

ИННОВАЦИОННЫЕ ПЕДАГОГИЧЕСКИЕ ТЕХНОЛОГИИ В ОБРАЗОВАНИИ

АНАЛИЗ ЛОГИЧЕСКОЙ СТРУКТУРЫ РОМАНА «ТИХИЙ ДОН» С ПОМОЩЬЮ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

В.А. Бубнов

Общеинститутская кафедра естественно-научных дисциплин
Московский городской педагогический университет
2-й Тульский пер., 4, Москва, Россия, 115191

В работе используется программа Microsoft Excel для расчета относительных частот служебных частей речи в романе М.А. Шолохова «Тихий Дон», на базе которых определяется логическая структура данного романа.

Ключевые слова: самостоятельные и служебные части речи, частоты, закон устойчивости частот.

Мышление — это процесс, в результате которого рождается мысль, выраженная средствами естественного языка либо языком символов или знаков. Основным средством языка для выражения мысли в так называемой рациональной форме является словарный состав языка, образующий лексическую систему языка. В рамках лексической системы слово является важнейшей номинативной единицей языка. Представление же о слове как об основной единице наименования явлений действительности складывается непосредственно в речевой практике людей.

Большинство слов называют предметы, их признаки, количество, действия, процессы и выступают как полнозначные, самостоятельные слова, выполняющие в языке номинативную функцию. Но есть и слова, которые формируют отношения между указанными словами. Такое различие между словами в теоретических аспектах русского языка позволяет различать самостоятельные и служебные части речи [4—5].

Самостоятельные (или знаменательные) части речи либо называют предметы, качество или свойства, количество, действие или состояние, либо указывают на них. К этим частям речи относятся семь разрядов слов: имя существительное, имя прилагательное, имя числительное, местоимения, глагол, наречие, категория состояния.

В мыслительной деятельности слова, относящиеся к самостоятельным частям речи, являются объектами мысли и в качестве объектов мысли являются равноценными в психологическом отношении

Служебные части речи лишены номинативной (назывной) функции. Они проявляются в отношениях и связях между словами и предложениями (предлоги, союзы), а также при передаче смысловых и эмоциональных оттенков значений, выраженных самостоятельными частями речи (частицы). К служебным частям речи относятся предлоги, союзы, частицы.

Слова, относящиеся к служебным частям речи, устанавливают отношения между объектами мысли, которые инвариантны к различным смысловым значениям предметов мысли, т.е. множеству значений самостоятельных частей речи.

Такая роль служебных частей речи позволяет в процессе мышления из слов естественного языка формировать единые формы построения умозаключений, которые свойственны всем естественным языкам. Такие типичные формы мышления, выраженные средствами языка, формируют то, что называют *логикой мышления*.

Очевидно, что формы эти разнообразны ввиду богатства средств естественных языков.

Первая такая форма логического мышления была введена два тысячелетия тому назад Аристотелем и названа им *силлогизмом*, что в переводе с греческого означает «согласование». Силлогизм состоит из трех частей: большой посылки, малой посылки и вывода, которые следуют в таком, например, порядке: *Все люди смертны. Сократ — человек. Следовательно, Сократ смертен*. Вывод, полученный путем силлогического рассуждения, считается достоверным, если все посылки истинны и верна его форма. Форма силлогизма представлена в табл. 1.

Таблица 1

Форма силлогизма

Большая посылка	Все M суть P	Все люди смертны
Малая посылка	Все S суть M	Сократ — человек
Вывод	Все S суть P	Следовательно, Сократ смертен

Здесь предикат вывода обозначается через P , а субъект вывода — через S . Большая посылка связывает предикат вывода (смертные) со средним термином M (люди). Малая посылка связывает субъект вывода (Сократ) со средним термином, а вывод связывает субъект с предикатом.

Данная форма силлогизма представляет один из первых опытов формализации мышления, причем в ней слова, относящиеся к самостоятельным частям речи, обозначаются символами P , S , M . Последнее означает, что, заменяя эти символы словами, мы получаем различные по содержанию суждения, но одинаковые по форме. В этом и есть сущность логических рассуждений. В процессе мышления встречаются суждения в форме простых предложений, в которых утверждается или отрицается принадлежность предмету известных свойств, состояний, видов деятельности. Схема этих суждений такова: « S есть P » или « S не есть P ».

В русском языке [4—5] предложения, принадлежащие к указанной схеме мышления, делятся на *утвердительные* (утверждается то, что высказывается о пред-

мете мысли) и *отрицательные* (отрицается то, что высказывается о предмете мысли). Грамматически отрицание обычно выражается частицей *не*, а утверждение — ее отсутствием. Полное отрицание достигается постановкой частицы *не* перед сказуемым, такое предложение называется *общеотрицательным*. В математической логике [3] постулируется, что утвердительные и отрицательные предложения описывают два состояния одного и того же объекта мысли. Использование же частицы *не* при построении отрицательного предложения эквивалентно выполнению логической операции, называемой операцией *отрицания*.

Данная операция является той моделью процесса мышления, когда мы рассуждаем: «*S есть P*» или «*S не есть P*». В математике необходимо не только назвать операцию, но и дать формальное правило выполнения этой операции. В алгебре высказываний, как разделе математической логики, неделимые простые предложения называются высказываниями, и в определенной схеме рассуждений предполагается, что все утвердительные высказывания суть *истинны*, а общеотрицательные — *ложны*. Для анализа истинности или ложности высказываний вводят семантическую область *S*, состоящую из двух символов «*И — истинно*» и «*Л — ложно*». Тогда в области (на множестве) $S = \{И, Л\}$ вводят одноместную операцию отрицания по следующему правилу: если высказывание *A* есть истинно, то его отрицание (читается не-*A*, обозначается \bar{A}) — ложно; и наоборот.

Если через *A* и *B* обозначить два простых высказывания и абстрагироваться от словестного содержания этих высказываний, и рассматривать символы *A* и *B* как объекты мысли, то процесс составления сложного высказывания (предложения) связан с выполнением неких действий над данными символами. В алгебре высказываний эти действия моделируются двухместными логическими операциями над элементами множества $S = \{И, Л\}$, в котором элементы *И* и *Л* являются значениями символов *A* и *B* в так называемой семантической области. Результат логической операции будем обозначать символом *X*, значения которого являются также элементами множества $S = \{И, Л\}$.

Формальное определение двухместной операции таково. Логическая двухместная операции — это правило, по которому любой упорядоченной паре значений символов *A* и *B* на множестве $S = \{И, Л\}$ ставится в соответствие значение на этом множестве результата операции *X*. Правила выполнения логических операций называются таблицами истинности, так как они определяют истинность сложного высказывания *X* в зависимости от истинности простых высказываний *A* и *B*. В данном исследовании рассмотрим четыре таких правила (табл. 2).

Таблица 2

Правила выполнения логических операций

Комбинации символов		Результаты операций			
<i>A</i>	<i>B</i>	$X_1 = A \cdot B$	$X_2 = A \vee B$	$X_3 = A \Delta B$	$X_4 = A \rightarrow B$
<i>И</i>	<i>И</i>	<i>И</i>	<i>И</i>	<i>Л</i>	<i>И</i>
<i>И</i>	<i>Л</i>	<i>Л</i>	<i>И</i>	<i>И</i>	<i>Л</i>
<i>Л</i>	<i>И</i>	<i>Л</i>	<i>И</i>	<i>И</i>	<i>И</i>
<i>Л</i>	<i>Л</i>	<i>Л</i>	<i>Л</i>	<i>Л</i>	<i>И</i>

Перечисленные в табл. 2 логические операции моделируют мышление, в процессе которого части сложного предложения объединяются при помощи союзов и союзных слов, являющихся средствами выражения синтаксических отношений между частями сложного предложения.

Первая логическая операция $X_1 = A \cdot B$ (точка — символ операции, которую можно и не использовать) называется операцией *конъюнкции* (конъюнкция в переводе с латинского — «соединение»). Она эквивалентна использованию союза *и* при установлении соединительных *отношений* между частями сложносочиненного предложения [3]. Логическая операция $X_2 = A \vee B$ называется операцией *дизъюнкции* (дизъюнкция в переводе с латинского — разбиение, обособление), которая моделирует процесс мышления, связанный с установлением *разделительных отношений* в сложносочиненных предложениях. Кроме операции дизъюнкции, разделительные отношения между частями сложного предложения моделируются в процессе мышления логической операцией $X_3 = A \Delta B$, называемой «*исключающим или*». Обе эти операции соответствуют использованию *союза или* в сложносочиненных предложениях [3]. Операция $X_4 = A \rightarrow B$ называется операцией *импликации* (импликация в переводе с латинского — сплетение, переплетение), которая моделирует процесс применения в сложных предложениях двойных союзов, один из которых причинный (*если ... то*), а другой временной (*когда ... то*).

Очевидно, что количественные данные по рассмотренным частям речи применительно к какому-либо литературному тексту не только определяют его логическую структуру, но и позволяют выделить определенные смыслы, отражающие общие законы бытия. Эти смыслы не зависят от содержания литературного произведения, определяемого самостоятельными частями речи, но позволяют различать индивидуальность стиля автора

С целью выявления индивидуальности стиля автора романа «Тихий Дон» в [2] производился подсчет относительных частот появления предлогов *в, на, с*, имеющих место в указанном тексте. Методика такого расчета изложена в [1].

При расчете относительных частот появления в тексте «Тихого Дона» частицы *не* и союзов *и, или, если, когда* воспользуемся методикой, изложенной в [1]. Для простоты поиска в тексте двойных союзов *если ... то* и *когда — то* их поиск с помощью программы Microsoft Excel будем производить по первым словам, что конечно внесет определенную погрешность в окончательные результаты.

Итак, следуя [1], постулируем следующую гипотезу. Поиск числа повторений той или иной служебной части речи среди выбранного количества слов исследуемого текста отождествим с известной задачей математической статистики о повторении испытаний, т.е. количество слов текста будем считать числом испытаний n_i , а число m_i повторений частицы — числом появления события. Тогда можно ввести понятие частоты

$$P_i = \frac{m_i}{n_i}, \quad (1)$$

как отношение указанных чисел.

В математической статистике известны случаи, когда при увеличении числа испытаний числовые значения частот колеблются около некоторой величины и отклонения частот от указанной величины уменьшаются с ростом числа испытаний. Как правило, в качестве таковой величины принимается среднее арифметическое P_{cp} частот P_i ; и это P_{cp} согласно известной теореме Я. Бернулли принимается в качестве вероятности появления данного события при одном испытании. Если в формуле (1) символом i будем обозначать номер серии испытаний, то P_{cp} необходимо вычислять так:

$$P_{cp} = \frac{\sum_{i=1}^N P_i}{N}, \quad (2)$$

где N — число серий.

Применительно к данному роману М.А. Шолохова расчет P_{cp} для указанных служебных частей речи производится следующим образом. Текст каждого из томов романа делится на фрагменты из десяти тысяч слов. Далее начинаем разыскивать число повторений каждой служебной части речи в первом фрагменте текста из десяти тысяч слов. В этом случае число испытаний $n_1 = 10000$ назовем первой серией испытаний, а получившееся число m_1 повторений данной служебной части речи следует считать числом появлений разыскиваемого события. Теперь по формуле (1) можно вычислить частоту P_1 для данной частицы, полученную в первой серии испытаний. Для получения частоты P_2 надо к первому фрагменту текста добавить второй и для $n_2 = 20000$ с учетом нового значения m_2 вычислить P_2 по (1). Указанный процесс продолжается до тех пор, пока таким анализом не будет охвачен весь исследуемый текст. Полученный таким образом набор чисел P_i покажет, имеет ли место закон устойчивости частот.

Если закон устойчивости частот будет иметь место, то P_{cp} вычисляется по формуле (2). В таблице 3 приведены результаты расчетов частот частицы *не* и выбранных ранее союзов, выполненные в программе Microsoft Excel, текста первого тома романа М.А. Шолохова «Тихий Дон». Числа этой таблицы, выделенные полужирным шрифтом, представляют средние частоты, вычисленные по формуле (2). По данным этой таблицы построены графики (рис. 1—5), иллюстрирующие поведение частот рассматриваемых служебных частей речи в зависимости от номера серии испытаний. Эти графики свидетельствуют о выполнимости закона устойчивости частот для рассматриваемых служебных частей речи.

Таблица 3

Частотный анализ первого тома «Тихого Дона»

Номер i -той серий	Количество n_i слов в серии	Частоты P_i союзов/частицы <i>не</i>				
		<i>и</i>	<i>или</i>	<i>не</i>	<i>если</i>	<i>когда</i>
1	10000	0,01410	0,00020	0,00760	0,00010	0,00020
2	20000	0,01510	0,00025	0,00945	0,00015	0,00045
3	30000	0,01697	0,00030	0,00930	0,00013	0,00057

Номер i -той серий	Количество n_i слов в серии	Частоты P_i союзов/частицы <i>не</i>				
		<i>и</i>	<i>или</i>	<i>не</i>	<i>если</i>	<i>когда</i>
4	40000	0,01828	0,00028	0,00853	0,00013	0,00053
5	50000	0,01808	0,00024	0,00854	0,00016	0,00052
6	60000	0,01765	0,00023	0,00868	0,00013	0,00047
7	70000	0,01733	0,00023	0,00874	0,00013	0,00051
8	80000	0,01719	0,00024	0,00888	0,00014	0,00054
9	90000	0,01747	0,00024	0,00886	0,00013	0,00053
10	100000	0,01776	0,00022	0,00864	0,00013	0,00054
11	110000	0,01800	0,00022	0,00855	0,00015	0,00061
12	120000	0,01781	0,00023	0,00851	0,00018	0,00064
13	130000	0,01789	0,00023	0,00865	0,00018	0,00064
14	140000	0,01759	0,00024	0,00862	0,00017	0,00062
		0,01723	0,00024	0,00868	0,00014	0,00053

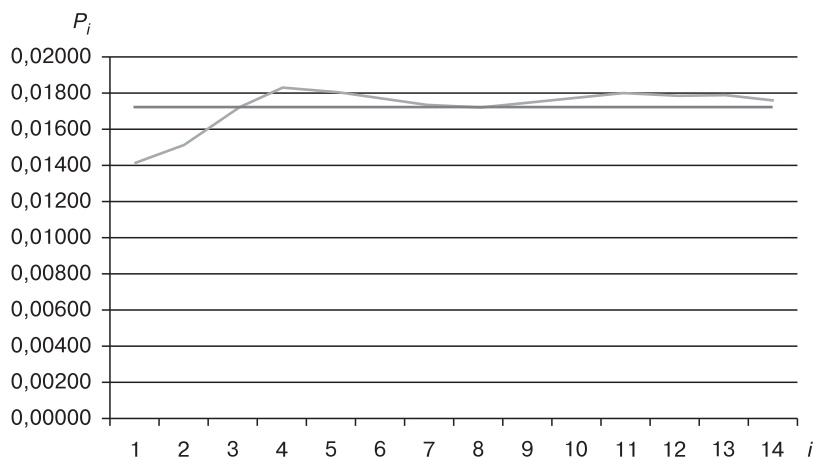


Рис. 1. Закон устойчивости частот для союза *и* («Тихий Дон», т. 1)

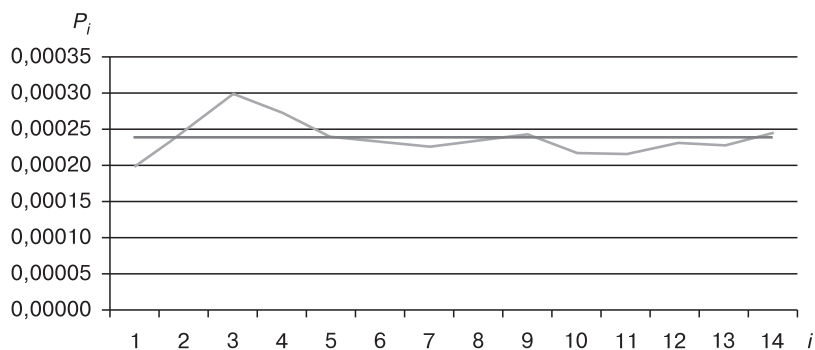


Рис. 2. Закон устойчивости частот для союза *или* («Тихий Дон», т. 1)

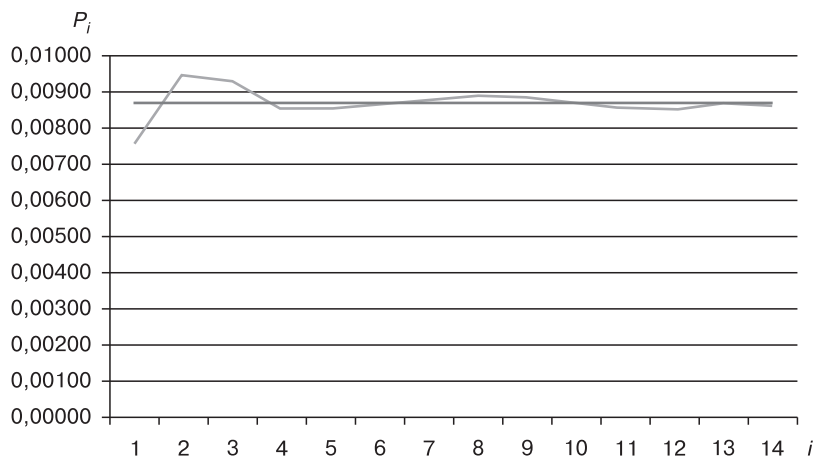


Рис. 3. Закон устойчивости частот для частицы *не* («Тихий Дон», т. 1)

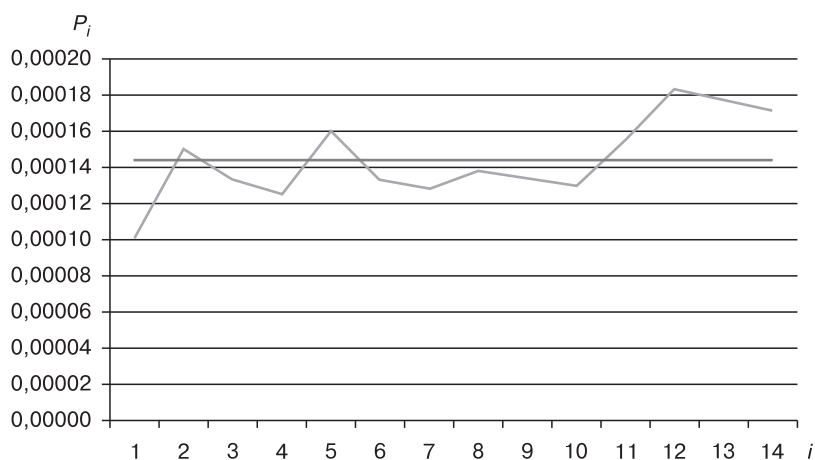


Рис. 4. Закон устойчивости частот для союза *если* («Тихий Дон», т. 1)

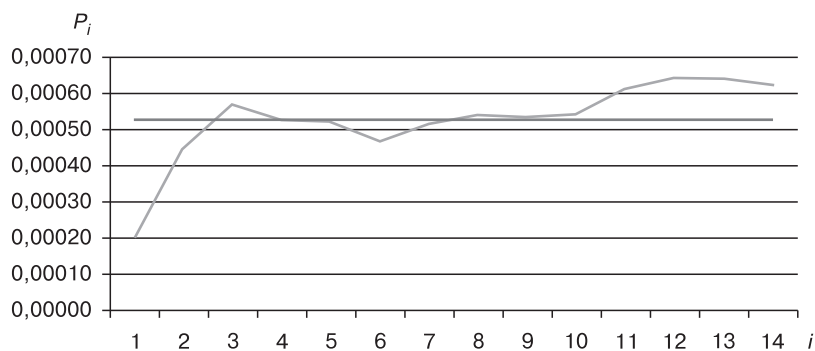


Рис. 5. Закон устойчивости частот для союза *когда* («Тихий Дон», т. 1)

Аналогичные расчеты выполнены и для остальных томов «Тихого Дона». По результатам расчетов по всем томам данного романа в табл. 4 представлены частоты, вычисленные по (2) рассматриваемых служебных частей речи.

Таблица 4

Средние арифметические частоты служебных частей речи по томам «Тихого Дона»

Номер тома	Союзы/частица				
	<i>и</i>	<i>или</i>	<i>не</i>	<i>если</i>	<i>когда</i>
1	0,01723	0,00024	0,00868	0,00014	0,00053
2	0,01740	0,00049	0,00884	0,00048	0,00076
3	0,02167	0,00038	0,00962	0,00030	0,00084
4	0,02342	0,00047	0,01494	0,00016	0,00105

Из данных табл. 4 следует, что от первого до четвертого тома частоты каждого союза/частицы отличаются незначительно. Это говорит о том, что оба тома написаны одним автором. Для контент-анализа всех четырех томов «Тихого Дона» введем понятие интегральной частоты \bar{P} как средней арифметической величины по четырем томам, вычисленной для каждого союза/частицы по следующей формуле

$$\bar{P} = \frac{1}{4} \sum_{i=1}^4 p_i, \quad (3)$$

где i — номер тома, а p_i берутся из данных табл. 4. Результаты расчетов по (3) сведены в табл. 5.

Таблица 5

Интегральные частоты служебных союзов/частицы для «Тихого Дона»

	<i>и</i>	<i>или</i>	<i>не</i>	<i>если</i>	<i>когда</i>
\bar{P}	0,01993	0,000395	0,01052	0,00027	0,000795

Данные табл. 5 представим графически так: по горизонтальной оси отложим название служебных частей речи, а по вертикальной — их частоты и полученные таким образом точки плоскости соединим ломанными линиями. В результате получим фигуру (рис. 6), которую назовем *интегральным спектром* служебных союзов/частицы *и, или, не, если, когда*.

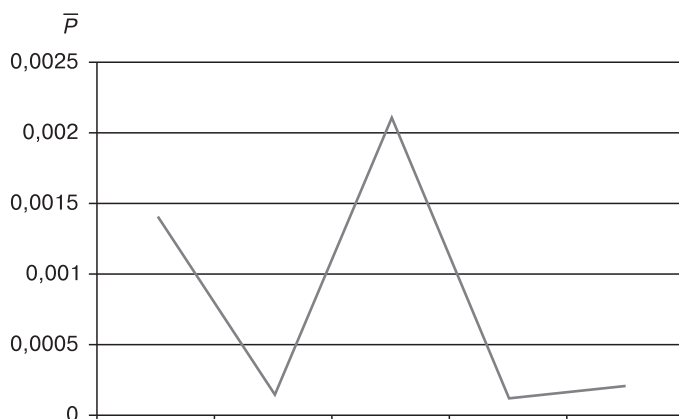


Рис. 6. Интегральный спектр служебных частей речи («Тихий Дон», т. 1)

В этом интегральном спектре преобладает частота употребления частицы не, которая логику мышления строит следующим образом.

Когда кто-нибудь думает об известном предмете, он составляет о нем известное суждение, тем самым утверждая, что «*S есть P*». Когда же начинают думать о нем же, сочетая его с другим свойством, то понятие «*S есть P*» в нашем мышлении расстраивается и это расстройство логически символизируется как отрицание «*S уже не есть P*». Теперь на место отвергнутого утверждения «*S есть P*» мышление каждого ставит новое утверждение «*S есть C*», за которым при смене мысли следует новое отрицание, за ним новое утверждение и так далее без конца.

Такая схема мышления есть не что иное, как попытка схематизации методами логики текущих явлений природы, которые представляются уму логически в виде ряда последовательных взаимно исключаящих друг друга состояний.

Проделанный анализ романа «Тихий Дон» свидетельствует о том, что законы логики сохраняются в любой деятельности человека и по ним можно определять индивидуальность этой деятельности.

ЛИТЕРАТУРА

- [1] Бубнов В.А., Сурвило А.В. Формальный анализ авторской индивидуальности литературного произведения // Информатика и образование. 2007. № 8. С. 83–87.
- [2] Бубнов В.А., Сурвило А.В. Опыт информатизации учебного процесса при обучении студентов гуманитарных специальностей // Вестник Московского городского педагогического университета. Серия «Естественные науки». 2014. № 2 (28). С. 82–91.
- [3] Бубнов В.А. Информатика и информатизация: знаково-символьный аспект: монография. М.: БИНОМ. Лаборатория знаний, 2015. 320 с.
- [4] Волгина Н.С. Современный русский язык. Синтаксис: учебник. М.: Высшая школа, 2003. 416 с.
- [5] Розенталь Д.Э., Голуб И.Б., Теленкова М.А. Современный русский язык. М.: АЙРИС-пресс, 2014. 448 с.

LITERATURA

- [1] Bubnov V.A., Survilo A.V. Formal'nyj analiz avtorskoj individual'nosti literaturnogo proizvedenija // Informatika i obrazovanie. 2007. № 8. S. 83 — 87.
- [2] Bubnov V.A., Survilo A.V. Opyt informatizacii uchebnogo processa pri obuchenii studentov gumanitarnyh special'nostej // Vestnik Moskovskogo gorodskogo pedagogicheskogo universiteta. Serija «Estestvennye nauki». 2014. № 2 (28). S. 82–91.
- [3] Bubnov V.A. Informatika i informatizacija: znakovo-simvol'nyj aspekt: monografija. M.: BINOM. Laboratorija znaniy, 2015. 320 s.
- [4] Volgina N.S. Sovremennyy russkij jazyk. Sintaksis: uchebnik. M.: Vysshaja shkola, 2003. 416 s.
- [5] Rozental' D.Je., Golub I.B., Telenkova M.A. Sovremennyy russkij jazyk. M.: AJRIS-press, 2014. 448 s.

ANALYSIS OF LOGICAL STRUCTURE OF THE NOVEL « AND QUIET FLOWS THE DON» BY MEANS OF INFORMATION TECHNOLOGIES

V.A. Bubnov

All-institute chair of natural-science disciplines
Moscow city pedagogical university
2-j Tul'skij per., 4, Moscow, Russia, 115191

In work the Microsoft Excel program for calculation of relative frequencies of auxiliary parts of speech in the novel «And Quiet Flows The Don» by M.A. Sholokhov on the basis of which the logical structure of this novel is defined is used.

Key words: independent and auxiliary parts of speech, frequencies, law of stability of frequencies.