

СОДЕРЖАНИЕ

Предисловие	5
Для чего упрощать текст закона, и как эту проблему решают в других странах?	7

Часть I

ОПРЕДЕЛЕНИЕ УРОВНЯ СЛОЖНОСТИ РОССИЙСКИХ ЗАКОНОВ

1. Как оценить сложность закона? Подходы к оценке сложности юридических текстов	13
1.1. Предыдущие исследования уровня сложности текста нормативных правовых актов	13
1.2. Метрические показатели, пригодные для оценки сложности юридических текстов	21
2. Что усложняет текст закона? Проведение эксперимента, направленного на выявление негативных лингвистических практик, усложняющих понимание юридических текстов	25
2.1. Подготовка к эксперименту	25
2.2. Результаты эксперимента	37
2.3. Определение коэффициентов значимости метрических параметров синтаксических конструкций в части их влияния на усложнение текста	44
3. Что говорят юристы? Анализ комментариев респондентов	54
4. Сколько в России законов? Составление корпуса федеральных законодательных актов	61

5.	Насколько сложный текст российского закона? Разработка методологии и расчеты	66
5.1.	Авторская методика оценки синтаксической сложности законодательных актов.....	66
5.2.	Результаты анализа синтаксической сложности законов и расчеты ИСЗ.....	70

Часть II

ПРЕДЛОЖЕНИЯ ПО УПРОЩЕНИЮ ЗАКОНОДАТЕЛЬНЫХ АКТОВ

6.	Как не надо писать законы? Ненадлежащие лингвистические практики разработки законопроектов	88
6.1.	Выявленные негативные практики использования слов в законодательных актах	88
6.2.	Выявленные негативные практики в использовании словосочетаний в законодательных актах	91
6.3.	Выявленные негативные практики в использовании предложений в законодательных актах	106
6.4.	Выявленные негативные практики в оформлении структурных элементов законодательного акта.....	112
7.	Как надо писать закон? Предложения по совершенствованию юридической техники вновь принимаемых законодательных актов	125
	Заключение.....	130
	Список использованных источников.....	132

ПРИЛОЖЕНИЯ

Приложение 1.....	137
Приложение 2.....	144
Приложение 3.....	220

ПРЕДИСЛОВИЕ

Данная публикация является результатом прикладного исследования «Анализ существующих юридико-технических и лингвистических подходов к разработке проектов законодательных актов, выработка предложений по их корректировке», проведенного в 2019 году Национальным исследовательским университетом «Высшая школа экономики».

Цель исследования — изучение массива законодательных актов Российской Федерации с проведением их юридико-технического и лингвистического анализа для выявления негативных практик (подходов) к законодательному закреплению юридических норм, затрудняющих восприятие и уяснение положений таких законодательных актов. Кроме того, по итогам исследования следовало сформулировать предложения по совершенствованию юридической техники вновь принимаемых законодательных актов.

Актуальность темы исследования обосновывается нерешенностью задач по достижению понятности нормативно-правового регулирования и возрастающей сложностью текстов законодательных актов. В условиях перехода к цифровой экономике начинает обсуждаться идея цифрового нормотворческого процесса. Переход к цифровому нормотворчеству невозможен без детального изучения текущего состояния юридической техники и синтаксических подходов в языке законодательных актов.

В основе исследования лежали два подхода к изучению предметной области — экспертный и машинный. В рамках экспертного анализа выявлялись конкретные синтаксические конструкции, которые по субъективным оценкам привлеченных экспертов вызывали затруднения в прочтении законодательных актов. Такие оценки обобщались в рамках экспертных панелей. Далее особенности применения подобных синтаксических конструкций в законодательных актах анализировались машинным методом. Для этого использовались методы компьютерной лингвистики, которые позволяли оценить применение исследуемых практик в законодательной технике в отношении всего корпуса законодательных актов, в редакциях, действующих на различные даты начиная с 1991 года. Это позволило изучить особенности синтаксиса законодательства в динамике.

Новизной настоящей работы является глубокий анализ синтаксиса российского законодательства и выявление негативных практик построения синтаксических конструкций, препятствующих пониманию правового регулирования. До настоящего исследования данный метод применялся очень ограниченно, и это не позволяло оценить синтаксические практики российского законодателя посредством точных метрических параметров и соответственно получить объективные выводы.

Кроме того, авторы разработали оригинальную методику определения интегрального индекса сложности конкретных законодательных текстов. Данный индекс основан на сравнении параметров исследуемого текста со средними и максимальными значениями таких параметров в целом по законодательству.

В рамках исследования проводился эксперимент по оценке уровня сложности законодательных актов. В нем приняли участие более 40 человек (студенты магистратуры юридических факультетов и юристы-практики). Для эксперимента были отобраны разные по сложности фрагменты законодательных актов, после прочтения которых респондентам предлагалось ответить на вопросы на понимание текста. Не все из приглашенных экспертов смогли пройти тест до конца. Ответы на вопросы на понимание только что прочтенного текста могли быть неочевидны, и уважаемые юристы-практики, по их отзывам, чувствовали себя в роли экзаменуемых. Авторы выражают глубокую признательность всем участникам эксперимента, которые прошли его до конца, в частности проф. М.А. Краснову, А.В. Кашанину, А.О. Лебедеву и др. Также выражаем благодарность Е.В. Ковтуну, поддержавшему исследование и внесшему неоценимые предложения по его содержанию.

Для чего упрощать текст закона, и как эту проблему решают в других странах?

Принимаемые в последнее время отдельные законодательные акты характеризуются сложной лингвистической структурой и юридической техникой, что усложняет их понимание даже профессионалами. Это может приводить к различным трактовкам закона и, как следствие, к необходимости толкования его положений ненормативными актами (письмами) регуляторов.

По мнению множества экспертов, язык законодательных актов становится от года к году сложнее. При этом данный процесс происходит на фоне универсальной тенденции к упрощению подачи информации, в том числе СМИ¹, и все большему преобладанию в медиапространстве так называемой клиповой культуры².

Таким образом, затрудняется диалог между государством и обществом.

В то же время принципы понятности, прозрачности заложены в Концепции открытости федеральных органов власти, утвержденной распоряжением Правительства Российской Федерации от 30.01.2014 № 93-р. В данной Концепции в качестве одного из механизмов реализации принципов открытости предусматривается обеспечение понятности нормативно-правового регулирования. Также в 2012 году в статье «Демократия и качество государства» В.В. Путин писал: «Необходимо улучшение языка правотворчества. Его надо сделать если не благозвучным (в древнем мире законы часто писали стихами для лучшего запоминания), то хотя бы понятным для адресатов норм»³.

¹ Кучаков Р., Савельев Д. Сложность правовых актов в России: Лексическое и синтаксическое качество текстов / под ред. Д. Скугаревского (Серия «Аналитические записки по проблемам правоприменения»). СПб.: ИПП ЕУСПб, 2018. 20 с.; «Понятное правительство»: шаги к открытости власти [Электронный ресурс]. URL: <https://open.gov.ru/events/5510188/> (дата обращения: 16.01.2019).

² Романов Н.А. Клиповая культура в современном медиапространстве // Человек. Культура. Образование. 2017. № 3 (25).

³ Путин В.В. Демократия и качество государства // Коммерсантъ. 06.02.2012. № 20. С. 1.

Упрощение языка официальных документов органов государственной власти является общемировым трендом. Эта проблема в международной практике рассматривается шире, чем определен предмет настоящего исследования. Упрощению подлежат все официальные документы, исходящие от государственных органов. Например, в Великобритании функционирует организация Plain English Campaign, миссией которой является продвижение использования различными организациями простого английского языка при их взаимодействии с обществом. Компания разработала собственный знак качества «Crystal Mark», проставление которого разрешается на документах, признанных максимально понятными для предполагаемой аудитории. Например, знак «Crystal Mark» проставлен на бланке заявления на получение паспорта Великобритании⁴.

Ниже приводятся краткие сведения о международном опыте по упрощению языка как в целом официальных документов (США), так и законодательных актов в частности (Ирландия, Австралия, Новая Зеландия).

Ирландия

В Ирландии упрощением языка законодательства занимается Комиссия по реформированию законодательства⁵, созданная законом от 1975 года⁶. Она назначается правительством и состоит из пяти человек — практикующих судей и юристов. Роль комиссии заключается в том, чтобы держать, контролировать и проводить исследования с целью реформирования законодательства. За время работы было пересмотрено более 350 законов⁷. Правовая реформа включает:

- развитие права;
- его кодификацию (в том числе упрощение и модернизацию);
- пересмотр и консолидацию статутного права.

В 1999 году был выпущен «Консультационный документ по вопросам разработки и толкования законов: простой язык и

⁴ URL: <http://www.plainenglish.co.uk/services/crystal-mark.html>.

⁵ URL: <https://www.lawreform.ie/>.

⁶ URL: https://www.lawreform.ie/_fileupload/RevisedActs/WithAnnotations/EN_ACT_1975_0003.PDF.

⁷ URL: <http://revisedacts.lawreform.ie/revacts/intro>.

закон»⁸. В нем описывается международный опыт и законодательство, принципы удобочитаемости в лексике, синтаксисе, структуре, форматировании и использовании дополнительных средств (примеров, формул, карт, диаграмм) в нормативных актах.

Австралия

Реформа по упрощению официального стиля также проводится в Австралии⁹. Изменения официальной письменной формы в Австралии находятся в ведении Офиса парламентского адвоката (Office of Parliamentary Counsel — OPC)¹⁰, основанного в 1970 году. Учреждение занимается непосредственным написанием законопроектов, поправок к ним и подзаконных актов. Ведомства должны использовать его редакционные услуги для всех правительственных законопроектов и распоряжений, постановлений и нормативных актов внешних территорий и территории Джервис-Бэй, а также других законодательных актов, разработанных или утвержденных Генерал-губернатором. Также Офис парламентского адвоката занимается публикацией нормативно-правовых актов, ведением Федерального реестра регулирования, обучением и консультированием.

На сайте опубликованы два документа¹¹: «OPC’s Guide to Reducing Complexity in Legislation» и «Plain English Manual». Первый посвящен улучшению структуры законодательства и написанию положений, сокращению длины нормативных актов, разделов и предложений, причинам усложнения применения законодательства и другим темам. Второй документ посвящен непосредственно синтаксическим и лексическим правилам удобочитаемости, а также планированию нормативного акта и вспомогательным средствам (таблицам, заголовкам, разделению на параграфы и т.д.).

⁸ URL: https://www.lawreform.ie/_fileupload/consultation%20papers/cpPlainLanguage.htm.

⁹ *Шашек В.В., Харченко Н.А.* Проблема ясности языка законодательства и множественности интерпретаций текстов законов (на материале статей Гражданского Кодекса Российской Федерации) // Молодой ученый. 2016. № 7. С. 1191–1196. URL: <https://moluch.ru/archive/111/27876/> (дата обращения: 20.02.2019).

¹⁰ URL: <https://www.opc.gov.au/>.

¹¹ URL: <https://www.opc.gov.au/drafting-resources/plain-language>.

Новая Зеландия

В Новой Зеландии также функционирует Офис парламентского адвоката (Parliamentary Counsel Office)¹². Он тоже занимается публикацией нормативных актов и участвует в написании законопроектов. Офис парламентского адвоката работает с другими ведомствами, чтобы сделать законодательство, соответствующим цели, конституционно обоснованным и доступным для всех пользователей¹³. Доступность означает, что законодательство легко найти, легко в нем разобраться и понять. Меры включают и написание законодательства простым языком. Ведомство разработало стандарты простого языка, представленные в специальном разделе¹⁴, а также чеклист и принципы понятного написания законопроектов. Они касаются и структуры текста, синтаксиса и лексики в предложениях, тона предложений, а также использования заголовков и навигации.

Еще одним ведомством, разработавшим стандарты простого языка в официальном письме, является законодательная комиссия¹⁵. В 1996 году она опубликовала свою инструкцию по написанию законопроектов «Legislation Manual Structure and Style»¹⁶.

США

В 2010 году в США был принят закон об упрощении письменной формы (Plain Writing Act)¹⁷. Простая письменная форма определяется как написание, которое должно быть ясным, кратким, хорошо организованным и следовать другим передовым методам, соответствующим теме или области и целевой аудитории. Цель этого закона заключается в повышении эффективности и подотчетности федеральных органов перед общественностью путем содействия четкой правительственной коммуникации, которую общественность может понять и использовать. Однако здесь имеются в виду

¹² URL: <http://www.pco.govt.nz/>.

¹³ URL: <http://www.pco.govt.nz/strategic-goal/>.

¹⁴ URL: <http://www.pco.govt.nz/plain-language/>.

¹⁵ URL: <https://www.lawcom.govt.nz/home/welcome-law-commission>.

¹⁶ URL: <https://www.lawcom.govt.nz/sites/default/files/projectAvailableFormats/NZLC%20R35.pdf>.

¹⁷ URL: <https://www.govinfo.gov/content/pkg/PLAW-111publ274/pdf/PLAW-111publ274.pdf>.

документы, которые нужны для получения пособий и услуг или оплаты налогов, объясняют гражданам соблюдение регулирования и представляют собой письма, уведомления, публикации, формы или инструкции, но не регулирование.

Этот закон требует, чтобы во всех ведомствах:

- назначили ответственных за его выполнение;
- сообщили о требованиях закона сотрудникам;
- обучили сотрудников простому письму;
- установили процесс контроля за постоянным соблюдением требований закона;
- создали на веб-сайте ведомства раздел об использовании простого языка для информирования пользователей, получения и реагирования на обратную связь;
- назначили одно или более контактное лицо, чтобы получать и реагировать на общественное мнение по поводу осуществления ведомством закона и его отчетов;
- опубликовали на сайте агентства отчеты в открытом виде;
- начали использовать простой язык в каждом охватываемом документе учреждения, который агентство выпускает или существенно пересматривает;
- ежегодно публиковать отчет о соблюдении этого закона.

Также имеется объединение The Plain Language Action and Information Network¹⁸, которое продвигает идею простого письменного языка в государственной службе. Были разработаны принципы¹⁹, которым нужно следовать, чтобы писать более простым и понятным языком. Рекомендации связаны с правильным выбором слов, использованием глаголов, существительных и местоимений, построением абзацев, предложений, организации всего текста. Описывается и написание текстов для сайта.

В июне 1998 года президент Билл Клинтон издал меморандум о простом письме в правительственной письменной речи²⁰. Он потребовал к октябрю того же года начать использовать простой язык во всех новых документах (кроме нормативного регулирования),

¹⁸ URL: <https://www.plainlanguage.gov/>.

¹⁹ URL: <https://www.plainlanguage.gov/media/FederalPLGuidelines.pdf>.

²⁰ URL: <https://www.govinfo.gov/content/pkg/WCPD-1998-06-08/pdf/WCPD-1998-06-08-Pg1010.pdf>.

которые разъясняют, как получить пособие/услугу или соблюсти требование регулирования. К 2002 году все такие документы, созданные до октября 1998 года, надлежало переделать под простой язык. К 1 января 1999 года предполагалось использование простого языка во всех выдвинутых и окончательных нормативных документах. Документы следовало писать в повседневной лексике, короткими предложениями и в активном залоге, они должны были иметь четкую логическую структуру и читабельное оформление.

В России в настоящее время широкой кампании по продвижению использования понятного русского языка в официальных документах не ведется. В рамках работы Открытого правительства с 2015 года продвигалась идея повышения прозрачности нормативных правовых актов за счет подготовки для ключевых законодательных актов, специальных материалов, содержащих положения законодательного регулирования в ясных схемах и составленных понятным языком, — в первую очередь для федеральных законов о федеральном бюджете на очередной год. Однако речь не шла об упрощении языка самих нормативных правовых актов, и в этом требовалось разобраться, что с таким языком «не так», в том числе, что именно является препятствием для их понимания.

Для достижения указанной цели необходимо разработать методологию юридико-технического и лингвистического анализа законодательных актов с целью оценки уровня понятности таких актов для целевых групп. При этом методология должна предусматривать использование методов вычислительной лингвистики. В противном случае потребуются значительное количество кадровых ресурсов для оценки уровня сложности текстов нормативных правовых актов. Данная попытка предпринята в настоящем исследовании.

Часть I

ОПРЕДЕЛЕНИЕ УРОВНЯ СЛОЖНОСТИ РОССИЙСКИХ ЗАКОНОВ

1. Как оценить сложность закона? Подходы к оценке сложности юридических текстов

1.1. Предыдущие исследования уровня сложности текста нормативных правовых актов

Анализом уровня сложности текстов нормативных правовых актов в России занимались разные группы исследователей.

А. Сотрудники Института проблем правоприменения Европейского университета в Санкт-Петербурге *Р. Кучаков и Д. Савельев* провели исследование корпуса правовых актов методами компьютерной лингвистики. Исследователи пришли к выводу, что в России наблюдается усложнение текстов правовых актов — падение лексического разнообразия и усложнение структуры предложений. В особенности такое усложнение фиксируется в 2014–2017 годах¹. Авторы использовали при анализе сложности нормативных актов две метрики: *лексическое разнообразие* (повторяемость одних и тех же слов) и *длину зависимых связей* в предложении (сложность структуры предложения для чтения).

Для определения лексического разнообразия был рассчитан коэффициент TTR (*англ.* Type-Token Ratio) по формуле (1.1):

$$TTR = \text{ЧУТ} / \text{ОЧТ}, \quad (1.1)$$

где ЧУТ — число уникальных токенов;

ОЧТ — общее число токенов.

¹ *Кучаков Р., Савельев Д.* Сложность правовых актов в России: Лексическое и синтаксическое качество текстов / под ред. Д. Скугаревского (Серия «Аналитические записки по проблемам правоприменения»). СПб.: ИПП ЕУСПб, 2018.

Под токеном (текстоформой) в исследовании понимался выделенный в текстовом потоке минимальный фрагмент для последующего анализа (слово, число, знак препинания и т.д.). Уникальными считались токены, использованные в тексте лишь однажды. При этом их лемматизация (приведение к начальной форме слова) не проводилась, т.е., например, слова «закон» и «закона» считались разными токенами. TTR измеряется от 0 до 1. Он равен единице, если все слова в тексте различны, если же в тексте много повторов одних и тех же слов, значение коэффициента стремится к нулю.

Для определения расстояния между зависимыми словами в предложении (maxDepLen) (*англ.* dependency length, DepLen) использовался синтаксический парсер ru-syntax (размещенный для общедоступного использования в сети Интернет по адресу: URL: <https://github.com/tiefling-cat/ru-syntax>). Эта метрика показывает, сколько слов отделяет зависимые слова друг от друга. MaxDepLen исчисляется от единицы до бесконечности. По тексту высчитывается среднее значение MaxDepLen.

Б. Иное исследование в отношении корпуса текстов решений Конституционного Суда России проведено *А.В. Дмитриевой*. Она пришла к выводу о постепенном усложнении текстов таких решений. Для исследования корпуса решений Конституционного Суда РФ использовались расчеты индекса Флеша—Кинкейда. Данный индекс на основе анализа среднего количества слов в предложении и среднего количества слогов в словах позволяет определить, сколько требуется лет обучения, чтобы понять содержание того или иного текста (более подробное описание методики расчета индекса приведено ниже). Основываясь на данных расчета индекса, *А.В. Дмитриева* пришла к выводу, что для понимания текстов решения Конституционного Суда России в 1992 году требовалось 20 лет обучения, а в 2017-м — 25 лет².

В статье *Дмитриевой* указывается на опыт преподавания курса юридического письма в США. К основным требованиям к понят-

² *Дмитриева А.В.* Искусство юридического письма: количественный анализ решений Конституционного Суда России // Сравнительное конституционное обозрение. 2017. № 3 (118). С. 130.

ному юридическому письму (Plain Legal English), которыми должен овладеть студент первого года обучения в юридических школах США, относятся: 1) короткие прямые фразы, содержащие только важную информацию; 2) прямой порядок слов (подлежащее — сказуемое — дополнение); 3) активный залог, если нет действительно большой необходимости употребить пассивный залог; 4) минимум отрицательных конструкций; 5) использование параллельных структур для сложных идей; 6) отсутствие не необходимых слов³.

Анализ таких конструкций тоже мог бы лечь в основу оценки уровня сложности юридических текстов, но в исследовании не использовался.

В. Т.В. Анисимова, И.А. Самуйлова провели анализ законодательных актов 10 субъектов РФ Северо-Западного федерального округа РФ, принятых в 2014 году (80 текстов). Анализ проводился на основе количественной оценки текстов по следующим параметрам: общая статистика текста, статистика удобочитаемости, первичные описательные статистики, частотный анализ словоупотребления. Анализ позволил сформулировать типичные для правовых текстов особенности, затрудняющие их восприятие аудиторией. К ним авторы отнесли использование:

- более четырех существительных в родительном падеже, идущих подряд;
- более четырех предложно-именных групп, идущих подряд;
- слишком длинных предложений, трудных для грамматического разбора и восприятия;
- длинных перечислений;
- большого количества причастных и деепричастных оборотов;
- перегруженность знаками препинания⁴.

Однако обоснования влияния именно таких конструкций на сложность текста в статье не приводится.

³ Там же.

⁴ *Reynolds R.J.* Russian Natural Language Processing and Computer-assisted language Learning: Capturing the benefits of deep morphological analysis in real-life applications: PhD. Thesis Tromsø: Universitet I Tromsø (цит. по: *Дмитриева А.В.* Указ. соч. С. 125–133.

Г. Другая группа ученых, занимающихся проблемой использования русского языка в официальных документах, — исследователи из Санкт-Петербургского государственного университета. В подготовленном ими сборнике «Государственный язык России: нормы права и нормы языка»⁵ сделан вывод о необходимости большей унификации требований к государственному языку, который должен создавать единое и коммуникативное информационное пространство. Авторы полагают, что требования к государственному языку должны быть системными и едиными. Следует установить требования понятности и доступности, касающиеся любых адресованных гражданам официальных документов и правовых актов. Сейчас законодательно лишь запрещается использование слов и выражений, не соответствующих нормам современного русского литературного языка, за исключением иностранных слов, не имеющих общеупотребительных аналогов в русском языке⁶. Однако строгие критерии соответствия или несоответствия стиля того или иного документа нормам «современного русского литературного языка» в принципе отсутствуют.

Кроме того, эксперты из СПбГУ неоднократно в интервью упоминали проблемы языка законодательства. Для «Российской газеты»⁷ декан юридического факультета С.А. Белов отметил, что ошибки и неопределенности в законах приводят к тому, что их применение зависит от конкретных чиновников и судей. За последние годы около 100 документов были обжалованы и отменены на основании нарушения правовых требований к языку. В целом нормативные акты пишутся языком сложных научных текстов, требующим для понимания как минимум ученой степени. В них встречаются предложения длиннее 15 слов с несколькими причастными и деепричастными оборотами. Законодатель часто использует специализированные термины либо, наоборот, разговорные выражения, что приводит к искажению смысла текста. Ректор СПбГУ

⁵ Государственный язык России: нормы права и нормы языка / под ред. С.А. Белова, Н.М. Кропачева. СПб.: Изд-во Санкт-Петерб. ун-та, 2018. 130 с.

⁶ Федеральный закон от 01.06.2005 № 53-ФЗ «О государственном языке Российской Федерации».

⁷ Лингвисты нашли лазейки для коррупции в каждом третьем документе // Российская газета. URL: <https://rg.ru/2016/01/15/obr.html> (дата обращения: 05.12.2019).

Николай Кропачёв отметил⁸, что язык законодательства понимают не более 5% россиян. Такие данные приводятся на основе исследования 36 тыс. законодательных актов, которое провел Институт проблем государственного языка при СПбГУ. Также в ходе работы было установлено, что если в предложении содержится более семи знаков препинания, то его понимание затруднено. В то же время в нормативных актах были обнаружены фразы длиной в несколько страниц с 36 и даже 66 знаками препинания.

Обзор указанных исследований позволяет констатировать, что у экспертного сообщества сложилось единое мнение по вопросу о все большем усложнении для понимания текстов правовых актов. Однако используемые для подтверждения такого усложнения показатели различны. Отсутствует единство по вопросу о методике оценки уровня сложности юридических текстов. Среди рассмотренных исследований для данных целей применялись следующие подходы:

- анализ различных метрических показателей, например, *лексического разнообразия* текстов (повторяемости одних и тех же слов), *длины зависимых связей* в предложении, количества знаков препинания в предложениях и фразах;
- универсальные индексы удобочитаемости (например, индекса Флеша—Кинкейда) для юридических текстов;
- оценка соблюдения норм русского языка в правовых актах;
- экспертная оценка юридических текстов на основе комплекса параметров.

При этом исследователи не объясняют, почему были выбраны конкретные параметры оценки. В работах только постулируется, что такие параметры влияют на уровень сложности текста. Почему были выбраны именно эти метрики, а также степень их влияния на сложность восприятия текстов НПА в исследованиях не объясняется.

Следует иметь в виду, что нормативный правовой акт — это специфический текст, и поэтому неизбежны, например, повторяемость слов (необходимость в одних и тех же терминах, единожды введенных в употребление без возможности применения их синонимов из-за необходимости соблюдения принципа право-

⁸ Николай Кропачев: ясность языка — это тоже путь к счастью. URL: <https://tass.ru/interviews/3981423> (дата обращения: 05.12.2019).

вой определенности), а также использование определенных слов, в том числе сложных и специфических терминов и длинных слов (связанность имеющейся терминологии в той или иной сфере при ее регулировании). Поэтому некоторые параметры оценки уровня сложности юридических текстов становятся непригодными. В первую очередь речь идет о попытках применения для оценки правовых актов распространенных индексов удобочитаемости.

Из подобных *индексов удобочитаемости* широко известны индексы Флеша, Флеша—Кинсайда, Колман—Лиау, Автоматический индекс удобочитаемости (Automated Readability Index), SMOG (Simple Measure of Gobbledygook), LIX (*швед.* Läsbarhetsindex) индекс удобочитаемости.

Также распространена формула Дэйла—Чейла (Dale—Chall readability formula), с помощью которой оценивают не только статистические параметры текста, но и лексический параметр. Таким признаком выступает доля слов, относящихся к определенному фиксированному списку (лексический минимум базовых слов).

Например, наиболее распространенный индекс удобочитаемости — индекс Флеша — основывается на анализе длины предложения и длины употребляемых слов. Позже Рудольф Флеш совместно с Питером Кинкайдом по заказу армии США для оценки удобочитаемости инструкций по эксплуатации военной техники в 1975 году переделал ранее разработанный индекс. Он был изменен таким образом, чтобы значение индекса указывало на необходимое образование для понимания текста. В 1978 году данная формула была впервые введена в работу для оптимизации текстов — инструкций для военных в штате Пенсильвания по приказу Министерства обороны США. Используя оценку удобочитаемости, удалось добиться уровня удобочитаемости инструкций для понимания даже учениками 9-х классов. После этого индекс начал использоваться во всех штатах США, а затем и Министерством образования США при составлении образовательных программ и учебного материала⁹.

Для русского языка данные индексы адаптированы И.В. Оборновой. С этой целью был проведен сравнительный анализ средней

⁹ Енин В.С., Гахов Р.П. Алгоритм оценки удобочитаемости текстового материала // Перспективы развития информационных технологий. 2016. № 31. С. 22–28.

длины слова в русском и английском языках. В ходе исследования были использованы: «Словарь русского языка» / под ред. Ожегова — 39 174 слова и «Англо-русский словарь» / под ред. Мюллера — 41 977 слов. В результате анализа было установлено, что средняя длина слова в русском языке составляет 3,29 слога, а в английском — 2,97 слога¹⁰. В итоге для русского языка была получена следующая формула (1.2) индекса Флеша:

$$\text{Индекс Флеша} = 206,835 - 1,3 \times ASL - 60,1 \times ASW, \quad (1.2)$$

где *ASL* (англ. average sentence length) — средняя длина предложения в словах;

ASW (англ. average number of syllables per word) — средняя длина слова в слогах.

В таблице 1.1 обобщены метрики, которые применяются в указанных выше индексах.

Практически все распространенные методики оценки удобочитаемости основываются на оценке небольшого круга статистических (количественных) показателей текста в их различных комбинациях: длина абзаца (в словах), длина предложения в словах, длина слова в буквах (или в слогах). И.Ю. Мизернов провел оценку взаимосвязей некоторых из указанных индексов (Флеша, Флеша—Кинсайда, Дэйла—Чейла, тест Пауэрса, SMOG). В результате он пришел к выводу, что «большинство индексов удобочитаемости характеризуются высокой степенью как прямой, так и обратной корреляционной зависимости, что обусловливается преимущественным использованием индексов одной математической линейно-регрессионной модели, переменными которой выступает узкий круг статических параметров текста»¹¹. Мизернов полагает, что ни один из рассмотренных индексов удобочитаемости не направлен на оценку восприятия текста взрослым человеком, работающим с информационно-аналитическими документами.

¹⁰ Солнышкина С.И., Кисельников А.С. Сложность текста: этапы изучения в отечественном прикладном языкознании // Вестник Томского гос. ун-та. (Серия «Филология»). 2015. Т. 6. № 38. С. 86–99.

¹¹ Мизернов И.Ю. Анализ взаимосвязей индексов сложности текста // Информационные системы и технологии 2015. Материалы III Международной научно-технической интернет-конференции. ФГБОУ ВПО «Государственный университет-учебно-научно-производственный комплекс». 2015.

Таблица 1.1. Оцениваемые метрики в рамках распространенных индексов удобочитаемости

Индексы	Длина предложения (в словах)	Длина слова (в слогах)	Количество букв	Количество предложений	Количество слов	Количество слов длиннее 6 букв	Количество polyslogov в предложениях*	Количество сложных слов**
Индекс Флеша	да	да						
Индекс Колман—Лиану			да	да				
Автоматический индекс удобочитаемости			да		да			
SMOG				да			да	
LIX				да	да	да		
Формула Дэйла—Чейла	да							да

* Слов из более трех слогов.

** К сложным словам относятся те, которые не входят в перечень простых слов из 3000 слов.

В основном различные индексы удобочитаемости разработаны для достижения педагогических целей: категорирование литературы по уровню ее сложности в рамках обучения детей чтению и литературе (по классам обучения), а также в рамках обучения русскому языку как иностранному (по степени освоения русского языка)¹². Применение данных индексов для других целей ограничено соответствующей предметной областью.

В сети Интернет функционирует сервис оценки читабельности русскоязычных текстов на основе применения некоторых перечисленных выше индексов (<http://ru.readability.io/>). Почти все законы

¹² Анисимова Т.В., Самуйлова И.А. Понимание социально-значимых текстов как основа диалога власти и общества // Интегративный подход к психологии человека и социальному взаимодействию людей: векторы развития современной психологической науки. Материалы VII Всероссийской научно-практической конференции. 2017. С. 118–125.

при анализе по шкале таких индексов (<http://ru.readability.io/>) относятся к самым сложным текстам, требующим в среднем 26 лет обучения или окончания образования в аспирантуре, получения второго высшего образования, Ph.D.

Таким образом, по мнению авторов исследования, применение распространенных индексов удобочитаемости для оценки уровня сложности законодательных актов неэффективно.

1.2. Метрические показатели, пригодные для оценки сложности юридических текстов

Для анализа юридических текстов необходимо использовать сложные метрики (в том числе синтаксические), позволяющие выявлять сложные для понимания конструкции, которые без потери для смысла регулирования могут быть заменены более простыми. Ниже рассмотрим укрупненные группы метрических показателей текста, которые могут быть использованы для оценки сложности юридических текстов.

В грамматических метриках в качестве оцениваемого признака выступает доля слов, относящихся к определенным грамматическим категориям, которые имеют разную сложность. Такими категориями могут быть причастия, деепричастия, глаголы в страдательном залоге и т.д. Доля может вычисляться как на уровне предложения, так и на уровне текста; а также этой грамматической категории по отношению ко всем словам данной части речи.

В семантических метриках в качестве оцениваемого признака выступает доля слов, относящихся к определенным семантическим категориям. В частности, доля абстрактных слов может увеличивать сложность текста. Метрика сложна для реализации, так как необходимо иметь возможности определить для каждого слова его семантический класс. Кроме того, специфика юридических текстов предполагает большую долю абстрактных слов.

В синтаксических метриках в качестве оцениваемого признака выступает доля определенных синтаксических конструкций на одно предложение:

- причастные и деепричастные обороты;
- вводные конструкции;
- придаточные предложения;

- грамматические основы предложения (подлежащее, сказуемое или одно из них);
- знаки препинания (точка, запятая, двоеточие, точка с запятой, тире, кавычки, скобки).

Например, значимым показателем является среднее расстояние между зависимыми словами в предложении (количество слов, отделяющих зависимые слова друг от друга). Если слова, связанные синтаксически, линейно располагаются близко, например, одно за другим, — это простой текст. Если расстояние между ними в среднем большое, т.е. связанные слова часто расположены далеко, — то это сложный текст. Связанные синтаксически слова могут, например, разделяться причастными оборотами, вводными конструкциями, придаточными предложениями и прочими вставными элементами. Таким образом, этот признак является наиболее универсальным для определения сложности и важен для анализа текстов законов. Расстояние между зависимыми словами использовалось для оценки сложности нормативных правовых актов в исследовании, проведенном Институтом проблем правоприменения Европейского университета в Санкт-Петербурге (Р. Кучаков и Д. Савельев)¹³.

Макропоказатели текста тоже могут влиять на уровень сложности текста (размеры текста: по количеству абзацев, слов или знаков)¹⁴. Так, значимым признаком может быть среднее количество предложений в абзацах, среднее количество абзацев в каком-либо разделе (например, для законов — в части статьи или в целом в статье).

До недавнего времени использование грамматических и синтаксических метрик для оценки сложности текста было затруднено трудоемкостью их подсчета и анализа и невозможностью применения для этого машинного метода. Однако развитие компьютерной лингвистики позволяет автоматизировать обработку текстов на предмет их сложности на основе таких метрик.

Например, для определения расстояния между зависимыми словами в предложении (`maxDepLen`) был разработан синтаксический парсер `ru-syntax`. Его создали аспиранты кафедры компьютерной лингвистики НИУ ВШЭ на основе машинного обучения на

¹³ Кучаков Р., Савельев Д. Указ. соч.

¹⁴ Дмитриева А.В. Указ. соч.

Конец ознакомительного фрагмента.
Приобрести книгу можно
в интернет-магазине
«Электронный универс»
e-Univers.ru