

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДЕЛЕЙ АРХИТЕКТУРЫ TRANSFORMER В ЗАДАЧЕ УПРОЩЕНИЯ ТЕКСТА

Н.А. Прокопьев (*nikolai.prokopyev@gmail.com*)

О.А. Невзорова (*onevzoro@gmail.com*)

Ф.М. Гафаров (*fgafarov@yandex.ru*)

А.А. Гафиатуллин (*arslan2911@mail.ru*)

А.Р. Зиастинев (*ziastinvalmaz@gmail.com*)

Казанский федеральный университет, Казань

В статье рассматривается задача упрощения текстов на русском языке с использованием моделей архитектуры Transformer. Выбор наилучшей модели производится с помощью стандартных оценочных метрик BLEU, ROUGE, SARI и уточненной метрики лексической сложности на основе оценки читаемости. Исследовались предобученные модели T5 и BART, обучение производилось на основе набора данных в области информатики. В результате было выявлено, что модель BART лучше справляется с задачей упрощения текста, а также генерирует тексты, более соответствующие по форме эталонным.

Ключевые слова: лексическая сложность текста, метрика, архитектура трансформер, набор данных.

Введение

Упрощение текста – это активно развивающаяся область обработки естественного языка (NLP), направленная на адаптацию текстов для различных групп читателей, в том числе людей с ограниченной грамотностью, носителей языка с низким уровнем владения, лиц с когнитивными нарушениями и широкую аудиторию, нуждающуюся в доступном изложении сложных материалов. Другой целевой аудиторией являются учащиеся средних школ и студенты высших учебных заведений. Усвоение учебного материала – это важнейшая составляющая процесса обучения, для успешного усвоения разрабатываются различные методики. Персонализированный подход к обучению включает этапы, связанные с оценкой уровня подготовки обучаемого и предъявления ему учебного материала в соответствии с уровнем подготовки.

Таким образом, актуальной является задача, связанная с подготовкой учебных материалов, отличающихся оценками лексической, синтаксической и терминологической сложности текста при сохранении его ключевого смысла. В статье рассматриваются основные технологии, применяемые для решения поставленной задачи снижения сложности учебных текстов, основные метрики оценивания сложности текстов, а также разработанный исследовательский прототип системы снижения сложности учебных текстов в области информатики.

Статья структурирована следующим образом: в разделе 1 приведен обзор основных методов, применяемых для решения задачи упрощения текста, в разделе 2 изложены основные метрики оценивания сложности текста, далее в разделе 3 описаны решения по разработке исследовательского прототипа системы снижения сложности текста и проведен сравнительный анализ применяемых вычислительных моделей, в заключении представлены основные выводы и направления будущих разработок.

1. Методы автоматического упрощения текста

В настоящее время разрабатываются методы автоматического упрощения текстов, включая модели на основе глубокого обучения [Sheang et al., 2021], [Das et al., 2025], гибридные архитектуры и специализированные подходы в обработке научных текстов [Anjum 2023].

Современные методы упрощения текстов основываются на предобученных моделях на основе архитектуры Transformer. Модель T5 (Text-to-Text Transfer Transformer) демонстрирует высокую гибкость благодаря унифицированному подходу "текст-в-текст", что позволяет настраивать степень упрощения [Sheang et al., 2021]. Модель BART (Bidirectional and Auto-Regressive Transformer) превосходит T5 по ряду метрик, особенно в задачах сохранения связности текста [Das et al., 2025].

Для улучшения качества упрощения текста разработаны также гибридные методы. Например, метод SIMPLEX [Truică et al., 2023], который сочетает Word2Vec и трансформеры (BERT, RoBERTa, GPT-2). Другой пример – MultiLS [North et al., 2024], первая многозадачная архитектура, поддерживающая прогнозирование сложности, генерацию и ранжирование замен. Управление упрощением через контрольные токены, позволяющие регулировать длину и сложность текста исследуется в [Dmitrieva, 2023]. Используя гибридные модели трансформеров, разрабатываются методы, которые учитывают потребности конкретных аудиторий, например, студентов с дислексией [Sukiman et al., 2023].

Для русского языка созданы специализированные наборы данных, такие как RuSimpleSentEval [Sakhovskiy et al., 2021] и корпус аннотированных предложений [Ivanov et al., 2023], что способствует развитию методов упрощения для славянских языков. Эксперименты с управляемым упрощением [Dmitrieva, 2023] подтвердили эффективность моделей mBART и T5 для решения задачи упрощения русских текстов.

2. Метрики оценивания сложности текста

2.1. Задача оценки результата упрощения текста

Одна из главных проблем в задаче упрощения текста – создание автоматической метрики оценки, которая могла бы заменить трудоемкую экспертную оценку. Хорошая метрика должна учитывать три ключевых свойства упрощенного текста: грамматичность, сохранение смысла и простоту (насколько легко воспринимается упрощенный вариант).

Исчерпывающий обзор по автоматическим метрикам оценки в задаче упрощения текста содержится в [Alva-Manchego et al., 2020]. Можно выделить несколько классов автоматических метрик оценки задачи упрощения текста, которые используются в настоящем проекте.

2.2. Оценочные метрики машинного перевода

2.2.1. Метрика BLEU (BiLingual Evaluation Understudy) оценивает качество перевода на основе точности совпадения n -грамм между текстом, полученным путем машинного перевода (кандидатом) и эталонным текстом [Papineni et al., 2002]. Задача упрощения текста в этом случае рассматривается как задача перевода сложного текста в более простой. Метрика принимает значения от 0 до 1, в некоторых источниках нормализуется в диапазон от 0 до 100, где больший показатель означает лучшее качество. Метрика BLEU ориентирована на точность (чем больше совпадающих n -грамм, тем выше оценка) и не учитывает порядок слов.

2.2.2. Метрика ROUGE (Recall-Oriented Understudy for Gisting Evaluation) представляет собой семейство метрик, предназначенных для оценки суммаризации текста [Lin, 2004]. Метрики вычисляют степень совпадения n -грамм, последовательностей и пар слов между алгоритмически сгенерированными аннотациями и эталонными аннотациями, составленными экспертами. Принимают значения от 0 до 1, в некоторых источниках нормализуются в диапазон от 0 до 100, где больший показатель означает лучшее качество.

В ROUGE-1 сравниваются единицы (слова) между сгенерированным и эталонным текстами. В ROUGE-2 сравниваются последовательности из двух слов, взятых из сгенерированного и эталонного текста. ROUGE-L ищет самую длинную последовательность, которая является общей для двух текстов, а затем измеряет её длину.

2.3. Оценочные метрики упрощения текста

Метрика SARI (System output Against References and Input sentence) сравнивает текст, полученный на выходе модели упрощения текстов (кандидат) с несколькими эталонными текстами и исходным текстом [Xu et al., 2016]. Идея SARI заключается в том, чтобы вознаграждать модели за добавление n -грамм, которые встречаются в любом из эталонных текстов, но не в исходных текстах; вознаграждать сохранение n -грамм как

в кандидатах, так и в эталонных текстах; вознаграждать сохранение важных n -грамм. Таким образом, метрика вычисляет среднее арифметическое оценок точности и полноты операций добавления, сохранения и удаления n -грамм. В настоящее время эта метрика стала стандартом для оценки и сравнения моделей упрощения текста. Она принимает значения от 0 до 100, где больший показатель означает лучшее качество.

2.4. Метрики, основанные на оценке читаемости текста

2.4.1. Метрика Флеша (*Flesch Reading Ease*) оценивает читаемость текста на основе двух ключевых параметров: средняя длина предложения (в словах) и средняя длина слова (в слогах). Диапазон: от 0 (очень сложно) до 100 (очень легко) [Flesch, 1948].

2.4.2. Метрика Флеша-Кинкайда (*Flesch-Kincaid Grade Level, FKGL*) Оценки Флеша-Кинкайда представляет собой пересчет метрики *FRE* с коэффициентами, полученными на основе процедур множественной регрессии в тестах по чтению [Kincaid et al., 1975].

2.4.4. Иные метрики. Помимо рассмотренных выше метрик, можно отметить более сложные метрики, например, метрику упрощения текста *SAMSA* [Sulem et al., 2018], которая использует семантический анализ исходного текста, а также метрику *Coh-Metrix* [McNamara et al., 2014], разработанную для анализа текстов по множественным характеристикам и уровням языка и дискурса.

3. Разработка исследовательского прототипа для системы снижения сложности текста

3.1. Общая схема последовательности этапов разработки

Для проведения сравнительного анализа моделей упрощения текста была проведена разработка исследовательского прототипа (общая схема разработки представлена на рис. 1), в рамках которой выполнено:

1. Подготовка общего набора данных в виде набора пар текстов, в которых один текст является сложным, а другой – упрощенной версией этого текста. В набор входят как данные на основе авторских текстов, так и полностью сгенерированные данные.

2. Разделение общего набора данных на набор эталонных данных и на набор данных для обучения моделей. Набор эталонных данных должен статистически соответствовать набору обучающих данных и состоять из проверенных экспертами вручную пар текстов.

3. Обучение выбранных моделей на основе обучающего набора. Этап включает предварительную подготовку конфигурации обучения, само обучение и проверку результатов обучения с использованием стандартной метрики потерь (loss). Отбираются наиболее удачные конфигурации и результаты обучения.

4. Оценка моделей на основе рассмотренных оценочных метрик, а также уточненной метрики лексической сложности.

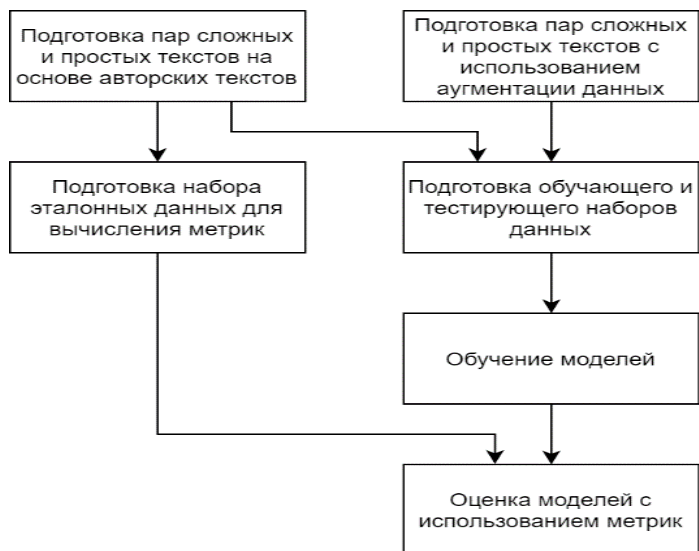


Рис. 1. Общая схема последовательности этапов разработки прототипа

3.2. Уточненная метрика лексической сложности

Для оценки сложности текстов была разработана формула для метрики лексической сложности (Lexical Complexity, LC) путем уточнения представленных ранее формул на основе метрики Флеша и иных оценок читаемости текста:

$$\begin{aligned}
 & \text{LC} = \frac{1}{N} \sum_{i=1}^N \left(\frac{L_i}{L_{\text{max}}} \right) \\
 & \text{LC} = \frac{1}{N} \sum_{i=1}^N \left(\frac{L_i}{L_{\text{max}}} \right) \cdot \left(\frac{L_i}{L_{\text{max}}} \right)
 \end{aligned}$$

В приведенных формулах RI – индекс читаемости, вычисляемый на основе метрик читаемости: FKGL (метрика Флеша-Кинкайда), FRE (метрика Флеша), CLI (метрика Колман-Лиану), SMOG (индекс SMOG), ARI (автоматический индекс удобочитаемости), LIX (индекс LIX).

LL – лексический уровень, вычисляемый на основе метрик лексического разнообразия: TTR (коэффициент лексического разнообразия, Token Type Ratio), RTTR (корневой коэффициент лексического разнообразия), CTTR (скорректированный коэффициент лексического разнообразия).

Параметры формул подробно описаны в статье [Поляков, 2024]. Все перечисленные параметры рассчитывались с использованием библиотеки Russian Texts Statistics¹, в которой все числовые коэффициенты адаптированы для русского языка на основе набора данных проекта Plain Russian². Формулы для вычисления параметров также представлены на сайте³.

Итоговая метрика лексической сложности не имеет установленной области допустимых значений, однако позволяет сравнивать сложность текстов, чем больше значение метрики, тем более сложным полагается текст.

3.3. Подготовка наборов данных

В качестве источника авторских текстов для общего набора данных использовалась Большая российская энциклопедия⁴ в области информатики. Экспертная оценка извлеченных текстов, а также оценка на основе метрики LC показала, что эти тексты являются простыми. Для получения усложненных версий этих текстов была использована большая языковая модель Deepseek R1-Distill-Llama-70B-Q8⁵ с промптом на усложнение текста (рис. 2).

Сгенерируй формальный, научный текст, содержащий сложные предложения и терминологию, который будет примерно равен по длине следующему простому тексту:

{target}

Ключевые требования:

- Текст должен быть формальным и научным по стилю,
- Текст должен содержать сложные предложения и терминологию,
- Длина сложного текста должна быть примерно равна длине простого текста.

Рис. 2. Промт на усложнение текста

Для аугментации данных использовалась та же модель: сначала с использованием отдельного промпта были сгенерированы простые тексты в энциклопедическом стиле в области информатики, затем были получены усложненные версии этих текстов. Полученный при помощи большой языковой модели набор текстов был валидирован экспертами.

Построение вручную эталонного набора пар текстов (простой-сложный) выполнялось на основе требований:

¹ <https://github.com/SergeyShk/ruTS>.

² <https://github.com/infoculture/plainrussian>.

³ https://sergeyshk.github.io/ruTS/stats/readability_stats_funcs/.

⁴ <https://bigenc.ru/>.

⁵ https://ollama.com/library/deepseek-r1:70b-llama-distill-q8_0.

- Средняя LC пар текстов эталонного набора должна соответствовать средней LC пар текстов набора данных для обучения,
- Средняя длина простых текстов эталонного набора должна соответствовать средней длине простых текстов набора данных для обучения.

Тексты, не попавшие в эталонный набор, были оставлены в наборе данных для обучения. В табл. 1 представлены характеристики полученных наборов данных: количество пар, средняя длина текстов в словах, средняя лексическая сложность.

Таблица 1

Характеристика	Значение
Общее количество пар текстов	568
Предметная область	информатика
Набор данных для обучения	
Количество пар текстов в наборе	468
Количество пар на основе авторских текстов	241
Количество пар на основе аугментации данных	227
Средняя длина простых текстов	155 слов
Средняя длина простых авторских текстов	67 слов
Средняя длина простых аугментированных текстов	288 слов
Средняя LC сложных текстов	105,9
Средняя LC простых текстов	86,4
Набор эталонных данных	
Количество пар текстов в наборе	100
Средняя длина простых текстов	145 слов
Средняя LC сложных текстов	108,0
Средняя LC простых текстов	86,3

3.4. Обучение моделей

Для эксперимента по сравнительному анализу моделей архитектуры Transformer в задаче упрощения текста были выбраны модели T5 и BART, предобученные версии для русского языка: FRED-T5-large⁶ и ru-bart-large⁷. Данные модели ранее были исследованы в решении этой задачи и показали хорошо применимые результаты [Das et al., 2025].

Модели были обучены на подготовленном наборе данных с различными конфигурациями обучения. В результате была выявлена наиболее удачная конфигурация, дающая наилучший результат обучения по стандартной метрике потерь (loss). В табл. 2 представлены характеристики этой конфигурации и полученные на ней результаты обучения для каждой модели.

⁶ <https://huggingface.co/ai-forever/FRED-T5-large>.

⁷ <https://huggingface.co/sn4kebyt3/ru-bart-large>.

Таблица 2

Характеристика	T5	BART
Конфигурация обучения		
Количество эпох	10	8
Разбиение данных	90/10%	90/10%
Размер батча	2	8
Скорость обучения	1e-4	3e-5
Стратегия оценки	по эпохам	каждые 50 шагов
Метрика обучения	eval loss	eval loss
Результаты обучения		
train loss в начале	4,01	4,02
train loss в конце	0,65	0,44
eval loss в начале	0,76	0,94
eval loss в конце	0,66	0,77

Результаты представлены в виде показателей метрики обучения в начале и в конце, после прохождения всех эпох. При успешном обучении происходит уменьшение метрики, чем ближе она к 0, тем успешнее результат.

3.5. Оценка моделей

Для обученных моделей была произведена оценка с использованием набора эталонных данных и оценочных метрик BLEU, ROUGE, SARI. Несмотря на то, что среди них непосредственно метрикой для задачи упрощения текста является SARI, метрика BLEU для машинного перевода и семейство метрик ROUGE для суммаризации текста тоже дают полезную информацию для сравнительного анализа, так как они иным образом проверяют соответствие сгенерированного моделью текста эталонному. Более подробно смысл сравнения моделей по метрикам BLEU и ROUGE раскрыт далее в разделе 4.

Кроме оценочных метрик была вычислена средняя разница лексической сложности LCdiff для сгенерированных и эталонных текстов:

,

где LCс – это LC сложного эталонного текста, LCs – это LC простого эталонного текста, LCm – это LC сгенерированного простого текста.

Таким образом, положительное значение LCdiff означает, что модель сгенерировала текст более сложный по метрике LC, чем эталонный, а отрицательное значение – что модель сгенерировала более простой текст.

В табл. 3 представлены результаты оценки моделей T5 и BART.

Таблица 3

Метрика	T5	BART
BLEU	34,1	41,7
ROUGE-1	53,9	53,9
ROUGE-2	41,8	42,4
ROUGE-L	52,5	52,4
SARI	47,1	56,2
LCdiff	+5,45	+1,55

4. Сравнительный анализ моделей

Сравнение моделей по оценочным метрикам показывает, что модель BART показывает лучшие результаты по метрикам BLEU и SARI и сопоставимые результаты по метрикам семейства ROUGE (см. табл. 3).

Лучший результат по метрике BLEU означает, что модель BART генерирует упрощенные тексты, которые более точно соответствуют эталонным простым текстам в аспекте полного соответствия n-грамм. Таким образом, BART генерирует тексты, которые более соответствуют эталонным по форме, чем T5. Этот результат соответствует результату, полученному в [Das et al., 2025] с оценками BLEU/T5 37,8 и BLEU/BART 41,5.

Сопоставимые результаты по метрикам семейства ROUGE означают, что модели генерируют тексты, содержащие одинаковые униграммы (отдельные слова), биграммы (пары слов) и последовательности слов наибольшей длины. Следовательно, обе модели в равной степени сохраняют лексику и формулировки текстов, одинаково качественно суммаризируют смысл исходных сложных текстов. Этот результат не соответствует исследованию [Das et al., 2025], в котором было выявлено, что модель T5 имеет меньшие оценки ROUGE (например, ROUGE-1/T5 48,2 и ROUGE-1/BART 52,7).

Лучший результат по метрике SARI показывает, что модель BART лучше справляется с задачей упрощения текста, чем T5, что также не соответствует результату [Das et al., 2025], где показано, что разница по этой метрике между моделями незначительна (SARI/T5 39,4 и SARI/BART 38,7).

Сравнение по средней разнице лексической сложности LCdiff соответствует сравнению по метрике SARI – обе модели генерируют более сложные тексты по сравнению с эталонными, однако BART генерирует более близкие по лексической сложности тексты.

В табл. 4 представлены показательные примеры упрощения текстов эталонного набора моделями T5 и BART.

Таблица 4

Эталонный текст	Выход T5	Выход BART
Наиболее важным примером сложности алгоритма является время его работы, измеряемое числом элементарных шагов ...	Алгоритмы, для которых максимальная временная сложность, наблюдаемая в экстремальных условиях...	Одним из ключевых показателей сложности алгоритма является время его работы...
Наиболее производительные универсальные микропроцессоры разрабатывают и производят компании Intel, AMD и IBM. В России пять основных разработчиков микропроцессоров...	В России основными производителями микропроцессоров являются МЦСТ, Baikal Electronics, Научно-технический центр «Модуль» ...	Компании Intel, AMD и IBM разрабатывают и производят самые мощные универсальные микропроцессоры. В России наиболее крупными производителями универсальных микропроцессоров являются...

Общие выводы сравнительного анализа следующие:

- На основе экспертной оценки сгенерированных моделями упрощенных текстов на основе набора эталонных данных и метрик ROUGE, SARI, LCdiff можно утверждать, что обе модели приемлемо справляются с задачей упрощения текстов на русском языке, сохраняя при этом смысл, лексику и формулировки, однако BART справляется несколько лучше.
- Результаты проведенного эксперимента отражены в табл. 3 и частично соответствуют результатам, полученным в исследовании [Das et al., 2025], в частности, в ключевом выводе относительно модели BART по оценке BLEU. Разница в выводах по метрикам ROUGE и SARI может быть объяснена различием в источниках данных (использовались тексты на английском языке из ресурсов Simple Wikipedia и Newsela), в объеме наборов данных для обучения, в конфигурации обучения.
- Разница в метрике LCdiff корректно предсказывает разницу в метрике SARI, что является обоснованием корректности используемой уточненной формулы лексической сложности LC.

Заключение

В результате исследования было проведено дообучение моделей T5 и BART архитектуры Transformer с последующим сравнительным анализом на основе оценочных метрик BLEU, ROUGE, SARI и уточненной метрики лексической сложности. Анализ показал, что обе модели приемлемо справляются с задачей упрощения текста, при этом BART показывает лучшие результаты, а также вносит меньше изменений в форму текста. Кроме того, анализ показал соответствие результатов существующим аналогичным исследованиям и обоснованность используемой уточненной

формулы лексической сложности. Свойства метрик: грамматичность, сохранение смысла, простота требуют дальнейшего оценивания в последующих исследованиях.

Дальнейшее направление исследований связано с разработкой и применением метрики терминологической сложности, а также с расширением другими источниками текстов, предметными областями и тематиками.

Список литературы

- [Alva-Manchego et al., 2020] Alva-Manchego F., Scarton C., Specia L. Data-Driven Sentence Simplification: Survey and Benchmark. *Computational Linguistics*. – 2020. – Vol. 46(1). – P. 135-187.
- [Anjum et al., 2023] Anjum A., Lieberum N. Automatic Simplification of Scientific Texts using Pre-trained Language Models: A Comparative Study // In: *Proc. CLEF Symposium*. – 2023. – URL: <https://ceur-ws.org/Vol-3497/paper-242.pdf> (дата обращения: 29.05.2025).
- [Das et al., 2025] Das S., Basak D., Bhattacharjee A. Text Simplification Using T5 Model and BART Model // In: *Proc. 8th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*. Kolkata, India, 2025. – P. 1-6. – doi: 10.1109/IEMENTech65115.2025.10959517.
- [Dmitrieva, 2023] Dmitrieva A. Automatic text simplification of Russian texts using control tokens // In: *Proc. 9th Workshop on Slavic Natural Language Processing (SlavicNLP 2023)*. Dubrovnik, Croatia, 2023. – P. 70-77. – doi: 10.18653/v1/2023.bsnlp-1.9.
- [Flesch, 1948] Flesch R. A new readability yardstick // *Journal of Applied Psychology*. – 1948. – Vol. 32(3). – P. 221-233.
- [Ivanov et al., 2023] Ivanov V., Gamal E.M. A new dataset for sentence-level complexity in Russian // In: *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2023”*. June 14–16, 2023.
- [Kincaid et al., 1975] Kincaid J.P., Fishburne R.P., Rogers R.L., Chissom B.S. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. Institute for Simulation and Training. – 1975. – URL: <https://stars.library.ucf.edu/istlibrary/56> (дата обращения: 30.05.2025).
- [Lin, 2004] Lin C. ROUGE: a package for automatic evaluation of summaries // In: *Proc. ACL 2004 Workshop on Text Summarization Branches Out*. Barcelona, Spain, 2004. – P. 74-81.
- [McNamara et al., 2014] Danielle S. McNamara, Arthur C. Graesser, Philip M. McCarthy, and Zhiqiang Cai. 2014. *Automated Evaluation of Text and Discourse with Coh-Metrix*. Cambridge University Press, USA.
- [North et al., 2024] North K., Ranasinghe T., Shardlow M., Zampieri M. MultiLS: An End-to-End Lexical Simplification Framework // In: *Proc. 3rd Workshop on Text Simplification, Accessibility and Readability (TSAR 2024)*. Miami, Florida, USA, 2024. – P. 1-11. – doi: 10.18653/v1/2024.tsar-1.1.
- [Papineni et al., 2002] Papineni K., Roukos S., Ward T., Zhu W. Bleu: A method for automatic evaluation of machine translation // In: *Proc. 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002)*. Philadelphia, Pennsylvania, 2002. – P. 311-318. – doi: 10.3115/1073083.1073135.

- [**Sakhovskiy et al 2021**] Sakhovskiy A., Izhevskaya A., Pestova A. et al. RuSimpleSentEval-2021 Shared Task: Evaluating Sentence Simplification for Russian // In: Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference “Dialogue”. – 2021. – P. 607-617.
- [**Sheang et al., 2021**] Sheang K.C., Saggion H. Controllable Sentence Simplification with a Unified Text-to-Text Transfer Transformer // In: Proc. 14th International Conference on Natural Language Generation. Aberdeen, Scotland, UK, 2021. – P. 341-352. – doi: 10.18653/v1/2021.inlg-1.38.
- [**Sukiman et al., 2023**] Sukiman S.A., Husin N.A., Hamdan H., Murad M.A.A. A Hybrid Personalized Text Simplification Framework Leveraging the Deep Learning-based Transformer Model for Dyslexic Students // Journal of Advanced Research in Applied Sciences and Engineering Technology. – 2023. – Vol. 34(1). – P. 299-313. – doi: 10.37934/araset.34.1.299313.
- [**Sulem et al., 2018**] Sulem E., Abend O., Rappoport A. Semantic Structural Evaluation for Text Simplification // In: Proc. 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2018). New Orleans, Louisiana, 2018. – P. 685-696. – doi: 10.18653/v1/N18-1063.
- [**Sun et al., 2023**] Sun R., Xu W., Wan X. Teaching the Pre-trained Model to Generate Simple Texts for Text Simplification. In: Findings of the Association for Computational Linguistics: ACL 2023. Toronto, Canada, 2023. – P. 9345-9355. – doi: 0.18653/v1/2023.findings-acl.595.
- [**Truică et al., 2023**] Truică C.O., Stan A.I., Apostol E.S. SimpLex: a lexical text simplification architecture // Neural Computing and Applications. – 2023. – Vol. 35. – P. 6265-6280. – doi: 10.1007/s00521-022-07905-y.
- [**Xu et al., 2016**] Xu W., Napoles C., Pavlick E. et al. Optimizing statistical machine translation for text simplification // Transactions of the Association for Computational Linguistics. – 2016. – Vol. 4. – P. 401-415.
- [**Поляков, 2024**] Поляков А.М., Зойдзе Э.А. Количественные критерии оценки сложности текста для методических целей // Наука в мегаполисе. – 2024. – № 3(59).