Частотный словарь русского языка (под редакцией Л. Н. Засориной)

Арина Лукаш.
13812.2

О словаре

Памяти Бориса Александровича Ларина

- **Словарь** - свод статистических данных о лексическом составе современного русского языка. Составлен на основании обработки средствами вычислительной техники 1 млн словоупотреблений, **40 тыс. единиц** словаря.
- Словарь составлен коллективом сотрудников филологического факультета Ленинградского государственного университета им. А. А. Жданова и Лаборатории семиотики Научно-исследовательского института прикладной математики и кибернетики при Горьковском государственном университете им. Н. И. Лобачевского.
- Для подготовительной работы над источниками словаря были привлечены около 100 человек: студенты, аспиранты и сотрудники кафедр **математической лингвистики** и **русского языка** филологического факультета ЛГУ.

Назначение словаря

- отражение **устойчивой, общеупотребительной, нейтральной** части лексики;
- более полная инвентаризация и систематизация словарного состава, разграничение активного и пассивного запаса;
- дать достаточно полные сведения о современной лексике с учетом жанровой ее дифференциации;
- составление базы для последующих лингвистических и лексикостатистических извлечений и интерпретаций;
- расчёт с помощью вычислительных машин общего количества инвентаризованных слов, статистических характеристик отдельного слова и целых словников, упорядочение частых и редких зон словника;
- последующее возможное усовершенствование графики и орфографии с учетом статистических сведений о словарном составе;
- использование данных для практической транскрипции и транслитерации, автоматизации печатного дела, распознавания и автоматического чтения буквенного текста;
- для составления различных словарей грамматик и учебников

Источники словаря

4 группы текстов-источников:

- художественная проза,
- драматургия,
- **научные и публицистические тексты, деловая речь**
- газетные и журнальные тексты.

Схема функционально-речевых сфер для словаря разработана известным лексикологом, академиком **Б. А. Лариным**. Каждой группе отведена $\frac{1}{4}$ часть выборки, около 250 тыс. словоупотреблений.

!

Функционально-речевые сферы	в % ко всей выборке
• художественная проза	25,4
• драматургия	27,2
• научные тексты	23,6
публицистические тексты	
• газетные тексты	23,8
журнальные тексты	

Формирование словаря

Элементы словаря:

- лексемы – единица словаря;
- словотипы, обладающие свойством «порождать» текстовые слова (слово-знаки).

Материалы для отбора элементов словаря:

- тексты, расчлененные на отрезки (ведущий принцип - относительная графическая эквивалентность словоформ, подчиненный - лексико-синтаксическая эквивалентность);
- букворяды: буквы русского алфавита, дефис и точка, не являющаяся концом предложения;
- транслитерированные слова, заимствованные из других языков (нкосаан, падаунги).

Аналитическая грамматика словаря:

- ограничение с учетом фонетической, графической, морфемной и синтаксической эквивалентности слов, без аспекта семантического варьирования;
- лексемы представлены в исходных формах, которые не сопровождаются дополнительными грамматическими пометами;
- переход от текста к словарю регулировался правилами грамматики,

3 части словаря

I. Алфавитно-частотный словник - наиболее важная и объемная часть словаря. Включает все лексемы, встретившиеся в текстах - 39268 разных слов.

Каждая лексическая единица словаря имеет характеристики:

- 1) частоту: а) общую частоту по всей выборке; б) частоты по подвыборкам – в 4-х жанровых группах текстов (I—газетно-журнальные тексты, II—драматургия, III—научные и публицистические тексты, IV—художественная проза);
- 2) количество текстов по жанрам, в которых встретилось данное слово

АЛФАВИТНО-ЧАСТОТНЫЙ СЛОВНИК

Словник	частота					КОЛ-ВО ТЕКСТОВ				
	общая	I	II	III	IV	I	II	III	IV	
аΔΔ	10719	1355	5203	1252	2909	15	21	43	13	
а-а	32	—	29	—	3	—	15	—	2	
аа-яй-яй	1	—	1	—	—	—	1	—	—	
абажур	24	—	17	6	1	—	5	2	1	
аббатство	2	—	2	—	—	—	1	—	—	
...	.	.	.	.	.	.	.	.	.	
столик	50	2	30	2	16	2	7	2	7	
столица	113	60	12	29	12	15	7	9	2	
стон	10	1	3	2	4	1	2	2	3	
стонать	16	—	3	3	10	—	2	3	4	

II. Частотный словник - содержит слова с частотой 10 и выше, всего 9044 единицы. Они расположены в порядке убывания частот. Наибольшую частоту имеет **в(во)**—42854.

По частотному словнику удобно отбирать списки частых слов для методических целей и других прикладных задач.

ЧАСТОТНЫЙ СЛОВНИК	
СЛОВНИК	частота
в (во)	42854
и	36266
не	19228
на	17262
я	13839
быть	13307

III. Статистическая структура словаря

По общей частоте слова можно определить его место в Частотном словнике и порядковый номер в Статистической структуре словаря.

СТАТИСТИЧЕСКАЯ СТРУКТУРА СЛОВАРЯ

№№ («ранг»)	Абсолютная частота	Число слов с данной частотой	Накопленная абсолютная частота	Накопленное число слов		Относительная частота	Накопленная относительная частота
				абсолютное	относительное		
556	55	31	828920	2380	0,0606091	0,0000521	0,7846782

Коэффициент лексического разнообразия словаря определяется отношением числа разных слов (V) к числу всех словоупотреблений

$$C = \frac{V}{N} = \frac{39\,268}{1\,056\,382} = 0,0371.$$

C зависит от размера выборки и размера полученного словаря разных слов, на который влияют тематические, жанровые и авторские особенности выбранных текстов.

Соотношение объема словника и объема текста можно оценивать и величиной, обратной C , т. е. средней частотой слова $F_{cp} = N/V$.

Источники словаря

Публицистические и научные тексты:

- **Ленин В. И.** Военная программа пролетарской революции. — Сочинения. Изд. 4-е. Т. 23. Госполитиздат; 1 Всероссийский съезд по внешкольному образованию 6—19 мая 1919 г. Там же, т. 29; Задачи союзов молодежи. Там же, т. 31; Карл Маркс. Там же, т. 21; Лучше меньше, да лучше. Там же, т. 33; Марксизм и восстание. Там же, т. 26; О государстве. Там же, т. 29.
- **Материалы XXII съезда КПСС.** М., Госполитиздат, 1962, с. 334—373.
- **Материалы XXIII съезда КПСС.** М., Госполитиздат, 1966, с. 3—63.
- **Калинин М. И.** Боевые задачи комсомольцев в колхозах.— В кн.: О коммунистическом воспитании. Избр. речи и статьи. [М.], «Молодая гвардия», 1956; О коммунистическом воспитании, (Докл. на собрании партийного актива г. Москвы 2 октября 1940 г.). Там же; О моральном облике нашего народа. Об искусстве и литературе. [Статьи, речи, беседы]. М., Гослитиздат, [Ленингр. отд-ние], 1957.
- **Вавилов С. И.** Глаз и Солнце. — О «теплом» и «холодном» свете. М., Изд-во АН СССР, 1961, с. II—45, 113—157.
- **Грибачев Н. М.** Оса из Лайфа: Оборотни; Туз из старой колоды; Лимонные корки; Штраус — новый троянский конь; Исландская сага с американской слезой; Шпейдель—тень над Европой; Карлос Ромуло—паяц с фирменной улыбкой; Голос нового века; Новая дистанция; Фаст—псаломщик ревизионизма; Человек и время; Время и цель.—Избр. произв. В 3-х т. Т. 3. М., Гослитиздат, 1960, с. 397—493.
- **Качалов Н. Н.** Стекло. М., Изд-во АН СССР, 1959, с. 9—39, 284—295, 308—338.
- **Леонов Л. М.** Горький сегодня.—Собр. соч. В 9-ти т. Т. 8. М., Гослитиздат, 1962; Наша Москва. Там же; О природе начистоту. Там же; Падение Зарядья. Там же; Слава России. Там же; Слово о Толстом. Там же.
- **Тарле Е. В.** Наполеон. — Сочинения. В 12-ти т. Т. 7 М. Изд-во АН СССР, 1959, с. 44—133.
- **Тимирязев К. А.** Жизнь растения. М., Изд-во АН СССР, 1962, с. 42-142.
- **Толстой А. Н.** Что мы защищаем?—Полн. собр. соч. [В 15-ти т.]. Т. 14. М., Гослитиздат, 1950; Москве угрожает враг. Там же; Разгневанная Россия. Там же; Народ и армия. Там же; Вековая сила. Там же.
- **Ферсман А. Е.** Занимательная минералогия. Изд. 2-е. М —Л Детгиз, 1945, с. 24-130.
- **Шолохов М. А.** Слово о Родине. — Собр. соч. В 8-ми т. Т. 8 М., Гослитиздат, 1960.

Источники словаря

Художественная проза:

- **Антонов С. П.** Весна.—В кн.: Весна. Рассказы. М., Гослитиздат, 1958; Поддубенские частушки.—В кн.: Деревенские повести. М., «Сов. Россия», 1957.
- **Гайдар А. П.** Голубая чашка. — Сочинения. В 2-х т. Т. 2. М—Л., Детгиз, 1949; Чук и Гек, Собр. соч. В 4-ч т. Т. 3. М., Детгиз, 1959.
- **Горький А. М.** В людях. - Собр. соч. В 30-ч т. Т. 13. М., Гослитиздат, 1951, с. 205—344.
- **Леонов Л. М.** Русский лес. Собр. соч. В 9-ч т. Т. 9. М., Гослитиздат, 1962, с. 121—242.
- **Овечкин В. В.** В том же районе.—В кн.: Трудная весна. (Районные будни). Л., Лениздат, 1957.
- **Паустовский К. Г.** Рождение моря. М., Воениздат, 1952, с. 14—129.
- **Пришвин М.М.** Кашеева цепь. — Собр. соч. В 6-ти М., Гослитиздат, 1956, с. 19—120.
- **Соболев Л. С.** Зеленый луч. — Избр. произв. В 3-х т, Т. 3 Гослитиздат, 1962, с. 7—141.
- **Толстой А. Н.** Хмурое утро.—Собр. соч. В 10-ти т. Т. 6. М., Гослитиздат, 1959, с. 7—126.
- **Федин К. А.** Первые радости. — Собр. соч. В 9-ти т. Т. 6. М., Гослитиздат, 1960, с. 123—242.
- **Шолохов М. А.** Поднятая целина.—Собр. соч. В 8-ми т. Т. 6. М., Гослитиздат, 1960, с. 7—75.

Журнальные и газетные тексты:

- «Вокруг света», 1958, № 5, с. 1—7, 17—18, 22—29, 33—49, 61—64; № 6, с. 5—41.
- «Наука и жизнь», 1961, № 7, с. 1—57.
- «Новое время», 1960, № 11, с. 1—25; № 13, с. 14—18.
- «Новый мир», 1959, № 12, с. 194—253.
- «Юность», 1962, № 11, с. 66—76, 80, 82—112.
- Газеты за 5 января 1968 г.: «Водный транспорт», «Известия», «Комсомольская правда», «Ленинградская правда», «Литературная Россия», «Правда», «Сельская жизнь», «Советская Россия», «Советский спорт», «Труд».

Источники словаря

Драматургия:

- **Алешин С. И.** Все остается людям.—В кн.: Пьесы. М., «Сов. писатель», 1962.
- **Арбузов А. Н.** Годы странствий.—«Театр», 1954, №3; Таня.—В кн.: Театр. [Пьесы]. М., «Сов. писатель», 1961.
- **Афиногенов А. Н.** Машенька.—В кн.: Пьесы. [М.], «Сов. писатель», 1956; Страх. Там же.
- **Булгаков М. А.** Дни Турбиных. М., «Искусство», 1955.
- **Горький А. М.** Враги.—Собр. соч. В 30-ти т. Т. 6. М., Гослитиздат, 1950; Достигаев и другие. Там же.
- **Зорин Л. Г.** Друзья и годы,—«Театр», 1962, № 8.
- **Иванов В. В.** Бронепоезд 14—69. — Собр. соч. В 8-ми т. Т. 1. М., Гослитиздат, 1958.
- **Лавренев Б. А.** За тех, кто в море. — Избр. произв. В 2-х т. Т. 2. М., Гослитиздат, 1958; Разлом. Там же.
- **Леонов Л. М.** Нашествие.—Собр. соч. В 9-ти т. Т. 7. М., Гослитиздат, 1961.
- **Погодин Н. Ф.** Кремлевские куранты.— Собр. драм. произв. В 5-ти т. Т. 2. М., «Искусство», 1960; Человек с ружьем. Розов В. С. В добрый час. М., «Искусство», 1959; В поисках радости. М., «Искусство», 1963.
- **Софронов А. В.** Московский характер.—Избр. произв. В 2-х т. Т. 2. М., Гослитиздат, 1955; Стряпуха.—«Театр», 1959, № 8.
- **Симонов К. М.** Русский вопрос. — Сочинения. В 3-х т. Т. 2. М., Гослитиздат, 1952.
- **Штейн А. П.** Океан.—В кн.: Драммы. М., «Сов. писатель», 1966.