Входными данными для всех алгоритмов является необработанный текст. Текст T состоит из n предложений $\langle s_1, s_2, \dots, s_n \rangle$, каждое предложение является набором слов $\langle w_1, w_2, \dots, w_m \rangle$. В качестве выходных данных каждый алгоритм возвращает список аспектов A заданного ритмического средства: A_{anph} — анафора, A_{eph} — эпифора, A_{sym} — симплока, A_{andp} — анадиплозис, A_{epnl} — эпаналепсис, A_{redpl} — редупликация, A_{epstr} — эпистрофа, A_{poly} — многосоюзиe, A_{aps} — апозиопеза.

Каждый аспект $a = \langle W, S \rangle$, $a \in A$, состоит из слова или словосочетания W , повторяющегося в нескольких предложениях $S = \langle s | s \in T, W \in s, |S| > 1 \rangle$. Список предложений S является контекстом для аспекта a . По контекстам эксперт-лингвист может сравнить оригинал текста и его перевод, чтобы проанализировать, как именно переводчик отражает ритмические средства автора текста. Также по контексту можно определить, верно ли выделен аспект.

Приведем пример появления ритмического средства в тексте и связанных с ним обозначений: *He does. He is fond of you. You are his favourite.*

Здесь имеется анадиплозис. Его появление считается аспектом $a \in A_{andp}$. $a = \langle W, S \rangle$, где повторяющиеся слова $W = \langle you \rangle$, контекст $S = \langle \text{«He is fond of you», «You are his favourite»} \rangle$.

Рассмотрим подробнее алгоритмы поиска ритмических средств.

3.1. Алгоритмы поиска анафоры, эпифоры и симплоги

На первом шаге составляется список кандидатов в анафоры: $K = \langle k_1, \dots, k_l \rangle$, где каждое $k \in K$ имеет вид $\langle w, S' \rangle$ и состоит из слова w и списка соседних предложений $S' = \langle s | s \in T, |s| > 1, s = \langle w, w_1, w_2, \dots, w_m \rangle \rangle$.

Для каждого k составляем аспект $a = \langle W, S \rangle$. Аспект составляется циклически: изначально $W = \langle w \rangle$, затем на каждом j -м шаге к W добавляется очередное слово w_j , такое, что $w_j \in s$ верно для нескольких предложений $s \in S'$. Если таких w_j оказывается несколько для разных предложений, берется w_j , соответствующее большему числу предложений. После добавления w_j в W фильтруется S' , а именно, остаются s , удовлетворяющие условию:

$$s = \langle w, w_1, w_2, \dots, w_j, w_{j+1}, \dots, w_m \rangle,$$

т. е. предложения, начинающиеся с набора слов W , но не состоящие из него полностью. Как только не удастся найти подходящее w_j , цикл заканчивается, $a_i = \langle W, S \rangle$, где W — список повторяющихся слов, а $S = S'$ после работы алгоритма.

Поиск эпифоры организуется аналогично, отличие заключается только в том, что формирование W ведется не с начала предложения, а с конца.

Список симплог A_{sym} , составляется из списков анафоры и эпифоры: A_{anph} и A_{eph} соответственно. Между собой сравниваются все пары аспектов $\langle b, c \rangle$ таких, что $b \in A_{anph}$, $c \in A_{eph}$. Если $b = c$, это значит, что в одном и том же контексте предложения имеют одинаковые начала и одинаковые окончания, значит, найдена симплога. Тогда b удаляется из A_{anph} , c — из A_{eph} , b добавляется в A_{sym} .

Алгоритм заканчивает работу, как только перебраны все пары $\langle b, c \rangle$.

3.2. Алгоритм поиска анадиплозиса

Представим текст как набор слов w и знаков препинания p : $T = \langle w_1, p_1, w_2, \dots, p_q \rangle$. Алгоритм ищет анадиплозис циклически, на каждом шаге выявляя аспекты с наибольшим $|W|$.

На первом шаге алгоритм поиска анадиплозиса перебирает текст и ищет в нем списки вида: $l = \langle w, p, w \rangle$, то есть два одинаковых слова, разделенных одним знаком препинания. Из l составляется аспект $a = \langle W, S \rangle$, где $W = \langle w \rangle$, а $S = \langle s \rangle$ и $l \in s$ или $S = \langle s_1, s_2 \rangle$, $s_1 = \langle w_{11}, \dots, w \rangle$, $s_2 = \langle w, w_{21}, \dots, w_{2m} \rangle$, если p обозначает конец предложения.

Следующий шаг практически полностью аналогичен предыдущему, за исключением того, что $l = \langle w_1, w_2, p, w_1, w_2 \rangle$, то есть ищутся уже пары одинаковых слов, разделенных одним знаком препинания. И тогда $W = \langle w_1, w_2 \rangle$.

В общем случае на i -м шаге алгоритм ищет список вида:

$$l = \langle w_1, w_2, \dots, w_i, p, w_1, w_2, \dots, w_i \rangle,$$

а $W = \langle w_1, w_2, \dots, w_i \rangle$. Работа алгоритма прекращается после шага с номером $i_{\max} = \max_{s \in T} |s|$.

3.3. Алгоритмы поиска эпаналепсиса, редупликации и эпистрофы

В алгоритмах поиска редупликации и эпистрофы перебираются все слова $w \in T$. Ищутся все наборы последовательных слов $W = \langle w_1, \dots, w_k \rangle$, $k > 0$, которые формируют список $l = \langle w_1, \dots, w_k, \dots, w_1, \dots, w_k \rangle$, $l \in T$, то есть W встречается в T несколько раз, при этом W не может полностью состоять из предлогов или союзов.

Если найден список вида $l = \langle w_1, \dots, w_k, w_1, \dots, w_k \rangle$, то есть W дублируется, и при этом $W \in s$ (W дублируется в рамках одного предложения), а для любого $a' \in A_{andp}$, $W \notin a'$ (найден не анадиплозис), то обнаружена редупликация. Тогда формируется аспект $a = \langle W, S \rangle$, где $S = \langle s \rangle$, и a помещается в A_{redpl} .

Если найден список вида $l = \langle w_1, \dots, w_k, w, \dots, w_1, \dots, w_k \rangle$, при этом $l = \langle s \rangle$ или $l = \langle s_i, s_{i+1} \rangle$, $i = 1, \dots, n - 1$, то есть W повторяется в начале и конце предложения или пары предложений, и, возможно, в их середине, то обнаружена эпистрофа. Тогда формируется аспект $a = \langle W, S \rangle$, где $S = l$, и a помещается в A_{epstr} .

Если найден список вида $l = \langle w_1, \dots, w_k, w, \dots, w_1, \dots, w_k \rangle$, то есть между повторяющимися W имеется хотя бы одно слово w , и при этом $W \in s$ или $W \in s_i$, $W \in s_{i+1}$, $i = 1 \dots n - 1$, и также для любого $a' \in A_{anph}, A_{eph}, A_{sym}, A_{andp}, A_{epstr}$, $W \notin a'$, то обнаружен эпаналепсис. Тогда формируется аспект $a = \langle W, S \rangle$, где $S = \langle s \rangle$ или $S = \langle s_i, s_{i+1} \rangle$, и a помещается в A_{epnl} .

3.4. Алгоритмы поиска синтаксических средств: многосоюзия и апозиопезы

Для поиска многосоюзия алгоритму необходимы списки союзов: список обычных союзов $C_1 = \langle \text{but, till, as, if, after, until, because, and, or, nor, so, before, since, that, till, until, unless, whether, while, where, when, why, what, how, whenever, although, though, once, than, whereas, thus} \rangle$ и список парных союзов $C_2 = \langle \langle \text{both, and} \rangle, \langle \text{either, or} \rangle, \langle \text{not only, but} \rangle, \langle \text{not only, but also} \rangle, \langle \text{rather, or} \rangle, \langle \text{just as, so} \rangle, \langle \text{neither, nor} \rangle, \langle \text{whether, or} \rangle, \langle \text{if, then} \rangle \rangle$, а также список союзных наречий $C_3 = \langle \text{after all, as a result, for example, in addition, in fact, in other words, on the other hand} \rangle$. Аналогичные списки имеются и для русского языка.

Каждый из списков C_i , где $i = 1, 2, 3$, обрабатывается следующим образом.

Для каждого союза или союзного наречия $c \in C_i$ и каждого предложения $s \in T$ проверяется условие $l \in s$, где l имеет один из видов:

$$l = \langle w_1, \dots, c, \dots, c, \dots, w_m \rangle,$$

$$l = \langle c, \dots, c, \dots, w_m \rangle,$$

$$l = \langle w_1, \dots, c, \dots, c \rangle,$$

$$l = \langle c, \dots, c \rangle,$$

то есть c появляется в s более одного раза, возможно, в начале или в конце предложения. Если условие выполняется, в список аспектов A_{poly} добавляется аспект $a = \langle W, S \rangle$, где $W = \langle c \rangle$, $S = \langle s \rangle$.

Алгоритм для апозиопезы, возвращает список аспектов A_{aps} , где аспект $a \in A_{aps}$ состоит только из предложения $s \in T$. Это самый простой из всех алгоритмов поиска ритмических средств, так как он заключается в поиске $s \in T$, которые содержат

многоточие, и добавлении $a = \langle \emptyset, s \rangle$ в список аспектов A_{aps} . Эти аспекты не содержат повторяющихся слов.

3.5. Программная реализация алгоритмов

Описанные алгоритмы поиска ритмических средств реализованы в веб-приложении ProseRhythmDetector. Оно разработано при помощи технологий HTML, CSS, JavaScript и опубликовано на GitHub⁵.

Приложение позволяет загрузить текст произведения на русском или английском языке, ищет в тексте ритмические средства при помощи описанных алгоритмов. Затем оно отображает текст и выделяет в нем найденные аспекты и их контекст различными цветами. Кроме того, ProseRhythmDetector подсчитывает частоту встречаемости ритмических средств в тексте и выводит список аспектов заданного средства в отдельной панели. Главный экран приложения с обработанной главой произведения Ш. Бронте «Городок» представлен на рис. 1.

The screenshot shows the main interface of the ProseRhythmDetector application. It features a menu on the left with options like 'Load File', 'Load ban file', 'Settings', and 'Download JSON'. The main area displays a text excerpt from 'Городок - глава 3' with highlighted rhythmic aspects. A 'Word's list' table is visible on the right, and a 'RESULTS' table at the bottom shows the frequency of various aspects.

Word	Count
и	170
она	138
не	132
я	110
в	98
на	84
что	79
с	66
он	63

Selected lexical aspects	Anaphora	Epiphora	Symploce	Anadiplosis	Epistrophe	Reduplication	Epanalepsis
<input checked="" type="checkbox"/> Anaphora	7	3	0	0	1	0	142
<input checked="" type="checkbox"/> Epiphora							
<input checked="" type="checkbox"/> Symploce							
<input checked="" type="checkbox"/> Anadiplosis							
<input checked="" type="checkbox"/> Epistrophe							

Anaphora's examples:
 1. "Грэм"
 Грэм отдал ей картинку, она же бросилась прочь, подобно преследуемому кредитору, и нашла убежище на коленях отца. Грэм вскричал, изображая ярость, и последовал за ней.
 2. "Она"
 Она тотчас же оказалась около него. – Лайте я посмотрю ваш глаз. – Я вовсе не собиралась попасть в него, я хотела ударить по губам, я не предполагала, что ударю так ужасно

Рис. 1. Главный экран приложения ProseRhythmDetector

Fig 1. Main page of the ProseRhythmDetector application

В результате работы с данным приложением лингвист-исследователь получает возможность оценить стиль текста произведения в целом, просматривая и анализируя встречающиеся в нем лексические и синтаксические средства. Также у него появляется возможность провести эксперименты с изучением особенностей нескольких произведений за гораздо меньшее время, чем потребовалось бы для исследования их вручную без автоматизации.

⁵<https://github.com/text-processing/html-tool>

Таблица 1. Количество ритмических средств в оригинале произведения
Ш. Бронте «Городок» и в его переводе

Table 1. The number of rhythmic features in the original text
of Ch. Bronte's "Villette" and in its translation

Средство	оригинал	перевод	разность	совпадения
Анафора	68	211	143	10
Эпифора	10	26	16	2
Симплока	0	2	2	0
Анадиплосис	12	27	15	4
Эпистрофа	0	24	24	0
Редупликация	0	486	486	0
Эпаналепсис	1649	4518	2869	119
Многосоюзиe	179	847	668	2
Апозиопеза	3	57	54	0
Всего	1 921	6198	4277	137

4. Эксперименты

Целью проводимых в данной работе экспериментов была оценка степени передачи при переводе перечисленных выше ритмических средств. Эксперименты проводились на материале двух произведений: Ш. Бронте «Городок» (Ch. Bronte "Villette") и в его переводе на русский язык Н. Михальской, и А. Мердок «Черный принц» (I. Murdoch "The Black Prince") и его переводе, выполненном И. Бернштейн и А. Поливановой.

В ходе экспериментов точность поиска инструментом ритмических средств составила от 62 % до 93 %. В таблицах 1 и 2 приведены количественные результаты поиска средств ритмики и их совпадений для этих произведений и их переводов. Ниже более детально рассмотрены выявленные ритмические средства и выполнено их сравнение в оригинальном тексте и его переводе.

4.1. Анафора

В тексте романа Ш. Бронте «Городок» и в его переводе из 68 оригинальных контекстов наблюдается 10 случаев полного совпадения анафорических повторов без изменения позиции в тексте, а также без изменения количества повторяющихся элементов. 8 из 10 случаев — употребление местоименной анафоры (местоимения в функции подлежащего I, she, it), которая наиболее точно передана в тексте перевода. Однако есть случаи замены одного местоимения другим и введения дополнительной конструкции в препозиции к его употреблению:

She had a pale face, hair like night, broad strong eyebrows, decided features, and a dark, mutinous, sinister eye: I noted that

Я внимательно взгляделась в нее: бледное лицо, иссиня-черные волосы, широкие выразительные брови, резкие чер-

Таблица 2. Количество ритмических средств в оригинале произведения
 А. Мердок «Черный принц» и в его переводе

Table 2. The number of rhythmic features in the original text
 of I. Murdoch's "The Black Prince" and in its translation

Средство	оригинал	перевод	разность	совпадения
Анафора	585	346	239	102
Эпифора	149	53	96	35
Симплока	9	3	6	3
Анадиплосис	37	37	0	2
Эпистрофа	78	0	78	0
Редупликация	23	8	15	2
Эпаналепсис	3340	2901	439	1001
Многосоюзие	219	366	147	98
Апозиопеза	0	0	0	0
Всего	4440	3966	474	1250

*she sat close by a little door, which door,
 I was well aware, opened into a small
 closet where books were kept. She was
 standing up for the purpose of conducting
 her clamour with freer energies.*

*ты лица и темные, мятежные, мрач-
 ные глаза. Я заметила, что она си-
 дит около небольшой двери, которая,
 как я знала, ведет в маленький чулан,
 где хранились книги. Ученица встала,
 дабы получить большую свободу дей-
 ствий.*

В тексте произведения А. Мердок «Черный принц» и в его переводе из 585 контекстов оригинальной анафоры 102 контекста воспроизведены с большей или меньшей количественной и позиционной точностью. В 45 контекстах количество и позиция повторяющихся элементов соответствуют полностью, например:

*I have never tried to please at the
 expense of truth. I have known, for long
 periods, the torture of a life without self-
 expression.*

*Я никогда не стремился к приятно-
 сти за счет правды. Я знал долгие му-
 чительные полосы жизни без самовы-
 ражения.*

В 22 контекстах позиция анафор сохранена, количество повторяющихся элемен-
 тов меньше на 1–4 элемента из 4–6, например:

*I waited. I tried to develop a new routine:
 monotony, out of which value springs. I
 waited, I listened.*

*Я ждал. Я снова постарался вы-
 работать упорядоченный образ жиз-
 ни, создать монотонность, из кото-
 рой рождаются всплески. Я выжидал,
 вслушивался.*

В 31 контексте незначительно изменена позиция анафоры (перенос на одно пред-
 ложение вперед или назад) и в 18 из них изменено количество повторяющихся эле-
 ментов, преимущественно в сторону их уменьшения, например:

And I was suddenly deeply frightened by the possibility of having my sister on my hands. I simply did not love her enough to be of any use to her, and it seemed wiser to make this plain at once. I waited for about ten minutes, trying to calm and clear my mind, and then went back to the bedroom door. I did not really expect that Priscilla would have got dressed and be ready to leave. I did not know what to do. I felt fear and disgust at the idea of mental breakdown, the semi-deliberate refusal to go on organizing one's life which is regarded with such tolerance in these days. I peered into the room.

Я боялся, как бы сестра не оказалась вдруг у меня на руках. Я просто-напросто не настолько любил ее, чтобы она могла на меня рассчитывать, и, видимо, лучше всего было сказать ей об этом прямо. Переждав минут десять, пока успокоятся мои нервы и прояснится голова, я встал и подошел к двери в спальню. В сущности, я и не надеялся, что застаю Присциллу одетой и готовой к уходу. Что мне делать, я не знал. Мне противна и страшна была сама мысль о «нервном расстройстве», этом наполовину сознательном уходе от упорядоченной жизни, к которому в наши дни принято относиться с такой терпимостью. Я заглянул в комнату.

В двух случаях в переводе число повторяющихся элементов больше, чем в оригинале (три против двух в оригинале).

Четыре контекста содержат меньшее количество элементов, повторяющихся не подряд, из чего следует, что повтор перестает восприниматься как анафорический, а переходит в эпаналепсис, например:

I went out of the room and closed the door quietly behind me. I heard a soft bound and then the key turning in the lock. I went down the stairs feeling very shaken and, yes, she had been right, disgusted.

Я спустился по лестнице с чувством растерянности и — да, она была права — отвращения. За это время стемнело, солнце больше не сияло на улице, и все в доме стало коричневым и холодным. Я вошел в гостиную, где сидели и беседовали Арнольд с Фрэнсисом.

В 483 контекстах оригинала анафорический повтор при переводе либо утрачен, либо заменен другим средством ритмизации, преимущественно эпаналепсисом, например:

Arnold Baffin wrote too much, too fast. Arnold Baffin was really just a talented journalist.

Арнольд Баффин писал слишком много, слишком быстро. По существу, Арнольд Баффин был всего лишь талантливым журналистом.

4.2. Эпифора

В тексте романа Ш. Бронте «Городок» и в его переводе из 10 выявленных случаев употребления эпифоры в оригинале отмечено 3 совпадения контекстов. Полного совпадения с точки зрения позиции в предложении и аналогичного употребления

ритмического средства в переводе не наблюдается, что связано в первую очередь со спецификой грамматической структуры языков:

“No, Missy,” said the nurse: “you are to share this young lady’s room,” designating me. Missy did not leave her seat, but I saw her eyes seek me. *И она указала на **меня**. Мисси не встала с места, но отыскивала **меня** глазами.*

В приведенном примере эпитетический повтор местоимения *me* (меня) в русском языке теряет финальную позицию в связи со сменой подлежащего в придаточном предложении (*ее глаза искали меня / Мисси отыскивала меня глазами*).

Замена инфинитива формой глагола в повелительном наклонении также не способствует сохранению эпитетического повтора в переводе:

Go to sleep. I cannot go to sleep. *Спи! – Я не могу спать.*

Более точный перевод данного фрагмента мог бы способствовать сохранению ритмического средства: *Иди **спать!** – Я не могу пойти **спать**.*

Переход из одной части речи в другую, в приведенном ниже примере из существительного *love* в формы: *влюблен, любят, любовь*, способствует переходу эпитеты в оригинальном тексте в полиптотон и деривацию в переводе:

He perhaps was not in love; but how many people ever do love, or at least marry for love, in this world. *Он, по-видимому, не был **влюблен** в нее, но разве так уж много людей в этом мире **любят** по-настоящему или женятся по **любви**?*

В этом случае не сохраняется ритмическое средство, оно переходит в иной статус.

В тексте произведения А. Мердок «Черный принц» и в его переводе на 149 случаев оригинальной эпитеты найдено 35 совпадений. В 24 контекстах оригинальная эпитета при переводе заменена другими видами повтора. Самой частотной является замена эпитеты на эпанафору (13 случаев):

You’d better get your own doctor tomorrow. Oh, I think I shall be better tomorrow. *Вам надо будет **завтра** самому вызвать доктора. – О, **завтра**, я думаю, мне будет лучше.*

Далее следует замена на смысловую повтор (замена на синоним), иногда в сочетании с изменением порядка слов (5 случаев):

Roger has become a devil. Some sort of devil. *В Роджера просто **дьявол** вселмлся. Какой-то **демон**.*

Обнаружено по 2 случая замены эпитеты на анадиплозис и градационный повтор. В переводе присутствуют контексты с эпитетой, которой нет в зеркальных оригинальных предложениях, что может рассматриваться как попытка частично компенсировать утрату эпитеты по сравнению с оригиналом.

4.3. Анадиплозис

Четыре случая анадиплозиса в тексте романа Ш. Бронте «Городок» точно воспроизведены в переводе. Повтор на стыке предложений или частей предложения довольно

сложно разграничить с редупликацией, которая рассматривается как повтор одного и того же слова в начале предложения, в связи с этим при переводе часто одно заменяется другим:

Her features worked, — “I am sorry; I am sorry!” *Прости, прости.*

При этом часто разработанный инструмент смешивает данные ритмические средства, что нарушает статистику. В этом случае задача разграничения данных средств лежит на исследователе.

При передаче анадиплозиса повтор может быть передан с изменением позиции его компонентов, а также некоторым изменением состава повторяющейся конструкции:

I grant I was not looking well, but, on the contrary, thin, haggard, and hollow-eyed; like a sitter-up at night, like an overwrought servant, or a placeless person in debt. In debt, however, I was not; <... >

К этому времени нервы у меня рашатались, но немного, совсем немного. Полагаю, и выглядела я неважно: худая, изможденная, с ввалившимися глазами, похожая на сиделку, проводящую ночи у постели больного, на переутомленную служанку или запутавшегося в долгах безработного. Однако я не запуталась в долгах и не оказалась в крайней бедности безработного <... >.

Также при переводе может осуществляться замена анадиплозиса анафорой:

He is fond of you. You are his favourite. *Ты ему нравишься. Ты его любимица.*

Количество повторяющихся элементов при этом не изменяется, их позиция в тексте также остается неизменной, но изменяется позиция элементов средства (в начале частей предложения).

Анадиплозис может быть утерян при переводе:

“I? I couldn’t do it!” (sinking into a chair). *«Я? Ни за что в жизни» (падает в кресло).*

Из 37 случаев анадиплозиса в оригинале произведения А. Мердок «Черный принц» два точно воспроизведены в переводе, например:

Art is imagination. Imagination changes, fuses. *Искусство – это воображение! Воображение пресуществляет, плавит в своем горниле.*

Так же как и в первом произведении анадиплозис может заменяться другим средством, например, эпаналепсисом:

I mean to indicate the curiously controlled rhythmic screaming which goes with a certain kind of hysterics.

Этим термином я обозначил странно ритмичные, рассчитанные вопли, сопровождающие некоторые истерические состояния. Истерика пугает

Hysterics is terrifying because of its willed and yet not willed quality.

тем, что она произвольна и непроизвольна в одно и то же время.

4.4. Эпаналепсис

Эпаналепсис — наиболее частотное ритмическое средство для текста романа Ш. Бронте «Городок». Количество повторяющихся элементов в переводе сохранено, однако их позиция несколько изменена, что в некоторой степени сказывается на характере ритма:

Bretton and I sat alone in the drawing-room waiting her coming; John Graham Bretton being absent on a visit to one of his schoolfellows who lived in the country.

В гостиной остались лишь миссис Бреттон и я, так как Джон Грэм Бреттон гостил в деревне у своего однокашника.

Диапазон между повторяющимися элементами в переводе меньше, чем в оригинале текста в два раза (12 и 6 слов), что усиливает ритм русского текста.

В следующем примере также несколько изменяется позиция повторяющихся элементов, однако оригинальный вариант сопровождается возрастанием количества слогов в синтагмах: 4–5–6, что способствует восприятию ритма на фонетическом уровне:

Tie my sash straight; make my hair smooth, please. Your sash is straight enough.

Пожалуйста, поправьте мне пояс и пригладьте волосы. — Но пояс у вас в порядке.

Сравнительно-сопоставительный анализ оригинальных и переведенных контекстов романа А. Мердок «Черный принц» позволил сделать заключение о том, что эпаналепсис полностью воспроизводится при переводе примерно в 25 % случаев (834 контекста):

I learnt later with abhorrence that he had set up in business as a self-styled psychoanalyst. Later still I heard he had taken to drink.

Позже я с отвращением узнал, что он завел себе практику в качестве самозваного «психоаналитика». Еще позже я слышал, что он пьет.

Иногда в переводе утрачивается один-два элемента, если количество повторяющихся элементов сравнительно велико (более 4).

Еще в 5 % случаев (167 контекстов) эпаналепсис заменяется другим средством ритмизации, что, тем не менее, способствует сохранению ритмического рисунка авторского текста. В проанализированных примерах были выявлены случаи замены эпаналепсиса на смысловой повтор в сочетании с синтаксическим параллелизмом и позиционной компенсацией эпаналепсиса, например:

He was stout (the raincoat failed to button) and not tall, with copious greyish longish frizzy hair and a round face and a slightly hooked nose and big very red lips and eyes set very close together.

Он был толст (макинтош явно не застегивался), невысок ростом, волосы густые и курчавые, давно не стриженные, с проседью, лицо круглое, со слегка крючковатым носом, толстыми, очень красными губами и уди-

вительно близко посаженными глазами.

Замена на анадиплозис с позиционной компенсацией эпаналепсиса:

Real bears, I believe, have eyes rather wide apart, but caricatured bears usually have close eyes, possibly to indicate bad temper or cunning.

Он походил, как я потом подумал, на карикатурного медведя. Не на настоящего — у настоящих медведей глаза, по-моему, расставлены широко, а вот на карикатурах их рисуют с близко посаженными глазами — вероятно, для того, чтобы выразить их свирепость и коварство.

4.5. Многосоюзиe

Совпадение контекстов в романе Ш. Бронте «Городок» составляет 24 из 179, однако точная передача ритмического средства — только в 2 случаях, при этом основным союзом, который обеспечивает употребление данного ритмического средства, является союз *and* (*и*):

My visits to her resembled the sojourn of Christian and Hopeful beside a certain pleasant stream, with “green trees on each bank, and meadows beautified with lilies all the year round.” The charm of variety there was not, nor the excitement of incident; but I liked peace so well, and sought stimulus so little, that when the latter came I almost felt it a disturbance, and wished rather it had still held aloof. One day a letter was received of which the contents evidently caused Mrs. Bretton surprise and some concern.

Жизнь моя не отличалась пленительным разнообразием и волнующими приключениями, но мне нравился этот покой, и, избегая всяческих перемен, я даже любое письмо воспринимала как нарушение привычного хода вещей и предпочитала, чтобы оно вовсе не приходило. Однажды миссис Бреттон получила письмо, содержание которого явно удивило и несколько обеспокоило ее.

Многосоюзиe вполне успешно воспроизводится при переводе произведения А. Мердок «Черный принц». Точная передача данного ритмического средства составила порядка 45 % (98 контекстов). Следует отметить, что практически во всех переведенных контекстах появлялись дополнительные средства многосоюзия: 219 примеров в оригинале и 366 в переводе, например:

Both taxman and dentist only too readily image forth the deeper horrors of human life: that we must pay, perhaps ruinously, for our pleasures, that our resources are lent, not given, and that our most irreplaceable faculties decay even as they grow.

И зубной врач, *и* налоговый инспектор, естественно, символизируют для нас подспудные ужасы жизни; они говорят о том, что мы должны платить, даже если цена разорительная, за все наши удовольствия, что блага даются нам в долг, а не дару-

ются, **что** наши самые невосполнимые богатства гниют уже в процессе роста.

Инструмент выделяет одинаковые союзы, расположенные в рамках одного абзаца, что несколько расширяет границы самого ритмического средства, которое определяется не только как повтор союзов при однородных членах предложения, но также выделяет различные группы однородных членов.

4.6. Симплока, эпистрофа и редупликация

Случаев употребления симплоки, эпистрофы и редупликации в романе Ш. Бронте «Городок» обнаружено не было.

Один из девяти случаев симплоки, найденных в оригинальном тексте произведения А. Мердок «Черный принц», имеет зеркальное совпадение в тексте перевода:

<i>I turned back and drove the other way, through the village, past the church. I even stopped and went into the church.</i>	<i>Я повернул назад и поехал через деревню мимо церкви. Я даже остановился и зашел в церковь.</i>
--	---

Другие средства ритмизации не были утрачены, но подверглись стилистическим преобразованиям. Таким образом, из восьми оставшихся случаев пять раз симплока заменена на эпаналепсис:

<i>She kept drawing it out without telling me and buying clothes. She went mad over buying clothes.</i>	<i>А она втихомолку тянула фунт за фунтом и покупала себе тряпки. Она на этих тряпках просто помешалась.</i>
---	--

Два раза симплока заменена на эпифору и сочетание эпифоры и эпаналепсиса:

<i>What did you say? What could I say?</i>	<i>А вы что ответили? Что я могла ответить?</i>
---	---

Один раз сохранен синтаксический параллелизм:

<i>I write whether I feel like it or not. I complete things whether I think they're perfect or not.</i>	<i>Я пишу независимо от того, легко мне или трудно. И завершаю любую работу, удалась она мне или не удалась.</i>
---	--

Три случая симплоки, зафиксированные в тексте перевода, употреблены переводчиком самостоятельно, возможно, в порядке компенсации утраченных средств ритмизации.

В оригинальном тексте выявлено 23 случая редупликации, из них два совпали при переводе. Случаев применения эпистрофы в переведенном тексте инструмент не зафиксировал.

5. Обсуждение результатов

Следует отметить, что детальный анализ большого количества ритмических аспектов (почти две тысячи аспектов в одном произведении и более четырёх тысяч в дру-

гом) стал возможен только с использованием автоматизации поиска и визуализации результатов. Разработанный программный инструмент позволил лингвистам сосредоточиться на аналитической части исследования, более детально сравнить каждый аспект ритмики с переводом и сделать выводы о причинах его утраты или замены.

Количественный анализ передачи ритмических средств в переводе романа Ш. Бронте «Городок» показал, что наибольшую степень близости к оригиналу имеет передача такого средства, как анадиплозис (33,3 %), далее следуют эпифора (20 %), анафора (14,7 %), эпаналепсис (7,2 %) и многосоюзиe (1,1 %). В целом процент совпадений при употреблении ритмических средств довольно низкий — 7,1 %. В переводе произведения А. Мердок «Черный принц» степень передачи ритмических средств отличается: многосоюзиe (44,7 %), симплока (33,3 %), эпаналепсис (30 %), эпифора (23,5 %), анафора (17,4 %), редупликация (8,7 %), анадиплозис (5,4 %). Общий процент совпадений существенно выше: 28,2 %.

Из всех рассмотренных средств анафора наиболее точно передана при переводе с точки зрения позиции в тексте и количества повторяющихся элементов. Большое количество наблюдаемых в тексте аспектов этого ритмического средства позволило обобщить и сформулировать следующие причины утраты анафоры.

1. Невозможность ее воспроизведения в силу объективного расхождения грамматических возможностей английского и русского языков. Для проанализированных примеров это часто разница в употреблении залоговых форм, например, для русского языка исключена возможность употребления страдательного залога с непереходными глаголами, глаголами с предложным управлением, в то время как для английского языка это норма:

I was upset and annoyed when my father once approached the subject, and although I could see that he had been made utterly miserable, I resolutely refused to discuss it.

Когда мой отец попытался однажды заговорить со мной о Присцилле, мне это было крайне неприятно, и я, хотя и видел, как он расстроен, решительно отказался обсуждать эту тему.

В этом примере один из элементов анафорического повтора *I* теряется.

2. Опасность нарушения узуса (общепринятого носителями данного языка употребление языковых единиц) переводящего языка. Например, перевод словосочетания *He needed* как *ему был нужен*, а не *он нуждался*.
3. Наличие формульных (клишированных) фраз:

I am 58. I still dream about it at least once a week.

Мне 58. Он и по сию пору снится мне примерно раз в неделю.

4. «Агрессивность» (назойливость) повтора в 5–6 элементах может приводить к нарушению стилистики речи. Кроме того, повтор местоимений-подлежащих (ими английский текст избыточен, особенно *I*) является избыточным для русского языка, где значение грамматических категорий дублируется через спря-

жение глаголов, изменение корневых гласных, согласных, свойственное для местоимений определенных лиц и чисел и др.

I waited, I listened.

Я выждал, вслушивался.

Утрата контекстов с эпитетическим повтором ожидаема с точки зрения разницы в позиции ремы (то новое, что сообщается о теме в данном высказывании) в английском и русском языках. Например, безударные личные местоимения в конце предложения или фразы на английском языке, которые могут составлять эпитету, не несут смысловой нагрузки, в то время как финальное слово в предложении или фразе русского языка автоматически получает фразовое ударение и воспринимается как рема. Для сохранения правильного тема-рематического членения предложения необходимо прибегать к перестановке элементов предложения или фразы, что неизбежно ведет к утрате эпитету. В значительном числе контекстов она компенсируется другими средствами ритмизации.

В большинстве случаев анадиплозис утрачен без какой-либо ритмической компенсации в составе предложений. Однако общее количество анадиплозисов в переведенном тексте (в тексте романа Ш. Бронте 12 оригинальных, 4 соответствуют оригиналу, 23 употреблены переводчиком самостоятельно; в тексте романа А. Мердок 37 оригинальных, 4 полностью или частично соответствуют оригиналу, 33 употреблены переводчиком самостоятельно) позволяет говорить о равновесном употреблении данного ритмического средства в оригинале и переводе, пусть и в порядке позиционной компенсации.

Эпаналепсис — наиболее часто встречающееся ритмическое средство для обоих исследуемых произведений. В большинстве случаев эпаналепсис при переводе утрачивается с последующей его компенсацией в предложениях, оригиналы которых эпаналепсиса в своем составе не имеют. Чаще всего в составе этих предложений имеются другие ритмические фигуры, в отдельных случаях предложения не содержат повтора вообще. Среди наиболее частотных причин утраты эпаналепсиса следует назвать невозможность с точки зрения узуса стопроцентного воспроизводства личного местоимения *I* в роли подлежащего, а также притяжательных местоимений, в особенности в препозиции к существительным, обозначающим части тела, предметы одежды, термины родства и др.

Многосоюзие наиболее полно передано в переводе произведения А. Мердок «Черный принц». Во многих контекстах это средство компенсируется смысловым повтором (через употребление синонимичного союза) или утрачивается, например:

The shop sold daily papers and magazines, writing paper and so on, and horrible gifts.

Они продавали газеты и журналы, бумагу всевозможных сортов, а также безобразные «подарки».

Остальные ритмические средства достаточно редко встречаются в оригинальном тексте или не встречаются вообще, что не даёт возможность сделать существенные выводы об особенностях перевода этих аспектов. Интересно, что хотя редупликация полностью отсутствует в оригинале романа Ш. Бронте «Городок», она выявлена в переводе в довольно большом количестве. Возможно это зависит от индивидуального стиля переводчика.

В целом основной причиной низкой степени передачи ритмических средств в переводе с английского на русский язык являются особенности лексико-грамматической структуры языков оригинала и перевода, относящихся к разным языковым группам. Кроме того, можно указать другие факторы утраты или замены аспектов ритмики текста, например, стратегии, избранные переводчиком, а также качество перевода и степень адекватности передачи ритмических характеристик. Детальный анализ этих факторов требует дальнейших исследований.

Следует отметить высокую трудоёмкость таких исследований и экспериментов. Особенно тяжело для человека сочетать трудоемкий поиск большого количества разнообразных ритмических средств, встречающихся в тексте с совершенно разной частотой, со сравнительным анализом оригинала и перевода. Разработанный авторами программный инструмент *ProseRhythmDetector* продемонстрировал высокую эффективность, позволяя осуществить трудоемкий процесс поиска таких ритмических средств, как анафора, эпифора, симплока, анадиплозис, эпистрофа, редупликация, эпаналепсис, апозиопеза, многосоюзие. Большое значение при этом имеет визуализация результатов поиска, позволяющая эксперту легко ориентироваться в тексте произведения. Применение инструмента дало возможность провести эксперимент с большими художественными произведениями, который практически невозможно выполнить вручную, и оценить степень ритмичности текста с точки зрения лексико-грамматического уровня реализации ритма.

Заключение

В данной работе авторы выполнили алгоритмизацию таких сложных лингвистических понятий, как средства ритмики прозаического текста. Разработанный в результате программный инструмент позволил автоматизировать значительную часть обработки текста на естественном языке, связанную с поиском и подсчётом целого комплекса ритмических аспектов художественных произведений. Использование инструмента лингвистами позволило провести сравнительный анализ оригиналов и переводов двух объемных романов и сделать выводы об особенностях передачи ритмики прозаического текста при переводе с английского на русский язык.

Степень совпадения или несовпадения передачи ритмических средств оригинала художественного произведения в его переводе является одним из маркеров адекватности выполненного перевода, но по большей части позволяет судить о специфике языковой структуры в целом, в особенности о характере передачи тех или иных ритмических средств, об их формах, вариациях, заменах. Это позволяет оценить специфику ритма как комплексного явления и обобщить возможные варианты способов перевода определенных ритмических средств.

Список литературы / References

- [1] Голякова Л. А., “Ритм художественного произведения: коммуникативно-прагматический аспект”, *Вестник Пермского университета. Российская и зарубежная филология*, 2011, №3, 94–99; [Golyakova L. A., “Ritm hudozhestvennogo proizvedeniya: kommunikativno-pragmaticheskij aspekt”, *Vestnik Permskogo universiteta. Rossijskaya i zarubezhnaya filologiya*, 2011, №3, 94–99, (in Russian).]

- [2] Голубева-Монаткина Н. И., “К проблеме ритма прозы”, *Известия Российской академии наук. Серия литературы и языка*, **76:2** (2017), 16–27; [Golubeva-Monatkina N. I., “K probleme ritma prozy”, *Izvestiya Rossijskoj akademii nauk. Seriya literatury i yazyka*, **76:2** (2017), 16–27, (in Russian).]
- [3] Голубева-Монаткина Н. И., “О ритме художественной прозы в свете идей М. М. Бахтина и переводе”, *Русский язык и культура в зеркале перевода*, 2016, № 1, 105–113; [Golubeva-Monatkina N. I., “O ritme hudozhestvennoj prozy v svete idej M. M. Bahtina i perevode”, *Russkij yazyk i kul'tura v zerkale perevoda*, 2016, № 1, 105–113, (in Russian).]
- [4] Иванова-Лукьянова Г. Н., “Ритм прозы в реальности и восприятии”, *Труды Института русского языка им. В. В. Виноградова*, **7:7** (2016), 365–382; [Ivanova-Luk'yanova G. N., “Ritm prozy v real'nosti i vospriyatii”, *Trudy Instituta russkogo yazyka im. V. V. Vinogradova*, **7:7** (2016), 365–382, (in Russian).]
- [5] Фененко Н. А., “Лингвокультурная адаптация текста при переводе: пределы возможного и допустимого”, *Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация*, **1** (2001), 70–75; [Fenenko N. A., “Lingvokul'turnaya adaptaciya teksta pri perevode: predely vozmozhnogo i dopustimogo”, *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Lingvistika i mezhkul'turnaya kommunikaciya*, **1:7** (2001), 70–75, (in Russian).]
- [6] Delmonte R., “Computing Poetry Style”, *Proceedings of 1st International Workshop ESSEM 2013/ CEUR Workshop Proc.*, 2013, 148–155.
- [7] Bobenhausen K., Hammerich B., “Literary metrics, Linguistic metrics, and the algorithmic analysis of German poetry using Metricalizer (2)”, *Langages*, 2015 3, 67–88.
- [8] Niculescu I., Trausan-Matu S., “Rhythm analysis of texts using Natural Language Processing”, *RoCHI*, 2016, 107–112.
- [9] Niculescu I., Trausan-Matu S., “Rhythm analysis in chats using Natural Language Processing”, *RoCHI*, 2017, 69–74.
- [10] Кишалова Л. В., “Анализ особенностей ритмической структуры текстов разных стилей речи”, *Вестник Брянского государственного университета*, 2016, № 1 (27), 257–261; [Kishalova L. V., “Analiz osobennostej ritmicheskoj struktury tekstov raznyh stilej rechi”, *Vestnik Bryanskogo gosudarstvennogo universiteta*, 2016, № 1 (27), 257–261, (in Russian).]
- [11] Boychuk E. et al., “Automated approach for rhythm analysis of French literary texts”, *Proceedings of 15th Conference of Open Innovations Association FRUCT*, IEEE, 2014, 15–23.
- [12] Damasevicius R., Kapociute-Dzikiene J., Wozniak M., “Towards Rhythmicity analysis of text using empirical mode decomposition”, *Proceeding of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, **1**, IC3K, 2017, 310–317.
- [13] Kelly A. R. et al., “Toward an ontology of rhetorical figures”, *Proceedings of the 28th ACM International Conference on Design of Communication*, ACM, 2010, 123–130.
- [14] Hromada D. D., “Initial experiments with multilingual extraction of rhetoric figures by means of PERL-compatible regular expressions”, *Proceedings of the Second Student Research Workshop associated with RANLP 2011*, 2011, 85–90.
- [15] Toldova S. et al., “Error analysis for anaphora resolution in Russian: new challenging issues for anaphora resolution task in a morphologically rich language”, *Proceedings of the Workshop on Coreference Resolution Beyond OntoNotes*, 2016, 74–83.
- [16] Dubremetz M., Nivre J., “Rhetorical Figure Detection: Chiasmus, Epanaphora, Epiphora”, *Frontiers in Digital Humanities*, **5**, 2018.

Lagutina N.S., Lagutina K.V., Boychuk E.I., Vorontsova I.A., Paramonov I.V., "Automated Search of Rhythm Figures in a Literary Text for Comparative Analysis of Originals and Translations Based on the Material

of the English and Russian Languages", *Modeling and Analysis of Information Systems*, **26:3** (2019), 420–440.

DOI: 10.18255/1818-1015-2019-3-420-440

Abstract. Analysis of the functional equivalence of an original text and its translation based on the achievement of rhythm equivalence is an extremely important task of modern linguistics. Moreover, the rhythm component is an integral part of functional equivalence that cannot be achieved without communication of rhythm figures of the text. To analyze rhythm figures in an original literary text and its translation, the authors developed the ProseRhythmDetector software tool that allows to find and visualize lexical and syntactic figures in English- and Russian-language prose texts: anaphora, epiphora, symploce, anadiplosis, epanalepsis, reduplication, epistrophe, polysyndeton, and aposiopesis. The goal of this work is to present the results of ProseRhythmDetector testing on two works by English authors and their translations into Russian: Ch. Bronte “Villette” and I. Murdoch “The Black Prince”. Basing on the results of the tool, the authors compared rhythm figures in an original text and its translation both in aspects of the rhythm and their contexts. This experiment made it possible to identify how the features of the author’s style are communicated by the translator, to detect and explain cases of mismatch of rhythm figures in the original and translated texts. The application of the ProseRhythm-Detector software tool made it possible to significantly reduce the amount of linguists-experts work by automated detection of lexical and syntactic figures with quite high precision (from 62 % to 93 %) for various rhythm figures.

Keywords: text rhythm, rhythm analysis, natural language processing, rhythm figures, automation

On the authors:

Lagutina Nadezhda Stanislavovna, orcid.org/0000-0002-6137-8643, PhD, associate professor, P.G. Demidov Yaroslavl State University, Sovetskaya str., 14, Yaroslavl, 150003, Russia, e-mail: lagutinans@rambler.ru

Lagutina Ksenia Vladimirovna, orcid.org/0000-0002-1742-3240, postgraduate student, P.G. Demidov Yaroslavl State University, Sovetskaya str., 14, Yaroslavl, 150003, Russia, e-mail: lagutinakv@mail.ru

Boyчук Elena Igorevna, orcid.org/0000-0001-6600-2971, PhD, associate professor, Yaroslavl State Pedagogical University named after K.D.Ushinsky, Respublikanskaya street, 108/1, Yaroslavl, 150000, Russia, e-mail: elena-boychouk@rambler.ru

Vorontsova Inna Alekseevna, orcid.org/0000-0001-5897-9299, Dr.Sc., associate professor, Yaroslavl State Pedagogical University named after K.D.Ushinsky, Respublikanskaya street, 108/1, Yaroslavl, 150000, Russia, e-mail: arinna1@yandex.ru

Paramonov Ilya Vyacheslavovich, orcid.org/0000-0003-3984-8423, PhD, associate professor, P.G. Demidov Yaroslavl State University, Sovetskaya str., 14, Yaroslavl, 150000, Russia, e-mail: Ilya.Paramonov@fruct.org

Acknowledgments:

The reported study was funded by RFBR according to the research project № 19-07-00243.