

З. К. ИВАНОВА (Тарту)

РАСПРЕДЕЛЕНИЕ ПРЕДЛОЖЕНИЙ С ОДНОРОДНЫМИ ЧЛЕНАМИ В МАРИЙСКОМ ЯЗЫКЕ

Грамматические явления в языке, имеющие количественный характер, подчиняются тем или иным вероятностно-статистическим законам, а статистика отражает закономерности их дистрибуции в речи, степень их употребления и дает основу для качественных оценок их реального развития. В. Г. Адмони (1963:57) писал, что «для подлинного постижения языка без более специализированного обследования количественной стороны языковых, а в частности грамматических, явлений обойтись невозможно. Причина этого заключается в том, что каждое языковое явление, помимо качественной определенности, говоря условно и образно, обладает и некоей «массой», то есть занимает в кругу других языковых явлений некое пространство и соизмеримо с ними в неких измерениях с разной степенью интенсивности».

В данной статье излагаются некоторые результаты исследования частотности употребления предложений с однородными членами (ОЧ) в современном марийском языке. Статистический анализ нами проводился сплошными выборками по 100 предложений из пяти произведений пяти жанров (подъязыков): 1) художественной прозы, 2) поэзии, 3) драматургии, 4) научной и 5) публицистической литературы.¹ Общая выборочная совокупность составила 2500 предложений (т. е. объем выборки $N = 2500$). За единицу организации выборки и подсчета приняты все встречающиеся в тексте простые предложения. Сложные пред-

¹ Для наблюдения использованы: 1) М. Шкетин, Эрегер, Йошкар-Ола 1953, стр. 3—5; Ш. Осып, Ёдырамаш корно, Йошкар-Ола 1957, стр. 203—207; Д. Орай, Чолга шудыр, Йошкар-Ола 1950, стр. 190—193; Н. Лекайн, Кугезе мланде, Йошкар-Ола 1960, стр. 50—52; В. Иванов, Ломберсолаште, Йошкар-Ола 1960, стр. 3—10; 2) М. Казаков, Поэзий, Йошкар-Ола 1951, стр. 28—32; С. Вишневецкий, Рвезылык сем, Йошкар-Ола 1973, стр. 83—88; И. Кырля, Муралтем мый йывырте, Йошкар-Ола 1969, стр. 11—23; В. Колумб, Эре чоҥешташ, чоҥешташ, Йошкар-Ола 1965, стр. 71—78; С. Николаев, 77 муру, Йошкар-Ола 1971, стр. 3—11; 3) С. Чавайн, Мукш отар, Йошкар-Ола 1957, стр. 187—189; А. Конаков, Тулык удыр, Йошкар-Ола 1976, стр. 3—5; С. Николаев, У саска, Йошкар-Ола 1941, стр. 5—7; К. Коршунов, Аксар ден Юлавий. — Ончыко 1976, № 1, стр. 12—14; А. Волков, Ксения, Йошкар-Ола 1952, стр. 46—48; 4) А. Асаев, С. Г. Чавайн, Йошкар-Ола 1951, стр. 102—106; С. Эман, Сылнымут нерген мут, Йошкар-Ола 1971, стр. 46—54; И. Иванов, Марла кнган тунгалтышыже. — Ончыко 1975, № 2, стр. 100—101; А. Александров, Романтизм да реализм. — Ончыко 1976, № 4, стр. 61—62; Т. Болло, Марий ойго — Венгрийште. — Ончыко 1976, № 4, стр. 74—76; 5) М. Казаков, Сылнымутын рудыжб-илыш. — Ончыко 1976, № 4, стр. 3—5; В. Ижболдин, Шочмо ялын тулжо. — Ончыко 1976, № 4, стр. 54—56; М. Сергеев, Палыме да палыдыме—влак, Йошкар-Ола 1976, стр. 138—144; В. Соловьев, Кызытсе илыш да калыкын культурыжо. — Ончыко 1976, № 1, стр. 71—74; М. Исметов (ред.), Паша геройна—влак, Йошкар-Ола 1974, стр. 5—9.

ложения считались как состоящие из нескольких простых предложений. Например, сложноподчиненное предложение с одним главным и двумя придаточными рассматривалось как три предложения.

В ходе анализа определена абсолютная частота (m_i) предложений с однородными членами в пяти подъязыках. Всего из 2500 предложений выявлено 586 с ОЧ. Отношением абсолютной частоты к объему выборки установлена относительная частота (%) исследуемых явлений:

$$p_i = \frac{m_i}{n_i} \cdot 100\%,$$

где n — объемы выборок подъязыков (в данном случае $n_i = n = 500$). Делением суммы абсолютных частот всех подъязыков ($\sum m_i$) на величину общей выборки ($\sum n_i = N$) установим вероятность употребления предложений с ОЧ в совокупности, т. е. среднюю относительную частоту исследуемых предложений, которая нами обозначена символом \bar{p} . Итак,

$$\bar{p} = \frac{\sum m_i \cdot 100}{N} = \frac{586 \cdot 100}{2500} = 23,4\%.$$

Результаты подсчета приведены в табл. 1.

Распределение предложений с ОЧ по подъязыкам

Таблица 1

Выборки	m_i	\bar{p} (%)	КФ
Проза	135	27,0	1,1
Поэзия	49	9,8	0,4
Драматургия	59	11,8	0,5
Научная литература	175	35,0	1,4
Публицистика	168	33,6	1,4
Итого	586	—	—
Среднее	117	23,4	—

Как видно из табл. 1, предложения с ОЧ в марийском языке являются закономерными, актуальными синтаксическими образованиями. Они обнаружены во всех исследуемых текстах (подъязыках), но частотность их употребления в речи неравномерна, что, по-видимому, объясняется стилевыми, тематическими факторами и авторской индивидуальностью. Например, в научной и публицистической литературе однородные члены употребляются чаще (соответственно 35 и 33,6%), чем в поэзии и драматургии (9,8 и 11,8%). Активность их в научной и публицистической речи, на наш взгляд, объясняется необходимостью точной и более полной передачи мысли, «для детального описания явления, предмета, для классификации» (Современный русский язык 1976: 108).

В поэтической речи наблюдается обратное явление дистрибуции однородных членов (9,8% против 35 и 33,6%). Причина этого заключается, по-видимому, в характере построения речи на синтаксическом уровне. Для поэзии характерна краткость, компактность, эпитетность, высокий темп речи-мысли, т. е. «чем меньше слов приходится на одно предложение, тем чаще предложения (а значит, и мысли) сменяют друг друга» (Головин 1971:144).

Низкая встречаемость предложений с ОЧ в драматургии (11,8%) также обусловлена особенностями стиля. В драматургии наблюдается преобладание разговорных элементов языка, а «разговорный язык не нуждается в тщательной обработке, в полноте и точности высказывания» (Гвоздев 1955:274). Речь персонажей оформляется структурно

простыми, в основном неполными предложениями в форме диалога, в которых однородные члены встречаются редко.

Неравномерность употребления синтаксических конструкций с ОЧ в различных подъязыках более наглядно отражают данные т. н. коррелятивной функции — КФ (см. Андреев 1967:22), т. е. отклонения выборочных абсолютных частот (m_i) от средней абсолютной частоты (\bar{m}).

КФ выражается отношением $\frac{m_i}{\bar{m}}$. Нейтральна для предложений с ОЧ

по всему языку зона значений КФ от 0,4 до 1,4 (см. табл. 1).

Для проверки надежности и точности определения средних частот употребления предложений с ОЧ в марийском языке нами выявлены абсолютные и относительные ошибки эксперимента. Абсолютная ошибка вычислена по формуле:

$$\varepsilon = t \sqrt{\frac{pq}{n}} \quad (\text{Tuldava 1969:14}),$$

где ε — абсолютная ошибка, p — относительная частота (%), $q = 100 - p$ (%), n — объем выборки, t — константа ($t = 2$ при уровне достоверности 95,5%). Для вычисления относительной ошибки использована формула:

$$\delta = \frac{\varepsilon}{p} \cdot 100\%.$$

Результаты вычислений абсолютных и относительных ошибок наблюдений и доверительных интервалов для относительных частот даны в табл. 2.

Таблица 2

Статистические ошибки наблюдений и доверительные интервалы распределения ОЧ

Выборки	p , %	ε , %	δ , %	Доверительные интервалы ($p \pm \varepsilon$)
Проза	27,0	3,9	14,4	23,1 ... 30,9
Поэзия	9,8	2,7	27,6	7,1 ... 12,5
Драматургия	11,8	2,9	24,6	8,9 ... 14,7
Научная литература	35,0	4,2	12,0	30,8 ... 39,2
Публицистика	33,6	4,2	12,5	29,4 ... 37,8

Как видно из табл. 2, для получения лингвистических прогнозов выборочная совокупность предложений с ОЧ достаточна, по крайней мере для трех подъязыков из пяти, так как «в настоящее время в лингвостатистике принято считать, что в грамматических исследованиях относительная ошибка не должна превышать 20%» (Бектаев, Пиотровский 1974:189). Для получения более точных показателей по поэзии и драматургии следовало бы увеличить выборку. Но для достижения наших целей это необязательно, ибо достаточно знать, что частота предложений с ОЧ в поэтической и драматической речи не превышает 15% (см. доверительные интервалы в табл. 2).

Подобным же образом в этих выборках выявлена дистрибуция однородных главных и второстепенных членов. Предложения с однородными подлежащими (ОП), сказуемыми (ОС), определениями (ОО), дополнениями (ОД) и обстоятельствами (ООБ) нами обнаружены во всех исследуемых подъязыках, но степень активности однородных членов различна. Числовые показатели вероятностей употребления одной и той же синтаксической категории в разных подъязыках представлены в табл. 3.

Таблица 3

Распределение однородных главных и второстепенных членов предложения по подъязыкам

Однородные члены	Проза		Поэзия		Драматургия		Научная литература		Публицистика		Вся выборка		Значения χ^2
	m_i	$p\%$	m_i	$p\%$	m_i	$p\%$	m_i	$p\%$	m_i	$p\%$	m_i	$p\%$	
ОП	15	12,0	11	22,5	2	3,4	27	15,4	36	21,4	91	15,5	13,0*
ОС	85	63,0	18	36,7	44	74,6	41	23,4	49	29,2	237	40,5	73,2*
ОО	5	4,0	3	6,1	5	8,4	53	30,2	35	20,8	101	17,2	39,1*
ОД	16	12,0	9	18,4	4	6,8	35	20,0	26	15,5	90	15,4	6,7
ООБ	14	11,0	8	16,3	4	6,8	19	11,0	22	13,1	67	11,4	3,9
Итого	135	100,0	49	100,0	59	100,0	175	100,0	168	100,0	586	100,0	—

На основании анализа числовых показателей табл. 3 можно утверждать, что во всех подъязыках, кроме научного, наиболее употребительны однородные сказуемые, которые составляют в среднем 40,5% от всех однородных членов. Далее следуют однородные определения (17,2%), однородные дополнения (15,4%) и однородные обстоятельства (11,4%). Среди подъязыков особенно высокая частотность распределения однородных сказуемых наблюдается в драматургии и художественной прозе (74,6 и 63%). Большой процент однородных сказуемых в этих подъязыках может свидетельствовать об эмоциональности и динамичности повествования. Как показывает анализ, в научной речи преобладают однородные определения (30,2%), которые служат целям уточнения и большего обобщения.

Расхождения вероятностей употребления однородных членов в разрезе подъязыков, особенно ОП, ОС, ОО, существенны. Это подтверждают статистические данные, полученные с помощью формулы:

$$\chi^2 = \sum \frac{(m_i - n_i \bar{p})^2}{n_i \bar{p}} \quad (\text{Tuldava 1970:168}).$$

В нашем случае число степеней свободы составляет 4; при уровне значимости 0,05 (5%) критическое значение χ^2 равняется 9,49. Лишь из-за малого объема общей выборки нам не удалось определить дифференцирующие показатели вероятностей для однородных дополнений и однородных обстоятельств, см. последнюю колонку в табл. 3, где значимость χ^2 , превышающая критическую величину, обозначена звездочкой.

ЛИТЕРАТУРА

Адмони В. Г. 1963, Качественный и количественный анализ грамматических явлений. — ВЯ, № 4, 57—63.
 Андреев Н. Д. 1967, Статистико-комбинаторные методы в теоретическом и прикладном языкознании. Ленинград.
 Бектаев К. Б., Пиотровский Р. Г. 1974, Математические методы в языкознании, ч. 2. Алма-Ата.
 Гвоздев А. Н. 1955, Очерки по стилистике русского языка. Москва.
 Головин Б. Н. 1971, Язык и статистика. Москва.
 Современный русский язык (под ред. Д. Э. Розенталя). 1976, ч. 2. Москва.
 Tuldava, J. 1969. Statistiline väljavõittemetod keeleteaduses. — Linguistica I, Tartu, 5—49.
 — 1970. Statistilised testid keeleteaduses. — Linguistica II, Tartu, 125—196.

ON FREQUENCY OF SENTENCES WITH HOMOGENEOUS MEMBERS IN MARI

This article reports on some of the results of a statistical linguistic study of the frequency of sentences with homogeneous members in different functional styles in Mari. Texts representing five styles (literary prose, poetry, drama, science fiction and publicistic essay writing) have been analyzed statistically in terms of frequency per one hundred of sentences of each of the texts under analysis, extracts being taken in continuous order. Table 1 shows that sentences with homogeneous parts occur in all the styles mentioned above but there is divergency in their frequency depending on the style they are used in as well as on differences in the subject-matter of the texts and the authors' individual manner of writing.

Alongside this the distribution of homogeneous subject, predicate, attribute, object and adverbial modifier members of sentences in the selected extracts is analyzed. Table 3 shows numerical readings or frequencies of each of the syntactical categories used in the five styles. Divergency in the frequency of usage of the homogeneous members, especially that of the subject, predicate and attribute members is significant (see the last column of readings in Table 3 where the meaning of χ^2 , that exceeds the critical index, is marked).