

Российская академия наук  
Институт русского языка им. В. В. Виноградова

НАЦИОНАЛЬНЫЙ КОРПУС  
РУССКОГО ЯЗЫКА:  
2006–2008

Новые результаты и перспективы

Санкт-Петербург  
«НЕСТОР-ИСТОРИЯ»  
2009

УДК 811.161.1

ББК 81.411.2

Издание подготовлено в Институте русского языка им. В. В. Виноградова РАН и осуществлено совместно Институтом русского языка и Институтом лингвистических исследований РАН в рамках Программы фундаментальных исследований Секции языка и литературы ОИФН РАН «Текст во взаимодействии с социокультурной средой»

*Утверждено к печати*

*Институтом русского языка им. В. В. Виноградова РАН*

*Рецензенты*

д. ф. н. Л. П. Крысин, к. ф. н. А. А. Пичхадзе

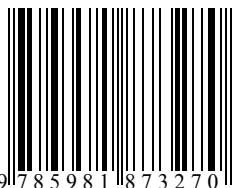
*Редакционная коллегия*

В. А. Плунгян, Е. В. Рахилина, Т. И. Резникова

**Национальный корпус русского языка: 2006–2009.** Новые результаты и перспективы / Отв. ред. В. А. Плунгян. – СПб.: Нестор-История, 2009. — 502 с.

**ISBN 978-5-98187-327-0**

В настоящем сборнике подводится итог второго этапа создания Национального корпуса русского языка ([www.ruscorpora.ru](http://www.ruscorpora.ru)) — современной и широко востребованной информационно-справочной системы по русскому языку на основе представительного электронного собрания текстов. Корпус, открытый для свободного доступа в Интернете в 2004 г., в период с 2006 по 2008 гг. был значительно увеличен в объеме и усовершенствован, в его состав вошли новые подкорпуса (поэтических текстов, устной речи и ряд других), появились новые возможности поиска и обработки данных. Все эти изменения, а также перспективы дальнейшего развития НКРЯ подробно описаны в статьях сборника. Специальный раздел посвящен использованию Корпуса при обучении русскому языку и литературе. В заключительной части сборника дается характеристика ряда других корпусов и корпусных проектов, представляющих интерес для специалистов и всех интересующихся современными технологиями в изучении языков.



© Коллектив авторов, 2009

© Институт русского языка  
им. В. В. Виноградова, 2009

© Издательство «Нестор-История», 2009

# Содержание

*Е. В. Рахилина.* Корпус как творческий проект . . . . . 7

## **I. ХРОНОЛОГИЧЕСКИЕ СРЕЗЫ РУССКОГО ЯЗЫКА В ФОРМАТЕ НКРЯ**

*С. О. Савчук.* Корпус текстов первой половины XX века:  
текущее состояние и перспективы . . . . . 27

*С. А. Оскольская.* Корпус письменных текстов XIX века:  
сферы употребления и жанровое разнообразие . . . . 46

*С. О. Савчук, Д. В. Сичинава.* Корпус русских текстов  
XVIII века в составе НКРЯ: проблемы и перспективы . . 52

## **II. ОСОБЫЕ ТИПЫ ТЕКСТОВ В СОСТАВЕ НКРЯ**

*Е. А. Гришина, К. М. Корчагин, В. А. Плунгян, Д. В. Сичинава.*  
Поэтический корпус в рамках НКРЯ:  
общая структура и перспективы использования . . . . 71

*А. Б. Летучий.* Диалектный корпус: состав  
и особенности разметки . . . . . 114

*Е. А. Гришина, С. О. Савчук.* Корпус устных текстов  
в НКРЯ: состав и структура. . . . . 129

### III. НОВЫЕ ПРОЕКТЫ В РАМКАХ НКРЯ

- Е. А. Гришина.* Корпус «История русского ударения» . . . 150  
*Е. А. Гришина.* Мультимедийный русский корпус  
(МУРКО): проблемы аннотации . . . . . 175

### IV. СЕМАНТИКА В НКРЯ

- Е. В. Рахилина, Г. И. Кустова, О. Н. Ляшевская,  
Т. И. Резникова, О. Ю. Шеманаева.*  
Задачи и принципы семантической разметки  
лексики в НКРЯ . . . . . 215  
*А. А. Кретов.* Анализ семантических помет в НКРЯ . . . 240  
*Г. И. Кустова, С. Ю. Толдова.* НКРЯ: семантические  
фильтры для разрешения многозначности глаголов . 258

### V. ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ КОРПУСНЫХ ЗАДАЧ

- А. А. Аброскин.* Поиск по корпусу: проблемы  
и методы их решения . . . . . 277  
*А. И. Зобнин, А. В. Сахарова.* Универсальная система  
разметки текста ObjectATE . . . . . 283  
*И. А. Пильщиков, А. С. Старостин.* Основные проблемы  
автоматизации базовых процедур ритмико-  
синтаксического анализа силлабо-  
тонических текстов . . . . . 298

### VI. ПЕРСПЕКТИВЫ ИСПОЛЬЗОВАНИЯ НКРЯ В РАЗЛИЧНЫХ ОБЛАСТЯХ

#### ОБРАЗОВАНИЕ

- С. О. Савчук, Д. В. Сичинава.* Обучающий корпус  
русского языка и его использование  
в преподавательской практике . . . . . 317  
*Н. Р. Добрушина.* Корпусные методики обучения  
русскому языку . . . . . 335

#### НАУКА

- М. Д. Воейкова.* Проблемы использования подкорпуса  
устной разговорной речи  
(на примере анализа русских диминутивов) . . . . 353

<i>Е. В. Падучева.</i> НКРЯ как ресурс при исследовании предметной соотнесенности имен . . . . .	374
<i>Д. О. Добровольский.</i> Корпус параллельных текстов в исследовании культурно-специфичной лексики . . .	383

## **VII. НКРЯ и другие корпуса**

<i>Т. И. Резникова.</i> Славянская корпусная лингвистика: современное состояние ресурсов . . . . .	402
<i>Б. В. Орехов.</i> Параллельный корпус переводов «Слова о полку Игореве»: итоги и перспективы . . .	462
<i>А. В. Костыркин.</i> Корпус японской разговорной речи . . .	474