

такому критерию. Какая же языковая реальность стоит за этой удобной для правил перехода от звуков к фонемам системой транскрипции? Возможно, что это именно набор звуков языка (в отличие от звуков речи).

Автор убедительно, с моей точки зрения, доказывает, что графика и орфография не являются разными разделами правописания. Значит ли это, что понятие графики он считает вообще излишним? Хотя прямо это не утверждается, но такое впечатление создается у читателя. Между тем понятие графики все-таки полезно, если понимать под графикой основные способы буквенного обозначения фонем в сильной позиции. Так, методисту, очевидно, полезно различать орфографические ошибки, нарушающие и не нарушающие принципы графики, поскольку они требуют разного подхода, а кодификатору — учитывать, затрагивает или нет предполагаемое изменение орфографии систему графики.

Таковы некоторые соображения, вызванные чтением книги С.М. Кузьминой. Естественно, что в небольшой рецензии нет возможности осветить все грани

этой работы, представляющей большой теоретический и практический интерес и паталкивающей на дальнейшие размышления.

Булатова Л. Н.

ЛИТЕРАТУРА

1. Авилова Н. С. Вид глагола и семантика глагольного слова. М., 1976.
2. Аванесов Р. И. Приставки на <з> — В кн.: О современной русской орфографии. М., 1964, с. 119.
3. Ильинская И. С., Сидоров В. Н. Современное русское правописание. — Уч. зап. МГПИ им. В. П. Потемкина, 1953, XXII, вып. 2, с. 8.
4. Булатова Л. Н. О специфике определения фонемной принадлежности безударных гласных в составе окончаний (на материале русских говоров). — В кн.: Общеславянский лингвистический атлас. Материалы и исследования. 1981. М., 1984.
5. Кузнецов П. С. Об основных понятиях фонологии. — ВЯ, 1959, № 2.

Казарян Б. К. Частотный словарь современного армянского языка (около 36 000 слов). Ред. Григорян В. М. — Ереван: Изд-во АН АрмССР, 1982. 758 с. (на арм. яз.).

В мировой статистической лексикографии существенное место принадлежит работам советских лингвостатистиков. Среди составляемых ими частотных словарей (ЧС) заметно выделяются ЧС языков народов СССР, поэтому внимание научной общественности будет привлечено к опубликованному недавно ЧС армянского языка. Он создан на базе коллективных исследований, выполнявшихся с 1967 г. в Отделе общей и сравнительной лингвистики, в Отделе структурной и математической лингвистики, а также в Лаборатории информационно-вычислительных систем Института языка им. Г. Ачаряна АрмССР.

Словарь примечателен не только тем, что представляет собой первый опыт составления ЧС данного языка, но и объемом подвергшихся расписыванию текстов: их общая длина равна 2 026 148 словоупотреблений. С этой точки зрения рецензируемый ЧС является самым обширным в отечественной статистической лексикографии. В нем приводятся 32,6 тыс. из 41 тыс. обнаруженных в выборке разных слов.

Словарь состоит из частотного списка 19,4 тыс. слов с частотами не менее 3, алфавитного списка 13,3 тыс. слов с частотами 1—2 и алфавитно-частотного списка слов с частотами не менее 10.

Во введении характеризуется качественный и количественный состав выборочной совокупности текстов, опубликованных в 1950—1973 гг., и методика их расписывания (вручную). Итоговые распределения в выборочной совокупности

следующим образом: художественная литература — 270 тыс.; поэзия — 230 тыс.; детская литература — 203 тыс.; драматургия — 155 424; лингвистическая литература — 178 тыс.; литературоведение — 230 тыс.; научно-популярная литература — 101 724 тыс.; учебники — 210 тыс., журналы — 200 тыс., газеты — 180 тыс. словоупотреблений соответственно. Кроме того, в выборку включены также 68 тыс. словоупотреблений, относящихся, согласно автору, к «нейтральному словарю» учеников 8—10 классов школы. Последняя часть корпуса, методика ее отбора и анализа никак не прокомментированы составителем. Судя по другим публикациям Б. К. Казаряна [4], она получена в результате ассоциативного эксперимента по выявлению «резервного» вокабуляра школьников. В таких экспериментах испытуемые называют слова, ассоциирующиеся в их сознании с той или иной заданной темой. Действительно, обращение хотя бы к началу частотного списка показывает высокую частотность и распространенность по всем 11 частям корпуса слов со значениями «быть», «говорить», «иметь», «приходить», «человек», «видеть», «идти», «много», «давать», «все», «большой», «время», тогда как слова со значениями «и», «он», «что», «я», «ты», «тот», «в», «сам», «если», «но» зарегистрированы лишь в 10 частях, и это совершенно естественно.

Все это, думается, следовало объяснить в предисловии, так же как и аргументировать возможность суммирования данных о словах, зафиксированных в

тексте (с их текстовыми частотами), и данных о «резервных» словах (уже не с частотами в тексте, а с числом информантов, назвавших каждое из них?).

В предисловии обсуждается проблема определения необходимого объема выборки. В качестве пороговой «достоверной» избрана частота 44 или, в относительном выражении, — 0,000022. По использованной автором формуле, которую он приводит в другой своей работе [2], именно такую частоту должно иметь последнее слово верхней зоны ЧС, обеспечивающее покрытие 90% текста. После определения пороговой относительной частоты 0,000022 ее значение подставляется в формулу Ципфа $p_r = k \cdot r^{-Y}$, откуда выводится величина ранга r этой частоты, равная 4545. На этом обсуждение проблемы заканчивается. Полученный результат используется для оценки эффективности определенной таким образом достоверной зоны ЧС. 4545 самых частых слов ЧС заняли от 89 до 92% всех словоупотреблений в 7 контрольных текстах (их список указывается) общей длиной 140 тыс. словоупотреблений.

На логарифмическом графике рангового распределения Ципфа представлены данные о 15 тыс. самых частых слов.

Далее в предисловии приводятся образцы карточек, на которые при расписывании текстов заносилось каждое слово в его словарной форме, а также карточек, на которые заносились составные глаголы, каждый из которых рассматривался как одно слово. При расписывании не фиксировались имена собственные и числительные (не сообщается, учитывались ли их употребления в общем объеме выборки).

Глаголы в форме страдательного залога, субстантивированные причастия, лексические омонимы и различные орфографические варианты регистрировались как самостоятельные словарные единицы.

Рассматривается средняя (графемная) длина самых частых слов, разбитых по 50 единиц. Она колеблется в пределах от 3,5 до 6,2 буквы, нарастая с уменьшением частоты. Более информативными, однако, были бы данные о длине не столько слов, сколько словоформ (как в словаре, так и в тексте).

Частотный список содержит в себе слова с частотами не менее 3. При каждом из них указывается его частота, распространенность по текстам (из 89) и по всем 11 частям выборочного корпуса. Список слов с частотами 1—2 упорядочен по алфавиту без какой-либо цифровой информации. В алфавитно-частотном списке слов с частотами не менее 10 при слове в виде дроби указывается частота (числитель) и распространенность по текстам (знаменатель) по отдельности для всего корпуса и для каждой из его 11 частей. Омонимы в пределах одной группы размещаются по их частоте.

Рецензируемый ЧС содержит богатейшие сведения о реальном употреблении лексики армянского языка в современных текстах различного содержания и жанрово-стилистической принадлежности. Он послужит ценным материалом для многочисленных собственно лингвисти-

ческих, социолингвистических, лингводидактических и других исследований. Доступ к этим данным должны иметь не только носители армянского языка или те, кто специализируется в его изучении. Он представляет большой интерес и для специалистов в области количественного языкознания. Именно поэтому в предисловии следовало в более явном виде представить минимальную статистическую информацию, а ее в нем недостает. Та же, которая там имеется, не всегда удовлетворительна.

Во-первых, отсутствует четкое указание на число слов, включенных в каждую часть словаря. Если из соображений экономии места пришлось пренебречь частотными рангами слов, то можно было бы дать их таблицей. Во-вторых, округленные числа (см., например, объемы 9 из 11 частей корпуса) с 3—5 нулями соседствуют с такими числами, как 2 026 148 (весь корпус) и 155 424 и 101 724 (объемы других частей), но при этом число всех слов ЧС дается округленно, как и число слов, приводимых в словаре. Отсюда, в-третьих, расхождения в числах, приводимых автором, и в числах, которые получает читатель, сам продельвающий соответствующие подсчеты по этому ЧС. На с. 6 объемом публикуемого словаря указан как 36 200 слов. Однако фактически в частотном списке приводятся 19,4 тыс. слов с частотами не менее 3 и 13,3 тыс. слов в списке с частотами 1—2. Кроме того, алфавитно-частотный словарь содержит 11 тыс. самых частых единиц, которые, естественно, не суммируются с предыдущими двумя объемами, а входят в первый из них. Таким образом, если в публикации приводится весь словарь, то в нем не 36,33 тыс. слов. И если в выборке обнаружены 41 тыс. слов, то почему слова исключены 8 тыс. остальных? На с. 14 говорится о распределении по длине 1 тыс. первых слов, но данные приводятся для 850 из них.

Можно было бы избежать некоторых погрешностей типографского характера. График на с. 11 оказался повернутым вправо на 90°; в формуле Ципфа на с. 9 не обозначен показатель степени; в выражении на с. 10 для расчета коэффициента «гамма» вместо знака деления стоит запятая.

Более тщательно построенное и отредактированное предисловие, сопровождаемое к тому же переводом, послужило бы той «визитной карточкой» серьезного и капитального ЧС армянского языка, которой он, без сомнения, заслуживает как своим содержанием, так и уровнем исполнения.

Манасян Н. С.

ЛИТЕРАТУРА

1. Казарян В. К. Частотный словарь как база словаря-минимума армянского языка: Автореф. дис. на соискание уч. ст. канд. филол. наук. Ереван, 1967.
2. Казарян В. К. Частотный словарь как основа словаря-минимума армянского языка. — В кн.: Вопросы лексикографии и лексикологии. Ереван, 1972 (на арм. яз.).