

влияние частотности коллокаций на эффективность их обработки учащимися.

---

1. *Nesselhauf N.* Collocations in a learner Corpus. Amsterdam : John Benjamins Publ. Company, 2005. 332 p.

2. *Tsai M.-H.* Teaching L2 collocations through concept-based instruction: The effect of L2 proficiency and congruency // *Intern. J. of Applied Linguistics*. 2020. Vol. 1. P. 23.

3. *Корсакова Е. А., Мухин М. Ю.* Преподавание академического английского студентам-оптикам на основе корпусного подхода: возможности, предоставляемые цифровой лингвистикой // *EdCrunch Ural: новые образовательные технологии в вузе : материалы конф. (Екатеринбург, 29–30 сентября 2020 г.)*. Екатеринбург : Изд-во Урал. ун-та, 2020. С. 137–149.

4. *Wolter B., Yamashita J.* Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: what accounts for L2 performance? // *Studies in Second Language Acquisition*. 2018. Vol. 40. № 2. P. 395–416.

5. *Ding C., Reynolds B. L.* The effects of L1 congruency, L2 proficiency, and the collocate-node relationship on the processing of L2 English collocations by L1-Chinese EFL learners // *Rev. of Cognitive Linguistics*. 2019. Vol. 17. № 2. P. 331–357.

**М. Ю. Мухин**

*Уральский федеральный университет  
им. первого Президента России Б. Н. Ельцина  
Екатеринбург, Россия*

### **Психолингвистические аспекты лексико-статистического анализа текста\***

В исследовании представлены возможности лексико-статистического анализа для интерпретации классической литературы. На материале

---

\* Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-012-00104 «Формализация индивидуальной лексической сочетаемости как средство описания идиостилей: корпусное сопоставительное исследование классической прозы XIX в.».

корпуса произведений Ф. М. Достоевского, Л. Н. Толстого, А. П. Чехова, И. А. Гончарова, И. С. Тургенева показаны авторские особенности употребления частотных слов и выражений. Сделаны выводы о специфике психоэмоциональных доминант в творчестве писателей-классиков.

*Ключевые слова:* лексическая статистика, корпус текстов, индивидуальный стиль, лексическая сочетаемость, стилометрия, психолингвистика

**Mikhail Yu. Mukhin**

*Ural Federal University*

*named after the first President of Russia B. N. Yeltsin*

*Yekaterinburg, Russia*

## **Psycholinguistic Aspects of Lexical and Statistical Text Analysis**

This study shows the possibilities of lexical and statistical analysis for the interpretation of classical literature. The authors' features of the frequency words use and expressions are shown on the material of the corpus of works by Fyodor Dostoevsky, Leo Tolstoy, Anton Chekhov, Ivan Goncharov, Ivan Turgenev. Conclusions have been drawn on the specifics of psycho-emotional dominants in the work of classical writers.

*Keywords:* lexical statistics, text corpus, individual style, lexical compatibility, stylometry, psycholinguistics

*Введение.* Объекты изучения лингвистики текста становятся все более объемными. Современные корпуса и новые источники материала (социальные сети, электронные СМИ, блоги и т. д.) измеряются сотнями миллионов и миллиардами слов, в то же время количественный анализ получает сегодня качественное приращение. Оно касается возможной интерпретации больших текстовых данных — особенно в тех случаях, когда объем ранее не позволял исследователю охватить своим обобщением текстовые массивы. Корпуса текстов все чаще становятся источниками материала для междисциплинарных исследований психо- и социолингвистических параметров речи [1–3], а роль корпусов как источника материала порождает порой жаркие дискуссии.

Новый импульс получило и филологическое обращение к литературной классике. Принципиально изменились возможности сопоставительного анализа разных текстов и авторских стилей, так как протяженность конкретных текстов — например, романа «Война и мир» (около полумиллиона слов) — уже не является непреодолимым препятствием для лингвистической разметки (см., например: [4]). Исследование, проводимое на материале больших текстов, может выявить стилевые признаки, которые автор не мог реализовать в огромном тексте намеренно. По ним, в частности, можно судить о психологических чертах самого автора или, по крайней мере, о его приоритетах в восприятии и отражении различных фрагментов окружающей действительности.

*Материалы и методы.* Исследование проводится на материале корпуса прозаических текстов XIX в. (более 4 млн слов). Авторы текстов широко известны: это Ф. М. Достоевский, Л. Н. Толстой, А. П. Чехов, И. А. Гончаров, И. С. Тургенев. У каждого из них своя индивидуальная манера письма, о чем написано в огромном количестве филологических источников.

Основными методами данного исследования являются сравнительный статистический анализ текстов русских классиков и его гуманитарная интерпретация, в первую очередь филологическая. Проводимый анализ основан на двух частных методиках. Первая связана с выборкой слов, которые часто (и в нескольких произведениях) использует один из авторов, но не употребляют другие. Эти индивидуально-авторские частотные слова отражают стилевые приоритеты отражения действительности. Вторая предполагает анализ особенностей авторской сочетаемости слов. Их можно выявить путем сравнения контекстов слов, часто употребляемых всеми писателями. Однако каждый из них делает это особым образом (см. подробнее: [5]). В результате применения предложенных методик в руках исследователя оказываются разные по характеру массивы данных — наборы слов и фразовых контекстов, которые в совокупности характеризуют авторские стили. Часть этого материала отражает психологические черты автора и персонажей литературного произведения.

*Результаты.* Рассмотрим авторские лексические группы, выявленные путем применения первой методики. В них нас будут больше всего интересовать слова, в той или иной степени характеризующиеся эмотивной семантикой или содержащие коннотации эмоционального типа. Списки упорядочены по алфавиту, а не по статистическим параметрам или значимости для каждого авторского стиля:

А. П. Чехов: *бешеный, всхлипывать, вяло и вялый, жутко, нервно, обидно, поморщиться, пошлость, скотина, томить, чепуха;*

Ф. М. Достоевский: *болезненный, бредить, виновный, вопль, горячка, дико, злиться, раздражаться, скандал, скверный, сострадание, струсить, тварь, унижение;*

И. А. Гончаров: *авторитет, бодрый, влечение, ворчать, дружеский, живость, лукаво, нехотя, суетиться, уныло, упираться, хитрость;*

И. С. Тургенев: *загадочный, несчастливый, неудовольствие, поспорить, предаваться, пристыдить, простонать, робость, сдержанный, смелость и смелый, унылый;*

Л. Н. Толстой: *веселье, интересовать, испуганно, отвратительный, неприятность, притворство, радовать, сочувствовать, счастливо, удивленно, удовлетворение.*

Общий взгляд на внеконтекстные параметры этих списков обнаруживает, с одной стороны, эмотивный полюс Достоевского — нервического характера, доходящего до болезненности, крайней степени эмоционального проявления. С другой стороны, полюс «эмоциональной устойчивости», представленный в творчестве Гончарова и Толстого. Ряды Чехова и Тургенева можно расположить между ними на условной шкале, причем первого — ближе к Достоевскому, а второго — к Гончарову. Эти соображения в целом подтверждают филологические представления об эмотивных доминантах творчества писателей-классиков.

Вторая методика обнаруживает особенности индивидуально-авторской лексической сочетаемости. Такие лексемы, как *лицо, вдруг, жизнь, человек, шаг (шаги)* и др., входят в число самых частотных слов русского языка — как в XIX, так и теперь в XXI в. Однако в творчестве разных авторов видны свои приоритеты в плане создания особого контекстного окружения этих слов.

Например, существительное *лицо* (и значит, характеристики самого лица) наиболее интересно употребляется в текстах А. П. Чехова. Лицо у его персонажей *колючее, поношенное, насмешливое, капризное, негодующее, смеющееся, заплаканное, томное, умоляющее*. Наречие *вдруг* с семантикой неожиданности (а в реальных контекстах подчеркивающее резкость проявления психоэмоциональных процессов) — в произведениях Ф. М. Достоевского. Оно оказывается рядом с глаголами типа *вскричать, завопить, отрезать, подхватить, вскинуться, озлиться, залепетать* и др. Если анализировать авторское словоупотребление в романах И. А. Гончарова, то, например, *жизнь* его персонажей метафорически предстает как *битва* или *страсть*, может характеризоваться как *артистическая* и иметь свой интересный и важный *узор*. Контексты слова *человек* у И. С. Тургенева показывают представления писателя о человеке как *нервическом*, который с трудом *тянет лямку жизни*, а его судьба *висит на нитке* или *ниточке*. Наконец, в творчестве Л. Н. Толстого интересно используется слово *шаг* (*шаги*): *быстрый шаг, тяжелые, большие, неслышные, падающие, ковыляющие шаги*. Эти определения нужны для характеристики не только походки персонажа, но и его внутреннего состояния, а их разнообразие демонстрирует важность именно этого признака человека для писателя.

*Заключение.* Выше приведены только некоторые примеры авторского употребления слов; у каждого писателя свой большой набор индивидуальных лексических средств, многие из них используются весьма часто и не обязательно намеренно. Статическое сопоставление, которое проведено на первом этапе исследования, позволяет выявить текстовые доминанты, которые не всегда легко уловить читателю, особенно если объем текста составляют сотни тысяч слов. Интерпретация полученных данных может иметь ценность для исследований на стыке психологии, социологии и лингвистики.

---

1. Борискина О. О. Корпусная vs психолингвистика: конфронтация или кооперация // Вестн. Воронеж. гос. ун-та. Сер.: Лингвистика и межкультурная коммуникация. 2009. № 2. С. 34–38.

2. Мухин М. Ю., Лозовская А. И. Социальная сеть как источник социолингвистических данных: лексико-статистический анализ // Вестн. ЮУрГУ. Сер.: Лингвистика. 2019. Т. 16. № 4. С. 38–44.

3. Хохлова Н. В. Корпус как инструмент социолингвистики // Вестн. гуманитар. ф-та Иванов. гос. хим.-технол. ун-та. 2014. № 6. С. 188–190.

4. Скоринкин Д. А. Семантическая разметка художественных текстов для количественных исследований в филологии: на примере романа «Война и мир» Л. Н. Толстого / Д. А. Скоринкин : дис. ... канд. филол. наук. М. : [Б. и.], 2018. 164 с.

5. Мухин М. Ю., Мухин Н. Ю. Авторский тезаурус и сочетаемость слов: лексико-статистические модели индивидуального стиля // Филологический класс. 2019. № 4 (58). С. 8–15.

**И. Г. Овчинникова**

*Первый Московский государственный  
медицинский университет им. И. М. Сеченова*

Москва, Россия

*Университет Хайфы*

Израиль

### **Лексические замены как отражение сбоя механизмов речи**

Сбои лексического выбора рассматриваются в норме и при болезни Альцгеймера (БА) на материале спонтанной речи и рецитации поэтических текстов. Замены исходной лексемы при ошибках лексического выбора отражают сбой парадигматического механизма, по ходу реализации которого возможны дефекты и других механизмов речи. В зависимости от дефекта определенного механизма появляются субституты, сходные с заменяемым словом семантически, фонетически или ассоциативно. Замены любого типа появляются как в норме, так и при патологии; существенное отличие патологии, обусловленной когнитивными нарушениями вследствие БА, от нормы обнаружено в частотности различных категорий замен. В норме преоблада-