



ПСИХОЛИНГВИСТИКА | PSYCHOLINGUISTICS

Научная статья | Original paper

Оценка сложности текстов и вызываемого ими интереса в сети Интернет

Н.А. Алмаев¹ ✉, Т.Д. Сальников¹

¹ Институт психологии Российской академии наук, Москва, Российская Федерация

✉ almaev@mail.ru

Резюме

Контекст и актуальность. Понимание текстов людьми приобретает все большую актуальность, особенно в контексте обучения всех видов и уровней, понимания инструкций, юриспруденции, а также разнообразных коммуникаций посредством сети Интернет. **Цель:** разработать адекватную систему оценки сложности текста и предложить схему пилотажного эксперимента, учитывающую фактор интереса к нему. **Гипотеза.** 1) Количество синтаксических отношений, приходящихся на одно предложение, определяет субъективную оценку сложности текста. 2) Сложность текста и интерес к нему связаны негативно. **Методы и материалы.** Предложен метод оценки сложности текстов на основе девяти статистических показателей — 4 исходных: количество букв, слов, синтаксических отношений (СинтО) и предложений, — и 5 производных от них. Была разработана схема эксперимента, включающая субъективную оценку сложности текста и интереса к нему. Для пилотажного исследования были взяты тексты 4 авторов с Яндекс Дзен. Из них было получено 8 стимулов путем либо упрощения, либо усложнения. Упрощение заключалось в уменьшении количества СинтО на предложение, а усложнение – в увеличении. Объем стимула — около 1000 знаков. Тексты и средства для их оценки предъявлялись в сети Интернет на специальном временном сайте. Участие было полностью анонимным, всего были зарегистрированы ответы с 33 уникальных IP-адресов. **Результаты.** В целом увеличение количества СинтО на предложение ведет к субъективному усложнению текстов, однако значительную роль играет лексика, тематика и представленность определенных СинтО в тексте. **Выводы.** Предложенная методика представляется перспективной для изучения процессов управления ресурсами, внимания и сосредоточения при условии контроля переменных в лабораторных условиях и привлечения таких методов психофизиологии, как айтрекинг и ЭЭГ.

Ключевые слова: сложность текста, синтаксические отношения, коммуникация в сети Интернет, интерес к тексту, субъективные оценки

Финансирование. Исследование выполнено при финансовой поддержке Российского научного фонда (РНФ) в рамках научного проекта № 23-28-10316, <https://ias.rscf.ru/user/doc/a.w.p.2023.76.legacy/290032>.

Благодарности. Авторы выражают благодарность С.О. Скорику за поиск испытуемых в соцсетях.

Для цитирования: Алмаев, Н.А., Сальников, Т.Д. (2026). Оценка сложности текстов и вызываемого ими интереса в сети Интернет. *Экспериментальная психология*, 19(1), 136–149. <https://doi.org/10.17759/exppsy.2026190109>





Evaluation of texts complexity and interest to them in Internet

N.A. Almayev¹ ✉, T.D. Salnikov¹

¹ Institute of Psychology, Russian Academy of Sciences, Moscow, Russian Federation

✉ almaev@mail.ru

Abstract

Context and Relevance. Human understanding of texts is becoming increasingly important, especially in the context of learning at all types and levels, understanding instructions, jurisprudence, and various online communications. **Objective:** to develop an adequate system for assessing text complexity and propose a pilot experiment design that takes into account the factor of interest in it. **Hypothesis.** 1) The number of syntactic relations per sentence determines the subjective assessment of text complexity. 2) Text complexity and interest in it are negatively related. **Methods and materials.** A method for assessing text complexity is proposed based on nine statistical indicators — four original ones: the number of letters, words, syntactic relations ('rels'), and sentences, and five derivatives from them. An experimental design was developed that includes a subjective assessment of text complexity and interest in it. Texts from four authors from Yandex Zen were used for the pilot study. From these, eight stimuli were obtained by either simplifying or complicating them. Simplification consisted of reducing the number of syntactic relations per sentence, while complexity consisted of increasing it. The stimulus length was approximately 1000 characters. Texts and assessment tools were presented online on a special temporary website. Participation was completely anonymous; responses from 33 unique IP addresses were recorded. **Results.** Overall, increasing the number of syntactic entities per sentence leads to subjective text complexity, although vocabulary, subject matter, and the presence of specific syntactic entities in the text play a significant role. **Conclusions.** The proposed method appears promising for studying resource management, attention, and concentration processes, provided that variables are controlled in a laboratory setting and psychophysiological methods such as eye tracking and EEG are incorporated.

Keywords: complexity of text, syntactic relations, communication on the Internet, interest, subjective assessments

Funding. The study was supported by the Russian Science Foundation, project number 23-28-10316.

Acknowledgements. The authors are grateful to Stanislav O. Skorik for assistance in data collection.

For citation: Almayev, N.A., Salnikov, T.D. (2026). Evaluation of texts complexity and interest to them in Internet. *Experimental Psychology (Russia)*, 19(1), 136—149. (In Russ.). <https://doi.org/10.17759/exppsy.2026190109>

Введение

В рамках дискуссий на общественно значимые темы в сети Интернет порождается огромное количество текстов. Возникает вопрос: каковы условия того, что они будут прочитаны людьми, а не только роботами в рамках различных поисковых алгоритмов и моделей машинного обучения? Выражение «прочитан людьми» предполагает, разумеется, не только формальное посещение страницы, но и понимание текста. Однако текст встречает человека не только своим расположением на странице и внешним видом (вопросы, которые исследуются, как правило, более всего в рамках темы “usability” (Ромич, Бороевич, 2025; Soegard, Dam, 2015), но и своим объемом, количеством букв, слов, предложений, лексикой и сложностью. Интуитивно понятно, что сложность играет важную роль, но как ее измерить? С другой стороны, сложности противостоит интерес, что мы также знаем из повседневной



жизни. Сильная мотивация заставляет людей читать весьма сложные тексты, не понимать, бросать, и снова, и снова приниматься читать, пытаясь составить адекватную репрезентацию написанного. В данной статье описывается применение методики, разработанной нами в 2024 году на основе компьютерной технологии регистрации синтаксических отношений. Сама по себе данная технология является относительно новой. В настоящее время благодаря развитию компьютерной лингвистики синтаксические отношения могут быть выделены автоматически в соответствии с международной системой, являющейся универсальной для множества языков, в том числе и для русского. Соответствие установлено в рамках проекта СинТагРус, разрабатываемого с 1998 года в ИППИ РАН, при участии ИРЯ РАН. Всего выделяется 49 синтаксических отношений (далее СинтО) (Савчук и др., 2024; Список, 2023). Нами их регистрация осуществляется при помощи пакета компьютерной обработки текста *Natasha* (<https://github.com/natasha/slovnet>) в экосистеме Python.

Явных признаков применения именно данной системы СинтО для задач психологии среди уже опубликованных статей нам обнаружить не удалось. Даже в наиболее свежей совместной работе сотрудников НЦПЗ РАМН и ИСА РАН (Девяткин и др., 2021) она еще не используется, не говоря уже о более ранних (Ениколопов и др., 2019 а, б, в, г).

Также не обнаружили мы признаков ее применения в таких научно-прикладных областях, как общая дидактика (Вахрушева и др., 2021) и обучение иностранным языкам (Виноградова и др., 2021; Лапошина, 2023), в которых задача определения сложности и удобочитаемости текстов имеет первостепенную важность. Пока, похоже, не применяется она и в области юриспруденции, где формулировки законов могут быть чрезвычайно громоздкими (Кнутов, Чаплинский, Алимпеев, 2022; Кнутов и др., 2021).

Подходы к анализу сложности текстов развиваются сейчас и в других странах (см. обзоры в Liu, Lee, 2023; Pan et al., 2024). Однако судить о том, какие конкретно показатели лингвистической статистики применяются на различных компьютерных ресурсах, по изложению в статьях зачастую затруднительно. Дело в том, что таких показателей может быть от нескольких десятков до нескольких сотен и взаимосогласование их происходит в ходе обучения нейронных сетей. При этом различные показатели работы нейросетей зачастую соотносятся между собой даже без экспериментов с участием людей.

Конечно, оценка сложности текстов в столь различных задачах, как интернет-коммуникация, чтение учебных текстов и разработка законов, имеет значительные отличия. Лица, участвующие в общественных обсуждениях, как правило, являются носителями языка, хотя и с сильно различающимися языковыми способностями. При этом они находятся в поиске релевантной информации, отвечающей их внутренним запросам. Соответственно, можно выделить два граничных уровня сложности текстов: 1) уровень «сканирования» в ходе интернет-серфинга для оценки релевантности встречаемых текстов; 2) уровень отвержения, запретительный уровень сложности, при котором дальнейшее чтение невозможно, даже при условии релевантности информации. Конечно, эти уровни постоянно колеблются, зависят как от множества индивидуальных различий, так и от состояния субъекта, но, тем не менее, они есть и могут быть оценены с помощью разнообразной статистики текста, в первую очередь синтаксических отношений.

Предварительно получить представление о границах этих значений и величине их вариации можно при помощи субъективных оценок испытуемыми специально созданных экспериментальных текстов. Фактор интереса к текстам является первостепенным, направ-



ляющим усилия читающего, — соответственно, он не может быть проигнорирован при исследовании интернет-коммуникаций и должен быть адекватно учтен.

Цель исследования — апробировать методику оценки сложности текста субъектами в условиях, приближенных к интернет-серфингу. Жанр исследования — пилотажное.

Гипотезы исследования: 1) количество синтаксических отношений, приходящихся на одно предложение, определяет субъективную оценку сложности текста; 2) сложность текста и интерес к нему связаны негативно.

Конечно, мы понимаем, что различные СинтО имеют различный вклад в сложность понимания текста. Поэтому и выдвижение гипотезы о решающем вкладе отношения СинтО/предложение имеет технический характер. Там, где оно нарушается, и следует искать усложняющие или, наоборот, облегчающие общее понимание СинтО.

Материалы и методы

Материал исследования — тексты, опубликованные на интернет-ресурсе Яндекс.Дзен. Поскольку многое в оценке сложности зависит от лексики, а в оценке интереса — от тематики, тексты унифицировались по тематическим направлениям. Всего было четыре тематических направления — публицистика, рассказы, научный и научно-популярный тексты: 1) публицистика депутата Государственной Думы М.Г. Делягина — сложный; 2) рассказы Махила, голландца, живущего в РФ, о путешествиях по ней — простой; 3) датировка Ригведы, научно-популярный, но ближе к научному, — сложный; 4) история языка иврит — научно-популярный в чистом виде — средний.

Затем для проверки гипотезы 1 сложные тексты подверглись упрощению путем простой разбивки предложений на более короткие, а короткие — усложнению путем объединения коротких предложений в более длинные. Таким образом, сами синтаксические отношения сохранялись, изменялось лишь их количество в одном предложении. Пара стимульных отрывков одного тематического направления представляла собой либо различные фрагменты одного и того же текста, либо отрывки из двух текстов одной авторской серии. Объем текстов, около тысячи знаков, был взят, чтобы не слишком утомлять испытуемых, но все же оставить этот фактор ощутимым. Предварительная качественная классификация текстов на простые, средние и сложные была следующей: от менее 10 до 15 СинтО на предложение — простой, 15—20 — средний, более 20 — сложный.

Таблица 1 / Table 1

Статистика по стимульным текстам
Stimuli texts statistics

Текст	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2
Количество знаков в тексте: / N signs	875	847	1133	1085	868	886	1036	1110
Количество слов в тексте: / N words	130	141	222	209	145	134	182	194
Количество предложений / N sentences	6	12	22	10	7	21	11	18



Текст	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2
Ср. длина слова (букв) / mean word lengths	6	6	5	5	5	6	5	5
Ср. длина предложения (слов): / mean sentence lengths	21	11	10	20	20	6	16	10
Количество СинтО: / N rels	148	170	262	255	196	193	216	222
Букв на СинтО / letters/rels	5	4	4	4	4	4	4	5
СинтО на слово / rels/words	1	1	1	1	1	1	1	1
СинтО/предл. / Rels/sentence	24	14	11	25	28	9	19	12
'acl'	5	2	2	0	1	0	12	1
'acl:reld':	2	0	0	3	2	1	1	9
'advcl'	0	0	1	2	2	0	1	1
'advmod':	12	7	32	29	12	8	14	21
'amod':	24	13	18	16	15	15	19	17
'appos':	0	6	0	1	4	9	1	1
'aux':	0	0	0	0	1	0	0	2
'aux:pass':	0	1	0	0	1	0	0	1
'case':	13	15	30	15	13	9	20	30
'cc':	5	7	12	13	5	5	8	4
'ccomp':	2	7	17	10	4	0	0	0
'compound':	0	0	0	0	0	0	1	0
'conj':	7	11	8	22	9	11	6	9
'cop':	0	0	1	1	1	0	1	1
'csubj':	1	2	0	2	0	0	2	2
'det':	6	3	7	4	7	1	8	8
'expl':	0	0	0	0	0	0	0	1
'fixed':	1	0	4	2	0	0	2	2
'flat:foreign':	0	0	0	0	2	4	0	0
'flat:name':	0	0	0	0	2	13	1	0
'list':	0	0	0	0	1	0	0	0
'iobj':	1	1	1	4	0	0	7	1
'mark':	2	1	6	6	5	0	5	1
'nmod':	12	19	8	3	12	18	9	11
'nsubj':	8	10	27	22	16	14	19	14
'nsubj:pass':	0	1	0	0	1	1	0	0
'nummod':	2	4	1	8	3	0	0	1



Текст	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2	Делягин 1 сложный	Делягин 2
'nummod:gov'	0	0	1	1	0	0	1	1
'obj':	7	4	5	16	2	1	14	6
'obl':	10	12	22	13	7	8	16	30
'obl:agent':	0	1	0	0	0	0	0	0
'orphan':	0	0	1	0	0	0	0	0
'parataxis':	3	5	6	5	8	11	7	7
'punct':	23	36	49	53	54	63	39	39
'xcomp':	2	2	3	4	6	1	2	1

Расшифровка СинтО содержится в Списке Синтаксических Отношений (2023).

Программный код для обработки текстов находится по адресу: https://github.com/Tim-Sa/text_diff_exp_notebook

Для проведения данного исследования был создан специальный сайт. Поскольку сайт более не поддерживается в связи с окончанием проекта, примеры текстов разной сложности даны в Приложении 2.

В целях удобства для испытуемых регистрации не производилось, тестирование было анонимным. Испытуемые не получали компенсационных выплат. Приглашение испытуемых производилось лично и через дружественные ТГ-каналы. Судя по аудитории каналов, возраст испытуемых порядка 30 ± 10 лет. Тексты предъявлялись в квазислучайном порядке.

Инструкция. Испытуемых просили оценить каждый текст по двум шкалам: 1) сложности и 2) интересу, вызываемому текстом. Оценки варьировались от 1 до 8, где 1 — наименьшая выраженность признака, а 8 — наибольшая. Шкалы предъявлялись в виде двух бегунков с демонстрацией текущего значения выраженности признаков в целых величинах.

Статистические процедуры: гипотезы и критерии

Обработка данных проводилась с помощью пакета Statistica 6.0. Применялись методы непараметрической статистики: Т-критерий Уилкоксона для оценки сдвига в результатах тестирования и критерий Фридмана (Friedman ANOVA) для соотнесения эффектов двух основных факторов интереса и сложности.

Отвержение гипотезы H^0 проводилось при значениях $p \leq .05$, в диапазоне значений $0,05 < p \leq 0,09$ определялись тенденции.

Результаты

Были собраны данные с 33 уникальных IP-адресов. Естественно, все результаты имеют предварительный и иллюстративный характер. Строго говоря, даже оценить полученные распределения при таком количестве наблюдений невозможно. Тем не менее, поскольку моды и медианы почти совпадают, можно предположить, что распределения близки к нормальным и, соответственно, к ним применимы базовые показатели среднего, дисперсии, стандартного отклонения и стандартной оценки.



Таблица 2 / Table 2

Субъективные оценки сложности текстов (N = 33)
Subjective evaluations of texts complexity (N = 33)

Текст	Деягин 1 сл.	Деягин 2 ср.	Махил 1 пр.	Махил 2 ср.	Ригведа 1 сл.	Ригведа 2 пр.	Иврит 1 ср.	Иврит 2 пр.
Медианы / Medians	2	4	1	2	4	3	3	3
Моды / Modes	2	4	1	1	4	3	3	3
Ср. знач / Means	2,82	4,09	1,76	2,32	3,82	3,15	3,12	3,21
Дисперсия / Variance	3,76	3,65	3,11	3,35	4,78	2,63	2,80	3,86
Станд. откл. / Std. Dev.	1,86	1,83	1,70	1,77	2,09	1,55	1,60	1,88
Станд. ошибк. / Std. Err.	0,32	0,32	0,30	0,31	0,36	0,27	0,28	0,33
СинтО/предлож. / Rels/ sentences	24	14	11	25	28	9	19	12

В целом можно видеть, что тексты, имеющие меньший показатель СинтО/предложение, воспринимаются как более простые, за исключением первого, наиболее сложного текста Деягина. Весьма вероятно, такая парадоксальная реакция испытуемых спровоцирована запретительной, вызывающей негативную реакцию сложностью текста. Не исключено также, что само название «сложность» воспринимается некоторыми испытуемыми как вызов их интеллектуальным способностям и провоцирует их на искажение субъективной оценки. Видимо, в будущем следует подумать о более нейтральной формулировке.

Таблица 3 / Table 3

Субъективные оценки интереса, вызываемого текстами (N = 33)
Subjective evaluations of interest caused by the texts (N = 33)

Текст	Деягин 1 сл.	Деягин 2 ср.	Махил 1 пр.	Махил 2 ср.	Ригведа 1 сл.	Ригведа 2 пр.	Иврит 1 ср.	Иврит 2 пр.
Медианы / Medians	2	3	4	3	3,5	3	4	5
Моды / Modes	2	3	4	3	3	3	4	5
Ср. знач / Means	2,85	2,91	3,38	3,32	3,68	3,24	4,35	4,76
Дисперсия / Variance	3,92	2,96	4,55	3,79	3,62	2,94	2,43	2,50
Станд. откл. / Std. Dev.	1,90	1,65	2,05	1,86	1,82	1,64	1,49	1,51
Станд. ошибк. / Std. Err.	0,33	0,29	0,36	0,32	0,32	0,29	0,26	0,26
СинтО/предлож. / Rels/ sentences	24	14	11	25	28	9	19	12

Таблица 4 / Table 4

Сложность, сравнение внутри тематических групп, Т-критерий Вилкоксона
Complexity, comparison within thematic groups, Wilcoxon T

	Valid	T	Z	p-level
Деягин сл. & Деягин ср.	33	64,0	2,7	0,008042
Махил пр. & Махил ср.	33	57,5	2,2	0,025083



Различия между текстами о Ригведе и иврите не достигают значимости по субъективной оценке сложности.

Таблица 5 / Table 5

Сложность, сравнение между тематическими группами, Т-критерий Вилкоксона
Complexity, comparison between thematic groups, Wilcoxon T

	Valid	T	Z	p-level
Деягин Ср. & Махил ср.	33	10,0	4,2	0,000026
Деягин Ср. & Иврит ср.	33	105,0	2,2	0,025642

Средний текст Деягина субъективно сложнее средних текстов как о Ригведе, так и об иврите.

Таблица 6 / Table 6

Сравнения текстов по вызываемому ими интересу,
между тематическими блоками
Comparison of texts by the interest caused by them,
between thematic blocs

	Valid	T	Z	p-level
Деягин ср. инт & иврит ср. инт.	33	58,00	3,15	0,0016
Деягин сл. инт & иврит пр. инт	33	24,50	3,45	0,0006
Ригведа сл. инт & иврит ср. инт.	33	95,50	2,03	0,0422
Ригведа пр. инт & иврит пр. инт	33	20,00	3,59	0,0003
Махил ср. инт & иврит ср. инт.	33	37,50	2,89	0,0039
Махил пр. инт & иврит пр. инт.	33	92,00	2,89	0,0039
Махил ср. инт & иврит пр. инт.	33	61,50	3,37	0,0007

По вызываемому интересу внутри тематических блоков различия незначимы, а между тематическими блоками — велики.

При этом, когда сравниваются тексты разных тематических блоков при отличающейся сложности этих текстов, уровень значимости различий возрастает. Вместе с тем, имеются исключения, по-видимому, продиктованные спецификой конкретных СинтО в текстах. Например, «Ригведа сложная» вызывает больший интерес, чем «Ригведа простая». Посмотрев на таблицу с СинтО (табл. 1), замечаем, что во втором тексте намного больше 'flat:name' (многосоставные имена) — 13 против 2. Вероятно, это и снижает интерес, при том что самих СинтО на предложение меньше. Разумеется, все эти отношения должны изучаться гораздо более подробно на различных выборках.

Таким образом, можно сделать предварительный вывод, что, хотя лексико-тематическая направленность имеет основное значение, сложность также играет немалую роль. В целом взаимоотношение сложности текста и интереса к нему — большая тема, нуждающаяся в дальнейших исследованиях.

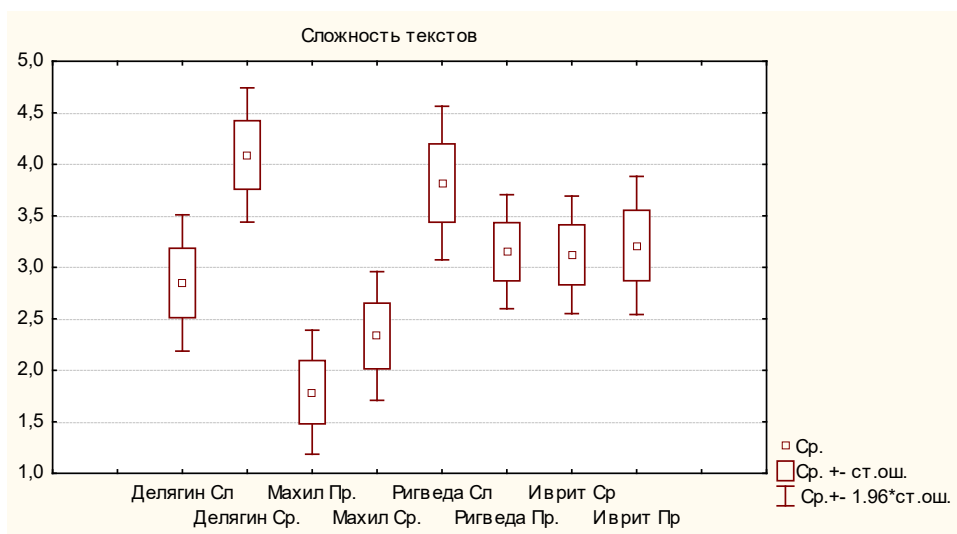


Рис. 1. Сложность текстов. Средние, средние ± стандартная ошибка, средние ± доверительный интервал (N = 33)

Fig. 1. Complexity of texts. Means, Means ± standard error, Means ± confidence interval (N = 33)

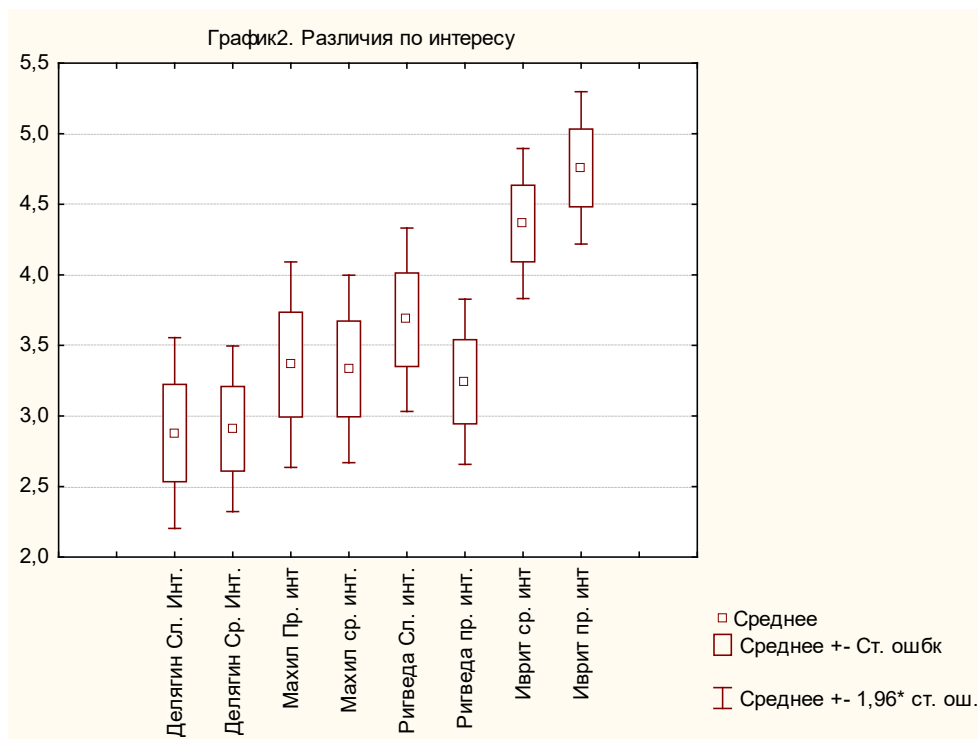


Рис. 2. Интерес, вызываемый текстами. Средние, средние ± стандартная ошибка, средние ± доверительный интервал (N = 33)

Fig. 2. Interest caused by the texts. Means, Means ± standard error, Means ± confidence interval (N = 33)



Обсуждение результатов

Предварительные оценки субъективной сложности сообщений в Интернете демонстрируют ограничения этой методики по надежности результатов. Причем дело даже не в количестве наблюдений. За дисперсией оценок может скрываться как различный уровень навыков чтения, общего и вербального интеллекта, состояния испытуемых во время чтения и оценки, так и произвольность оценок, и недобросовестность со стороны испытуемых. Ведь в целом мы даже не можем быть уверены в том, прочитал ли испытуемый текст или вопрос теста! И это общая проблема любого тестирования в Интернете и, шире, любого тестирования без применения средств психофизиологической регистрации. Вопросы регистрации движений глаз при тестировании обсуждаются в (Алмаев, Бессонова, Мурашева, 2020), а связь шкал лжи и показателей ЭЭГ — в (Алмаев, Мурашева, Петрович, 2020; Almayev, Murasheva, Petrovich, 2024). Частично решить проблему с достоверностью прочитывания текста можно при помощи задания: дать обратную связь, ответить на вопросы, указать, встречалось ли данное слово в тексте и т. п. По мере возрастания сложности все большее значение приобретает и фактор мотивации участия в эксперименте.

Тем не менее связка качеств «сложность — интерес» на материале чтения текстов потенциально может представлять значительную ценность в рамках лабораторного психофизиологического исследования. Такое исследование, включающее в себя, например, регистрацию движений глаз, кожно-гальваническую реакцию и ЭЭГ, выглядит весьма перспективным. Как меняются паттерны движений глаз при возрастающем интересе и, напротив, в случае скуки? Что происходит с различными показателями ЭЭГ, когда сложность текста преодолевает когнитивную активность чтения? Такой момент можно, например, регистрировать при помощи нажатия кнопки и отметки соответствующего события на ЭЭГ. Какие ЭЭГ и глазодвигательные паттерны соответствуют поиску когнитивных ресурсов, позволяющих вчитываться в сложный текст, как при наличии, так и при отсутствии интереса? Также чтение текстов различной сложности может выступать как модель для изучения нейрофизиологических процессов внимания, интереса и отвлечения. Можно, например, организовать отвлечения звуковыми стимулами с несколькими параметрами. Сложностью текста можно хорошо управлять на основе СинтО, подбирая и варьируя соответствующие стимулы. Интересом управлять несколько сложнее, но тоже, в общем-то, возможно, особенно в мелкосерийных лабораторных исследованиях — например, просить самих испытуемых подобрать текст на интересующую их тематику. С другой стороны, методы психологии (время реакции) и психофизиологии (негативность вызванных потенциалов, пупиллометрия, соотношения мощности в различных отведениях и проч.) могут использоваться для замера емкости когнитивных ресурсов, потребляемых СинтО, их комбинациями и структурами вложений одних СинтО в другие.

В целом, тема выглядит интересной и перспективной для психологии и психофизиологии. Возможность программного выделения СинтО позволяет разделить их на несколько больших классов, например коммуникативные и предметные, и изучать их объединения в различные кластеры по отдельности, внутри этих классов. Особо стоит отметить, что программные решения на основе Natasha относительно нетребовательны к вычислительным мощностям. СинтО — психологически осмысленный, «живой» параметр; многообразное изучение и использование его в целях психологических исследований, например индивидуальных различий, связи с мотивацией, с личностными чертами и т. п., представляется перспективным.



При этом уже сейчас, на основе самых предварительных данных, можно дать некоторую первичную рекомендацию. При прочих равных для легкости понимания текст должен характеризоваться величиной порядка 10—15 СинтО на предложение. Учитывая, что в среднем количество СинтО равно количеству слов (см. табл. 1), можно ориентироваться на соответствующее количество слов. Меньшее значение этого отношения может начать восприниматься как бессвязность, или требовать погружения в тему. Естественно, этот вопрос, в свою очередь, также нуждается в эмпирическом изучении.

Заключение

Найдена и самостоятельно освоена на основе программных средств, находящихся в свободном доступе, новая мера оценки сложности текстов, расширяющая существующие возможности в данной области. Разработана базовая методика, позволяющая учитывать тематический интерес и лексическую знакомость для испытуемых тех или иных тем в Интернете. Сложность текста находится в отношениях противоборства с интересом к нему. Хотя интерес к тексту определяется в основном интересом к теме, однако сложность текста способна снизить интерес к нему, в то время как относительная простота — повысить. Субъективные оценки могут не отличаться большой надежностью, особенно на данном этапе предварительных исследований. Тем не менее, данные методики могут развиваться и улучшаться как в онлайн-подходах, так и, особенно, с использованием лабораторных психофизиологических методов с регистрацией движений глаз и различных показателей ЭЭГ.

Ограничения. Данное исследование является пилотажным, оно ограничено как количеством наблюдений, так и контролем условий. В частности, не было проконтролировано, какие фрагменты текста испытуемые в действительности прочитали.

Limitations. This study is a pilot one and is limited by both the number of observations and the control of conditions. Specifically, it was not possible to control which text fragments the subjects actually read.

Список источников / References

1. Алмаев, Н.А., Бессонова, Ю.В., Мурашева, О.В. (2020). *Текст. Взгляд. Мотивация. Методические вопросы*. М.: Изд-во ИП РАН.
Almaev, N.A., Bessonova, Yu.V., Murasheva, O.V. (2020). *Text. View. Motivation. Methodological issues*. Moscow: Publishing house of the Institute of Sociology of the Russian Academy of Sciences. (In Russ.)
2. Алмаев, Н.А., Мурашева, О.В., Петрович, Д.Л. (2020). ЭЭГ корреляты личностных опросников: ОФДСИ Русалова, ОСТ-77 Трофимовой и ТТХ-140 Клонинджера. *Вестник психофизиологии*, 3, 135—138.
Almaev, N.A., Murasheva, O.V., Petrovich, D.L. (2020). EEG correlates of personality questionnaires: Rusalov's OFDSI, Trofimova's OST-77, and Cloninger's TTX-140. *Bulletin of Psychophysiology*, 3, 135—138. (In Russ.)
3. Вахрушева, А.Я., Солнышкина, М.И., Куприянов, Р.В., Гафиятова, Э.В., Климагина, И.О. (2021). Лингвистическая сложность учебных текстов. *Вопросы журналистики, педагогики, языкознания*, 40(1), 89—99.
Vahrusheva, A.Ya., Solnyshkina, M.I., Kupriyanov, R.V., Gafiyatova, E.V., Klimagina, I.O. (2021). Linguistic Complexity of Educational Texts. *Issues of Journalism, Pedagogy, and Linguistics*, 40(1), 89—99. (In Russ.)
4. Виноградова, О.А., Смирнова, Е.А., Вилкова, А.В., Пантелеева, И.М. (2020). Синтаксическая сложность академического текста: корпусное исследование работ русскоязычных студентов, изучающих английский язык, в сравнении с текстами профессиональных авторов. *Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология»*, 7.



- Vinogradova, O.A., Smirnova, E.A., Vilkova, A.V., Panteleeva, I.M. (2020). Syntactic Complexity of Academic Texts: A Corpus-Based Study of Works by Russian-Speaking Students Learning English in Comparison with Texts by Professional Authors. *RSUH Bulletin. Series: Literary Studies. Linguistics. Cultural Studies*, 7. (In Russ.)
5. Девяткин, Д.А., Ениколопов, С.Н., Салимовский, В.А., Чудова, Н.В. (2021). Речевые реакции на фрустрацию: автоматическая категоризация. *Психологические исследования*, 14(78).
Devyatkin, D.A., Enikolopov, S.N., Salimovskij, V.A., Chudova, N.V. (2021). Speech frustration responses: an automatic categorization. *Psychological research*, 14(78). (In Russ.)
 6. Ениколопов, С.Н., Кузнецова, Ю.М., Смирнов, И.В., Станкевич, М.А., Чудова, Н.В. (2019а). Создание инструмента автоматического анализа текста в интересах социо-гуманитарных исследований. Часть 1. Методические и методологические аспекты. *Искусственный интеллект и принятие решений*, 2, 28—38.
Enikolopov, S.N., Kuznetsova, Yu.M., Smirnov, I.V., Stankevich, M.A., Chudova, N.V. (2019a). Creating an Automatic Text Analysis Tool for Social Sciences and Humanities Research. Part 1. Methodological Aspects. *Artificial Intelligence and Decision Making*, 2, 28—38. (In Russ.)
 7. Ениколопов, С.Н., Ковалев, А.К., Кузнецова, Ю.М., Чудова, Н.В., Старостина, Е.В. (2019б). Признаки, характерные для текстов, написанных в состоянии фрустрации. *Вестник Московского университета. Серия 14: Психология*, 3, 66—85.
Enikolopov, S.N., Kovalyov, A.K., Kuznetsova, Yu.M., Chudova, N.V., Starostina, E.V. (2019b). Signs characteristic of texts written in a state of frustration. *Moscow University Psychology Bulletin*, 3, 66—85. (In Russ.)
 8. Ениколопов, С.Н., Медведева, Т.И., Воронцова, О.Ю. (2019в). Лингвистические характеристики текстов при депрессии и шизофрении. *Медицинская психология в России: электронный научный журнал*, 11(5).
Enikolopov, S.N., Medvedeva, T.I., Vorontsova, O.Yu. (2019c). Linguistic text characteristics in depression and schizophrenia. *Medical Psychology in Russia*, 11(5). (In Russ.)
 9. Ениколопов, С.Н., Медведева, Т.И., Воронцова, О.Ю., Иванова, П.О. (2019г). Оценка текстов, написанных больными эндогенными психическими заболеваниями. *Психиатрия*, 81(1), 56—68.
Enikolopov, S.N., Medvedeva, T.I., Vorontsova, O.Yu., Ivanova, P.O. (2019d). Evaluation of texts written by patients with endogenous mental illnesses. *Psychiatry*, 81(1), 56—68. (In Russ.)
 10. Кнутов, А.В., Чаплинский, А.В., Алимеев, Д.Р. (2022). Язык нормативных правовых актов: пора ли бить тревогу? *Вестник Пермского университета. Юридические науки*, 57, 399—426. <https://doi.org/10.17072/1995-4190-2022-57-399-426>
Knutov, A.V., Chaplinsky, A.V., Alimpeev, D.R. (2022). The Language of Regulatory Legal Acts: Is It Time to Sound the Alarm? *Perm University Herald. Juridical Sciences*, 57, 399—426. (In Russ.). <https://doi.org/10.17072/1995-4190-2022-57-399-426>
 11. Кнутов, А.В., Плаксин, С.М., Чаплинский, А.В., Григорьева, Н.Л., Синятуллин, Р.Х., Успенская, А.М. (2021). *Сложность российских законов. Опыт синтаксического анализа*. М.: Издательский дом Высшей школы экономики.
Knutov, A.V., Plaksin, S.M., Chaplinsky, A.V., Grigoryeva, N.L., Sinyatullin, R.H., Uspenskaya, A.M. (2021). *The Complexity of Russian Laws: An Experiment in Syntactic Analysis*. Moscow: Higher School of Economics Publishing House. (In Russ.)
 12. Лапошина, А.Н. (2023). *Лингводидактическое обоснование применения Автоматической оценки сложности учебного текста в преподавании РКИ: Дис. ... канд. пед. наук*. М.
Laposhina, A.N. (2023). *A linguodidactic rationale for the use of automatic assessment of educational text complexity in teaching Russian as a foreign language: Diss. Cand. Sciences (Ped.)*. Moscow. (In Russ.)
 13. Ромич, М., Бороевич, С. (2025). Типографские характеристики и обработка латинских и кириллических слов. *Экспериментальная психология*, 18(1), 108—118.
Romich, M., Boroevich, S. (2025). Typographic characteristics and processing of Latin and Cyrillic words. *Experimental Psychology (Russia)*, 18(1), 108—118. (In Russ.)
 14. Савчук, С.О., Архангельский, Т.А., Бонч-Осмоловская, А.А., Донина, О.В., Кузнецова, Ю.Н., Ляшевская, О.Н., Орехов, Б.В., Подрядчикова, М.В. (2024). Национальный корпус русского языка 2.0: новые возможности и перспективы развития. *Вопросы языкознания*, 2, 7—34.



- Savchuk, S.O., Arhangelsky, T.A., Bonch-Osmolovskaya, A.A., Donina, O.V., Kuznetsova, Yu.N., Lyashevskaya, O.N., Orekhov, B.V., Podryadchikova, M.V. (2024). The National Corpus of the Russian Language 2.0: New Possibilities and Development Prospects. *Linguistics Issues*, 2, 7—34. (In Russ.).
15. Список Синтаксических Отношений. (2023). URL: <https://ruscorpora.ru/media/uploads/2023/12/29/rajadw.pdf> (дата обращения: 10.04.2025).
List of Syntactic Relations (2023). (In Russ.). URL: <https://ruscorpora.ru/media/uploads/2023/12/29/rajadw.pdf> (viewed: 10.04.2025).
16. Almayev, N.A., Murasheva, O.V., Petrovich, D.L. (2024). Rusalov STQ-150 and Rusalov-Trofimova STQ-77 in Relation to the EEG Spectral Power. *Natural Systems of Mind*, 4(1), 6—16. https://doi.org/10.38098/nsom_2024_04_01_01 (viewed: 19.03.2026).
17. Kovriguina, L., Shilin, I., Shipilo, A., Putintseva, A. (2017). Russian Tagging and Dependency Parsing Models for Stanford CoreNLP Natural Language Toolkit. In: *Knowledge Engineering and Semantic Web — 8th International Conference, KESW 2017, Szczecin, Poland, November 8—10, 2017, Proceedings* (pp. 101—111). https://doi.org/10.1007/978-3-319-69548-8_8
18. Liu, F., Lee, J.S. (2023). Hybrid models for sentence readability assessment. In: *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications, BEA, 2023* (pp. 448—454).
19. Natasha/Slovnet (Электронный ресурс): репозиторий. URL: <https://github.com/natasha/slovnet> (viewed: 15.05.2024).
20. Ortega, L. (2003). Syntactic complexity measures and their relationship to l2 proficiency: A research synthesis of college-level l2 writing. *Applied Linguistics*, 24(4), 492—518.
21. Pan, W, Li, X, Chen, X, Xu, R. (2024). Textual form features for text readability assessment. *Natural Language Processing*, 1—42. <https://doi.org/10.1017/nlp.2024.50>
22. Soegaard, M., Dam, R.F. (Eds.). (2015). *The Encyclopedia of Human-Computer Interaction*. Aarhus, Dansk: The Interaction Design Foundation. URL: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed> (viewed: 19.03.2026).

Приложение / Appendix

Приложение 1. Статистика по критерию Фридмана (таблицы 1 и 2).

Приложение 2. Примеры стимульных текстов разной сложности.

<https://doi.org/10.17759/exppsy.2026190109>

Appendix 1. Statistics on Friedman' Anova (Tables 1 and 2).

Appendix 2. Examples of stimulus texts of varying complexity.

<https://doi.org/10.17759/exppsy.2026190109>

Информация об авторах

Николай Альбертович Алмаев, доктор психологических наук, главный научный сотрудник, Институт психологии Российской академии наук (ФГБУН ИП РАН), Москва, Российская Федерация, ORCID: <http://orcid.org/0000-0001-9381-7842>, e-mail: almaev@mail.ru

Тимофей Дмитриевич Сальников, инженер-исследователь, Институт психологии Российской академии наук (ФГБУН ИП РАН), Москва, Российская Федерация, ORCID: <http://orcid.org/0009-0007-2068-1172>, e-mail: salnikovtimofeyof@gmail.com

Information about the authors

Nickolay A. Almayev, Doctor of Sciences (Psychology), Chief Researcher, Institute of Psychology, Russian Academy of Sciences, Moscow, Russian Federation, ORCID: <http://orcid.org/0000-0001-9381-7842>, e-mail: almaev@mail.ru



Timofey D. Salnikov, Research Engineer, Institute of Psychology, Russian Academy of Sciences, Moscow, Russian Federation, ORCID: <http://orcid.org/0009-0007-2068-1172>, e-mail: salnikovtimofeyof@gmail.com

Вклад авторов

Алмаев Н.А. — литературный поиск; идея статистики; идея исследования; технические задания на программные продукты; планирование исследования; выбор стимулов; обработка данных; написание и оформление рукописи.

Сальников Т.Д. — литературный поиск; написание программного кода для анализа текстов; написание сайта со стимулами; размещение сайта в сети Интернет; получение данных.

Оба автора приняли участие в обсуждении результатов и согласовали окончательный текст рукописи.

Contribution of the authors

Nickolay A. Almayev — literature search; idea of statistics; research idea; technical specifications for software products; research planning; selection of stimuli; data processing; writing and formatting of the manuscript.

Timofey D. Salnikov — literature search; writing software code for text analysis; writing the website with stimuli texts; allocation of the website in the Internet; obtaining data.

Both authors participated in the discussion of the results and approved the final text of the manuscript.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest

The authors declare no conflict of interest.

Поступила в редакцию 12.12.2024

Поступила после рецензирования 11.05.2025

Принята к публикации 20.06.2025

Опубликована 30.03.2026

Received 2024.12.12

Revised 2025.05.11

Accepted 2025.06.20

Published 2026.03.30