

---

**IV Международный конгресс  
исследователей русского языка**

**РУССКИЙ ЯЗЫК:  
исторические судьбы  
и современность**

**ТРУДЫ  
и МАТЕРИАЛЫ**



**Москва, МГУ  
20–23 марта 2010 г.**

---



# **Русский язык: исторические судьбы и современность**

**IV Международный конгресс  
исследователей русского языка**

*Москва, МГУ имени М. В. Ломоносова,  
филологический факультет*

*20-23 марта 2010 года*

**Труды и материалы**

Lomonosov Moscow State University (MSU)  
Faculty of Philology

# **Russian Language: Its Historical Destiny and Present State**

**The Fourth International Congress  
of Russian Language Researchers**

Moscow, Lomonosov Moscow State University,  
**FACULTY OF PHILOLOGY**

March 20-23, 2010

**Proceedings and materials**

Collected by

*Marina L. Remneva, Anatoliy A. Polikarpov*

Moscow University Press

2010

Московский государственный университет имени М. В. Ломоносова  
Филологический факультет

# **Русский язык: исторические судьбы и современность**

**IV Международный конгресс  
исследователей русского языка**

Москва, МГУ имени М. В. Ломоносова,  
ФИЛОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ

20-23 марта 2010 года

**Труды и материалы**

Составители

*М. Л. Ремнёва, А. А. Поликарпов*

Издательство Московского университета

2010

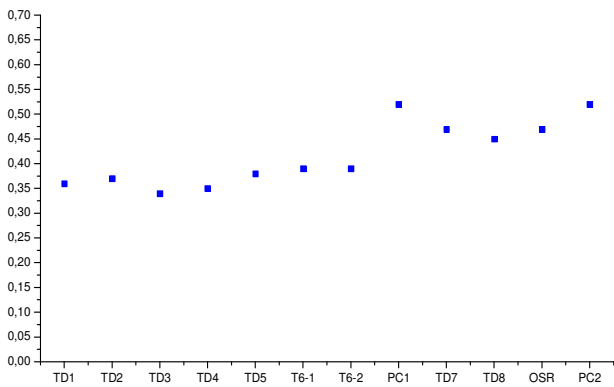


Рис. 1. Доля прямой речи в текстах Михаила Шолохова.

Хорошо видны существенные отличия объемов прямой речи в текстах первых шести частей ТД и в последних двух частях.

Для первого интервала мы видим близкие значения доли прямой речи в тексте (35–38%). Заметим, что все эти тексты были опубликованы Шолоховым в короткий промежуток времени с 1928 по 1932 г. Для сравнения мы определили долю прямой речи в некоторых произведениях Федора Крюкова, одного из предполагаемых претендентов на авторство ТД: в рассказе «Казачка» — 44%; «В родных местах» — 32%; Зыбь — 27%; «На речке лазоревой» — 41%; «Картинки школьной жизни» — 32%; «Неопалимая купина» — 38%. Можно заключить, что исследованный параметр (доля прямой речи в художественном тексте) достаточно устойчива на протяжении первого текстового интервала ТД и не сильно отличается от характеристики текстов Ф. Д. Крюкова, у которого мы наблюдаем определенное постоянство доли прямой речи в его произведениях — примерно 35% от общего объема текста. Для очерков и коротких рассказов это значение возрастало до величины примерно 40%. Причем в более значительных по объему произведениях, в которых художественный материал, очевидно, прошел более глубокую и основательную авторскую обработку, мы отмечаем соответствующее снижение доли прямой речи.

Если теперь рассмотреть тексты седьмой и восьмой частей ТД, то видно, что ситуация меняется, объем прямой речи заметно возрастает (до 45%). Такие же и даже большие значения объемов прямой речи (50–55%) мы встречаем в 1-й книге «Поднятой целины» (1932 г.) и в послевоенных текстах Шолохова — «Поднятая целина», кн. 2-я (1960 г.) и «Они сражались за Родину». Одним из первых такое изменение стиля шолоховских произведений обнаружил еще в 2000 г. ростовский исследователь А. Венков [2: 30–34], который отметил появление в тексте, начиная с конца 6-й части, длинных монологов отдельных персонажей.

Исследование относительного вклада в прямую речь шолоховских текстов «диалогов» разного объема показало, что увеличение общих объемов прямой речи в последних частях ТД, в ПЦ и в ОСР происходит за счет резкого возрастания доли «длинных» диалогов-монологов perso-

нажей: с нескольких процентов до 10–15% и более. Их относительный вклад в общий объем прямой речи представлен на рис. 2.

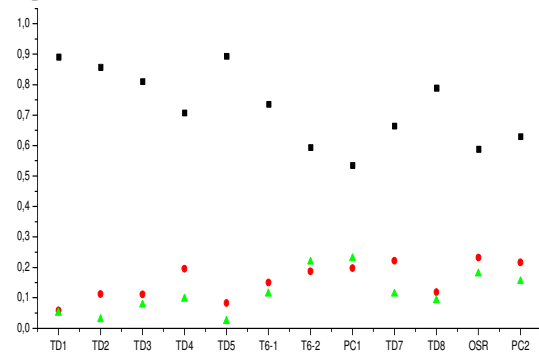


Рис. 2. Доля коротких (число букв в непрерывном фрагменте «прямой речи» менее 200 — ■), среднего объема (от 200 до 350 букв — ●) и длинных «диалогов» (▲) в текстах Крюкова, Шолохова и в «Тихом Доне»/

Мы видим, что в тексте «Тихого Дона» имеет место четко выраженная тенденция: к шестой части романа доля коротких диалогов заметно снижается при одновременном увеличении диалогов среднего и большого объемов. Начиная с середины 6-й части ТД мы можем сказать, что в текстах и «Тихого Дона», и обеих книг «Поднятой целины», и в главах романа «Они сражались за Родину» не менее трети объема прямой речи приходится на «протяженные» диалоги, а сама общая доля прямой речи в тексте возрастает с 32–35% до 45–50% и более (что соответствует появлению в тексте пространственных рассуждений и баек в духе деда Щукаря).

Таким образом, проведенное изучение шолоховских текстов подтверждает полученные другими авторами наблюдения о заметном изменении количественных характеристик авторского стиля (в данном случае — доли прямой речи) в последних частях романа «Тихий Дон», а также в последующих произведениях М. А. Шолохова — в романе «Поднятая целина» и в романе «Они сражались за Родину».

#### Литература

1. Аксенова Л. З. (Сова), Вертель Е. В. О скандинавской версии авторства «Тихого Дона» // «Вопросы литературы». 1991. Февраль. С. 68–81.
2. Венков А. В. «Тихий Дон»: источниковая база и проблема авторства. Ростов-на-Дону, 2000.
3. Загадки и тайны «Тихого Дона». Итоги независимых исследований текста романа. 1974–1994. Самара, 1996.
4. Макаров А. Г., Макарова С. Э. Цветок-Татарник. В поисках автора «Тихого Дона»: от Михаила Шолохова к Федору Крюкову. М., 2001.
5. Марусенко М. А., Бессонов Б. А., Богданов Л. М., Аникин М. А., Мясоедова Н. Е. Темные воды «Тихого Дона» // В поисках потерянного автора. СПб., 2001.
6. Фоменко В. П., Фоменко Т. Г. Авторский инвариант русских литературных текстов. Приложение. Кто был автором «Тихого Дона»? // Новая Хронология Руси и Рима. Т. 2. М., 1995.
7. Хьетто Г., Густавссон С., Бекман Б., Гил С. Кто написал «Тихий Дон»? М., 1989.

## Анализ количественных характеристик авторского стиля романа «Тихий Дон» и его соотношение с другими текстами М. А. Шолохова на основе иерархической кластеризации

А. Г. Макаров, А. А. Поликарпов

Московский государственный университет имени М. В. Ломоносова  
anatpoli@mail.ru

В. В. Поддубный, О. Г. Шевелев

Томский государственный университет  
pvv@inet.tsu.ru, oshevelyov@gmail.com

Как показали ранее проведенные эксперименты на текстах второй половины XX века, метод иерархической кластеризации текстов, основанный на использовании гипергеометрического критерия и интегральной меры рассогласования, дает осмысленную группировку текстов по различным стилям, в том числе по авторским [1]; [3]; [5]. Одними из наиболее успешно проявивших себя признаков являются частоты появления 55 служебных слов из списка А. Т. Фоменко.

Для анализа авторства Тихого Дона нами для начала был проведен ряд экспериментов по кластеризации с помощью данного метода и данных признаков текстов самого «Тихого Дона», других произведений М. А. Шолохова, а также произведений авторов, предположительно связанных с романом «Тихий Дон». В качестве первого шага исследования мы провели анализ текста «Тихого Дона» и других произведений М. А. Шолохова на «макроуровне», то есть выбрали

разбиение текстов на крупные блоки: для «Тихого Дона», как уже упоминалось выше, за основу были взяты тексты восьми частей романа, причем 6-й часть разбивалась на две половины — гл. 1–29 и гл. 30–65 (ТД-1 — ТД-5, ТД-6–1, ТД-6–2, ТД-7, ТД-8). К ним были добавлены тексты 1-й и 2-й книг «Поднятой целины» (ПЦ1, ПЦ2), «Они сражались за Родину» (ОСР) и послевоенный рассказ М. Шолохова «Судьба человека» (СЧ). «Донские рассказы», в силу жанровых отличий, а также незначительности объемов текста отдельных рассказов для надежной статистической обработки стиливых характеристик, на данном этапе исследования были из анализа исключены.

В качестве исследуемых признаков нами были выбраны служебные слова (предлоги, союзы, частицы) по списку В. П. Фоменко и Т. Г. Фоменко [4]. На рис. 1. представлено полученное дерево кластеризации для вышеупомянутых признаков для текста «Тихого Дона», разделенного на девять частей. На рис. 2 представлены результаты анализа текстов для случая, когда к исследуемому корпусу текстов были добавлены тексты ПЦ1, ПЦ2 и ОСР.

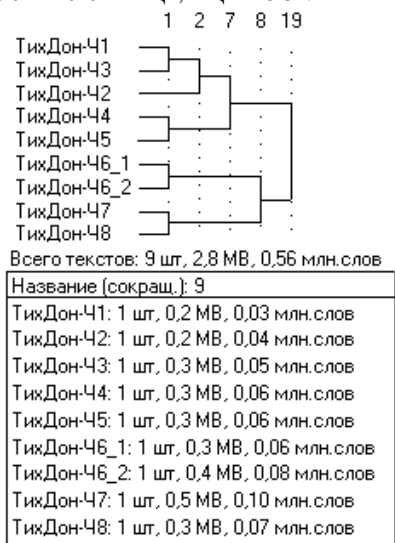


Рис. 1.

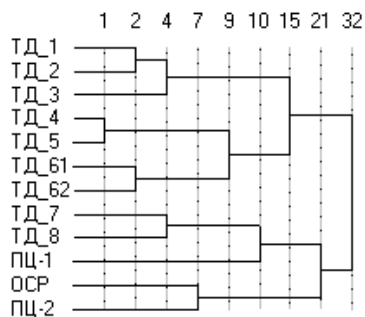


Рис. 2.

Полученные результаты количественной кластеризации выявляют характерную особенность: тексты частей ТД объединяются, группируются в **четыре** относительно самостоятельных кластера: первый кластер включает в себя первые три части ТД, второй — 4-ю и 5-ю части, третий — первую и вторую половины 6-й части. Наконец, четвертый кластер объединяет 7-ю и 8-ю части ТД. При этом наблюда-

ется сближение образовавшихся кластеров в две группы, объединяющие *первую* и *вторую* пару кластеров.

Анализ результатов кластеризации по всем указанным текстам М. А. Шолохова (ТД, ПЦ и ОСР) показал добавление к наблюдаемым четырем кластерам нового, в который вошли «послевоенные» произведения М. А. Шолохова (ПЦ2 и ОСР). Текст ПЦ1 подсоединился к четвертому кластеру текстов ТД (ТД-7 и ТД-8). Заметим, что три эти текста создавались и были опубликованы примерно в одно время начиная с 1932 по 1940 гг. В конечном счете, по результатам анализа кластеризации можно констатировать группирование исследованных текстов в три ярко выраженные стиливые группы. **Первая группа** включает тексты ТД первых шести частей романа. При этом имеет место выраженное структурирование и внутри этой группы: в отдельные подгруппы кластеры устойчиво объединяются тексты первых трех частей ТД, четвертой пятой частей ТД и, наконец, первая и вторая половины шестой части ТД. Отметим здесь, что все эти тексты были опубликованы в течение короткого промежутка времени в 1928–1932 гг. **Вторая группа** включает в себя тексты Шолохова 30-х годов: кластер из последних двух частей романа — ТД-7 (1937 г.) и ТД-8 (1940 г.) и примыкающей к ней, но заметно отличной по характеристикам, 1-й кн. ПЦ (1932 г.). **Третья группа** образуется послевоенными произведениями М. Шолохова (2-я кн. ПЦ, ОСР) и резко выделяется по частотным характеристикам из остальных текстов.

Полученные результаты оказываются весьма интересными. Характер частотных характеристик подтверждает существенную неоднородность текста романа, которая была в свое время выявлена при текстологическом анализе ТД, когда именно с середины 6-й части ТД отмечалось нарастающее соавторское вмешательство в текст, при котором нарушалась и хронология описываемых событий, и историческая точность и достоверность, менялись язык и мировоззренческая направленность авторского текста [2]. Что касается текстов ПЦ и ОСР, то они по своим стиливым количественным характеристикам заметно дистанцированы от характеристик текстов ТД. При этом стиливые характеристики послевоенных произведений Михаила Шолохова (2-я кн. ПЦ и ОСР) резко отличаются не только от стиля ТД, но и от стиля 1-й кн. ПЦ. И, наконец, следует отметить еще раз, что характеристики 1-й кн. ПЦ сближаются, но не объединяются, со стиливыми характеристиками последних частей ТД (7-й и 8-й).

## Литература

1. Кукушкина О. В., Поддубный В. В., Поликарпов А. А., Шевелев О. Г. Автоматическая классификация текстов корпуса русских текстов конца XX века по жанровым типам и источникам // Русский язык: исторические судьбы и современность. Международная конференция. Труды и материалы. Москва, МГУ, 20–23 марта 2007 г. М., 2007. С. 391–392.
2. Макаров А. Г. и Макарова С. Э. Цветок-Татарник. В поисках автора «Тихого Дона»: от Михаила Шолохова к Федору Крюкову. М., 2001.
3. Поддубный В. В., Шевелев О. Г. Сравнение стилей текстовых произведений по частотному признаку на основе гипергеометрического критерия // Теоретическая и прикладная информатика. Томск, 2004. Вып. 1. С. 101–110.
4. Фоменко В. П., Фоменко Т. Г. Авторский инвариант русских литературных текстов. Приложение. Кто был автором «Тихого Дона»? // Новая Хронология Руси и Рима. Т. 2. М. 1995.
5. Polikarpov A., Kukuškina O., Poddubnyj V., Shevelov O. Optimization of feature space in tasks of classification and clustering of texts of various types / Abstracts from Qualico-2009. Graz, 2009.

## Числовой профиль сюжета

Г. Я. Мартыненко, Т. Ю. Шерстинова

Санкт-Петербургский государственный университет

g.martynenko@gmail.com, sherstinova@gmail.com

Русский рассказ, Серебряный век, сюжет, динамика, размер предложения, временной ряд, динамический контур

**Summary.** The paper concerns dependence of quantitative text parameters on linear text structure. Russian short stories of the beginning of the XXth century are investigated. For each short story a dynamic contour of sentence length is built. Common types of dynamic contours are determined for the given text corpus. The most typical contours for Russian short stories of the Silver Age are presented.

1. В последние годы возрос интерес к ритмической организации непозитического текста. Причем ритм понимается до-  
524

статочно широко — как проявление симметричных свойств на различных уровнях структуры текста. Чем вызван этот ин-