

3. Selevko G.K. "Modern educational technologies" Moscow "Scientific education" 1998.
4. Kudaibergenova KS The center of innovative practice is a source of pedagogical technology. «Almaty», 2001.-75p.
5. Kudaibergenova KS The nature of competence is in the self-development of the individual. - Almaty, -2006.-86 p.
6. Ways to organize methodical work to improve the professional level of teachers // Scientific and practical collection. -Almaty, 2008.

С.С. Татиева

Рациональное использование информационных технологий на занятиях по казахскому языку

Аннотация. В данной статье рассмотрена роль информационных технологий в образовательном процессе. Выделены преимущества использования информационно-компьютерных технологий.

Ключевые слова: информационные технологии, учебный процесс, технические средства

S.S. Tatiyeva,

Rational use of tools based on information technology in Kazakh language lessons

Summary. The article describes the use of information technology in the educational process. The advantages of the use of information and computer technologies.

Key words: information technology, learning process, technical means.

Сведения об авторе:

Татиева Светлана Сериковна, магистр, сениор-лектор кафедры языков Международного университета информационных технологий.

About author:

Svetlana S. Tatiyeva, master, senior-lector at Language Department, International Information Technology University.

УДК 81-13

Шетиева А.Т.

Международный университет информационных технологий, Алматы, Казахстан

ПРИМЕНЕНИЕ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ ПРИ ОПРЕДЕЛЕНИИ СМЫСЛОВОГО СОДЕРЖАНИЯ СПОРНОГО ТЕКСТА

Аннотация: В настоящей статье на примере привлечения онлайн seo-анализа для выявления смыслового содержания текста демонстрируется возможное использование в лингвистическом исследовательском процессе в качестве дополнительных процедур информационных технологий. Материалом послужили спорные тексты, которые могут поступать на исследование специалисту-лингвисту в разном объеме и количестве.

Ключевые слова: анализ текста, смысловое содержание, seo-анализ текста, ключевые слова, семантическое ядро, спорный текст.

Введение

В эпоху глобальной цифровизации информационные технологии активно внедряются во все сферы жизни и деятельности человека. Этот процесс захватывает и исследовательскую деятельность как самую передовую сферу. В лингвистике интеграция языка и технологий началась еще во второй половине 20-го века. Сейчас имеются не только отдельные программы, но и целые направления, созданные на стыке языка и информационных технологий – компьютерная лингвистика, корпусная лингвистика, машинный перевод, информационно-поисковые технологии. Однако данная статья посвящена применению одной частной методики в рамках исследовательского процесса в лингвистике.

Перед тем, как приступить непосредственно к изложению хода исследования считаем необходимым привести определение двух основных понятий. В качестве материала исследования был взят такой тип текстов, как *«спорные тексты»*. Данное название является условным и используется по отношению ко всем речевым произведениям, «представленным на судебную экспертизу или лингвистическое исследование» [1, с.62]. Акцент здесь делается на природу текста – это речевое произведение, т.е. текст, созданный человеком, результат речевой деятельности человека.

При исследовании спорных текстов первым и важным является вопрос определения его *смыслового содержания*. Фактически, без данного шага невозможно приступить к дальнейшему изучению того или иного аспекта, поскольку между определением содержания и последующими действиями есть связь, проявляющаяся через отношение «общее – частное». Между тем под *содержанием текста* понимается «1. внутренняя (понятийная, смысловая) сторона языковых единиц, по отношению к которой внешняя (звуковая, графическая и т.п.) их сторона приобретает свойства выражения; 2. внутренняя сторона построения высказывания, его смысловое наполнение; 3. описание какой-то ситуации» [2]. Первые два понимания связаны с лингвистическим контекстом и семантикой имеющихся в тексте языковых единиц. Третье понимание требует выхода за рамки лингвистического контекста и перехода к более сложному явлению – ситуации (или экстралингвистическому контексту), в которой был создан текст. Анализ речевой ситуации предполагает обращение к форме речевой ситуации, месту, времени, участникам коммуникации, их взаимоотношениям, продиктованным правилами и условиями, принятыми в обществе, намерениям и мотивам говорящего, интерпретации высказываний слушающим и пр. С алгоритмом такого анализа можно ознакомиться в источнике [3]. Настоящая же статья охватывает первые два понимания.

Методы исследования

Ход (алгоритм) исследования содержания текста

Содержание текста (то, о чем в нем говорится) представляется в виде наиболее частотных понятий, выраженных, как правило, словами, повторяющимися в тексте, поскольку «коммуниканты как бы «ходят кругами» вокруг определённых тем, неоднократно возвращаясь к ним в течение всего разговора» [4, с.26]. Такие слова составляют *семантическое ядро* текста. На первом этапе исследования выявляются наиболее частотные в плане употребления слова.

После чего производится определение внутритекстовых связей между выявленными словами. Через установление отношений между ключевыми понятиями, составляющими семантическое ядро текста, определяется объект разговора, поскольку «смыслы отображают не столько сами явления действительности, сколько связи и отношения между ними» [5]. Иначе говоря, смысл текста, структура смысла «определяются собственно семантическими единицами и отношениями между ними» [5].

Таким образом, на этапе определения основных смыслов, выраженных конкретными лексическими единицами, пока без учета прагмалингвистических характеристик спорного текста, происходит подсчет всех слов в тексте, устанавливаются наиболее часто встречающиеся слова, затем определяются лексико-грамматические связи между ними.

Как уже отмечалось выше, чтобы определить повторяющиеся слова в тексте, производится количественный подсчет слов, репрезентирующих то или иное понятие. Для этого в настоящей статье привлекается seo-анализ на онлайн-ресурсе <https://advego.com> [6].

Хотя seo-анализ имеет совершенно иные цели и связан с такими сферами, как маркетинг, продвижение продукта, в данном случае для лингвистического исследования важными являются следующие моменты:

1) в ходе seo-анализа ведется статистический подсчет, когда выявляются наиболее частотные слова, поскольку «семантический анализ призван показать соотношение общего количества слов в тексте и значимых слов, составляющих семантическое ядро» [7]; данный шаг оптимизирует работу исследователя, поскольку машина, как правило, быстро выводит результат, экономя время;

2) использование программы полностью исключает человеческий фактор, иначе говоря, содержание текста определяется объективно.

Результаты

В качестве примера в статье используется текст из материалов уголовного дела из архива «Экспертно-оценочного центра по г. Алматы», представленный специалисту-лингвисту в рамках лингвистического исследования. Данный текст был избран поскольку его объем (количество слов) позволяет рассмотреть текст в качестве примера в рамках статьи. В целях сохранения конфиденциальности имена собственные из данного текста полностью исключены.

Исследуемый текст:

«Говорящий: Так, я, следователь [называются фамилия, имя, отчество]. Сегодня у нас 5-ое сентября 2016 года. Мною будет проведен допрос в качестве свидетеля В настоящий момент у нас 10 часов 00 минут вечера, 5-ое сентября шестнадцатого года. Так, [называются фамилия, имя, отчество], да? Я Вам разъясняю, что в настоящий момент Вы будете допрошены в качестве свидетеля. В связи с чем я Вам разъясняю Ваши права согласно статье 78-ой, я Вам разъясняю, что в качестве свидетеля для дачи показаний может быть вызвано и допрошено любое лицо, которому могут быть известны какие-либо обстоятельства, имеющие значение для дела. Также я Вам разъясняю в порядке статьи, в соответствии со статьей 110, 214, я Вам также разъясняю, что Вы имеете право отказаться от дачи показаний. Которые могут повлечь для Вас самой или Ваших близких родственников, сейчас. Фонограмма окончена...»

Текст вводится в соответствующее окно (рис. 1).

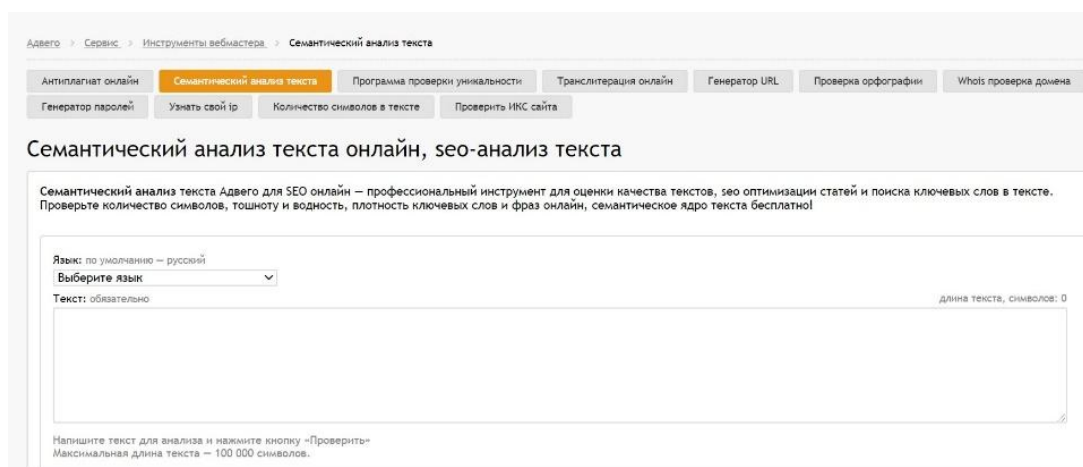


Рисунок 1 – Интерфейс онлайн-ресурса advego.com

Далее программа производит подсчет и выводит статистические данные (рис. 2).

Статистика текста

Наименование показателя	Значение
Количество символов	871
Количество символов без пробелов	737
Количество слов	127
Количество уникальных слов	68
Количество значимых слов	25
Количество стоп-слов	58
Вода	80.3 %
Количество грамматических ошибок	0
Классическая тошнота документа	2.24
Академическая тошнота документа	15.1 %

Рисунок 2 – Результаты статистического анализа текста

Как видно из Рисунка 2, статистические данные охватывают следующие параметры: количество символов, количество символов без пробелов, количество слов, количество уникальных слов, количество значимых слов, количество стоп-слов.

Остановимся лишь на представляющих интерес с лингвистической позиции составляющих. Итак, интересующим нас первым параметром является количество слов. Анализируемый текст состоит из 127 слов.

Следующий параметр – «стоп-слова». Здесь указывается «количество в тексте слов, не несущих смысловой нагрузки (предлоги, союзы, местоимения и др.)» [7]. Во время поиска по определенному запросу такие слова не учитываются поисковыми системами, что связано с их грамматическими и лексическими особенностями: предлоги, союзы и местоимения служат дополнением знаменательных частей речи, не обозначают как такового основного лексического значения. В анализе текста в силу лексических и грамматических особенностей данные слова также не учитываются, так как семантическое ядро могут составлять слова знаменательных частей речи.

Третий параметр, представляющий интерес для лингвистического исследования – «уникальные слова». Это – имена собственные, представляющие имена и фамилии людей, названия мест, организаций, улиц, городов и пр.

При подсчете наиболее частотных слов статистические программы показывают, сколько раз употреблялось одно слово в конкретной форме (см. рис. 3).

Семантическое ядро

Фраза/слово	Количество	Частота, %
разъяснить	5	3.94
качестве свидетеля	3	2.36 / 4.72
качество	3	2.36
свидетель	3	2.36
статья	3	2.36
5-ое	2	1.57
5-ое сентября	2	1.57 / 3.15
ваш	2	1.57
дача	2	1.57
дачи показаний	2	1.57 / 3.15
допросить	2	1.57

Рисунок 3 – Семантическое ядро текста

Именно на данном этапе и только по отношению к русскому языку возникают затруднения. Если английский язык уже полностью подчинен искусственному интеллекту, с русским языком, относящимся к флективным языкам, дело обстоит иначе. Поэтому программа одно и то же слово либо словосочетание, но в разных грамматических формах, т.е. с разными окончаниями, считает, как разные слова, и выдает каждый результат отдельно. Поэтому далее после статистического подсчета исследователю необходимо проанализировать результаты,

объединить одни и те же словоформы и определить семантическую связь ключевых слов между собой, исходя из их языковых характеристик, а также контекста, в котором они были употреблены.

После обработки статистических данных были выявлены ключевые слова, составляющие семантическое ядро текста: *в качестве свидетеля, разьяснять права, в соответствии со статьей, 5-ое сентября, дача показаний*. Распределение ключевых слов в соответствии с их лексическим и контекстуальным значениями позволило определить смысловое содержание текста: *допрос 5-го сентября в качестве свидетеля человека для дачи показаний. Следователь, которому принадлежат все высказывания в тексте, перед началом допроса разьясняет права*.

Заключение

Таким образом, при предоставлении на лингвистическое исследование больших по объему и количеству текстов применение информационных технологий оптимизирует исследовательский процесс, что проявляется в экономии времени и усилий лингвиста, исследующего материал. Однако с учетом указанных в статье сложностей, вызванных структурными особенностями русского языка, остается надежда, что в дальнейшем данные задачи решатся.

ЛИТЕРАТУРА

1. Нелюбин Л.Л. Толковый переводоведческий словарь, 3-е издание, переработанное. – М., Флинта: Наука, 2003, https://pervodovedcheskiy.academic.ru/1608/смысловое_содержание.
2. Создание эффективного медиатекста. Теория прагматики, http://pragmatica.ucoz.ru/publ/teorija_pragmatiki_sredstv_massovoj_kommunikacii/osnovnye_kategorii_pragmatiki_tekstov_smk/pragmaticheskoe_soderzhanie_teksta/3-1-0-5
3. Shetiyeva, A.T. Algorithm for pragmalinguistic analysis of speech // ART-SANAT, Специальный выпуск: SI. – Стамбул: Istanbul Univ, Research Inst Turkology, Dept Art History, Istanbul Univ, Research Inst Turkology, Dept Art History, 2016. – С.286-292
4. Власова Л.А., Попова И.В. Выявление смыслового содержания корпуса текстов СМС в судебной лингвистической экспертизе // Современная теоретическая лингвистика и проблемы судебной экспертизы: сборник научных работ по итогам Международной научной конференции «Современная теоретическая лингвистика и проблемы судебной экспертизы». – М., Государственный институт русского языка им. А.С. Пушкина, 2019. – С.24-34.
5. Бондарко А.В. Лингвистика текста в системе функциональной грамматики // Текст. Структура и семантика. – Т.1. М., 2001. – С.4-13.
6. <https://advego.com/text/seo/>
7. https://promopult.ru/library/Семантический_анализ
8. Энж Э., Спенсер С., Фишкин Р., Стрикчиола Д. SEO – искусство раскрутки сайтов, Пер. с англ., 2-е изд, перераб. и доп., СПб.: БХВ-Петербург, 2014. – 688 с.

REFERENCES

1. Neljubin L.L.. Tolkovyj perevodovedcheskij slovar', 3-e izdanie, pererabotannoe, M., Flinta: Nauka, 2003, https://pervodovedcheskiy.academic.ru/1608/smyslovoe_soderzhanie.
2. Sozdanie jeffektivnogo mediateksta. Teorija pragmatiki, http://pragmatica.ucoz.ru/publ/teorija_pragmatiki_sredstv_massovoj_kommunikacii/osnovnye_kategorii_pragmatiki_tekstov_smk/pragmaticheskoe_soderzhanie_teksta/3-1-0-5
3. Shetiyeva, A.T. Algorithm for pragmalinguistic analysis of speech // ART-SANAT, Special'nyj vypusk: SI. – Sтамбул: Istanbul Univ, Research Inst Turkology, Dept Art History, Istanbul Univ, Research Inst Turkology, Dept Art History, 2016. – S.286-292

4. Vlasova, L.A., Popova, I.V. Vyjavlenie smyslovogo sodержaniya korpusa tekstov SMS v sudebnoj lingvisticheskoj jekspertize // Sovremennaja teoreticheskaja lingvistika i problemy sudebnoj jekspertizy: sbornik nauchnyh rabot po itogam Mezhdunarodnoj nauchnoj konferencii «Sovremennaja teoreticheskaja lingvistika i problemy sudebnoj jekspertizy», M., Gosudarstvennyj institut russkogo jazyka im. A. S. Pushkina, 2019. – S.24-34.
5. Bondarko A.V. Lingvistika teksta v sisteme funkcional'noj grammatiki // Tekst. Struktura i semantika. T.1., M., 2001. – S.4-13.
6. <https://advego.com/text/seo/>
7. https://promopult.ru/library/Semanticheskij_analiz
8. Jenzh, Je., Spenser, S., Fishkin, R., Strikchiola, D. SEO – iskusstvo raskrutki sajtov, Per. s angl., 2-e izd, pererab. i dop., SPb.: BHV-Peterburg, 2014. – 688 s.

Шетиева А.Т.

**Ақпараттық технологияларды даулы мәтіннің
мағыналық мазмұнын анықтауда қолдану**

Аңдатпа: Берілген мақалада мәтіннің мағыналық мазмұнын анықтау үшін онлайн seo-анализді пайдалану мысалында, лингвистикалық зерттеуде қосымша ақпараттық технологияларды қолдану мүмкіндігі көрсетілген. Зерттеу материалы маман-лингвистке әртүрлі көлемде келетін даулы мәтіндер болып табылады.

Түйінді сөздер: мәтін талдауы, мағыналық мазмұн, мәтіннің seo-талдауы, түйін сөздер, семантикалық ядро, даулы мәтін

Assel T. Shetiyeva

**Application of information technologies in determining the
sense content of controversial text**

Summary. This article illustrates the possible use of information technologies as an additional procedure in the linguistic research process using an example of attracting online seo-analysis to identify the semantic content of the text. The material of the scientific research was controversial texts that can be provided to a linguist in different volumes and quantities.

Key words: text analysis, semantic content, text seo-analysis, keywords, semantic core, controversial text.

Сведения об авторе:

Шетиева Асель Тлевлеевна, PhD, ассистент-профессор кафедры языков Международного университета информационных технологий.

About author:

Assel T. Shetiyeva, PhD, assistant-professor, Languages Department, International Information Technology University.