

UDC 81-139

DOI: 10.34671/SCH.SVB.2020.0403.0011

## АНАЛИТИЧЕСКИЙ ОБЗОР СОВРЕМЕННЫХ МЕТОДОВ ЛИНГВИСТИЧЕСКОГО АНАЛИЗА ТЕКСТОВ: ТЕНДЕНЦИИ И ПЕРСПЕКТИВЫ

© 2020

SPIN-код: 3233-9838

AuthorID: 305354

**Гудкова Светлана Анатольевна**, кандидат педагогических наук, доцент кафедры  
«Теория и практика перевода»

SPIN-код: 8847-5986

AuthorID: 807978

**Дайнеко Марина Вадимовна**, старший преподаватель кафедры  
«Теория и практика перевода»

*Тольяттинский государственный университет*

(445020, Россия, Тольятти, ул. Белорусская, 14, e-mail: elena01-05@mail.ru)

**Аннотация.** Современный этап развития общества, его цифровизация и интеллектуализация, внедрение искусственного интеллекта в сферу деятельности, которая ранее считалась гуманитарной, внесли свои коррективы в методы и приемы работы с текстом, его анализом и синтезом. Возникли проблемы формирования требуемых современным обществом компетенций у специалистов XXI века, связанных как с созданием текстов, так и с его анализом и интерпретацией. Это породило потребность поиска инструментов для ускорения процесса лингвистического анализа текстов различного дискурса с учетом тенденций использования гибридных методов научного исследования в интеграции с информационными технологиями. В статье представлен аналитический обзор современных методов, приёмов анализа и синтеза текста. Увеличение частотности использования способов лингвистического анализа информации, установление корреляции между единицами анализируемого текста и ментальными процессами подготовило благодатную почву для исследования обширного спектра инструментов текстовой интерпретации, которые необходимо структурировать и систематизировать. По результатам проведенного исследования представлен теоретический обзор трудов лингвистов, филологов и ученых из смежных наук, а также авторское видение развития прикладного аспекта лингвистического анализа текстов. Показана авторская интегративная модель этапов проведения лингвистического анализа и синтеза информации, на основе гибридных методов лингвистического исследования, структуризации и систематизации.

**Ключевые слова:** цифровизация, методы анализа текста, интент-анализ, дискурс-анализ, контент-анализ, нарративный анализ, семантический анализ, синтаксический анализ, морфологический анализ, анализ тональности, кластерный анализ, гибридные методы, философия, синергетика, семиотика, системология.

## ANALYTICAL REVIEW OF MODERN METHODS OF LINGUISTIC TEXT ANALYSIS: TRENDS AND PROSPECTS

© 2020

**Gudkova Svetlana Anatolyevna**, candidate of pedagogical sciences, associate professor  
of the department «Theory and Practice of Translation»

**Dayneko Marina Vadimovna**, senior lecturer of the department  
«Theory and Practice of Translation»

*Togliatti State University*

(445020, Russia, Togliatti, st. Belorusskaya, 14, e-mail: elena01-05@mail.ru)

**Abstract.** The present scientific article is devoted to an analytical review of modern text analysis methods and techniques, as well as related approaches. The increasing frequency of using the text analysis techniques and methods, establishment of correlations between text analysis units and mental processes, as well as social phenomena have provided a fertile field for researching a wide range of methods and approaches designed for specialists to work with text. It is underlined that the methods and approaches in question need to be structured and systematized. Text analysis is the responsibility of many specialists that succeed in the modern labor market. It should be noted that the digitalization of society and the integration of artificial intelligence into different spheres that were previously considered to be humanitarian have changed the methods and techniques of work with text and has modified the competences that specialists are supposed to possess. The article highlights that the competences related to both creating and analyzing texts are highly appreciated by modern society. The modern stage of society development allows specialist to accelerate the process of linguistic text analysis referred to different discourses when taking into account the modern tendencies that involve using scientific research hybrid methods. The article provides a theoretical review of the works written by linguists, philologists and scientists dealing with related areas, as well as the author's vision of the development of the linguistic text analysis applied aspect. The research also presents the authors' integrative model of linguistic text analysis stages that is professional activity visualization of modern specialists which are charged with creating and analyzing texts.

**Keywords:** digitalization, text analysis methods, intent analysis, discourse analysis, content analysis, narrative analysis, semantic analysis, syntactic analysis, morphological analysis, tone analysis, cluster analysis, hybrid methods, philosophy, synergy, semiotics, systemology.

## ВВЕДЕНИЕ

Актуальность данной статьи в том, что представлен прикладной аспект теоретического исследования, связанного с обзором различных современных средств, обеспечивающих эффективную работу специалистов, работающих с текстом и связанных с лингвистическим анализом текста. Среди данных специалистов можно выделить переводчиков, журналистов, специалистов по связям с общественностью, редакторов, корректоров, бизнес-аналитиков и т.д. Современные автоматизированные системы анализа текста увеличивают скорость работы над ним, выполняя достаточно стандартизован-

ные задачи и, таким образом, позволяют специалистам сосредоточиться на решении более трудоёмких, сложных профессиональных вопросов и задач. Увеличение частотности использования приёмов и способов анализа текста, установление корреляций между единицами анализа текста и ментальными процессами, а также общественными явлениями подготовило благодатную почву для исследования обширного спектра методов и подходов к работе с текстом, которые необходимо структурировать и систематизировать. Анализ текстов входит в зону ответственности многих специалистов, востребованных на современном рынке труда. Необходимо

отметить, что цифровизация общества и интеграция искусственного интеллекта в сферу деятельности, которая ранее считалась гуманитарной, внесла свои коррективы в методы и приемы работы с текстом и в проблематику формирования требуемых современным обществом компетенций у специалистов различных профессий, связанных как с созданием текстов, так и с его анализом и интерпретацией. Современный этап развития общества позволяет ускорить процесс лингвистического анализа текстов различного дискурса с учетом современных тенденций использования гибридных методов научного исследования

## МЕТОДОЛОГИЯ

Анализ теоретических источников показал, что в качестве центрального понятия интерпретации текста выступает единица анализа, которая является неделимым и неразложимым звеном единого целого. Среди примеров единицы анализа текста можно выделить фонему, морфему, слово, словосочетание, предложение (лингвистический анализ); речевое действие и связанные с ним операции, концепты, фреймы (психолингвистический анализ); мотивы, намерения (психологический анализ) (О.В. Митина, А.С. Евдокименко [1, с. 29]; Е.В. Головина [2, с. 10-20]).

На толкование определенных единиц анализа текста и связей, зависимостей между ними оказывает влияние определенная модель, опирающаяся на теоретические и опытные или эмпирические знания о языке и порождении речи. Каждая модель позволяет рассмотреть ограниченный круг характеристик и свойств текста и допускает некоторые погрешности. Наиболее известными моделями считаются когнитивные модели, модели ядерных утверждений и психолингвистические модели. Необходимо также заметить, что оценка соотношения единиц языка, а также интерпретация текста на основе тех или иных моделей позволяет воссоздать конкретные аспекты и нюансы коммуникативных ситуаций (О.В. Митина, А.С. Евдокименко [1, с. 30]).

Область диагностики приёмов, подходов анализа текста включает в себя:

- отправителя информации, который обладает определенными знаниями, опытом, целеустановками, чувствами, эмоциями, мотивами, стереотипами, предрасудками;
- текст, которому свойственны контент и определённый дискурс;
- реципиент, который наделён определенными знаниями, умениями, навыками интерпретации и на которого оказывается влияние.
- экстралингвистические факторы (историко-культурные события и явления, социальная и политическая обстановка) (О.В. Митина, А.С. Евдокименко [1, с. 31]; Е.В. Полухина, Д.В. Просянюк [3, с. 106]).

Исходя из специфики организации области диагностики приёмов и подходов к анализу текста, большинство исследователей признают комплексность приёмов и методов лингвистического анализа и интерпретации текста (О.В. Митина, А.С. Евдокименко [1, с. 31-35]; Е.В. Полухина, Д.В. Просянюк [3, с. 107-108]; Т.А. Терехова, С.К. Малахаева [4, с. 143-146]; Э.В. Хилханова [5, с. 136-137]; О.Л. Чурашева [6, с. 32]; С.С. Шляхова [7, с. 100-106]).

## РЕЗУЛЬТАТЫ

В результате проведенного анализа, классификации анализируемой информации, ее структуризации, была сформирована табличная структура, отражающая авторское видение проблемы исследования.

Таблица 1 «Современные методы анализа текста» отображает аналитический обзор современных подходов и средств информационных технологий для эффективного лингвистического анализа текста. Авторская табличная модель, включающая не только методы и средства лингвистического анализа, но и конкретизирует решаемые на основе их задачи, отражая сферу прикладной

деятельности различного ряда специалистов.

Таблица 1 - Современные методы анализа текста

Метод	Задачи	Профессии	Программные средства
Фоносемантический анализ	- выявить звуковой символизм, его многозначность, иконические знаки; - оценить воздействие фонетики текста на подсознание человека и психоэмоциональное состояние; - изучить культурные и когнитивные процессы с научной точки зрения.	лингвисты, психологи, специалисты в сфере рекламы, шар, лингвокриминалисты.	CATPAC, DIATON, Vaal
Интент-анализ	- выявить авторские мотивы, намерения и установки, манипуляцию; - определить степень и характер эффекта СМИ на сознание и мнение адресатов информации; - оценить природу неясностей в спорных текстах.	Пиар, бизнес-аналитики, специалисты интернет-компаний, лингвокриминалисты-лингвисты, переводчики, психологи, маркетологи.	Minnesota Contextual Content Analysis, Leximancer, Ethnograph
Контент-анализ	- осуществить качественно-количественную интерпретацию содержания массивов текста, - оценить рассредоточение и концентрацию слов, словосочетаний в пределах текста, - генерировать тезаурус и семантическое поле; - сопоставлять списки слов, поиск по ключевым словам; - интерпретация визуального ряда: рисунки, фотографии, видеозаписи, песни, произведения искусства, публицистические статьи и т.д.	Программисты, лингвисты, литературоведы, политологи, юристы, журналисты, социологи, маркетологи, специалисты в сфере рекламы, шар-менеджеры	Yoshukoder, Kwalitan, PROTAN, INTENT, Crawdad Desktop.
Дискурс-анализ	- использование текстов с учётом определенной социальной, политической ситуации и культурных, исторических факторов.	Лингвисты, социологи, психологи, специалисты по искусственному интеллекту, этнографы, литературоведы, философы.	Smart CAT
Экспертная оценка текста	- установить адресанта текста или целевую аудиторию; - определить авторство текста: возраст, пол, национальность, место рождения, место проживания, образование, профессия, психоэмоциональное состояние; - установить обстоятельства создания изучаемого текста, выявить речевые акты (оскорбление, унижение, угроза и т.д.) в различных видах дискурса.	Лингвисты, лингвокриминалисты, переводчики, контент-менеджеры, политики, политологи, специалисты в области торгового и патентного права.	Smart CAT
Нарративный анализ	- проанализировать фактическую информацию; - интерпретировать текст, не реконструируя историческое прошлое; - определить слова какой категории используются в текст чаще всего, например, слова с положительной или отрицательной коннотацией.	Сотрудники СМИ, программисты, специалисты по связям с общественностью, социологи, политологи, психологи, управленцы, культурологи, философы.	LIWC, PC-ACE
Морфологический анализ текста	- выделить из текста словоформы; - распознать слова или их сочетания; - привести слова к словарному виду; - распознать грамматические признаки словоформ (часть речи, падеж, число, склонение и т.д.); - распознать интенцию автора в спорных текстах.	Лингвисты, лингвокриминалисты.	Textanz, ATLAS.ti, TextArc
Синтаксический анализ	- разделить текст на составные части речи формы, функции и структурных отношений внутри предложения; - составить линейную последовательность лексем с их формальной грамматикой; - построить древо зависимостей.	Лингвисты, переводчики, прагматисты	Profiler Plus и DictaScope
Семантический анализ	- сегментировать текст; - отобрать ключевые слова; - обнаруживать смысловые паттерны слов, главные, ключевые идеи текста; - осуществлять тематический анализ, сравнительный анализ, и анализ смежности.	Лингвисты, журналисты, лингвокриминалисты, психологи, политологи.	T-LAB Tools for Text Analysis

Таблица имеет также и практическую значимость, поскольку её можно использовать с целью выделения умений и навыков применительно к специалистам, чья работа предполагает проведение анализа текстов. Авторы считают, что успешное достижение поставленных перед филологами целей и задач по анализу представленного текста обеспечивается за счёт продуманной и грамотно выстроенной стратегии. Стратегии современного лингвистического исследования помогают определить, какое семантическое содержание должно быть отражено в тексте, какие риторические, стилистические и грамматические средства будут приемлемы в нём.

Данная таблица также показывает общие методы лингвистического анализа текста, но многие специалисты отмечают наличие частных методов лингвистического исследования (Е.В. Полухина, Д.В. Просянюк [3, с. 110-114]). Например, кластерный анализ подразумевает, что текст рассматривается как «мешок слов». Данный анализ допускает, что:

- 1) порядок слов в тексте не имеет значения;
- 2) часто употребляемые в тексте слова не учитываются при анализе;
- 3) текст воспринимается как вектор и хаотичное, диффузное множество слов;
- 4) разные формы одно и того же слова имеют одинаковое значение. Анализ тональности относится к формализованному подходу; предполагает оценку эмоциональной окраски текста, заложенных в нём коннотаций. Неформализованный анализ подразумевает рассмотрение текста как конгломерата смыслов. Предполагается,

что исследователь сосредотачивается на эксплицитно и имплицитно транслируемых автором смыслах, способах репрезентации информации, воссоздании авторских точек зрения и аргументов. Неформализованный анализ опирается на метод индукции, носит описательный характер и ориентирован на решение аналитических проблем. В ходе данного анализа детерминируется семантическая структура текста и исследователем генерируются коды-маркеры выявленных тем.

В любом случае авторы отмечают, что и общие, и частные методы лингвистического анализа базируются на грамотной стратегии использования достижений не только филологии и лингвистики, но и смежных дисциплин. Стратегии диктуют использование различных выразительно-образительных средств (тропов), например, эпитетов, метафор, сравнений, метонимии, гиперболы, литоты (Э.В. Хилханова [5, с. 138-139]; М.В. Катинская, А.В. Замяткина [8, с. 162-164]).

Учитывая в своем исследовании, что многие исследователи сходятся во мнении, что лингвистический анализ текста базируется на гибридных методах исследования, то есть на комплексных методах, заимствованных из разных метадисциплин, прежде всего, из философии (осознание всеобщей взаимосвязи объектов и явлений), системологии (рассмотрение текста как динамично и перманентно развивающейся системы с учетом иерархии её элементов), синергетики (объект рассматривается как система, которой присущи саморазвитие и саморегуляция за счёт воздействия внешних факторов и появления точек бифуркации) и семиотики (изучение языка как знаковой системы), авторы предлагают интегративную модель поэтапного применения гибридных методов лингвистического анализа для формирования требуемого набора компетенций.

Представленная ниже авторская модель (рисунок 1) визуализирует этапы и уровни лингвистического исследования на основе взаимосвязи гибридных методов, обеспечивающих проведение тщательного и всестороннего анализа текстов (С.А. Гудкова, А.И. Яницкий [9, с. 213-215]; В.П. Даниленко [10, с. 150-180]; З.И. Комарова [11, с. 134-147]; Р.Г. Пиотровский [12, с. 100-120]). Данный вопрос рассматривался ранее в работе Гудковой С.А. в рамках апробации вышеуказанной модели при обучении студентов различных направлений подготовки работе с иноязычным текстом [9, с. 213-215].

Приведённая выше модель наглядно демонстрирует устойчивые корреляции между различными гибридными методами, которые способны обеспечить пошаговую, равномерную и осознанную интерпретацию текста для продуктивного решения профессиональных задач как на родном, так и на иностранном языке.

Указанные в данной научной статье примеры методов анализа текстов не исчерпывают всей области их изучения с разных ракурсов. В результате проведенного нами исследования становится очевидным, что отправной точкой интерпретации текстов является определение единицы анализа. В зависимости от поставленных целей, коммуникативных установок, специфики профессиональной деятельности от специалиста требуется выбирать разные методики, базирующиеся на программных средствах. Развивая концепцию о релевантности автоматизированных систем анализа текста, мы считаем важным заметить, что они, безусловно, являются катализаторами процесса грамотного толкования текста, повышают продуктивность и эффективность работы специалистов, в обязанности которых входит интерпретация текстов, однако автоматизированные системы не могут полностью заменить человеческий интеллект. Программные средства могут успешно реализовывать шаблонные задачи, основанные на принципе алгоритма, но они не способны принимать во внимание все имеющиеся нюансы и различать интенцию автора при выполнении лингвистического анализа текста.

Следовательно, аналитические способности и навыки

интерпретации, а также резюмирования текста преувеличивают у специалиста, отсюда следует, что зачастую качественный и комплексный анализ текста должен быть прерогативой человека, а не систем искусственного интеллекта.

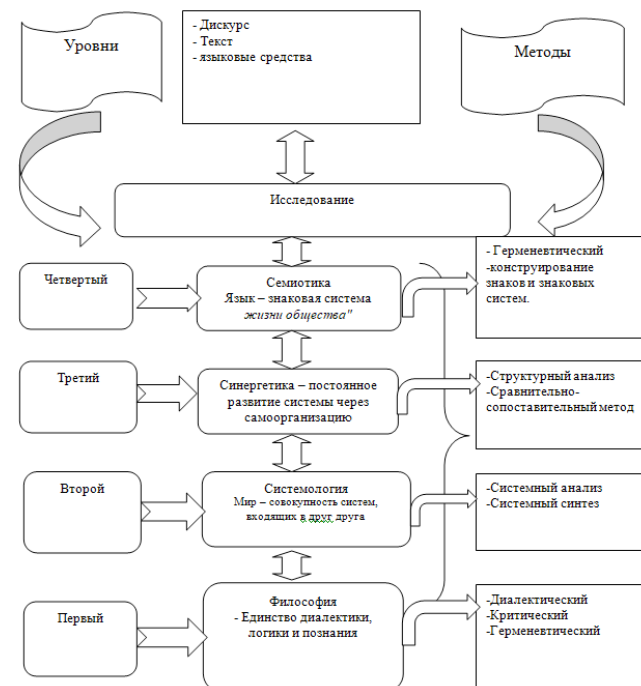


Рисунок 1 - Концептуальная интегративная модель методов лингвистического анализа текста: гибридные методы исследования

## ВЫВОДЫ

Таким образом, авторы приходят к выводу, что в настоящее время искусственный интеллект помогает успешно реализовывать шаблонные задачи в филологии и лингвистике, основанные на принципе алгоритма, но он не способен принимать во внимание все имеющиеся нюансы и различать интенцию автора при выполнении лингвистического анализа текста. Для решения профессиональных задач специалисту необходимо развивать аналитические способности и навыки интерпретации, так как качественный анализ текста до сих пор остается прерогативой человека, а не систем искусственного интеллекта.

В свете всего вышесказанного необходимо подчеркнуть, что конкурентоспособный филолог и лингвист должен владеть современными автоматизированными системами, действуя в коллаборации с ними, но, вместе с тем, владеть навыками использования гибридных научных методов для проведения качественного лингвистического анализа текста.

Перспективы дальнейшего исследования включают подбор наиболее эффективных инструментальных средств лингвистического анализа текста, выявление ключевых компетенций и критериев их оценки в разрезе каждого метода исследования и построение компетентностной модели специалиста, работающего с разными видами текстов.

## СПИСОК ЛИТЕРАТУРЫ:

1. Митина О.В., Евдокименко А.С. Методы анализа текста: методологические основания и программная реализация // Вестник Южно-Уральского государственного университета. Серия: Психология. 2010. № 40 (216). С. 29-38. URL: <https://cyberleninka.ru/article/n/metody-analiza-teksta-metodologicheskie-osnovaniya-i-programmnaya-realizatsiya> (дата обращения: 26.06.2020).
2. Головина Е.В. Лингвистический анализ текста: учебное пособие. — Оренбург: Изд-во Оренбургский государственный университет, 2012. — 106 с.
3. Полухина Е.В., Просянюк Д.В. Методы анализа текста в смешанном дизайне исследования // Политическая наука. 2015. № 2. С.



41-48. URL: <https://cyberleninka.ru/article/n/metody-analiza-teksta-v-smeshannom-dizayne-issledovaniya> (дата обращения: 30.06.2020).

4. Терехова Т.А., Малахаева С.К. Нарративное интервью как методологический конструкт междисциплинарного качественного исследования // Психология в экономике и управлении. 2013. №1. С. 91-95. URL: <https://cyberleninka.ru/article/n/narrativnoe-intervyu-kak-metodologicheskij-konstrukt-mezhdistsiplinarnogo-kachestvennogo-issledovaniya> (дата обращения: 30.06.2020).

5. Хилханова Э.В. Критический анализ дискурса: принципы, методы и практика (на примере дискурса СМИ) // Вестник Бурятского государственного университета. 2012. № SA. С. 136-140. URL: <https://cyberleninka.ru/article/n/kriticheskij-analiz-diskursa-printsipy-metody-i-praktika-na-primere-diskursa-smi> (дата обращения: 30.06.2020).

6. Чурашева О.Л. Применение метода контент-анализа в исследовании «Публичная библиотека в обеспечении информационно-психологической безопасности читателей» // Библиосфера. 2007. №3. С. 31-37. URL: <https://cyberleninka.ru/article/n/primeneniye-metoda-kontent-analiza-v-issledovanii-publichnaya-biblioteka-v-obespechenii-informatsionno-psihologicheskoy-bezopasnosti> (дата обращения: 01.07.2020).

7. Шляхова С.С. О состоянии фоносемантики в России. Часть первая. Проблемы в области исследования лингвистического иконизма // Вопросы психолингвистики. 2018. №1 (35). С. 99-104. URL: <https://cyberleninka.ru/article/n/o-sostoyanii-fonosemantiki-v-rossii-chast-pervaya-problemy-v-oblasti-issledovaniya-lingvisticheskogo-ikonizma> (дата обращения: 30.06.2020).

8. Катынская М.В., Замятина А.В. К вопросу о комплексном подходе к изучению политической публичной речи // Вестник Костромского государственного университета. 2017. Т. 23. №3. С. 160-165. URL: <https://cyberleninka.ru/article/n/k-voprosu-o-kompleksnom-podhode-k-izucheniyu-politicheskoy-publichnoy-rechi> (дата обращения: 01.07.2020).

9. Гудкова С.А., Яницкий А.И. Методология лингвистического исследования: теоретические и практические аспекты применения гибридных методов // Балтийский гуманитарный журнал. 2020. Т. 9 № 1 (30). С. 211-215. URL: [https://elibrary.ru/download/elibrary\\_42550059\\_29740731.pdf](https://elibrary.ru/download/elibrary_42550059_29740731.pdf) (дата обращения: 01.07.2020).

10. Даниленко В.П. Методы лингвистического анализа : курс лекций / В.П. Даниленко. – М. : ФЛИНТА: Наука, 2011. – 280 с.

11. Комарова З. И. Методология, метод, методика и технология научных исследований в лингвистике: учебное пособие. – Екатеринбург: Изд-во УрФУ, 2012. – 818 с. (электронная версия).

12. Пиотровский, Р.Г. Лингвистическая синергетика: исходные положения, первые результаты, перспективы / Р.Г. Пиотровский. - СПб.: Филологический факультет СПбГУ, 2006. - 160 с.

*The article was received by the editors 26.06.2020*

*The article was accepted for publication 27.08.2020*