

относится к лексикологии, подразделяемой на морфологию (учение о словообразовании) и семасиологию.

Соответственно этому и синтаксическая единица — синтагма — изучается синтаксической стилистикой (рассматривающей индивидуальные конструкции и формы словоизменения) и частью науки о языке — синтаксисом, в котором К. Балдингер выделяет синтаксическую морфологию (науку о формах словоизменения) и синтаксическую семасиологию.

Однако, несмотря на ряд интересных построений, некоторые положения, выдвинутые К. Балдингером, вызывают замечания. Прежде всего, это взгляд автора на индивидуальное происхождение всякой языковой инновации, выработавшийся под непосредственным воздействием идей К. Фосслера. Трудно также согласиться и с неогумбольдтианским взглядом на язык как на своеобразный национальный аспект миропонимания; подобной точки зрения, вслед за Л. Вейсгербером, придерживается и К. Балдингер.

Несколько искусственной кажется и таблица «Структура языкознания», где, на наш взгляд, преувеличенное значение придается науке о речи, в частности индивидуальной морфологии и синтаксису, выделенным как самостоятельные лингвистические дисциплины. Некоторые возражения вызывает «Синоптическая таблица развития языкознания...». В ней почти нет указаний на работы многих ученых, в частности полностью отсутствуют сведения о современных работах по компаративистике. Она слишком кратка для того, чтобы представить хотя бы в сжатом виде историю языкознания, и излишне подробна для истории семасиологии.

Этот недостаток характерен для всей книги, которая, как нам кажется, мало специализирована. Несмотря на то, что автор подробно говорит о смежных дисциплинах — лингвофилософии, психологии и логике, в книге мало анализируются чисто семасиологические труды — Й. Трира, Х. Шпербера, Э. Дорнзейфа, Г. Стерна и многих других. Однако в целом «Семасиология» К. Балдингера представляет несомненный интерес как попытка сжатого изложения достижений современной семасиологии.

Г. М. Николаева

Harry H. Josselson. The Russian word count and frequency analysis of grammatical categories of standard literary Russian.— Detroit, Wayne univ. press, 1953. 274 стр.

После незаслуженно забытой работы В. Ф. Чистякова и Б. К. Крамаренко¹ книга Гарри Йосселсона является новой попыткой систематического обследования лексики современного русского литературного языка при помощи лингвистической статистики. Задачи исследования состоят в том, чтобы, с одной стороны, составить частотный словарь наиболее употребительных русских слов (вторая часть книги), а с другой, проанализировать употребительность различных грамматических категорий (первая часть книги).

Расчеты проводились при помощи счетной машины. В связи с этим каждое словоупотребление зашифровывалось на особой карточке. При достаточной экономности (до 22 знаков) кодовое число передает основные характеристики слова. Оно указывает на текст, из которого взято слово, на его место в словаре Ушакова, на жанрово-стилистические и хронологические приметы слова и, наконец, на его грамматические характеристики (от 2 до 7). Такой метод кодирования дает возможность пользоваться карточкой Йосселсона и при машинном переводе русских текстов. Всего подсчитано 506 044 словоупотреблений, охватывающих 41 115 разных слов, из которых в частотный словарь внесено 5230 слов, размещенных по мере убывания частотности в 6 списках.

Таким образом, по объему обследованного материала словарь Йосселсона уступает словарям Кединга (11 млн. словоупотреблений) и Вандер-Беке (1200 тыс. словоупотреблений)², приближаясь к словарям Хенмона, Оса и Брауна, опирающимся на 400—500 тыс. словоупотреблений³. Словарь Йосселсона охватывает разные хронологические этапы, а также стили и жанрово-тематические слои русского литературного языка. В списках II—VI частотность каждого слова дается относительно трех периодов (XIX в., 1900—

¹ В. Ф. Чистяков и Б. К. Крамаренко, Опыт приложения статистического метода к языкознанию, вып. I, Краснодар, 1929.

² F. W. Kaeding, Häufigkeitwörterbuch der deutschen Sprache, Steglitz, 1897—1898; G. E. Vander Beke, French word book, New York, 1929.

³ V. A. C. Henmon, A french word book, based on a count of 400 000 running words, Univ. of Wisconsin, 1924; V. G. Holz, Vocabulario usual, vocabulario común y vocabulario fundamental, Madrid, 1953; Ch. B. Brown, W. M. Carr and M. L. Shane, A graded work book of Brazilian Portuguese, New York, 1945.

1918 г., с 1918 г. до наших дней), двух форм речи (диалогической и монологической) и трех жанровых типов литературного языка (язык публицистики, язык художественной литературы, научно-деловая речь). Несомненно, указанная схема заслуживает внимания, поскольку она позволяет проследить статистическое поведение слова в разных стилях (жанрах) и в различные периоды истории литературного языка. Однако из-за недостаточного объема статистического материала автор вынужден пользоваться здесь небольшими величинами¹, что противоречит основному постулату статистики, находящему свое выражение в законе больших чисел (теорема Бернулли—Чебышева), и, таким образом, в значительной степени снижает ценность наблюдений автора.

Стремясь к возможно более широкому охвату лексики русского литературного языка, Йосселсон обращается к выборочной обработке материала. С одной стороны, из 148 обследованных текстов, длиной в 1 млн. словоупотреблений, статистическому анализу были подвергнуты лишь отдельные отрывки, составляющие около половины общего объема этих текстов (ср. выше). С другой стороны, практиковался метод исключения слов из дальнейшего подсчета в тех случаях, когда эти слова набирали, по мнению автора, достаточное для включения их в ту или иную группу количество словоупотреблений (ср. стр. 41). При этом число, указывающее на количество текстов, в которых была подвергнута подсчету частотность (F) слова, определяется автором как ранг слова. Такое определение ранга (R) слова вряд ли можно признать удачным, поскольку в этом случае величина R не соотносится с величиной F (ср. функциональную зависимость R и F в системах Циффа и Гиро). Что касается выборочности подсчетов, то прием этот неприемлем с точки зрения основных принципов лингвостатистики. Ведь связанный текст образует единую систему информации, которая может быть вскрыта только путем сплошного подсчета входящих в него слов².

Нельзя полностью согласиться с выбором, а иногда и использованием анализируемых текстов. Во-первых, не до конца продуманы жанрово-тематические и хронологические принципы их отбора. Внешне дело обстоит как будто вполне благополучно: 25% текстов падает на XIX в., 25% — на начало XX в., 50% текстов охватывает литературу советского периода (после 1917 г.). В действительности среди писателей XX в. слишком большой удельный вес занимают авторы, для языка которых характерно подчеркнутое стилистическое новаторство, выражающееся то ли в словотворчестве, то ли в необычном словоупотреблении, то ли в использовании архаизмов, просторечных и диалектных слов (например, Андреев, Белый, Зощенко, Леонов, Пильник, Панферов, Шишков). Более точную картину частотной структуры русской лексики дает преимущественный анализ языка таких авторов, как Лермонтов, Чехов, Кузрин, Павленко, В. Некрасов (произведения последних двух авторов вообще не использованы при составлении словаря).

В словаре слабо отражена советская драматургия, которая служит автору основным источником для знакомства с современным разговорным языком. Отсутствуют поэтические тексты, между тем как статистический анализ стихов таких авторов, как Некрасов, Симонов, Сурков, Антокольский, ориентирующихся на литературно-разговорную речь, был бы в такой работе вполне уместен.

Во-вторых, случается, что шкала частотности выводится на основании статистического анализа малого количества текстов — так, например, 1-й список (204 наиболее употребительных слова) составлен в результате анализа всего лишь 10 произведений.

Указанные недостатки в лингвистической и статистической методике отразились, по нашему мнению, в следующих погрешностях словаря:

1. Ожидаемая и реальная частотность многих слов дает расхождения, значительно превышающие допустимые в лингвостатистике отклонения. На это указывает сам автор (ср. его цифровые данные на стр. 28—29).

2. В частотном словаре Йосселсона немало таких слов, которые трудно отнести в разряд «ходовых слов русского литературного языка». Ср. *господин, сей* (2-й список, включающий слова с наибольшей частотностью), *мужик, царь* (3-й список), *верста*,

¹ Ср., например, частотную характеристику некоторых слов в списке VI:

Слово	Общая частотность	Хронология			Жанры					
		1830—1900	1900—1918	1918	Публицистика		Художеств. литература		Научно-деловые тексты	
					разг.	неразг.	разг.	неразг.	разг.	неразг.
	14	3	5	6	1		1	11		1
<i>воск</i>	13	5	2	6	1	1	2	6		3

² Ср. B. Mandelbrot, Structure formelle des textes et communication, «Word», vol. 10, № 1, 1954.

господи! (4-й список), *ежели, купец* (5-й список), *башмак, жалование, ихний, побрать* (6-й список) и т. д., всего мы насчитали около 200 слов этого типа. Вместе с тем в словаре отсутствуют такие действительно ходовые русские слова, как *атака, ботинки, ввиду, величина, включать, всемирный, зарплата, сантиметр* и т. д.

Хотя при анализе употребительности грамматических категорий Йосселсон шел по пути, намеченному его предшественниками — Чистяковым и Крамаренко, — проведенная статистическая систематизация позволила ему сделать ряд наблюдений, которые представляют интерес для грамматической стилистики русского языка. Ср. следующие жанровые соотношения в употреблении некоторых грамматических категорий¹.

Грамматическая категория	Разговорная (диалогическая речь)	Неразговорная (монологическая речь)
Именные формы глагола	1.5	6.0
Частицы	18.3	4.5
Сравнительная и превосходная степени прилагательных	4.9	2.1
Глаголы в настоящем времени	68.2	33.6
Глаголы в прошедшем времени (соверш. вид)	56.6	91.7
Глаголы в будущем времени (соверш. вид)	43.4	8.3

Несомненно, что в тех случаях, когда подобный статистико-стилистический анализ² проводится на значительном по объему и тщательно отобранном языковом материале, представляется возможным выявить твердые критерии как для стилистической квалификации отдельных языковых единиц, так и для определений границ речевых и языковых стилей, в которых так нуждается стилистика³.

Интересен математический анализ распределения плотности слов, обладающих малой частотностью в «Капитанской дочке». Для слов с частотностью 3 функция плотности монотонно возрастает в начале и в конце текста, причем перегиб кривой (наименьшая плотность) происходит на средних участках текста. Напротив, плотность слов с частотностью 4 и 5 постепенно убывает к концу текста. Эти факты в определенной степени перекликаются с характерными для стиля и композиции Пушкина статистическими закономерностями, которые были обнаружены Д. Благим⁴.

В заключение следует отметить, что, несмотря на имеющиеся недостатки, работа Г. Йосселсона несомненно привлечет внимание историков русского литературного языка, а также специалистов в области машинного перевода, стилистики и методики преподавания. Вместе с тем она является стимулом к дальнейшему статистическому исследованию русского языка.

Л. А. Носак и Р. Г. Пиотровский

¹ Приведенные цифры являются процентными величинами. Они заимствованы из разных таблиц и поэтому соотносятся лишь по горизонтали. Последние два ряда соотношены также и по вертикали.

² О методах стило-статистики см. G. P e r d a n, *Language as choice and chance*, Groningen, 1956, стр. 12 и сл.; ср. также: P. G u i r a u d, *Les caractères statistique du vocabulaire*, Paris, 1954; W. F u c k s, *Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen*, Köln, 1955.

³ В. В. В и н о г р а д о в, *Итоги обсуждения вопросов стилистики*, ВЯ, 1955, № 1, стр. 72—82; ср. также его рецензию на книгу А. Н. Гвоздева «Очерки по стилистике русского языка» (ВЯ, 1952, № 6, стр. 142).

⁴ Д. Б л а г о й, *Мастерство Пушкина*, М., 1955, стр. 247—266.