

*Е.А. Гришина, С.О. Савчук*

**НАЦИОНАЛЬНЫЙ КОРПУС РУССКОГО ЯЗЫКА  
КАК ИНСТРУМЕНТ ДЛЯ ИЗУЧЕНИЯ  
ВАРИАТИВНОСТИ ГРАММАТИЧЕСКИХ НОРМ**

Вариативность – это фундаментальное свойство языка, которое сообщает ему гибкость, избыточность, возможность множественного выражения одного и того же смысла. Будучи следствием языковой эволюции, вариативность становится почвой для дальнейшего развития языка. Постоянным источником вариантов является живое употребление. Новое явление появляется, вступает в конфликт с уже существующим и либо побеждает и вытесняет старое, либо уходит из употребления. В каждый период мы наблюдаем соотношение нового и старого, и это соотношение также меняется во времени. Изучение конкуренции вариантов является важным этапом в определении тенденций в развитии языка, в определении живых активных процессов в языке.

В отношении понимания и трактовки самого понятия «языковой вариант» существуют некоторые разногласия. Согласно одной точке зрения, языковыми вариантами называются формальные разновидности одной и той же языковой единицы, которые при тождестве значения различаются частичным несовпадением своего звукового состава. В этом случае вариантными могут быть слова или формы слов, причем варианты могут различаться следующими признаками: произношением, местом ударения, составом формообразовательных аффиксов. Более широкое понимание вариативности позволяет отнести к вариантам словообразовательные модификации типа *туристский – туристический*, *накат – накатывание*, *омич – омичанин*; параллели, этимологи-

чески восходящие к разным языкам, типа *глас – голос; врата – ворота; ночь – ноцъ* и др.<sup>1</sup>

Проблема вариативности тесно связана с вопросами ортологии (нормализации языка), поскольку существование нескольких вариантов обычно сопровождается их оценкой с точки зрения сложившихся норм употребления. Вариант может быть кодифицирован в литературном языке, оценен как более или менее предпочтительный, а может остаться за пределами литературной нормы.

Понятно, что оценки и рекомендации должны опираться на научное представление о норме, которое отвечает следующим критериям.

1. Норма должна соответствовать системе языка, возможности которой реализует.
2. Норма определяется не только системными закономерностями языка, но складывается под воздействием социальных оценок, как итог взаимодействия общественного выбора и индивидуальных вкусов.
3. Норма должна соответствовать речевому узусу, то есть реализации языковой системы в определенных условиях общения.
4. Статистические показатели должны использоваться в качестве объективных выразителей употребительности, распространенности языковых единиц<sup>2</sup>.

---

<sup>1</sup> См. об этом: *Горбачевич К.С.* Вариантность слова и языковая норма. Л., 1978; *Валгина Н.С.* Активные процессы в современном русском языке. М.: «Логос», 2001. С. 26–40; *Граудина Л.К.* Вопросы нормализации русского языка: Грамматика и варианты. М.: «Наука», 1980. С. 100–121.

<sup>2</sup> *Граудина Л.К.* Вопросы нормализации русского языка... С. 63.

Как представляется, надежным инструментом для исследования нормы и вариантов может служить корпус текстов, поскольку он предоставляет в распоряжение лингвиста зафиксированное в текстах состояние речевого узуса определенной эпохи и позволяет провести количественное и качественное изучение этого состояния. Корпус дает возможность при разработке нормативных оценок и рекомендаций освободиться от вкусовых предпочтений, выйти за пределы своего индивидуального языкового сознания и опереться на более объективную эмпирическую базу – статистические данные, полученные при обследовании представительного материала. Какие характеристики корпуса свидетельствуют о его представительности?

Во-первых, объем. Для Национальных корпусов языков своеобразным ориентиром или порогом (который многие корпуса уже переступили) служит цифра в 100 млн словоупотреблений. Общий объем НКРЯ составляет в настоящее время около 160 млн словоупотреблений<sup>1</sup>. В нем можно условно выделить две части – современную и диахроническую. Корпус современных текстов составляют тексты, период создания которых укладывается в рамки 1951–2007 гг. Объем этой части корпуса – 97,5 млн словоупотреблений, причем половина текстов относится к периоду после 2000 г. Диахроническая часть составляет около 53 млн словоупотреблений и объединяет тексты XVIII в. (1,1 млн с/у), XIX в. (23,3 млн с/у прозаических текстов и 2,5 млн с/у в поэтическом корпусе) и 1-й пол. XX в. (25,4 млн с/у).

Во-вторых, говоря о представительности корпуса, обычно имеют в виду прежде всего жанровое разнообразие составляющих его текстов, которые относятся ко всем основным сферам использования русского языка (научной, официально-деловой, публицистической, церковно-богословской, художественной, разговорно-бытовой, включая устную и электронную коммуникацию).

---

<sup>1</sup> По данным сайта [ruscorpora.ru](http://ruscorpora.ru) на январь 2008 г.

Еще один аспект репрезентативности корпуса – социолингвистический: состав авторов, чьи произведения вошли в корпус, чрезвычайно разнообразен по основным социологическим параметрам. Поскольку формирование корпуса еще не завершено, анализ по этим характеристикам не проводился, однако по самым грубым оценкам (точные, вероятно, вообще невозможны, в силу принципиальной невозможности установить авторство многих текстов), в корпусе собраны тексты не менее 20000 авторов. Это носители русского языка, различающиеся по возрасту, уровню образования и владения языком, профессиональной принадлежности, являющиеся представителями разных типов речевых культур<sup>1</sup>. Интересно сравнить эти цифры с объемом экспериментального материала, который был в распоряжении исследователей в «докорпусную» эпоху. В ходе анкетирования, проводившегося в 1959–1964 гг. по программе «Русский язык и советское общество», было получено 4000 заполненных вопросников по морфологии, а коллективом Л.К. Граудиной при создании частотного словаря грамматических вариантов была обследована выборка из 500 газет за 1968–1972 годы общим объемом 2 млн слов. Даже этого материала оказалось достаточно, чтобы оценить и уточнить многие нормативные предписания.

Важный аспект репрезентативности Национального корпуса русского языка – наличие в нем текстов, относящихся к разным периодам создания, что позволяет проследить изменения в употреблении языковых явлений, а возможно, и установить динамику этих изменений.

Тексты Национального корпуса распределены по временным периодам следующим образом:

---

<sup>1</sup> О разных типах речевых культур см. работы О.Б. Сиротининой, В.Е. Гольдина и др., например: *Гольдин В.Е., Сиротинина О.Б.* Внутринациональные речевые культуры и их взаимодействие // Вопросы стилистики: проблемы культуры речи. Саратов, 1993.

Период	Млн с/у	Период	Млн с/у	Период	Млн с/у
18 век	1,1	1911–1920	3,9	1961–1970	5,8
1800–1830	2,1	1921–1930	8,6	1971–1980	7,0
1831–1870	11,3	1931–1940	4,9	1981–1990	7,2
1871–1900	10,3	1941–1950	3,3	1991–2000	23,2
1901–1910	3,4	1951–1960	3,5	2001–2007	46,8

Основной массив текстов, собранных в НКРЯ, охватывают период в 200 лет, поэтому он наиболее приспособлен для изучения коротких (несколько десятилетий) и средних (1–2 столетия) языковых изменений. Объем корпуса, как кажется, позволяет изучать вариативность и изменчивость достаточно частотных языковых явлений. Отсюда следует, что можно надеяться на получение надежных результатов по таким направлениям, как:

- изучение морфологических вариантов имен, глаголов и т.д. и их эволюции;
- исследование словообразовательных вариантов и связанной с ними проблемы паронимов, продуктивности словообразовательных моделей и словообразовательных средств;
- исследование изменения вариантов управления, согласования и других конструкций;
- исследование акцентологических вариантов и изменений в акцентной системе русского языка;
- исследование лексической вариативности, в частности, изменения состава и семантических соотношений в синонимических рядах и тематических группах и т.д.

Принято различать широкое (дескриптивное) и узкое (прескриптивное) понимание нормы<sup>1</sup>. Норма в широком смысле (или

<sup>1</sup> Косериу Э. Синхрония, диахрония и история // Новое в лингвистике. Вып. 3. М.: «Прогресс», 1963. С. 143–343; Гавранек Б. Задачи литературного языка и его культура // Пражский лингвистический кружок. Сб. ст. М.: «Прогресс», 1967. С. 338–377; Граудина Л.К. Вопросы нормализации русского языка: Грамматика и варианты. М.: «Наука», 1980.

объективная норма) близка к понятию узуса, речевой практики, то есть общепринятых, устоявшихся способов употребления языковых единиц. Норма в узком смысле (или субъективная норма) – это результат целенаправленной деятельности лингвистов по кодификации языка, система предписаний и правил, отраженных в словарях, грамматиках и справочниках. Если по отношению к некодифицированным сферам языка (территориальным диалектам, профессиональным и социальным жаргонам) термины «норма» и «узус» можно использовать недифференцированно, то в литературном, кодифицированном языке норма и узус не всегда совпадают, речевая практика не всегда следует нормативным предписаниям.

Что же можно изучать с помощью корпуса – узус или норму? Ответ на этот вопрос можно получить, проанализировав состав корпуса.

Тексты, формирующие корпус, можно разделить на три группы. Основную группу составляют тексты, опубликованные в авторитетных изданиях, то есть написанные профессиональными авторами и прошедшие жесткую редакторскую подготовку. Ко второй группе относятся изданные тексты, но, как правило, написанные «непрофессиональными» авторами и прошедшие менее тщательную редакторскую подготовку (дневники, записные книжки, официальные документы, массовая литература, региональные или узкопрофессиональные издания). К третьей группе принадлежат тексты, рукописные или машинописные, не прошедшие редактирования (бытовая и деловая переписка, дневники, канцелярские документы, электронная коммуникация, записи устной речи). По очень грубым подсчетам, доля текстов,

---

С. 63–72; *Крысин Л.П.* Языковая норма в проекции на современную речевую практику // *Русский язык сегодня*. Вып. 4 Языковые нормы. М., 2006. С. 296–297; *Лантева О.А.* Теория современного русского литературного языка. М.: ВШ, 2003. С. 58–65.

прошедших слабую редакционную подготовку или вовсе ее лишенных, составляет около 20%.

Таким образом, в текстах основной части корпуса мы наблюдаем реализацию норм и правил в практике письменной речи высококвалифицированного носителя языка (профессионального автора и редактора), а примерно пятая часть корпуса предоставляет нам возможность наблюдать некодифицированный или слабо кодифицированный узус, или норму, отраженную в речевой практике рядового носителя языка. Так что возможен сопоставительный анализ распространенности вариантов в разных типах изданий и источников.

Таким образом, Национальный корпус русского языка представляется нам весьма надежным инструментом для исследования языковой вариативности и состояния норм. Покажем это на примере из области морфологии.

Слово *бездарь* относительно новое. Принято считать, что его ввел в обиход Игорь Северянин: *Вокруг талантливые трусы И обнаглевшая бездарь... И только Вы, Валерий Брюсов, Как некий равный государь...* (сборник «Громокипящий кубок», «поэза» номер 136, написанная в 1912 г.). В Толковом словаре Д.Н. Ушакова (1935 г.) это слово приводится с ударением на втором слоге: **БЕЗДА'РЬ**, и, мн. нет, ж., собир. Бездарные люди. || То же, что бездарность (лишенный таланта человек, тупица).

В Толковом словаре русского языка С.И. Ожегова и Н.Ю. Шведовой (1992 г.) представлены уже другой акцентологический вариант этого слова: **БЕ'ЗДАРЬ**, -и, ж. (*разг. пренебр.*). То же, что бездарность (во 2 знач.). [Ни к чему не способный, лишенный таланта человек]. Орфоэпический словарь русского языка под редакцией Р.И. Аванесова (1997 г.) отмечает оба варианта: **Б'ЕЗДАРЬ**, -и, мн. -и, -ей; *не рекомендуется безда'рь*. То же в Словаре трудностей К.С. Горбачевича, но вариант **безда'рь** снабжен пометой *устаревшее*. Грамматический словарь А.А. Зализняка разграничивает два значения слова *бездарь*, по-

скольку в разных значениях оно имеет разный набор грамматических признаков: в собирательном значении ‘бездарные люди’ – неодушевленное, в конкретном значении ‘бездарный человек’ – одушевленное.

Однако все словари относят это слово к женскому роду<sup>1</sup>. Между тем данные Национального корпуса свидетельствуют о другом. Анализ контекстов употребления этого слова (всего 121 вхождение) выявил среди них такие, в которых представлены словоформы разной родовой принадлежности. Форм женского рода единственного числа значительно больше (21), среди них два раза встретился вариант *безда'рь* (у А. Мариенгофа и Н.И. Бухарина, цитирующего И. Северянина), причем контексты представляют оба значения слова, отмеченные в словарях, – собирательное и конкретное:

Кстати, такая литература в то время и не имела корней и два-три (да обчелся) даровитых писателя беспомощно металась по безвкусным широтам *бездари*. [М.М. Пришвин. Дневники (1927)]. Пожилая *бездарь* и молодая глупость жгут классические учебники. [В. Дудинцев. Белые одежды (1987)]. И значит, если я место придумаю, все равно вы там никого не найдете, потому что это уже будет придумано *бездарью*, вашим служащим. [А. Кабаков. Сочинитель (1990–1991)].

В то же время отмечено 5 контекстов употребления разных падежных форм мужского рода: Им. ед. (2), Вин. ед. (*бездаря*), Тв. ед. (*бездарем*), Пр. ед. (о *бездаре*), все они относятся к текстам, созданным не ранее 1998 года:

Был такой деятель в ЦК хрущевских и брежневских времен, бывший секретарь Украины. Жуткий *бездарь*. [Н. Амосов. Голоса времен (1999)]. А последним был Киркоров, тот самый, которого лет за

---

<sup>1</sup> В словаре Л.К. Граудиной и др. «Грамматическая правильность русской речи» (2004 г.) оно не отмечено в списке из 50 существительных, с основой на мягкий согласный, *ж* и *ш*, род которых колеблется между женским и мужским.



семь до этого Алла Борисовна в сердцах обозвала "бездарем". [А. Беляков. Алка, Аллочка, Алла Борисовна (1998)]. Даже если у профессора и феноменальная память – лет десять назад это было, – все равно *о бездаре* столько подробностей не накопишь. [О. Новикова. Мужской роман (1999)]

Этот факт, как представляется, свидетельствует о появлении в конце XX-го века у слова **бездарь**, -и, *жс.* варианта **бездарь**, -я, *м.*, и эта пара пополняет группу существительных, имеющих колебание в роде (*толь, тюль, псалтырь, лебедь* и пр.). Можно предположить, что носителями «младшей нормы» вариант мужского рода осознается как основной: «Но вот женский род этого слова привел меня в замешательство» (из записи на форуме в интернете).

В докладе будут приведены другие примеры корпусного анализа отдельных «слабых участков» нормы, уточняющие рекомендации пособий по практической стилистике и культуре речи.