

РАСЧЕТ И ОБОСНОВАНИЕ ВЫБОРКИ МАССОВОГО ЭКСПЕРИМЕНТАЛЬНОГО ОБСЛЕДОВАНИЯ НОСИТЕЛЕЙ РУССКОГО ЯЗЫКА, НАПРАВЛЕННОГО НА ДИАГНОСТИКУ РАЗВИТИЯ СИСТЕМЫ ЦВЕТООБОЗНАЧЕНИЙ

На основе имеющихся данных государственной статистики о численности населения на 1 января 2017 года [1, 2] в численном выражении генеральная совокупность (N) составляет 121 255 699 человек.

Расчет выборочной совокупности производился с применением формулы для случайного бесповторного отбора респондентов [3, с. 44]:

$$n = \frac{N \cdot t^2 \cdot p \cdot q}{N \cdot \Delta^2 + t^2 \cdot p \cdot q}$$

где n – объем выборочной совокупности; N – объем генеральной совокупности, Δ – предельная ошибка репрезентативности, p и q – выборочные доли, t – коэффициент доверия.

Предельная ошибка репрезентативности Δ составляет для данного исследования 0,05, следовательно, значение доверительной вероятности равно 0,95. Это позволяет говорить о верности выводов по результатам исследования относительно параметров генеральной совокупности в 95% случаев из возможных 100%. Соответствующий выбранному значению предельной ошибки репрезентативности Δ коэффициент доверия t составляет 1,96 [3, с. 44]. Для удобства расчетов его значение может быть увеличено до 2 (в этом случае значение доверительной вероятности составит 0,954) [4, с. 60]. Предельная ошибка выборки, таким образом, не превысит 5%, что допускается в исследованиях, проводимых в области общественных наук [4, с. 63]. Значение выборочных долей p и q принимается за 0,5 для каждой [3, с. 44].

$$n = \frac{121255699 \cdot 4 \cdot 0,5 \cdot 0,5}{121255699 \cdot 0,0025 + 4 \cdot 0,5 \cdot 0,5} = \frac{121255699}{75785,811875} = 1599,9$$

Таким образом, репрезентативная выборочная совокупность для исследования при предельной ошибке выборки, не превышающей 5%, составит 1600 человек. В целях повышения достоверности результатов исследования целесообразно предусмотреть запас в размере 400 человек. С учетом этой корректировки итоговая выборочная совокупность составит 2000 человек.

Поскольку носители языка представляют собой объект исследования, крайне неоднородный в возрастном, территориальном, социальном и других отношениях, критерии отбора респондентов должны определяться в строгом соответствии с рабочей гипотезой о корреляции структуры системы цветообозначений русского языка с полом, возрастом и местом проживания носителей.

Квоты выборочной совокупности ($n=2000$) рассчитываются в соответствии с составом генеральной совокупности.

В первую очередь, в генеральной совокупности определяются процентные доли мужского и женского населения старше 15 лет (по возрастным группам) в городах и сельской местности (табл. 1).

Таблица 1.

СОСТАВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ [8, с. 21; 9, с. 92]

возраст	городское население				сельское население			
	мужчины		женщины		мужчины		женщины	
	N	%	N	%	N	%	N	%
15-19	2454068	6	2369727	5	964073	7	902121	6
20-24	2937833	7	2908866	6	1055540	7	925492	6
25-29	4608721	11	4653103	9	1425577	10	1191029	7
30-34	4798310	12	4952620	10	1471851	10	1314667	8
35-39	4208161	11	4449282	9	1297025	9	1240011	8
40-44	3771836	9	4130368	8	1235914	9	1242920	8
45-49	3246864	8	3627995	7	1191824	8	1212787	7
50-54	3205367	8	3887340	8	1339644	9	1402543	9
55-59	3451046	9	4579719	9	1496065	10	1628295	10
60-64	2807763	7	4226204	9	1155755	8	1420110	9
65-69	2151124	5	3590868	7	791372	6	1103859	7
70-74	898256	2	1703481	3	295709	2	518658	3
75-79	1087541	3	2624347	5	426763	3	962777	6
80-84	482622	1	1358257	3	208999	1	544253	3
85 +	344886	1	1182461	2	129702	1	461328	3
<i>всего</i>	40454398	100	50244638	100	14485813	100	16070850	100

Следующим шагом является выявление соотношения мужчин и женщин в генеральной совокупности, а также соотношения городского и сельского населения (табл. 2–3).

Таблица 2.

СОСТАВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ
(НАСЕЛЕНИЕ РФ В ВОЗРАСТЕ 15 И СТАРШЕ; МУЖЧИНЫ И ЖЕНЩИНЫ)

	N	%
Мужчины	54940211	45
Женщины	66315488	55
Всего	121255699	100

Таблица 3.

СОСТАВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ
(ГОРОДСКОЕ И СЕЛЬСКОЕ НАСЕЛЕНИЕ В ВОЗРАСТЕ 15 И СТАРШЕ)

	N	%
Городское население	90699036	75
Сельское население	30556663	25
Всего	121255699	100

После этого в генеральной совокупности вычисляются процентные доли мужчин и женщин старше 15 лет в городском и сельском населении (табл. 4–5).

Таблица 4.

СОСТАВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ
(ГОРОДСКОЕ НАСЕЛЕНИЕ В ВОЗРАСТЕ 15 И СТАРШЕ; МУЖЧИНЫ И ЖЕНЩИНЫ)

	N	%
Мужчины	40454398	45

Женщины	50244638	55
Всего	90699036	100

Таблица 5.

**СОСТАВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ
(СЕЛЬСКОЕ НАСЕЛЕНИЕ В ВОЗРАСТЕ 15 И СТАРШЕ; МУЖЧИНЫ И ЖЕНЩИНЫ)**

	N	%
Мужчины	14485813	47
Женщины	16070850	53
Всего	30556663	100

Дальнейшие расчеты проводятся с целью дифференциации в генеральной совокупности процентных долей мужского населения старше 15 лет в городах и сельской местности и женского населения старше 15 лет в городах и сельской местности (табл. 6–7).

Таблица 6.

**СОСТАВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ (ГОРОДСКОЕ И СЕЛЬСКОЕ
НАСЕЛЕНИЕ; МУЖЧИНЫ В ВОЗРАСТЕ 15 ЛЕТ И СТАРШЕ)**

	N	%
Городское население	40454398	74
Сельское население	14485813	26
Всего	54940211	100

Таблица 7.

**СОСТАВ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ (ГОРОДСКОЕ И СЕЛЬСКОЕ
НАСЕЛЕНИЕ; ЖЕНЩИНЫ В ВОЗРАСТЕ 15 ЛЕТ И СТАРШЕ)**

	N	%
Городское население	50244638	76
Сельское население	16070850	24
Всего	66315488	100

С опорой на произведенные расчеты внутри генеральной совокупности и в соответствии с полученными процентными долями рассчитывается количество респондентов, которых необходимо опросить для реализации выборочной совокупности (табл. 8–13).

Таблица 8.

**СОСТАВ ВЫБОРОЧНОЙ СОВОКУПНОСТИ
(N=2000; МУЖЧИНЫ И ЖЕНЩИНЫ В ВОЗРАСТЕ 15 И СТАРШЕ)**

	N	%
Мужчины	900	45
Женщины	1100	55
Всего	2000	100

Таблица 9.

**СОСТАВ ВЫБОРОЧНОЙ СОВОКУПНОСТИ
(N=2000; ГОРОДСКОЕ И СЕЛЬСКОЕ НАСЕЛЕНИЕ В ВОЗРАСТЕ 15 И СТАРШЕ)**

	N	%
Городское население	1502	75,1
Сельское население	498	25,0
Всего	2000	100

Таблица 10.

**СОСТАВ ВЫБОРОЧНОЙ СОВОКУПНОСТИ (N=1502;
ГОРОДСКОЕ НАСЕЛЕНИЕ В ВОЗРАСТЕ 15 И СТАРШЕ; МУЖЧИНЫ И ЖЕНЩИНЫ)**

	N	%
Мужчины	676	45
Женщины	826	55
Всего	1502	100

Таблица 11.

**СОСТАВ ВЫБОРОЧНОЙ СОВОКУПНОСТИ (N=498;
СЕЛЬСКОЕ НАСЕЛЕНИЕ В ВОЗРАСТЕ 15 И СТАРШЕ; МУЖЧИНЫ И ЖЕНЩИНЫ)**

	N	%
Мужчины	234	47
Женщины	264	53
Всего	498	100

Таблица 12.

**СОСТАВ ВЫБОРОЧНОЙ СОВОКУПНОСТИ (ГОРОДСКОЕ И СЕЛЬСКОЕ
НАСЕЛЕНИЕ; МУЖЧИНЫ В ВОЗРАСТЕ 15 ЛЕТ И СТАРШЕ)**

	N	%
Городское население	666	74
Сельское население	234	26
Всего	900	100

Таблица 13.

**СОСТАВ ВЫБОРОЧНОЙ СОВОКУПНОСТИ (ГОРОДСКОЕ И СЕЛЬСКОЕ
НАСЕЛЕНИЕ; ЖЕНЩИНЫ В ВОЗРАСТЕ 15 ЛЕТ И СТАРШЕ)**

	N	%
Городское население	836	76
Сельское население	264	24
Всего	1100	100

В заключение рассчитанные квоты распределяются между возрастными группами в городской и сельской местности в полном соответствии с долями генеральной совокупности (ср. табл. 1 и табл. 14).

Таблица 14.

СОСТАВ ВЫБОРОЧНОЙ СОВОКУПНОСТИ (N=2000)

возраст	городское население				сельское население			
	мужчины		женщины		мужчины		женщины	
	N	%	N	%	N	%	N	%
15-19	40	6	42	5	16	7	16	6
20-24	47	7	49	6	16	7	16	6
25-29	74	11	74	9	23	10	18	7
30-34	81	12	83	10	23	10	21	8
35-39	74	11	74	9	21	9	21	8
40-44	61	9	66	8	21	9	21	8
45-49	54	8	58	7	19	8	18	7
50-54	54	8	66	8	21	9	24	9
55-59	61	9	74	9	23	10	27	10
60-64	47	7	74	9	19	8	24	9
65-69	34	5	58	7	14	6	18	7
70-74	13	2	25	3	5	2	8	3
75-79	20	3	42	5	7	3	16	6
80-84	8	1	25	3	3	1	8	3
85 +	8	1	16	2	3	1	8	3
<i>всего</i>	676	100	826	100	234	100	264	100
<i>ИТОГО</i>	<i>1502</i>				<i>498</i>			

Выборочная совокупность формируется с помощью комбинированной стратегии. На первом этапе ($n < 1000$) используется простая случайная бесповторная выборка. Далее применяется стратегия типической выборки, стратифицированной по полу, возрасту и месту проживания носителей русского языка, и осуществлялся направленный рекрутинг участников исследования до выполнения всех необходимых условий репрезентативности выборки.

Предлагаемая схема формирования выборочной совокупности в настоящее время используется при проведении массового социолингвистического обследования носителей русского языка, направленного на диагностику развития системы цветообозначений. Данные собираются в ходе он-лайн эксперимента. Предполагается, что в ходе исследования будет получено не менее 40 000 валидных цветообозначений от по меньшей мере 2 000 участников с различными социальными характеристиками (полом, возрастом, профессией, местом проживания). Собранные данные пройдут многоступенчатую «очистку», включающую проверку написания и орфографии, которая будет проводиться с использованием компьютерных программ и вручную, и представлены в форматах OpenOffice и Excel, удобных для последующего количественного анализа данных.

Список литературы:

1. Демографический ежегодник России. 2017: статистический сборник / Росстат. М.: Федеральная служба государственной статистики, 2017. 263 с.
2. Российский статистический ежегодник. 2017: статистический сборник / Росстат. М.: Федеральная служба государственной статистики, 2017. 686 с.
3. Ротман Д. Г. Оперативные социологические исследования в комсомольской работе. Мн.: Университетское, 1988. 113 с.
4. Могильчак Е. Л. Методика социологического исследования. Выборочный метод: учебное пособие для бакалавриата и магистратуры. М.: Юрайт, 2018. 117 с.