

УДК 81'373.611

Н.Д. Голев

ЛЕКСИЧЕСКОЕ ФУНКЦИОНИРОВАНИЕ СЛОВООБРАЗОВАНИЯ И ЛЕКСИКО-СЛОВООБРАЗОВАТЕЛЬНАЯ СИСТЕМНОСТЬ РУССКОГО ЯЗЫКА В СВЕТЕ СТАТИСТИКИ ИНТЕРНЕТА

Исследуется взаимодействие словообразовательной и лексической систем русского языка, а именно лексическое функционирование русских производных слов, которые оцениваются с точки зрения их частотности в текстах Интернета. Количественные характеристики извлекаются автором из поисковых систем Яндекс, Рамблер и др. Их данные используются для выявления степени «лексичности» производного слова, описания словообразовательных парадигм на шкале лексического функционирования, переоценки границ между потенциальными и реальными словами, лексикографической репрезентации в частотных, толковых и словообразовательных словарях. Значительное место в работе занимает обсуждение методологической проблемы соотносимости показаний языкового сознания, полученных в эксперименте, и объективных данных статистики Интернета.

Ключевые слова: словообразование, лексика, словообразовательная парадигма, внутрисловообразовательное слово, потенциальное слово, окказионализм, узуальное слово, отглагольный дериват, статистика.

В статье реализуется лексикоцентрический взгляд на образование и функционирование производного слова, в соответствии с которым сама потребность в новообразовании лежит в лексике, а ее реализация подчиняется законам лексического функционирования. При таком взгляде на влияние словообразования на неологический процесс подобное влияние относится к сфере следствий, а не причин, в частности, такие функции словообразовательной системы, как категоризация или пропозиция, оцениваются как следовые [1].

Объектом описания выступают словообразовательные дериваты, вошедшие (или входящие) в сферу лексического функционирования, где они подчиняются его законам и, как следствие, приобретают собственно лексические свойства. Этот процесс достаточно длительный, трудно отслеживаемый, поскольку границы его этапов размыты, поэтому долгое время он не был предметом внимания дериватологов. В предлагаемом исследовании данный процесс рассматривается на материале, позволяющем в какой-то мере вывести его в «наблюдаемую зону» и описать, опираясь на достаточно надежные количественные данные. Источником такого материала являются поисковые системы Интернета и сопровождающие результаты поиска отдельных слов статистические показатели.

1. Идея исследования и ее методологические презумпции

Взаимодействие словообразования и лексики начинается уже с этапа возникновения слова, с процесса трансформации производного слова, традиционно называемого потенциальным, в окказиональное слово. Окказиональные слова весьма противоречивы по своему языко-речевому статусу. С одной стороны, они отчетливо несут на себе отпечаток внутреннего устройства виртуального словообразовательного механизма, представляясь «правильными»

словами, не осложненными индивидуально-лексическими наслоениями, присущими узואльным словам. С другой стороны, говоря даже о начальных этапах деривации слова, мы должны отметить, что они возникают не в виртуальных пространствах, а в реальных речевых актах вследствие вполне конкретных потребностей контекста и речевой ситуации, поэтому уже первые употребления деривата так или иначе индивидуализированы и конкретизированы. Поэтому они одновременно «лексичны» и «алексичны». Читатель может в предыдущем предложении найти иллюстрацию развиваемого тезиса: несмотря на то, что лексические новообразования *лексичны* и *алексичны*, самим фактом употребления их в приведенном контексте «перешли» из разряда потенциальных в окказиональные и, казалось бы, насквозь прозрачны и выводимы, они на самом деле впитали в себя определенные смыслы контекста, и их значения выводимы лишь с учетом этих смыслов, и (если позволительно предположить, что за ними есть будущее в плане узואлизации) эти смыслы в дальнейшем могут лексикализаться на уровне словарных единиц, обрести какие-либо модификации и еще более отдалиться от «чистого» (=выводимого без остатка) лексико-словообразовательного значения.

За такой постановкой вопроса имеется и более общий теоретический вопрос, уже не однажды поднимавшийся нами (см., например, работы [2; 3]), – предметом какой науки является функционирование производного слова: словообразования (куда его относят по сложившейся традиции) или лексикологии. Мы показали выше, что уже с первого акта реализации (явленности) потенциального слова в речевом материале оно «лексично», тем более что его дальнейшее бытие неизбежно усиливает его «лексичность», ослабляя (переводя в план генетического) его связь со словообразованием.

Лексичность окказионализма отчетливо проявляется и в том, что его индивидуально-ситуативный характер относителен: никакое слово, образованное по действующим словообразовательным моделям, не может быть априори квалифицировано как окказиональное, поскольку ничто не препятствовало ему быть образованным ранее в другой ситуации и другим носителем языка. Зафиксированность в текстах (тем более в словарях) никоим образом не может служить критерием окказиональности (новизны), хотя, по сути, именно он служит мерилом для ее квалификации в работах большинства дериватологов. Несостоятельность данного критерия хорошо иллюстрируют, в частности, поисковые системы Интернета, в котором мы легко находим единицы, не зафиксированные в словарях, например, такие, как *кидальщик*, зафиксированные на 12000 страницах, *кидальщица* – 857, *кивальщик* – 56, *кивальщица* – 3, *недокинуть* – 156, *недокомплектовать* – 614 и т.п. К приведенным примерам добавим слова, квалифицированные Е.А. Земской как потенциальные: *возражатель*, *спрашиватель*, *повторитель*, *разлучитель* [4. С. 218–220]; первое из них представлено в Интернете на 2452 страницах, второе – 934, третье – 477, четвертое – 538. Тот факт, что возникновение слова не сводится к элементарным критериям, отмечал еще Ф. де Соссюр, ср. его высказывание по этому поводу: «Таким образом, значительная часть процесса образования по аналогии протекает еще до того, как появляется новая форма... Ошибочно думать, что процесс словотворчества приурочен точно к моменту возникновения новообразования; элементы нового слова были даны уже раньше. Им-

провизируемое мной слово, например, *in-décor-adle* «такой, которого невозможно украсить», уже существует потенциально в языке <...> его реализация в речи есть факт незначительный по сравнению с самой возможностью его образования» [5. С. 200]. Все сказанное лишний раз подтверждает, что вопросы о сущности основополагающих понятий словообразования, находящихся на стыке с лексикологией, остаются открытыми; в их числе понятия словообразовательного процесса, словообразовательной интенции, нового слова (лексической инновации), потенциального слова, окказионального слова и др.

Одну из попыток конкретно-исследовательского решения проблемы постепенного «вырастания» окказионального слова из потенциального, а узального – из потенциального находим в статье Н.Б. Лебедевой [6], выполненной на статистической базе. Исследования Н.Б. Лебедевой заключаются в разработке и исследовательской реализации эксперимента, основанного на создании ситуации, активизирующей у испытуемых языковую интуицию; такая ситуация формирует у них установку на интроспекцию по отношению к результатам языко-интуитивной деятельности, что позволяет экспериментатору статистически фиксировать их рефлексии на те или иные дериваты в рамках направленной установки. Статистика таких рефлексий и давала основания для квалификации дериватов на шкале «степень лексичности». Таким образом, в идеях анализируемой статьи Н.Б. Лебедевой лежит лексикоцентрическое в своей основе представление об эволютивном характере словообразовательного процесса и нахождении зарождающейся лексической единицы на одном из этапов этого процесса. Данная эволюция проецируется в системную плоскость и обнаруживается в возможности типологии дериватов по признаку «степень лексичности» (или, от обратного, по признаку «степень отдаления от этапа потенциальности и виртуальности»).

В данной методике реализован принцип единства системного и ментального бытия языка, основы которого мы находим еще у И.А. Бодуэна де Куртенэ и позднее у Л.В. Щербы в его знаменитой статье «О тройном аспекте языковых явлений и об эксперименте в языкознании» [7]. В частности, Л.В. Щерба трактовал языковую систему, языковой материал и речевую деятельность (соответственно, как компонент последней – языковое сознание) как аспекты (видения) **единого феномена**, каковым является язык, не придавая каждой из них онтологически самостоятельного статуса. Таким образом, в данном подходе языковая системность достаточно адекватно проявляется в показаниях языкового сознания, что дает основания исследователю рассчитывать на ее успешное моделирование при достаточно массовом характере их (показаний) проявления. Коснувшись методологических установок такого рода исследований, считаем важным заметить, что единство системного и ментального бытия языка в классических работах, открывавших эпоху системной лингвистики, рассматривалось как естественное, само собой разумеющееся. В частности, это проявлялось в использовании термина «ассоциация» (Н.В. Крушевский, Ф. де Соссюр, И.А. Бодуэн де Куртенэ). Так, характеризуя ассоциативные идеи Н.В. Крушевского, И.А. Бодуэн де Куртенэ пишет: «Если в результате закона ассоциации (соединения представлений) по сходству слова должны укладываться в нашем уме в системы или гнезда (семьи слов), то благодаря закону ассоциации по смежности те же самые слова

должны строиться в ряды» [8. С. 184]. В разделе «Ассоциативные отношения» Ф. де Соссюр совершенно естественным образом переходит от «ассоциативных групп в нашем сознании» к словоизменительным и словообразовательным парадигмам как проявлениям собственно языковой системности [5. С. 158–159].

Применительно к настоящей статье приведенные методологические тезисы не отдаленные абстракции, они достаточно непосредственно включаются в квантитативную методику, применяемую как в статье Н.Б. Лебедевой, так и далее в нашем непосредственном исследовании фрагмента языкового материала.

2. Методика квантитативного измерения функционального потенциала лексических дериватов

Предметом рефлексии испытуемых в эксперименте Н.Б. Лебедевой стали четыре глагола: *нахвастаться*, *перехвастаться*, *забодриться*, *нахрабриваться*. В результате эксперимента получено 300 анкет. Обобщая результаты, автор отмечает, что степень нормативности этих глаголов различна – от безусловно недопустимого, неприятие которого указывается большинством информантов, до понимания их как «естественных». Сопоставление словообразовательной и лексической нормы дает возможность выяснить, в какой степени образованное слово соответствует словообразовательной системе и в какой – лексической норме. Предполагается, что есть определенная (нежесткая) корреляция между первым и вторым.

Продолжим линию оценки перспективности исследований данного направления. Обратимся к практическим значимостям. Они заключены в лексикографических потребностях. Уже давно существует потребность отделения собственно узуальных слов от незузуальных. Нерешенность этой проблемы бросается в глаза всякому, кто рассматривал с этих позиций любой толковый словарь (как и словари многих других типов): вопросы типа, почему в словарях нашло отражение слово *наболтаться*, но нет слова *натрепаться*, есть *наливочка*, но отсутствует *настоечка*, есть наипрекраснейший, но нет наикрасивейший¹? – остаются до сих открытыми для теоретической лексикографии, а в лексикографической практике решаются на основе интуиции составителей словарей. Было бы весьма упрощенно трактовать интуицию такого тонкого знатока языка, каким, несомненно, является лексикограф, как ненадежный и неприемлемый инструмент познания семантики языка. Интуиция даже рядового носителя языка достаточно адекватно ее отражает, что – в числе прочего – доказывает и эксперимент Н.Б. Лебедевой в указанной статье, результаты которого подтверждают наши квантитативные данные², и тем не менее потребность в получении более объективных (желательно формализованных) оценок функционально-семантического статуса слова неизбежно возникает по мере развития как теории семантики, так и лексикографической практики.

¹ Примеры сверены с данными «Сводного словаря современной русской лексики» [9].

² В ходе заполнения граф таблицы, приведенной далее, мы неизбежно проводили мысленный прогностический эксперимент, который оказывался в значительном числе случаев вполне реалистичным. Читатель может легко в этом убедиться, повторив таковой эксперимент.

Далее мы осуществляем попытку апробации для этой цели одной из простейших методик количественного представления функционально-семантического статуса слова – использование его частотных характеристик. При этом речь идет не об известных частотных словарях, «мощность» которых не соответствует потребностям такого исследования. Их выборка не позволяет системно фиксировать столь тонкий лексический материал, каковым являются едва лишь зарождающиеся лексические единицы, эпизодически выходящие за пределы статуса потенциальных слов и высвечивающиеся в речевом материале как словообразовательные окказионализмы. В качестве такого инструмента мы предлагаем использовать статистические и иллюстративные данные поисковых систем Интернета, превосходящие по количественной мощности частотные словари в сотни тысяч раз (см. другие штудии автора на эту тему: [3; 10]). Например, сделаем запрос в поисковой системе Яндекс на потенциальное слово *вболеться*, которое оказалось зафиксированным на 4 марта 2007 г. в Интернете на четырех страницах и в трех сайтах и, следовательно, потенциальность которого уже не может быть оценена как абсолютная (но по сравнению с другими окказионализмами она тем не менее весьма высокая). При необходимости качественной оценки окказионализма Интернет позволяет обратиться к иллюстративным данным, в частности, по отношению к слову *вболеться* они таковы: *«Рискну в белизну страницы – седьмое небо, в чистый слог **Вболеться**, вслышаться, вродиться. Вмолчать-ся. Как всему итог...»* (персональный сайт поэта В. Петрова); *Понимаю тебя, я еще не сильно **вболелся** в Кими, а уже категорически против его возможного перехода в Феррари.* (Форум «Переход Монтой»); *Нужно было добиться инвалидности 1-й группы, и я добилась, отказавшись ходить и есть, имитируя падучую. Я так **вболелась** в свою роль, что уже не могла из нее выйти.* (И. Лиснянская. Хвастунья). Как не трудно убедиться, каналом перехода слова *вболеться* с уровня потенциальных на уровень реальных является художественная речь, что легко объяснимо, учитывая высокую степень интенсивности-выразительности, которую несет словообразовательная модель «В...СЯ» в сочетании с глаголами чувства.

Теперь вернемся к методологическим тезисам из подраздела 1. Сопоставим качественную сторону данных, полученных в эксперименте, и статистических данных Интернета. Рассмотрим в этой связи вопросы, предлагаемые испытуемым Н.Б. Лебедевой. Информанты должны ответить на вопрос 1: Данное слово Вам: а) понятно, обычное слово; б) понятно, но как-то непривычно; в) понятно, но так не говорят, режет слух; г) непонятно, нет такого слова и на вопрос 2: «Можно ли так сказать?», отмечая один из пунктов вопросника: а) да; б) скорее да, чем нет; в) скорее нет, чем да; г) нет. Какой механизм рефлексивной деятельности предполагается этими вопросами? Естественно, что отвечающий должен сконцентрировать свои усилия на формировании и затем оценке концентрированного образа, обобщающего в чувственной (чувственно-языковой?) форме свой коммуникативный опыт (в первую очередь – в позиции слушающего и читающего) использования данного слова, т.е. оценку языкового материала, по терминологии Л.В. Щербы в указанной выше статье. Большую роль в этом образе играет количественная сто-

рона такого функционирования, что означает, что в языковое чувство входит и такой компонент, как частотность.

Насколько точен внутренний «счетчик» рядового носителя языка? Исследование Н.Б. Лебедевой расположило анализируемые ею глаголы в следующей последовательности на оси «степень лексичности» (начиная с более высокой) или, напротив, «степень потенциальности (начиная с наиболее низкой)»: *нахвастаться* – *перехвастаться* – *забодриться* – *нахрабриваться*. Краткий комментарий исследователя. Глагол *нахвастаться* (вдоволь похвастаться) – узуальный, словарный. Глагол *перехвастаться* (похвастаться сверх меры) – узуальный (хотя не отмечен в словаре), входит в «ближнюю» потенцию, но еще не вошел в лексическую норму; потенциальное слово. Глагол – *забодриться* (начать бодриться) – незуальный, создан с некоторыми деформациями словообразовательной системы, входит в «дальнюю» потенцию, лексической норме не соответствует; среднее между потенциальным и окказиональным словом. Глагол *нахрабриваться* (длительно, интенсивно совершать действие «храбриться») – незуальное слово, созданное со значительными деформациями словообразовательной нормы; составляет «сверхдальнюю» потенцию, лексической норме не соответствует; появление его в речи маловероятно, оно находится на грани окказионального и невозможного. Эта цепочка на оси частотности в Интернете выглядит следующим образом (табл. 1).

Таблица 1

Частотные показатели слов по данным поисковых систем Интернета

Глагол	Поисковая система		
	Яндекс	Rambler	Google
Нахвастаться	3143	493	636
Перехвастаться	292	9	10
Забодриться	319	62	62
Нахрабриться	151	6	2

Как видим, имеется несоответствие оценок, оно касается второго и третьего члена анализируемой цепочки. Причина такого несоответствия, на наш взгляд, заключается в том обстоятельстве, что глагол *забодриться* относительно широко вошел в узуальный фонд, который, однако, по тем или иным обстоятельствам не отразился в языковом сознании участников эксперимента. Обращение к конкретному речевому материалу, представленному в Интернете, позволяет принять такое предположение как достаточно адекватное: среди примеров, иллюстрирующих употребление глагола *забодриться* в Интернете, подавляющее число составляют, во-первых, жаргонные употребления («*Забодрился твой причесон, плотненький такой завертончик*», «*Я забивал играть попросил чтобы мне потерли логи 5 раз просил но не помогло потом меня забодрили играть я начал!*» и т.п.), которые оказались вне сфер общения испытуемых (студентов-филологов) и потому не могли быть адекватно оценены ими в плане узуальности); во-вторых, значительное число примеров является отсылкой на профессионализм, зафиксированный еще В. Далем («о лошадях, *забодриться*, *заиграть*, *затопать*»), который вообще в современной речи вышел из узуса. На основании одного ряда слов, разуме-

ется, нельзя говорить о высокой степени коррелятивности количественных оценок языковой интуиции и объективной частотности, но и он дает предпосылки для постановки вопроса о степени их коррелятивности и, следовательно, о необходимости (а также и возможности) его решения. Можно априори предположить как наличие тенденций к их сближению, так и возможность системных (не случайных) расхождений.

3. Семантические отношения между словообразовательными и лексико-словообразовательными единицами сквозь призму их частотности

Непосредственным объектом изучения в данном, конкретно-исследовательском, разделе выступают результаты реализации продуктивных словообразовательных моделей в речевом материале, отраженном в русском Интернете (Рунете). Таковыми результатами являются внутриглагольные модификационные образования от глаголов состояния. Исходным пунктом для отбора выступили сами глаголы, которые проанализированы в статье Н.Б. Лебедевой (*хвастаться, хвастаться, бодриться*), а также их синонимы, антонимы и гипонимы: *хвалиться, бахвалиться, трепаться, болтать, молчать, грустить, тосковать, печалиться, горевать, болеть, хворать*. Данные глаголы выступали в качестве мотивирующих для модификаторов, образуемых путем присоединения формантов. Их список: В-, В...СЯ, ВЗ-, ВЫ-, ДО-, ДО...СЯ, ЗА-, ЗА...СЯ, ИЗ...СЯ, НЕДО-, О-, ОТ-, ОТ...СЯ, ПЕРЕ-, ПО-, ПОД-, ПРИ-, ПРО-, НА-, НА...СЯ, С-, У- и под. Таким образом мы испытываем возможность вписывания «экспериментальных глаголов Н.Б. Лебедевой» в более широкий контекст. Результаты сведены в табл. 2 и 3¹.

Таблица 2

Префиксальные дериваты невозвратных глаголов

Глагол	Префикс						
	до-	пере-	за-	недо-	на-	по-	вз-
Болеть	3 075	323 224	4 232 888	245	1 010 883	303 283	54
16 431 568	475	144	299	0	1971	50	14
Горевать	245	879	6 973	11	315	44 519	423
819 553	383	5	84	0	720	1	277
Грустить	548	705	177 923	14	228	97 976	446
1 958 382	286	87	1 066	0	595	28	37
Молчать	1 794	2 542	1 512 773	414	2 355	1 133 869	3
11 450 221	2 081	4R	29R	40R	3 408	23R	0R
Тосковать	7	436	84 690	1	51	10 283	176
665 009	198	50	546	0	521	305	703
Хвастать	133	472	286	2	1 834	327 125	0
184 557	57R	9R	636R	2	2 959	1 744 440	46
Хворать	57	917	60 289	0	278	895	0
947 736	47R	0	37R	0	175R	0R	0
Болеть	54	230	72	378	28 902	197 823	292
16 431 568	14	0	19	163	651	96	4
Горевать	423	261	100	3	1 458	4	0
819 553	277	18	239	1	56	0	0
Болеть	3 075	323 224	4 232 888	245	1 010 883	303 283	54
16 431 568	475	144	299	0	1 971	50	14

¹ В составлении таблицы принимала участие студентка 3-го курса КемГУ Я. Дударева. Автор выражает ей благодарность за помощь.

Таблица 3

Префиксальные дериваты невозвратных глаголов (продолжение)¹

Глагол	Префикс						
	<i>до-</i>	<i>пере-</i>	<i>за-</i>	<i>недо-</i>	<i>на-</i>	<i>по-</i>	<i>вз-</i>
1 958 382	446 37	7 0	69 44	22 11	1 231 0	633 0	60 1
Молчать 11 450 221	3 0R	9 0R	122 429 0R	1 046 427 16R	919 220 735	418 444Y	63 64R
Тоско- вать 665 009	176 703	4 0	74 23 592	0 11	208 48	10 0	1 22
Хвастать 184 557	0 46	2 3	798 0	163 51	246 100	210 4	1 0
Хворать 947 736	0 0	55 0R	0 0	0 0	238 0	3 843 3	0 0

Далее мы покажем некоторые наиболее общие «выходы» такого рода построений.

1. В нашем материале не нашлось ни одного приставочного производного глагола, который был бы частотнее исходного, бесприставочного, что не требует комментариев: все сложное (а производное, как правило, сложнее и по форме, и по семантике) существует в меньших количествах, чем простое. Однако некоторые приставочные глаголы существенно приближаются в количественном отношении к бесприставочным, резко выделяясь по этому признаку среди прочих приставочных образований (*помолчать*, *похвастаться*, *погрустить*, *заболеть*, *загрустить*, *замолчать*), и причины здесь тоже понятны лингвисту-«системщику» – так или иначе такое сближение по частотности отражает приближение пары бесприставочный / приставочный глагол к корреляциям собственно видового содержания² (известно, что в сфере вида идет своеобразная конкуренция за первенство на оси семантической маркированности / немаркированности, и частотность отражает ее); нет сомнения в том, что в области так называемых чисто видовых корреляций мы вправе ожидать дальнейшее сближение и обнаружение случаев с обратной корреляцией, ср.: *стараться* – 18 млн, *постараться* – 10 млн, но уже *писать* – 103 млн, *написать* – 122 млн, *делать* – 93 млн, *сделать* – 169 млн. Как видим, тут есть материал для системной прогностики в области динамики грамматической системы русского языка.

Таким образом, в конкуренцию лексики и словообразования самым решительным образом вторгается грамматика, и мы можем наблюдать процессы их грамматикализации в «квантитативном зеркале».

¹ Курсивом внизу указаны показатели возвратных глаголов, которые большей частью являются производными с составными формантами (типа ДО...СЯ, например, ДОСТУКАТЬСЯ, ДОМЕЧАТЬСЯ).

R – Рамблер. Переход на эту поисковую систему связан с тем, что по неизвестным нам причинам Яндекс по данному запросу не дифференцирует невозвратный и возвратный глаголы, дублируя во втором случае показатели по невозвратному глаголу.

Y – Яндекс. Указано в том случае, когда Rambler показывает «0», тогда как в показателях Яндекса данные возвратные глаголы отмечаются (вкупе с невозвратными).

² К примеру, *заболеть* по отношению к *болеть* – уже не просто «начинательный глагол», а именно глагол, который претендует на видовой коррелят *болеть* – процесс, *заболеть* – акт, событие).

2. Очевидно, что такие таблицы дают исследователю возможность измерения потенциала функционирования словообразовательных моделей (типов). К примеру, хорошо видно, что порождающая мощность моделей с формантом НЕДО- явно «недоотягивает» до мощности модели с ПЕРЕ- или ОТ-; очевидно слабы модели с У-, С-, ВЗ-, ПОД-; напротив, сильны модели с ЗА-, НА-; модели с ДО- и ИЗ- близки по количественной мощности, но ДО- все-таки несколько «сильнее».

3. Весьма интересен факт конкуренции лексики и словообразования, обнаруживаемой в горизонтальных рядах. Сравнение показывает наличие двух каналов пополнения речевого материала: наращивание мощности словообразовательной модели и, напротив, – лексикализация. Так, в сфере действия слабых (и не только слабых) моделей мы всегда можем найти «всплески» активности отдельных лексем с явными лексическими наслоениями: *приболеть, изболеть, отболеть, умолчать, смолчать, отмолчиться, взгрустнуть* (ср.: *взгрустнулось*), *наболеть* (ср.: *наболело*). Их суммарные показатели явно не уступают показателям более многочисленных регулярных и стандартных производных.

4. Горизонтальные ряды, выстроенные в порядке частотности (убывающей или возрастающей), показывают мотивирующий потенциал исходных глаголов. В определенном аспекте он коррелятивен частотности самих этих исходных глаголов, что не требует особых комментариев: частотное – значит активное во всех отношениях, в том числе и в отношении деривационного развития. По этой причине больше всего «нулей» в табл. 2 – у глагола *хворать*, а в табл. 3 – у глагола *бахвалиться*. Мы попытались найти коррелятивность в этом плане на уровне общих частотностей мотивирующего глагола и суммарного выхода всех производных от него, однако закономерностей здесь не оказалось: слишком сильно «возмущающее» влияние лексикализированных производных глаголов.

Гораздо больше проявлений коррелятивности на уровне ранговых коэффициентов частотности: во многих строках порядок следования глаголов по частотности такой же, как в исходном столбце. И такая тенденция, несомненно, присутствует, но с такой же закономерностью она нарушается воздействием особенностей лексического функционирования производных слов как единиц лексической системы. Например, в табл. 3 находим полное соответствие в первой строке; во второй, третьей и четвертой поменялись местами (3-м и 4-м) глаголы *храбриться* и *бахвалиться*, в пятой – глаголы *бахвалиться* и *бодриться* и т.п.

5. Достаточно хорошо проявляется в табл. 1 тенденция к количественному преобладанию невозвратных глаголов над возвратными, что объясняется специфическим (узкофункциональным) содержанием составных формантов. Известно (например, по работам И.А. Мельчука), что при чистой постфиксации семантические отношения часто обратные: *веселить* – «каузировать *веселиться*»; ср.: *печалиться* – 2,0 млн, *печалить* – 1,8 млн, *веселить* – 2,2 млн, *веселиться* – 2,8 млн. Соотношение глаголов *хвастать* (184 тыс.) и *хвастаться* (1, 02 млн) определяется не семантикой, а стилистикой, поскольку это формальные варианты непереходного глагола (*стучать* / *стучаться*, *мчать* / *мчаться*); прогноз результатов их конкуренции на основе данных Интернета – явно в пользу возвратных глаголов.

Таким образом, мы видим хорошие возможности у статистики Интернета участвовать в решении проблем системной лингвистики. На наш взгляд, они компенсируют (хотя здесь все достаточно относительно) имеющиеся в этом плане ограничения. Большая часть их связана с высокой степенью формализации количественного коэффициента слова и, как следствие (недостатки – продолжение достоинств!), отвлечения от вариативности семантики, поэтому возникает эффект неоднозначности моделей: полисемия производных всегда остается фактором, снижающим «чистоту» получаемых характеристик системных отношений единиц. Очевидны и сугубо программные ограничения поисковых систем Интернета, связанные с дифференциацией слов и словоформ, слов и вариантов слова. В частности, в ходе нашей работы в отдельных запросах не различались возвратные и невозвратные глаголы, глаголы совершенного и несовершенного видов, дифференцируемые суффиксами. Оценка значимости такого рода помех для решения конкретных задач, видимо, еще долго будет присутствовать при обсуждении результатов, полученных на основе статистик поисковых систем Интернета. И тем не менее для решения многих задач данный источник, несомненно, перспективен. Количественная мощность, доступность, проверяемость (как фактор достоверности), возможность обращения к непосредственным речевым произведениям – существенные лингвистиколингвистические достоинства статистических данных Интернета. Для словообразования они важны возможностью переходить от спорадических фиксаций производных слов к системным, в частности возможностью верификации любого потенциального слова как реального.

Литература

1. Голев Н.Д. Словообразование как эволютивный синергетический процесс // Актуальные проблемы современного словообразования. Вып. 2 / Под ред. Л.А. Араевой. Кемерово, 2008. С. 21–31.
2. Голев Н.Д. О возможностях Интернета в количественном описании лексической семантики // Славянская филология: исследовательский и методический аспекты: Материалы I Междунар. науч. конф., 23–24 марта 2006 г. / Под ред. Н.Б. Лебедевой. Вып. 1. Кемерово, 2006. С. 41–49.
3. Голев Н.Д. Поисковые системы Интернета как лингвистический источник (на примере решения некоторых теоретических и прикладных вопросов русского словообразования) // Новые явления в славянском словообразовании: Система и функционирование: Докл. XI Междунар. науч. конф. Комиссии по славянскому словообразованию при Междунар. комитете славистов / Под ред. Е.В. Петрухиной. М., 2010. С. 197–212.
4. Земская Е.А. Современный русский язык: Словообразование. М., 1973.
5. Соссюр Ф. де. Труды по языкознанию. М., 1977.
6. Лебедева Н.Б. Экспериментальное исследование производного слова: соотношение выводимости как словообразовательного признака и узуральности как лексического признака // Новые явления в славянском словообразовании: Система и функционирование: Докл. XI Междунар. науч. конф. Комиссии по славянскому словообразованию при Междунар. комитете славистов / Под ред. Е.В. Петрухиной. М., 2010. С. 197–212.
7. Щерба Л.В. О трех аспектах языковых явлений и об эксперименте в языкознании // Щерба Л.В. Языковая система и речевая деятельность. Л., 1974.
8. Бодуэн де Куртенэ И.А. Николай Крушевский, его жизнь и научные труды // Бодуэн де Куртенэ И.А. Избранные труды по общему языкознанию. М., 1963. Т. 1. С. 146–202.
9. Сводный словарь современной русской лексики. М., 1991.
10. Голев Н.Д. Лингвистический статус и некоторые особенности функционирования мотивированных единиц русского языка // Номинативные единицы языка и их функционирование: Межвуз. сб. науч. тр. Кемерово, 1987. С. 105–115.