

Метод и алгоритм обнаружения признаков лингвистических дефектов в научно-технических текстах¹

А.В. Швец, Ю.М. Кузнецова, Г.С. Осипов, А.В. Латышев

Аннотация. Рассматривается метод автоматического выявления признаков, свидетельствующих о наличии лингвистических дефектов в предложениях русского языка. Предлагаемый метод учитывает семантические, синтаксические, морфологические и лексические характеристики элементов текста, их контекст и взаимную сочетаемость. Результаты проведенных экспериментов показывают применимость метода к обнаружению признаков нарушений грамматических норм, правил согласования и управления.

Ключевые слова: семантический анализ, лингвистический анализ, нарушение грамматических норм, дефектность текста.

Введение

Основные цели научной публикации – закрепление процесса познания и хранение знания, сообщение информации и доказательство ее истинности, – определяют характерные особенности научного стиля. По определению А.Л. Пумпянского, «основная задача научной и технической литературы – предельно ясно и точно довести определенную информацию до читателей. Это достигается логически обоснованным изложением фактического материала» [1]. В.И. Карасик показывает, что ясность и точность составляют базу интерпретируемости текста, то есть обеспечивают читателям возможность извлечь содержащуюся в нем информацию; при этом «ясность научного текста определяется четкостью понятийно-терминологического аппарата, логичностью изложения, иллюстративным материалом и простым и строгим литературным языком», а его точность «заключается в развертывании и уточнении характеристик понятия» [2]. Именно той

частью, которая будет воспринята потенциальным читателем, а не общим количеством всей информации, содержащейся в научном тексте, определяется его информационная ценность [3]. Вследствие этого, стремясь сделать свое сообщение как можно более информационно эффективным, автор научной публикации должен специально заботиться о выборе адекватных языковых средств, ясно, точно и полно передающих подразумеваемый им смысл. Выражение научных понятий и умозаключений, выявление, описание и интерпретация новых данных и закономерностей, оценка результатов исследования, формулирование выводов, инструкций, рекомендаций и т. п. – эти виды речевых действий требуют специального оформления в соответствии с выработанными в языковой культуре нормами и традициями. Если требования ясности и точности не выполняются, снижается информационная ценность текста, и он оказывается на периферии либо вообще вне границ научной коммуникации [3-7].

¹ Работа выполнена при финансовой поддержке Минобрнауки России по государственному контракту № 14.514.11.4018 в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2013 годы».

Научная публикация, в отличие, например, от устного выступления, является продуктом не спонтанным, а подготавливаемым в процессе целенаправленной деятельности. Создавая научный текст, автор имеет возможность тщательно продумать композицию, отобрать наиболее точные слова и однозначные грамматические конструкции, возможно полно удовлетворяющие требованиям, предъявляемым к качеству изложения [7]. Однако, как показывает практика, «установка на определенное качество речи не всегда создает желаемое качество» [7]. Приходится констатировать, что «современная массовая научная литература, бесспорно, страдает от речевых погрешностей, затрудняющих понимание содержания» [8]. В потоке публикаций немалую долю составляют тексты, в которых обнаруживаются те или иные отступления от норм научного изложения – отступления, рассматриваемые далее как *ошибки*, или *дефекты*. Качество текста, связанное с наличием в нем ошибок, в рабочем порядке предлагается именовать *дефектностью*, текст же, которому это качество присуще в выраженной степени, будем называть *дефектным*.

Рассматривая причины, по которым человек может вообще допускать в своей речи ошибки, В.М. Алпатов указывает:

- на недостаточное владение установленной нормой,
- на сознательное отклонение от нормы (в языковой игре, поэзии и т.п.),
- на бессознательное стремление удовлетворить ту или иную потребность [9]. По мнению Д.Б. Эльконина, основные сложности, возникающие при порождении письменного сообщения, определяются такими качествами мысли, как ее целостность, психологическая слитность, взаимосвязанность отдельных элементов [10]. Похожим образом М.П. Котюрова считает, что обнаруживаемые нарушения связаны с такими разнонаправленными свойствами мышления и речи, как дискретность смыслов и линейность речевого потока. При создании научного текста внимание автора сосредоточено на «макросмыслах», их целостности; такое внимание обуславливает «крупноблочный» взгляд на содержание и его выражение. Отсюда «скольжение», «скачки» мысли, приводящие к неточно-

сти или неправильности выражения связи компонентов одного словосочетания, словосочетаний в предложении, а также предикативных единиц в сложном предложении [8]. А.В.Казанская предлагает рассматривать мотивационные факторы в качестве основной причины нарушения каузальности высказывания, выражающегося в синтаксически неправильном построении фраз (меняются местами субъекты и объекты, действительный и страдательный залого), неправильного употребления существующих понятий, изобретения непонятных для адресата сообщения неологизмов [11].

Ошибки в устной и письменной речи выступают предметом изучения в целом ряде дисциплин. Один аспект изучения связан с интересом к процессу отражения в ошибке порождающих ее *ментальных особенностей автора*: согласно известной формулировке Л.С. Выготского, мысль не выражается, а совершается в слове [12], следовательно, уровень владения речевыми средствами *не столько отражает, сколько определяет уровень протекания когнитивных процессов*. Поэтому ошибкам посвящаются работы, выполненные в русле психопатологического подхода [13-15], возрастной психолингвистики [16-19], а также психолингвистики обучения иностранному языку [20-22]. Другой аспект исследования ошибок связан с интересом к *коммуникативной* (в частности, *информационной*) *эффективности текста* в рамках теории и практики литературного редактирования [23-25].

Исследователи выделяют разнообразные виды вербальных ошибок и предлагают различные их классификации. Наш опыт показал, что определенные группы ошибок в научных текстах встречаются систематически, в то время как другие оказываются нетипичными. Дефекты, характерные для научных текстов, можно распределить на следующие группы.

1. **ФАКТИЧЕСКИЕ ОШИБКИ** – искажения фактического материала по небрежности или по незнанию.

2. **ЛОГИЧЕСКИЕ ОШИБКИ** (алогизмы) – наличие в тексте рассуждений, противоречащих логике, например:

- Non liquet – «неясно». По происхождению это формула римского судопроизвод-

ства. Судьи, голосуя за приговор, выражали одно из трех мнений: оправдываю, осуждаю и неясно (то есть воздерживаюсь).

- *Qui pro quo* – «одно вместо другого». Смещение понятий, путаница, недоразумение.

- *Contradictio in adjecto* – «противоречие в определении». Например, *сухая влага*. Внутреннее противоречие, противопоставляемое внешнему противоречию – *contradictio in subiecto* – «противоречию в предмете».

- *Post hoc, ergo propter hoc* – «после этого – значит по причине этого». Временная последовательность событий принимается за причинную зависимость.

- *Idem per idem* – «то же посредством того же». Доказательство какого-либо положения посредством самого этого положения.

- *Petitio principii* – «предвосхищение оснований». Аргумент, основанный на выводе из положения, которое само требует доказательства.

- *Circulus vitiosus* – «порочный круг». Приведение в качестве доказательства того, что еще само нуждается в доказательстве.

- *Ignotum per ignotius* – «неизвестное через более неизвестное».

- *Consequentia non valet* – «вывод не имеет силы». Из правильных посылок делается вывод, который из них не вытекает.

3. ТЕКСТОВЫЕ ОШИБКИ

- Грамматические – ошибки в структуре языковой единицы (слова, словосочетания или предложения), нарушения грамматической нормы – словообразовательного, морфологического или синтаксического оформления мысли в соответствии с требованиями сочетаемости, согласования и управления. К этой группе относятся: ошибочное словообразование; ошибки в образовании форм слов; нарушение видовременной соотнесенности глаголов; ошибки в согласовании и управлении; нарушение согласования подлежащего и сказуемого; ошибки в употреблении причастных и деепричастных оборотов; ошибки в построении предложения с однородными членами; ошибки в построении сложных предложений.

- Речевые ошибки – неправильное использование языковой единицы; в отличие от грамматических, речевые ошибки определяются только в контексте. К этой группе относятся: неточность словоупотребления (употребление слова в несвойственном ему значении, например, неверный выбор паронимов); нарушение функционально-стилевой нормы научного стиля (использование слов и фразеологизмов разговорно-просторечной и эмоционально-экспрессивной окраски); тавтология и плеоназм; нарушения лексической сочетаемости; неудачное употребление личных и указательных местоимений в анафорических конструкциях (в результате чего создается двусмысленность); лексическая неполнота высказывания (пропуск необходимого в предложении слова); неудачный порядок слов; ошибки в употреблении устойчивых сочетаний (например, нарушение фразеологизмов и клише) [26-30].

4. КОМПОЗИЦИОННЫЕ ОШИБКИ – несоответствие структуры текста требованиям, предъявляемым первичной научной публикации [31-33].

Все перечисленные группы ошибок могут критически снижать информационную ценность научной публикации.

Основной идеей настоящей работы является представление о том, что степень лингвистической дефектности научного текста может быть предметом автоматического анализа, основанного на методе, описанном ниже. Представляется очевидным, что в первую очередь проверке на возможность формализации подлежат способы выявления таких видов дефектов, которые не связаны со значением языковых единиц, а такими, согласно приведенной выше типологии, являются грамматические и композиционные ошибки. Ранее нами были получены определенные результаты в направлении разработки средств, позволяющих производить автоматическое установление структурной полноты научной публикации [34]. Здесь мы обращаемся к проблеме автоматизации поиска признаков грамматических нарушений. Выявление грамматических ошибок, определение их общего количества на текст (то есть оценка де-

фектности текста), выявление склонности автора к определенному виду ошибок (за которой может стоять специфика протекания когнитивных процессов), отслеживание динамики дефектности и создание обобщенного описания ошибок, характерных для публикаций автора или научного коллектива – решение таких задач становится возможным при разработке соответствующего метода анализа дефектности научно-технических текстов.

Одной из характерных черт научной коммуникации в современной России является существование двух противоречивых тенденций. С одной стороны, проблема качества научных публикаций признается на государственном уровне – специальные курсы культуры научной речи являются обязательным компонентом профессиональной подготовки для любой специализации. С другой стороны, наблюдается увеличение количества изданий и издательств, публикующих научные работы при отсутствии обязательной научной редактур в большинстве из них. По сообщению специалистов, переход на электронные издательские технологии неожиданно явился дополнительным фактором, снижающим общий уровень правильности современной русской научной речи, поскольку оказалось, что на экране дисплея ошибки и опечатки выявляются хуже, чем на бумаге [26]. Судя по тому, что количество научных публикаций, не соответствующих требованиям ясности и точности вследствие отступления от стандартов письменной речи вообще и научной в частности, растет, вторая – негативная – тенденция пока оказывается более сильной. В этих условиях представляется весьма актуальным создание средств, которые можно использовать как при экспертизе научных текстов, так и в процессе их редактирования или саморедактирования.

1. Метод автоматического обнаружения дефектов

В основе метода автоматического обнаружения дефектов лежит некоторое множество правил, с помощью которых можно выявить в тексте признаки, свидетельствующие о наличии лингвистических дефектов, таких как нарушения согласования, управления, разрывность

фраз и др. Такое множество правил было получено следующим способом:

- 1) выбирается одно из правил русского языка;
- 2) исследуются примеры предложений, удовлетворяющих данному правилу;
- 3) извлекаются условия, выполнение которых свидетельствует о наличии дефекта; при формировании условий степень обобщения ограничивается множеством правильных предложений;
- 4) в текстах выделяются предложения, для которых выполняются полученные условия;
- 5) если среди выделенных предложений содержатся правильные предложения или при просмотре текста обнаруживаются дефектные предложения, которые не были выделены, то условия уточняются, и выполняется действие 4. Правило является результатом последовательного итерационного уточнения условий. Рассмотрим более подробно примеры дефектов и полученные правила для их обнаружения.

Одной из распространенных ошибок, встречающихся в научных текстах, является отсутствие согласования причастия с определяемым словом, стоящим перед причастным оборотом. Приведем правило, позволяющее выявлять такие ошибки.

Правило 1: если предложение содержит причастный оборот, выделенный запятыми, и причастие не согласуется ни с одним из существительных, местоимений, прилагательных и числительных, стоящих перед оборотом, в роде, числе и падеже в ед. ч. и в числе и падеже во мн. ч. (и не согласуется с однородными членами в падеже во мн. ч.), то такое предложение содержит признак нарушения и является подозрительным с этой точки зрения.

Приведем пример предложения, удовлетворяющего данному правилу: «Существует возможность превращения идиолекта в некий субстрат с аморфным *содержанием и «экономной» формой, не дающих* реальных шансов для диагностики говорящего». Как видно, отсутствует падежное согласование причастия с однородными определяемыми словами.

Для обозначения порядка явлений (мыслей) и связей между ними используются парные элементы, указывающие на последовательность изложения. В качестве таких элементов могут выступать, например, словосочетания «с одной

стороны, с другой (стороны)», пара слов «либо, либо», вводные слова «во-первых, во-вторых» [35]. Отсутствие одного из этих элементов является признаком дефекта в тексте. Опишем правило для выявления такого нарушения.

Правило 2: имеет место признак дефектности текста, если выполнено, по крайней мере, одно из условий:

1) второй элемент пары встретился в тексте раньше, чем первый;

2) между однотипными элементами определенной пары отсутствует элемент другого типа, т.е. был пропущен один из элементов;

3) после первого элемента пары в оставшейся части текста отсутствует второй элемент.

При выполнении условия правила выявляется предложение, содержащее найденный элемент, для которого отсутствует парный. Далее можно вручную проверить контекст данного предложения в тексте. Отметим, что наличие одновременно обоих элементов не является необходимым требованием, однако при отсутствии одного из них стоит обратить внимание: возможно, если автор не использовал других средств, заменяющих эти элементы, нарушен порядок изложения мысли, например: «Нашим восприятием знаков, напоминающих нам об истории, репрезентирующих те или иные события в актуальном настоящем, управляют несколько важных механизмов. **Во-первых**, это «распознавание имени»....». Другие «важные механизмы» не отмечены явно в тексте, что затрудняет его целостное восприятие.

При подсчете количества одинаковых слов, употребляемых в пределах одного предложения, можно выявить плеоназмы, содержащиеся в тексте. Под плеоназмом понимается дублирование некоторого элемента смысла; наличие нескольких языковых форм, выражающих одно и то же значение, в пределах законченного отрезка речи или текста – а также само языковое выражение, в котором имеется подобное дублирование [36]. Плеоназм принято подразделять на обязательный, т.е. обусловленный языковой системой, и факультативный, т.е. не обусловленный языковой системой; факультативные плеоназмы бывают конвенциональные (закрепленные языковой нормой) и неконвен-

циональные, т.е. создаваемые заново говорящим или пишущим [37]. Плеоназмы часто встречаются в научно-технических документах, однако высокая степень содержания факультативных, в особенности, неконвенциональных, плеоназмов может говорить о дефектности научного текста. Экспериментальным путем установлено, что дефектными могут быть предложения с плеоназмами, содержащими более двух слов с совпадающими нормальными формами. Таким образом, можно сформулировать следующее правило.

Правило 3: если предложение содержит, по крайней мере, три слова, имеющие совпадающие нормальные формы, то оно является подозрительным (содержит признак дефектного плеоназма).

Приведем пример предложения, удовлетворяющего данному правилу: «На **первых** этапах обучения на **первый** план выходят **первые** два аспекта».

При построении конструкций управления важен правильный выбор падежных форм. Например, предлоги «согласно», «вопреки» управляют дательным падежом (согласно, вопреки приказу) и творительным (согласно с требованиями), постановка существительного в родительном падеже недопустима [38]. При нарушении падежа в результате лингвистического анализа предлог «согласно» не будет связан с управляемым словом, и часть речи для слова «согласно» может быть установлена как наречие. Поэтому признаком дефекта является наличие несвязанного с другими элементами предложения слова «согласно» и следующего за ним слова в родительном падеже. Ниже представлено правило, с помощью которого можно выявлять такое нарушение управления.

Правило 4: если предложение содержит синтаксический элемент «согласно» или «вопреки», не связанный с другими элементами предложения, и первое слово, которое расположено после этого элемента и непосредственно следующих за ним наречий, союзов и числительных в числовой записи, имеет характеристику «падеж» и стоит в родительном падеже, то такое предложение содержит признак нарушения управления.

Такая сложная структура правила необходима для того, чтобы не пропустить предложения, в которых слово с родительным падежом не является первым, например, «*Согласно 4 пряных* законов...», и в то же время, чтобы не отнести к подозрительным правильные предложения, например, «*Согласно* [5], нет *необходимости* выполнять...».

Еще одним признаком нарушения грамматической нормы является следование за словами «более» и «менее» сравнительной или превосходной степени прилагательного. Наличие большого числа таких сочетаний может являться существенным недостатком текста. Опишем правило для выявления такого дефекта.

Правило 5: если предложение содержит слово «более» или «менее» и следующее за ним слово является прилагательным в сравнительной или превосходной степени, то предложение обладает признаком дефекта.

Примерами дефектов являются, например, словосочетания «более энергичнее», «более оптимальный».

Опишем далее общий алгоритм обнаружения дефектов.

Предлагаемый алгоритм обнаружения дефектов в научно-технических текстах заключается в последовательном применении сформированных правил к результатам лингвистического анализа, которые представляют собой полуструктурированные данные, т.е. текст с установленными свойствами его элементов. Такие данные позволяют учитывать семантические, синтаксические, морфологические и лексические характеристики элементов текста, их контекст и взаимную сочетаемость. Лингвистический анализ проводится на первом шаге алгоритма обнаружения дефектов с использованием существующих синтаксического и семантического анализаторов [39-40]. В ходе работы алгоритма при выполнении условия какого-либо правила предложение, содержащее признак нарушения, добавляется в структурированный список подозрительных предложений вместе с меткой типа дефекта. Одновременно увеличивается показатель количества выявленных признаков дефектов соответствующего типа. Такие показатели далее планируется применить к вычислению общей степени дефектности

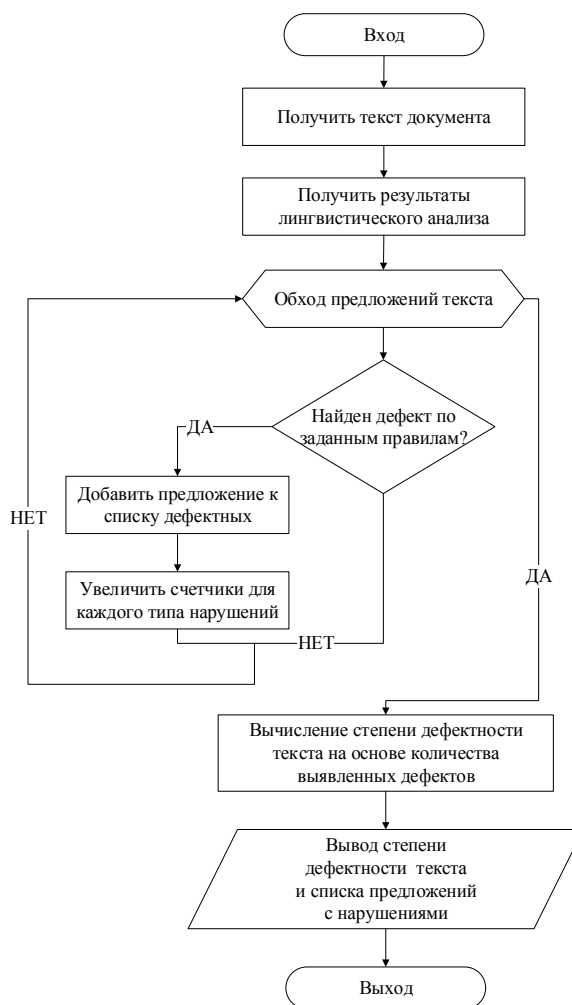


Схема алгоритма обнаружения дефектов в научно-технических текстах

научного текста. В общем виде алгоритм обнаружения дефектов в научно-технических текстах представлен на рисунке

2. Результаты применения метода автоматического обнаружения дефектов

Для проверки эффективности выявления предложений, содержащих нарушения правил русского языка, с помощью разработанного метода анализа дефектности и описанного выше алгоритма было проанализировано свыше 600 публикаций, которые в основном представляют собой статьи научных студенческих конференций. Результат выполнения алгоритма показал,

что среди выделяемых по правилам предложений действительно содержатся дефектные. Примеры дефектов, выявленных автоматически в соответствии с предложенными правилами, приведены в таблице. Средняя колонка « k/N » содержит два показателя: N – общее число автоматически выявленных предложений, содержащих признаки дефектов, k – количество тех выявленных предложений, дефектность кото-

рых подтверждена экспертами.

Поясним численные данные второго столбца в первой строке таблицы. При проверке выполнения условия первого правила было автоматически проанализировано свыше 12 тыс. предложений, содержащих причастный оборот. Из этого множества предложений было отобрано всего 98 предложений, удовлетворяющих условию правила, что составляет около 1% от всех

Примеры автоматически выявленных дефектов

Тип нарушения	k/N	Примеры дефектных предложений
1. Отсутствие согласования причастия с определяемым словом	15/98	С учетом ранее полученных данных, о том, что пассивные дети чаще встречаются в семьях недостаточно стимулирующего типа, можно предположить, что экспериментатор пытается компенсировать тип взаимодействия, сложившейся в семье, побуждая ребенка к активным действиям. Была составлена анкета для опроса жителей, которая включала спектр вопросов по отраслям (ЖКХ, потребительский рынок, образование, здравоохранение, культура, спорт, социальная поддержка населения и т. д.), оценивающую инфраструктуру данного города.
2. Нарушение последовательности вводных слов «во-первых, во-вторых»	15/15	...перед испытуемыми стоит задача, во-первых , методом проб и ошибок найти значимые клавиши, удерживать их в памяти, а затем , согласно инструкции, осветить ячейки в порядке возрастания цифр. Стремление к общению приглушено по двум причинам. Во-первых , высокая критичность к другим не способствует накоплению позитивного опыта общения... (вторая причина не указана).
3. Нарушение последовательности «с одной стороны, с другой (стороны)»	40/58	Отсутствие реального опыта собственного материнства, осмысление опыта родительской семьи ориентирует девушек, с одной стороны , на формирование с будущим ребенком доверительных, близких, товарищеских отношений. (продолжение мысли отсутствует) В данном контексте не следует рассматривать Европейский союз как иерархичную систему управления. С одной стороны , сама логика возникновения... Однако Европейский союз, в отличие от национального государства, — это скорее переговорная система...
4. Дефектные плеоназмы	не определено / 4225	Генерация музыкальной структуры может применяться как для простого музыкального представления самоподобия, так и для создания сложных музыкальных структур в рамках современного музыкального творчества. То, что я назвал понятием, в этих школах обычно называют содержанием понятия, хотя содержание этого содержания может несколько варьироваться от школы к школе и соответственно отличаться от моего.
5. Неправильный выбор падежной формы после предлогов «согласно», «вопреки»	9/12	Все работы по расконсервации скважины проводятся согласно типовых правил и инструкций , с противовыбросовым оборудованием и герметизирующей головкой, установленными на скважине для предотвращения аварийного выброса нефти. Урожайность кукурузы в большей степени, согласно наших расчетов , зависит от осадков за май месяц ($r > 0,5$), потом идет сумма осадков за май и июнь месяцы.
6. Сравнительная степень прилагательного после слов «более» и «менее»	2/4	Динамика разрушения ПДС и восстановления подвижности воды происходят в карбонатах более медленное , чем в кварцевых моделях. Более ярче это проявилось в карбонатных пористых средах.

предложений с причастными оборотами, встречающихся в текстах данной выборки. Таким образом, для выявления дефектных предложений такого типа требуется рассмотреть лишь 98 предложений, оставшиеся же 12 тыс., благодаря методу, исключаются автоматически, что значительно упрощает работу по выявлению таких нарушений.

В связи с большим количеством выявленных плеоназмов, экспертная оценка числа дефектных предложений не проводилась. Вместо этого была установлена оценка уровня содержания плеоназмов в тексте, которая показала, что количество слов, образующих плеоназмы, не превышает 0.75% от количества всех слов текста. Значит, для проверки плеоназмов на дефектность, в одном тексте потребуется проанализировать лишь несколько автоматически выделенных предложений.

При анализе остальных полученных численных данных видно, что уровень согласованности данных автоматического анализа и экспертной оценки достаточно высок. Так, при выявлении нарушений последовательности вводных слов «во-первых, во-вторых» заключения экспертов во всех рассматривавшихся случаях подтвердили корректность результатов работы программы (соответствующие показатели – 15 и 15). В отношении нарушений при выборе падежной формы существительного после предлогов «согласно» и «вопреки» также достигается достаточно высокий уровень согласованности (9/12).

Заключение

В работе предложен метод и алгоритм автоматического обнаружения признаков лингвистических дефектов в научно-технических текстах. Показано, что лежащие в основе метода правила позволяют находить нарушения речи в предложениях русского языка. Таким образом, предложенный метод применим к обнаружению дефектов, что обуславливает целесообразность его дальнейшего развития.

Далее планируется определить способ вычисления степени дефектности текста на основе эмпирически полученных статистических данных о пороговых значениях количества определенного рода ошибок, которое не является кри-

тическим для полноценного функционирования публикации. С помощью автоматического анализа двух групп научных публикаций – предварительно оцененных экспертами как дефектные и как качественные – будет установлено, какие нарушения, в каком количестве и сочетании могут становиться препятствием для адекватного понимания содержания и создавать у читающего человека ощущение общей некачественности текста.

Литература

1. Пумпянский А.Л. Введение в практику перевода научной и технической литературы на английский язык. М.: Наука, 1965. 304 с.
2. Карасик В.И. О категориях дискурса // Языковая личность: социолингвистические и эмотивные аспекты: Сб. науч. тр. Волгоград: Перемена, 1998. С. 185-197.
3. Сенкевич М.П. Стилистика научной речи и литературное редактирование научных произведений. М.: Высшая школа, 1984. 320 с.
4. Валеева Н.Г. Жанрово-стилистическая характеристика научных текстов. Введение в переводоведение. М.: РУДН, 2006. 85 с.
5. Селезнева Н.А. Использование модальных глаголов для осуществления функций научного текста // Актуальные проблемы языкознания и литературоведения. Университетские чтения ПГЛУ. Пятигорск, 2008. [Электронный ресурс] http://www.pglu.ru/lib/publications/University_Reading/2008/II/uch_2008_II_00019.pdf.
6. Чернявская В.Е. Коммуникация в науке: нормативное и девиантное. Лингвистический и социокультурный анализ. М.: Книжный дом «ЛИБРОКОМ», 2011.
7. Троянская Е.С. К общей концепции понимания функциональных стилей // Особенности стиля научного изложения / Отв. ред. Е.С. Троянская. М.: Издательство «Наука», 1976. С. 23-82.
8. Котюрова М.П. Речевые погрешности и их причины (к вопросу о культуре письменной научной речи) // Речевое общение (Теоретические и прикладные аспекты речевого общения). Специализированный вестник КрасГУ. Вып. 8-9 (16-17). 2006.
9. Алпатов В.М. Вступительная статья // Фрей А. Грамматика ошибок: Пер. с англ. М.: КомКнига, 2007. С. 8-12.
10. Эльконин Д.Б. Развитие устной и письменной речи учащихся / Под ред. В.В. Давыдова, Т.А. Нежной. М.: ИНТОР, 1998. 112 с.
11. Казанская А.В. Речевые ошибки в мотивационном аспекте. Дис. ... канд. психол. наук. Москва, 1998. 152 с.
12. Выготский Л.С. Собр. соч.: В 6 т. М.: Педагогика, 1982–1984. Т. 2.
13. Журавлёв И.В. Семиотический анализ расстройств речемыслительной деятельности. М.: Издательство ЛКИ, 2007.

14. Микиртумов Б.Е., Ильичёв А.Б. Клиническая семантика психопатологии. СПб.: СПбГПМА, 2007.
15. Пашковский В.Э., Пиотровская В.Р., Пиотровский Г.Р. Психиатрическая лингвистика. М.: Книжный дом «ЛИБРОКОМ», 2009.
16. Восейкова М.Д. Ранние этапы усвоения детьми именной морфологии русского языка. М.: Знак, 2011.
17. Путь в язык: Одноязычие и двуязычие. М.: Языки славянских культур, 2011.
18. Седов К.Ф. Онтопсихоллингвистика: становление коммуникативной компетенции человека. М.: Лабиринт, 2008.
19. Цейтлин С.Н. Речевые ошибки и их предупреждение. М.: Книжный дом «ЛИБРОКОМ», 2009.
20. Выборнов А.В. Словарь типичных ошибок английского языка. М.: Книжный дом «ЛИБРОКОМ», 2012.
21. Грамматические аспекты перевода. М.: Издательский центр «Академия», 2012.
22. Судовцев В.А. Научно-техническая информация и перевод. М.: «Высшая школа», 1989.
23. Стилистика и литературное редактирование. М.: Гардарики, 2007.
24. Козлова М.М. Редактирование материалов массовой информации. СПб.: СПБИНВЭСЭП, 2009.
25. Котюрова М.П., Баженова Е.А. Культура научной речи. Текст и его редактирование. Учебное пособие для вузов. Издание 2 М.: Флинта 2008.
26. Беззубов А.Н. Введение в литературное редактирование. СПб.: СПбГУ, 1997.
27. Владимирова Т.Л. Язык и стиль научного текста: учебное пособие. Томск: Изд-во Томского политехнического университета, 2010.
28. Иванова Т.Б., Баженова Е.А., Дускаева Л.Р. Орфографические, пунктуационные, речевые нормы русского языка в таблицах и тестах: Учебное пособие. Пермь: Перм. ун-т., 2006.
29. Культура деловой речи. Составление и оформление документов служебного и личного характера // [Электронный ресурс] <http://do.gendocs.ru/docs/index-51478.html>.
30. Титова Е.В. Методологические ошибки в педагогических исследованиях // Северо-запад России: педагогические исследования молодых ученых / Под ред. А.Г. Козловой, Т.С. Буториной, А.П. Тряпицыной. СПб: ООО «Нестор», 2005.
31. Колкер Я.М. Анализ текста путем выделения коммуникативных блоков (с опорой на коммуникативные блоки) как один из приемов проверки его большей или меньшей доступности для восприятия // Смысловое восприятие речевого сообщения в условиях массовой коммуникации / Под ред. Т.М. Дридзе, А.А. Леонтьева. М.: Издательство «Наука», 1976.
32. Крижановская Е.М. Коммуникативно-прагматическая структура научного текста. Дис. ... канд. филол. наук. Пермь, 2000.
33. Салимовский В.А. Жанры речи в функционально-стилистическом освещении (русский академический текст). Дисс. ... док. филол. наук. Екатеринбург, 2002.
34. Кузнецова Ю.М., Осипов Г.С., Чудова Н.В., Швеца А.В. Автоматическое установление соответствия статей требованиям к научным публикациям // Труды ИСА РАН. 2012. Т. 62. Вып. 3. С. 132-138.
35. Бабайцева В.В., Чеснокова Л.Д. Русский язык. Теория. 5-9 классы. М.: Дрофа, 2012.
36. Лебедева Л. Плеоназм. В кн.: Русский язык: Энциклопедия. М., 1979.
37. Ляховецкая О.Я. Виды плеонастических выражений в разноструктурных языках. В кн.: Семантические процессы и их проявление в языках разного типа. Саратов, 1985. - 129 с.
38. Розенталь Д.Э., Джанджакова Е.В., Кабанова Н.П. Справочник по правописанию, произношению, литературному редактированию. М.: ЧеРо, 1999.
39. Осипов Г. С., Смирнов И. В., Тихомиров И. А., Соченков И. В. Система интеллектуального вертикального поиска // Труды двенадцатой национальной конференции по искусственному интеллекту с международным участием КИИ-2010. - М.: Физматлит, 2010. Т. 4. С. 45-54.
40. Сокирко А. В. Семантические словари в автоматической обработке текста: По материалам системы ДИАЛИНГ: Дисс. ... канд. тех. наук. Москва, 2001. 120 с.

Швец Александр Валерьевич. Инженер-исследователь ООО «Технологии системного анализа». Окончил Сибирский федеральный университет в 2011 году. Автор 12 печатных работ. Область научных интересов: компьютерная лингвистика, математическое моделирование, методы оптимизации, искусственный интеллект. E-mail: alexandershvets@mail.ru.

Кузнецова Юлия Михайловна. Старший научный сотрудник ИСА РАН. Окончила МГУ в 1991 году. Кандидат психологических наук. Автор 6 печатных работ. Область научных интересов: психология развития, психоллингвистика. E-mail: kuzjum@ya.ru.

Осипов Геннадий Семенович. Заместитель директора по научной работе ИСА РАН. Автор 166 печатных работ. Доктор физико-математических наук, профессор. Область научных интересов: представление знаний, приобретение знаний интеллектуальными системами, динамические интеллектуальные системы, семантический поиск. E-mail: gos@isa.ru

Латышев Андрей Валерьевич. Директор по инновационным технологиям ЗАО «РосИнтернет Технологии». Автор 9 печатных работ. Кандидат технических наук. Область научных интересов: управление нагрузкой серверов в распределенных системах, методы извлечения данных, поиска и аналитической обработки неструктурированной информации. E-mail: andrey.latyshhev@gmail.com