

В ходе пропозиционального анализа вторичных текстов нам удалось выявить, что (1) в среднем, воспроизводится не более 40 % информации; (2) главные пропозиции воспроизводятся при помощи глаголов и глагольных конструкций, при пересказе, как правило, большинство из них сохраняются. Ученики также воспроизводили агенты и объекты; (3) в пересказах чаще опускаются инструменты, сирсконстанты времени, места, цели.

Список использованных источников:

1. Арутюнова, Н. Д. Понятие пропозиции в логике и в лингвистике / Н. Д. Арутюнова // Изв. АН СССР. Сер. лит. и яз. — № 1, Т. 35. — М.: Наука, 1976. — С. 46–54.
2. Падучева, Е. В. О семантике синтаксиса / Е. В. Падучева. — М, 1974. — 292 с.
3. Солнышкина, М. И. Пропозициональное моделирование для оценки информативности текста / М. И. Солнышкина, Е. В. Мартынова, М. И. Андреева // Ученые записки Национального общества прикладной лингвистики. — № 3 (31), 2020. — С. 47–57.
4. McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. Automated evaluation of text and discourse with Coh-Matrix. Cambridge, M.A.: Cambridge University Press. 2014. — 270 p.

УДК 81'373.46

Л. А. Мухаметзянова

Науч. рук. — д-р филол. наук, профессор *Солнышкина М. И.*
Казанский (Приволжский) федеральный университет
Институт филологии и межкультурной коммуникации

ТЕРМИНОЛОГИЧЕСКАЯ ПЛОТНОСТЬ КАК ПРЕДИКТОР СЛОЖНОСТИ ТЕКСТА

Понимание научных текстов предполагает, что читатель должен обладать знаниями соответствующей — специализированной — лексики. Исследования показывают, что терминологическая плотность текста — хороший предиктор сложности текста. Однако кроме специальных терминов существует пласт лексики общенаучного характера. Как отмечает Л. И. Борисова, терминологическая лексика занимает только 25 % лексики, в то время как лексика общего характера, то есть общенаучные и общеупотребительные слова, составляет 75 % [1, с. 117–121].

Сложность текста — валидированный параметр восприятия и понимания, детерминирующий мотивацию к учению и, в конечном счете, успех обучения. В данной статье предлагается метод оценки сложности текста, лексическим предиктором ко-

того является доля общенаучной лексики. В рамках данной научной проблематики создание словаря общенаучной лексики для автоматического анализа сложности научных текстов является актуальной задачей. Дополнительную актуальность создания данного словаря обеспечивает отсутствие современного словаря общенаучной лексики русского языка. Предметом исследования явились сложность общенаучной лексики и технология создания словаря общенаучной лексики.

Созданию словаря общенаучной лексики предшествовало исследование, подтвердившее, что слова общенаучного характера представляют большую для восприятия трудность, чем общеупотребительные слова. Материалом данного этапа исследования стали 20 предложений из НКРЯ [2]. В исследовании использован метод айтрекинга, а также программы Experiment Builder и Data Viewer. Айтрекер — устройство, представляющее собой систему регистрации движения глаз. В процессе чтения камера фокусируется на зрачке реципиента и передает некоторую информацию на компьютер. Айтрекер записывает время, координаты и длительность фиксаций [3].

Исследование осуществлялось для подтверждения/опровержения следующей гипотезы: общенаучные слова имеют большую эксплицитность фиксаций взглядов, а частоты их использования в речи ниже частот общеупотребительных слов. Алгоритм исследования включал следующие этапы:

1. Отбор списка общенаучных слов разной длины и частотности.
2. Отбор 20 предложений, содержащих одно общенаучное слово и одно общеупотребительное слово одной длины. Гипотезой для каждого предложения является утверждение о том, что общенаучное слово представляет большую сложность для восприятия.
3. Чтение текста каждым из респондентов списка из 20 отобранных предложений с экрана с использованием айтрекера.
4. Выгрузка и статистическая обработка данных, интерпретация результатов исследования, подтверждение или опровержение гипотезы.

Продемонстрируем алгоритм исследования на примере 17 предложения, прочитанного участником 17. Гипотеза к предложению 17 сформулирована следующим образом: слово «обобщить», являясь общенаучным, будет иметь большую эксплицитность в предложении «Я решил обобщить некоторые свои знания», поскольку его частота ниже частот всех других слов в предложении. Исследование показало, что среднее значение

длительности фиксации на слове «обобщить» — 454 мс., длительности фиксации на слове «знание» — 354,9 мс. Таким образом, слово «обобщить» имеет большую эксплицитность, что выше среднего значения по данному предложению. Следовательно, данный вывод дает основания считать гипотезу подтвержденной: общенаучное слово имеет большую длительность фиксации, чем общеупотребительное слово одинаковой длины.

Анализ данных показал, что гипотезы к 14 предложениям из 20 оказались подтвержденными, что составляет 70 %. Это дает право считать общую гипотезу подтвержденной.

На втором этапе исследования был разработан и реализован алгоритм создания словаря общенаучной лексики.

Второй этап исследования включал следующие шаги:

1. Создание авторского корпуса учебных текстов, включающий тексты учебников ФГОС для 8 класса по 5 предметным областям. Объем корпуса составил около 195 тысяч слов.

2. Создание референциального списка слов, зафиксированных в ранее изданных словарях. В качестве источника референциального списка при отборе слов были использованы оригинальные или переводные словники следующих словарей общенаучной лексики: Academic Word List от Averil Coxhead [4]; Частотный словарь общенаучной лексики под ред. Е. М. Степановой [5]; Словарь науки. Общенаучные термины и определения, науковедческие понятия и категории Сарыбекова М. Н., Сыдыкназарова М. К. [6]; Русско-китайский словарь общенаучной лексики С. С. Нестерова, пер. на кит. яз. Н. В. Донских (разд. «География», «Экономика», «История России»); А. Х. Юсупова, пер. на кит. яз. Лю Янь [7]. Словники иностранных изданий были переведены на русский язык и объединены в общем списке.

3. Разработка программы поиска общенаучных слов в Корпусе учебных текстов. Для удобства мы разграничили словники uni- и n-грамм. Каждый словник содержал информацию о частотности входящих в него языковых единиц. Для написания программы автоматического поиска общенаучной лексики в нашем корпусе мы импортировали три библиотеки: rymorphy2, nltk и spacy.

4. Первичный словник имел в своем составе 5125 слов, после удаления языковых единиц-дубликатов в словнике осталось 3848 слов. В финальный словник вошли лемматизированные термины (uni- и n-грамм) и их абсолютная частота.

Разработанный алгоритм создания словаря общенаучной лексики может быть использован на материале текстов других

предметных областей. Перспектива исследования видится в валидации созданного словника общенаучной лексики на полном Учебном корпусе русского языка (УКРЯ), созданном на материале текстов 12 предметных областей [8].

Список использованных источников

1. Борисова Л. И. Об опыте создания переводческого словаря-справочника общенаучной лексики // Слово. Словарь. Термин. Лексикограф [Электрон. текстовые дан. (15,6 Мб)]: Сб. статей по материалам Междунар. науч.-практич. конф. памяти д-ра филол. наук, проф. Юрия Николаевича Марчука (г. Москва, 1–2 марта 2019) / Л. И. Борисова. — М.: ИИУ МГОУ, 2019. 1 электрон. опт. диск (CD-ROM). — С. 117–121.
2. Национальный корпус русского языка [Электронный ресурс]. — Режим доступа: <http://ruscorpora.ru/> — Дата доступа: 31.05.2021.
3. Барабанщиков, В. А. Айтирекинг в психологической науке и практике / отв. ред. В. А. Барабанщиков. — М.: Когито-Центр, 2015. — 410 с.
4. Coxhead, A. J. (1998). An academic word list (English Language Institute Occasional Publication No. 18). — Wellington, New Zealand: Victoria University of Wellington.
5. Степанова, Е. М. Частотный словарь общенаучной лексики / Е. М. Степанова. — М.: Изд-во МГУ, 1970. — 88 с.
6. Сарыбеков М., Сыдыкназаров М. Словарь науки. Общенаучные термины и определения, науковедческие понятия и категории: учебное пособие / М. Сарыбеков, М. Сыдыкназаров. — 2-ое изд-е, доп. и перераб. — Алматы: ТРИУМФТ, 2008. — 504 с.
7. Нестерова, С. С. Русско-китайский словарь общенаучной лексики / сост. С. С. Нестерова, Н. В. Донских, А. Х. Юсупова, Лю Янь. — Хабаровск: Изд-во Тихоокеан. гос. ун-та, 2010. — 160 с.
8. Солнышкина, М. И. Сложность текста как междисциплинарная проблема / М. И. Солнышкина, В. Д. Соловьев, Э. В. Гафиятова, Е. В. Мартынова // Вопросы когнитивной лингвистики. — 2022. — № 1. — С. 18–39.

УДК 811.161.3'367.4:398.91(=111)

Ю. А. Петрушевская

доцент, канд. филол. наук
МГУ имени А. А. Кулешова, г. Могилев

**БЕЛОРУССКИЕ ПОСЛОВИЦЫ В СРАВНЕНИИ
С АНГЛИЙСКИМИ ПОСЛОВИЦАМИ (К ПРОБЛЕМЕ
ПАРЕМИОЛОГИЧЕСКИХ ПАРАЛЛЕЛЕЙ)**

В статье рассматриваются белорусские пословицы, которые имеют аналоги в английском языке на фоне других языков (как европейских, так и не европейских). Обращается внимание, что наличие таких пословиц де-