

УДК 81'32

ББК 81.1

Л59

Рецензенты: канд. филол. наук., вед. науч. сотр., Л. Л. Иомдин (ИППИ РАН им. А. А. Харкевича), канд. филол. наук, доц. Д. А. Кочаров (С.-Петерб. гос. ун-т)

*Рекомендовано к печати*

*научной комиссией в области наук о языках и литературе  
Санкт-Петербургского государственного университета*

**Лингвистический анализатор: преобразование текста  
Л59 в метаязыковую структуру данных** / С. А. Кузнецов, Т. Г. Скребцова, С. Г. Суворов, А. В. Клементьева. — СПб.: Изд-во С.-Петерб. ун-та, 2019. — 238 с.  
ISBN 978-5-288-05927-8

В пяти главах монографии, упорядоченных согласно последовательности этапов обработки текста, рассматриваются графематический анализ, морфологический (атрибуция токенов), конструктивно-синтаксический и коммуникативно-синтаксический анализ предложения, функционально-прагматический анализ текста. Продемонстрирована работа лингвистического анализатора, процесс преобразования понятийно-риторических единиц текста и информационного представления текста в метаязыковую структуру данных. В приложениях даны анализ предложений из повести Ю. В. Трифонова «Обмен», примеры статей семантико-синтаксического словаря, примеры семантического разбора законодательных текстов.

Книга предназначена для специалистов в области автоматической обработки текста и будет полезна в практике преподавания компьютерной лингвистики.

**УДК 82.09  
ББК 83.3**

*Работа выполнена по проекту*

*«Автоматическое преобразование текстовой информации  
в метаязыковую структуру данных» в рамках мероприятия*

*«Проведение фундаментальных научных исследований  
по приоритетным направлениям Программы развития СПбГУ»*

ISBN 978-5-288-05927-8

© Санкт-Петербургский  
государственный  
университет, 2019  
© Авторы, 2019

## ОГЛАВЛЕНИЕ

Введение .....	5
Глава 1. Графематический анализ .....	16
Глава 2. Атрибуция токенов .....	27
Морфологический анализ буквенных токенов .....	—
Алгоритм морфологического предсказания .....	32
Атрибуция других типов токенов .....	38
Глава 3. Конструктивно-синтаксический анализ предложения .....	44
Семантико-синтаксический словарь .....	—
<i>Понятие семантико-синтаксического варианта</i> .....	—
<i>От морфологической атрибуции словоформы к ее семантико-семантическому варианту</i> .....	47
<i>Условия реализации</i> .....	49
<i>Собственное значение</i> .....	55
Алгоритм конструктивно-синтаксического анализа .....	67
Глава 4. Коммуникативно-синтаксический анализ .....	91
Теоретические предпосылки .....	—
Общая концепция информационной структуры текста .....	95
От синтаксических структур к информационному представлению текста .....	97
Накопление информации: контроль за объектами и признаками .....	102
Представление информации о пространстве и времени .....	107
Определение временной структуры текста .....	113
Автоматическое распознавание типа текста .....	135
<i>Понятие типа текста</i> .....	—
<i>Научный стиль</i> .....	139
<i>Официально-деловой стиль</i> .....	142
<i>Публицистический стиль</i> .....	144
<i>Функционально-смысловые типы речи: описание, повествование, рассуждение</i> .....	146
Разграничение диктума и модуса .....	163

Глава 5. Функционально-прагматический анализ текста.....	167
Выявление пропозиций и их анализ .....	—
Функционально-понятийные типы пропозиций .....	176
<i>Факт</i> .....	—
<i>Мнение</i> .....	183
<i>Оценка</i> .....	191
<i>Волеизъявление</i> .....	194
Композитное высказывание как единица риторической структуры текста.....	196
Синтез понятийно-риторической структуры текста.....	200
Закключение .....	205
Источники.....	207
Словари и справочники .....	208
Литература.....	209
Список сокращений в примерах .....	221
Приложение 1. Анализ предложений из повести Ю. В. Трифонова «Обмен» .....	222
Приложение 2. Примеры статей семантико-синтаксического словаря ...	228
Приложение 3. Примеры семантического разбора законодательных текстов .....	235