

Компьютерная лингвистика и интеллектуальные технологии:
по материалам международной конференции «Диалог 2018»

Москва, 30 мая — 2 июня 2018 г.

ЛЕКСИКАЛИЗАЦИЯ КАК ОГРАНИЧЕНИЕ НА СТАНДАРТНУЮ ВТОРИЧНУЮ ИМПЕРФЕКТИВАЦИЮ РУССКОГО ГЛАГОЛА (ПО ДАННЫМ НКРЯ)

Горбова Е. В. (e.gorbova@spbu.ru)

Санкт-Петербургский государственный
университет, Санкт-Петербург, Россия

LEXICALIZATION AS A CONSTRAINT ON THE STANDARD SECONDARY IMPERFECTIVATION OF RUSSIAN VERB (BASED ON THE RUSSIAN NATIONAL CORPUS)

Gorbova E. V. (e.gorbova@spbu.ru)

Saint Petersburg State University, Saint Petersburg, Russia

The paper deals with the phenomenon of lexicalization (degrammaticalization) as a possible constraint on the standard secondary imperfectivization of Russian verb, which is regarded as the unique regular morphological mechanism of aspectual inflection. Several cases of lexicalization are discussed on the basis of an analysis of the material found in the Russian National Corpus — two cases of lexicalized plural noun forms (*chasy* 'clock' and *kapli* 'eyedrops') and two cases of secondary imperfectives (*zaigryvat'* 'flirt' and *zamalchivat'* 'hush smth up'). These cases of using lexicalized forms are discussed in comparison with corresponding non-lexicalized forms, if any, of standard plural (for the nouns above — 'hours' and 'droplets') and standard secondary imperfectives (for the verbs above — 'be starting to play' and 'be starting to get silent'). The analysis reveals two different possibilities: the parallel use of lexicalized and non-lexicalized forms (with lexicalized or non-lexicalized forms predominating) and the occurrence

of lexicalized forms only (the case of *zamalchivat'*). The discussion leads to the conclusion that lexicalization significantly hinders standard secondary imperfectivation of prefixed Russian verbs.

Key words: lexicalization, Russian aspect, inflectional category, secondary imperfectivation

Отдельным сюжетом статьи [Храковский 2018], в целом посвященной критическому анализу предложенной в [Горбова 2017] словоизменительной модели русского вида, является обсуждение лексикализации (деграмматикализации)¹. Поясним, что лексикализация в [Там же] использовалась в качестве возможного объяснения ограничения на стандартную соотносительность формы (единожды) префигированного перфектива и вторичного имперфектива (НСВ2)². Целью предлагаемой работы является уточнение этого ограничения, базирующееся на проведенных локальных корпусных (на материале НКРЯ³) исследованиях наличия (или отсутствия) нелексикализованной формы и, в случае наличия, частотности ее употребления по сравнению с лексикализованной.

Кратко представим имеющиеся по вопросу лексикализации позиции применительно к дискуссии о статусе русского вида как словоизменительной vs. словоклассифицирующей грамматической категории (ГК). В [Храковский 2015: 326–327] была предложена классификация (единожды) префигированных перфективов, 4-я и 5-я группы которых отличаются отсутствием стандартного семантического соотношения между СВ и НСВ2 либо у части лексем многозначной вокабулы (это 4-я группа, ср. *пережить (кризис)* и (*сильно*) *переживать*), либо у вокабулы целиком (5-я группа: *заиграть (марш)* ~ *заигрывать (с девушкой)*, *раздумать (говорить)* ~ *раздумывать (над дилеммой)*, (*наконец*) *замолчать* ~ *замалчивать (трудности)*). Наличие случаев с семантической девиацией в парах СВ ~ НСВ2 — один из аргументов против словоизменительной трактовки вида.

В [Горбова 2017: 32] в качестве контраргумента формулируется утверждение: «отсутствие стандартного семантического соотношения между префигированным СВ и НСВ2 является следствием лексикализации (деграмматикализации) грамматической формы». Однако в цитируемой работе дело не ограничивается данным утверждением, как можно было бы понять из [Храковский

¹ Параллельные термины лексикализация и деграмматикализация используются здесь как синонимы вслед за [Плунгян 2011] для обозначения явления, суть которого сводится к тому, что в некоем конкретном случае граммема (вместе с выражающей ее словоформой) становится отдельной лексемой. При этом одновременно имеет место лексикализация (граммема порождает лексему) и деграмматикализация (граммема теряет свой статус). В литературе представлено и несинонимическое использование этих терминов как, например, в [Norde 2009].

² Используемые здесь и далее обозначения НСВ2 для вторичного имперфектива (и НСВ₁ — для первичного), а также понятие бимперфективной тройки, были введены в работе [Анна Зализняк, Микаэлян 2010].

³ Национальный корпус русского языка — URL: <http://www.ruscorpora.ru>

2018: 109]. В публикации [Горбова 2017] приводятся реальные примеры употребления формы *заигрывать* не как лексикализованного, а как семантически стандартного имперфектива от инхоатива *заиграть*. См. языковые иллюстрации в [Там же: 33] под номерами (6) и (7), последний повторен здесь как (1), а также количественные данные о наличии в выдаче НКРЯ таких примеров: 20 вхождений, что составило 3% выборки [Там же]. Таким образом, речь шла отнюдь не о том, что лексикализация является жестким ограничением (или запретом) на стандартную вторичную имперфективацию, а о том, что при появлении лексикализованной формы употребление говорящим стандартной имперфективной формы может быть (прагматически) затруднено. Причина возможных затруднений — образование самостоятельной лексической единицы, отдельного слова — лексемы и вокабулы и, как крайний случай, — формирование омонима.

- (1) *Минут через десять ярко пылал небольшой костерчик на берегу, грелась в открытых банках свиная тушенка, утки проснулись и летали парами над самой водой, далеко где-то заигрывали тетерева, будто бы накручивал кто-то ручку детской балалаечки, взошло солнце, крупная рыба всплескивала всюю, птицы заливались в лесу, но и ноги наши гудели, и сон нас обрывал...* [Юрий Казаков. Долгие крики (1966–1972)]

Именно к выводу о затрудненной реализации стандартного значения подводит анализ выдачи НКРЯ на запрос «заигрывать»: 94,5% (617 из 653 вхождений выдачи) оказались реализацией значения ‘шутливо, весело затрагивать кого-л., вызывая на ответ; кокетничать’ (МАС), омонимичной формам имперфектива как от инхоативного (не зафиксированного в словаре), так и от комплетивного перфектива *заиграть* (подробнее см. в [Горбова 2017: 32–33]). Однако весьма примечательным является то, что несмотря на не слишком благоприятную конкуренцию с деграмматизированной формой имперфектива (уже самостоятельной словарной единицей), говорящие, хотя и существенно реже, но все же используют НСВ₂ как от *заиграть*-комплетива⁴, так и от *заиграть*-инхоатива, см. количественные данные и языковые иллюстрации [Там же].

Поэтому, как представляется, указание В. С. Храковского на то, что «и лексикализованная форма *часы*, и нелексикализованная форма мн. ч. *часы* употребляются свободно без всяких ограничений..., т. е. у говорящего нет никаких затруднений с использованием и той, и другой формы» [Там же: 5] некоторым образом «бьет мимо цели» в том случае, если текст из [Горбова 2017: 32] (см. цитату выше) был понят как утверждение о том, что лексикализация в обязательном порядке приводит к отсутствию стандартного семантического соотношения между префигурованным СВ и образованным от него НСВ₂. Если же воспринимать приведенное ограничение не так жестко, а как импликацию: «если между вторичным имперфективом и соответствующим

⁴ Кстати, именно *заигрывать* как парная имперфективная форма к комплетиву *заиграть* (с двумя лексемами — ‘часто употребляя при игре, играя на чем-л., или чем-л., привести в негодность’ и ‘часто исполняя (сценическое или музыкальное произведение) сделать избитым, надоевшим’) фигурирует в МАС как омоним к лексикализованному *заигрывать*.

префигированным перфективом в речевой практике не обнаруживается стандартного семантического соотношения или же стандартный в семантическом отношении имперфектив находится на периферии, то такое положение дел может быть вызвано имевшей место лексикализацией (деграмматикализацией) грамматической формы», то из этого вытекает следующее. Для подтверждения сдерживающего влияния лексикализации на способность реализации стандартной формой стандартной же семантики нужны дополнительные исследования частотности употребления в речи некой грамматической формы без омонимии с деграмматикализованной формой и при наличии таковой.

За возможность более подробного обсуждения посвященного лексикализации утверждения в [Горбова 2017: 32] мне остается только поблагодарить В. С. Храковского. Что же касается дополнительных исследований, то на данный момент для обсуждения может быть предложен анализ выданных из НКРЯ по двум различным часам и каплям — с лексикализацией и без нее, а также по двум (не)лексикализованным имперфективам — *заигрывать* и *замалчивать*. Однако прежде хотелось бы обсудить еще один тезис из [Храковский 2018], который, будучи принят безоговорочно, ведет к бессмысленности сопоставления пар лексикализованных и нелексикализованных форм существительного (например, *часы1* и *часы2*; *капли1* и *капли2*), с одной стороны, и глагола (например, *заигрывать1* и *заигрывать2*) — с другой.

Имеется в виду указание на то, что при обсуждении феномена лексикализации (деграмматикализации) имперфективных форм обращение (вслед за [Маслов 1963: 4] и [Плунгян 2011: 214–215]) к аналогичному явлению в сфере субстантивной категории числа как к параллели неправомерно. В. С. Храковский специально отмечает два обстоятельства: а) в качестве примеров лексикализации приводились исключительно случаи реализации этого явления в сфере субстантивного числа; б) причина этому — отсутствие лексикализации в сфере «других глагольных словоизменительных категорий, у которых такого явления, очевидно, нет» [Храковский 2018: 4]. Далее приводятся примеры из НКРЯ, призванные самим фактом своего существования показать отсутствие затруднений при использовании как лексикализованной, так и нелексикализованной формы *часы*.

Думается, что здесь имеется противоречие: если аналогия для обсуждаемой лексикализации форм НСВ2 в виде той же лексикализации, но в сфере субстантивного числа (как бы ни характеризовалась последняя категория — как словоизменительная или как словоклассифицирующая) неправомерна, то, занимаясь глагольным видом, нет смысла обсуждать (не)затрудненность использования нелексикализованной формы *часы* на фоне лексикализованной. Верно, по-видимому, и обратное: если свободное употребление нелексикализованной формы *часы* наряду с лексикализованной при обсуждении лексикализации форм НСВ2 может служить аргументом, то правомерна и апелляция к феномену лексикализации в сфере субстантивного числа при обсуждении лексикализации НСВ2 у глагола.

Впрочем, несмотря на неочевидность постулирования нежелательности (запрета?) апелляции к субстантивной грамматической категории при

обсуждении лексикализации в сфере глагольного вида⁵, все же есть возможность сослаться и на «примеры лексикализации грамматической формы других глагольных словоизменительных категорий», ставя под сомнение утверждение, что у них «такого явления, очевидно, нет». Назовем, в частности, лексикализацию форм 3 л. ед. ч. настоящего времени глаголов *стоять* и *идти*: *стоит* 'имеет смысл; следует, надо' и *идет* 'ладно, согласен' (МАС). Оба случая лексикализации зафиксированы в МАС в виде отдельных значений (лексем) в рамках вокабул соответствующих глаголов. Невозможно не признать, что это довольно редкие, практически «штучные», случаи. Однако, по всей видимости, лексикализация (деграмматикализация) — это, в общем случае, феномен скорее именно «штучной» природы. Причем это равным образом характерно и в случае лексикализации существительных (типа *часы* 'прибор для измерения времени' и *капли* 'лекарственная форма', а также *утром, днем, вечером, ночью, зимой, весной, летом, осенью* — где имеет место лексикализация падежной формы с переходом в наречие). Заметим, что последний пример обнаруживает несколько бóльшую регулярность, оставаясь, тем не менее, в рамках невысокой, характерной в целом для области словообразования, регулярности.

Итак, будем считать лексикализацию (деграмматикализацию) явлением, устроенным принципиально одинаково в сфере как субстантивных, так и глагольных ГК.

Приведем результаты обращений к НКРЯ с целью сравнить частотность использования лексикализованной и нелексикализованной форм мн. ч. существительных. Начнем с формы *капли*. При запросе *капли* с указанием семантических признаков «предметное имя, вещества и материалы» (то есть при желании получить реализацию значения 'лекарственная форма') выдача в рамках пользовательского подкорпуса со снятой грамматической омонимией (объемом 2170 документов, 519 725 предложений, 6 003 398 слов, дата обращения — 19.01.2018) составила 26 вхождений в 16 документах. При втором запросе в рамках того же пользовательского подкорпуса, но на этот раз с введением слова в словарной форме (ед. ч.) и выбором грамматического признака «мн. ч.» и семантического признака «предметное имя», была получена выдача в 156 вхождений из 91 документа. Таким образом, случаев употребления нелексикализованной формы мн. ч., как в (2), оказалось в 6 раз больше, чем лексикализованной, как в (3).

(2) *Он ел, и губы его лоснились от жира, капли стекали по подбородку.*

[Токарева Виктория. Своя правда // «Новый Мир», 2002]

(3) *Марья Семеновна напоила гостя горячим молоком и сердечными каплями, уложила на диван, и он там проспал день, затем ночь, так тихо, не шевелясь, спал, что я подходил посмотреть — живой ли?*

[Виктор Астафьев. Затеси (1999)]

Обратимся теперь к сравнению употребительности лексикализованной и нелексикализованной формы *часы*. Так же, как и в предыдущем случае,

⁵ Причина, по которой при обсуждении явления лексикализации следует ограничиться исключительно глагольными категориями, остается не проясненной.

зададим пользовательский подкорпус со снятой омонимией (дата обращения — 19.01.2018, объем корпуса совпадает с указанным выше). Запрос *часы* (с реализацией значения ‘прибор для измерения времени’) дает выдачу в 533 вхождения в 163 документах. Запрос *час* с выбором грамматического признака «мн. ч.» и семантических признаков — «непредметное имя, время, период», имел результатом выдачу в 1148 вхождений в 360 документах. Таким образом, и здесь количество вхождений нелексикализованной формы превысило количество вхождений лексикализованной, на этот раз — более чем в 2 раза (точнее, в — 2,15). Сплошная проверка этих двух выдач позволила выявить и отсеять «мусор» — случаи включения форм с иной семантикой: в выдачу с лексикализованной формой *часов-прибора* оказались включены 13 вхождений нелексикализованной формы (*часов-периода времени*), а в выдачу с нелексикализованной формой — 4 вхождения лексикализованной (по-видимому, вследствие ошибок в разметке). С учетом этого соотношение таково: 1144 вхождения нелексикализованной формы на 520 вхождений лексикализованной, что несколько увеличивает преобладание нелексикализованной формы — до 2,2 раза.

Однако можно предположить, что преобладание *часов* как обозначения временного периода, причем чрезвычайно важного в жизни современного человека, у которого все «расписано по часам», обусловлено именно этой внеязыковой особенностью лексемы. Чтобы нейтрализовать воздействие этого фактора, «очистим» от данного перекоса выборку примеров, изъяв из нее все вхождения в нумеративных словосочетаниях, таких, как (4) и (5), и оставив только подобные (6).

- (4) *Сегодня в 11 часов утра мы проводим акцию протеста на привокзальной площади Владивостока.* [Даниил Володин. Рыбакам надоело правительство (2003) // «Ежедневные новости» (Владивосток), 2003.01.14]
- (5) *Скажу только, что за несколько часов специальных поисков мы так и не смогли понять, что же так не устраивает тех, кто плохо отзывается о современном состоянии зоопарка.* [О. Г. Баринев. Зоологический сад? 2003]
- (6) *Лишь в предрассветные сумерки (впрочем, иногда и в закатные часы) морские коньки парочками разбредаются по подводным зарослям — водорослям, свившимся в леса.* [Александр Голяндин. Рассказы о животных, и не только о них: А у морского конька что за конек? // «Знание — сила», 2003]

После проведения данной операции картина существенным образом меняется. Метод сплошной выборки *часов* в нумеративных контекстах показал, что в выдаче НКРЯ они преобладают: на них приходится 73,25% (838 на 306 вхождений; последние составляют 26,75% случаев реализации всех остальных контекстов, условно обозначенных как «нумеративные»). Если теперь сравнить количество вхождений лексикализованной и нелексикализованной формы *часы* (напомним — в одном и том же пользовательском подкорпусе НКРЯ), то соотношение будет выглядеть уже как 533 на 306 вхождений, то есть с преобладанием лексикализованных *часов* (‘прибора для измерения времени’) в 1,7 раза.

Итак, на трех рассмотренных примерах (лексикализованные и нелексикализованные существенно различные по семантике *заигрывать*, *капли* и *часы*) мы получили в высокой степени отличные друг от друга картины употребительности (не)лексикализованных форм. Они демонстрируют почти весь теоретический спектр возможностей, но без полярных случаев. Для *заигрывать* обнаружено сильное преобладание по употребительности лексикализованной формы над стандартной, семантически соотносимой с инхоативом *заиграть* (617 на 20 вхождений, то есть превышение в 31 раз). Для двух форм *капли*, наоборот, в 6 раз более употребительной оказалась нелексикализованная (156 на 26 вхождений). *Часы* же с лексикализацией (как прибор) оказались в 2,2 раза менее частотны, чем *часы* без лексикализации. Этот результат был получен в условиях отсутствия дифференциации контекстов на нумеративные и ненумеративные, однако при изъятии нумеративных контекстов соотношение изменилось практически на противоположное: лексикализованная форма в 1,7 раза превзошла нелексикализованную.

Полярным случаем будет такой, при котором в случае изменения семантики (тракуемом здесь как лексикализация) окажется заблокированным употребление нелексикализованной формы. Именно это, очевидно, и имеется в виду, когда в [Храковский 2015] вводятся две группы префигированных перфективов, 4-я и 5-я: в обеих имеются лексемы с отсутствием нормального семантического соотношения между СВ и НСВ₂. Наиболее показательный и чистый случай — это 5-я группа: она «включает перфективы, у которых формально есть соотносительные вторичные имперфективы, лексическое значение которых, однако, не детерминируется семантикой исходных перфективов и стандартными правилами вторичной имперфективации»; имеются в виду пары *заиграть* (*марш*) ~ *заигрывать* (*с девушкой*), *раздумать* (*говорить*) ~ *раздумывать* (*над дилеммой*), (*наконец*) *замолчать* ~ *замалчивать* (*трудности*) [Там же: 327].

Понятно, что основания для скольких-нибудь надежных выводов появятся только после проверки по НКРЯ более широкого перечня (или же всех) подобных случаев. Отметим, что по завершении анализа НСВ₂ *заигрывать* и обнаружения вхождений как лексикализованной, так и вполне стандартной в семантическом отношении формы, разумной представлялась гипотеза, что и для остальных упомянутых выше вторичных имперфективов найдутся случаи реализации стандартного для НСВ₂ значения. Однако аналогичное обращение к НКРЯ по поводу *замалчивать* привело к неожиданному результату. Запрос *замалчивать V, ipf*, сделанный по основному корпусу (115 645 документов, 23 803 881 предложение, 283 431 966 слов; дата обращения — 10.02.2018), дал выдачу в 328 вхождений из 273 документов. Сплошной ручной анализ этой выборки привел к результату, который не оправдал наших ожиданий: ни одного случая реализации стандартного значения вторичного имперфектива от инхоатива *замолчать*, то есть полное отсутствие нелексикализованной формы. Все полученные вхождения НСВ₂ *замалчивать* оказались имперфективами от *замолчать*² ‘умышленным молчанием скрыть, не дать узнать что-л.’ (МАС), как в (7).

- (7) *Лжи, которая замалчивает* — как *замолчаны* были планы декабристов или как *замолчала* вся наша историография роль русских царей.
[И. Л. Солоневич. Миф о Николае II (1949) // «Русский Миръ», 2001]

Тот же отрицательный результат был получен и при обращении к поисковым системам: выдача Google (дата обращения — 10.02.2018) содержала 390 вхождений (+ скрыты 400 похожих), среди них также не было обнаружено ни одного случая реализации нелексикализованного НСВ₂ от инхоативного *замолчать*. Таким образом, здесь мы получили тот самый полярный случай, применительно к которому оказывается верным ограничение, сформулированное в [Горбова 2017: 32]: «[о]тсутствие стандартного семантического соотношения между префигурованным СВ и НСВ₂ является следствием лексикализации (деграмматикализации) грамматической формы».

Более масштабное обследование случаев с лексикализацией вторичного имперфектива и изучение взаимоотношений лексикализованной и нелексикализованной (при ее наличии) форм входит в мои ближайшие исследовательские планы, наряду с изучением явления лексикализации в сфере глагольного словоизменения (упомянутые выше *стоит, идет*).

Подчеркнем, что приведенное в [Горбова 2017: 32] и подвергнутое критике в [Храковский 2018] утверждение было сформулировано именно как объяснение для феномена, обозначенного в [Храковский 2015] как специфика двух упомянутых выше групп глаголов, характеризующихся тем, что префигурованным перфективам (при наличии полисемии — в части лексем) формально сопоставлены вторичные имперфективы, «лексическое значение которых, однако, не детерминируется семантикой исходных перфективов и стандартными правилами вторичной имперфективации» [Там же: 327]. То есть утверждение о лексикализации как ограничении на стандартную имперфективацию релевантно для теоретически допустимого и реализованного в случае НСВ₂ *замалчивать* полярного случая, о частотности которого на данный момент нет надежных данных. При этом очевидно, что именно он был описан в [Храковский 2015] как формальная, но не семантическая, соотносительность НСВ₂ для префигурованных перфективов двух выделенных групп. Порождаемая лексикализацией новая лексема, либо затрудняющая, либо, в крайнем случае, блокирующая употребление стандартного и по форме, и по семантике вторичного имперфектива, не препятствует словоизменительной трактовке вида.

Источники языкового материала

1. Russian National Corpus [Natsional'nyj korpus russkogo yazyka], available at: <http://www.ruscorpora.ru>.

Литература

1. *Evgenieva A. I.* (ed.) (1981–1984), Dictionary of Russian Language in 4 volumes [Slovar' russkogo yazyka v 4 tomah], Academy of Sciences of URSS, Moscow, available at: <http://www.slovari.ru/default.aspx?s=0&p=240>
2. *Gorbova E. V.* (2017), Aspectual formation of Russian verbs: inflection, derivation, or a set of quasigrammemes? ("sore points" of Russian aspectology revisited) [Russkoe vidoobrazovanie: slovoizmenenie, (slovo)klassifikatsija ili nabor kvasigrammem? (eshcho raz o bolevyh tochkah russskoj aspektologii)], *Voprosy yazykoznanija*, No 1, pp. 24–52.
3. *Khrakovskij, V. S.* (2015), The category of aspect in Russian: Problematic issues [Kategoriya vida v russskom yazyke: bolevye tochki], The aspectual semantic zone: typology of systems and scripts of diachronic progresses: Proceedings of The fifth International Conference of the Commission on Aspectology of the International Committee of Slavists [Aspektual'naya semanticheskaya zona: tipologiya sistem i ssenarii diakhronicheskogo razvitiya: Sbornik statei V Mezhdunarodnoi konferentsii Komissii po aspektologii Mezhdunarodnogo komiteta slavistov, Kioto, 13–15 noyabrya 2015 g.], Tanaka Print, Kioto, pp. 321–334.
4. *Khrakovskij, V. S.* (2018). Problematic issues of the category of Russian aspect revised [I opiat' o bolevyh tochkah kategorii vida], *Voprosy yazykoznanija*, No 1, (forthcoming).
5. *Maslov Ju. S.* (1963), Morphology of verbal aspect in Modern Standard Bulgarian [Morfologiya glagol'nogo vida v sovremennom bolgarskom literaturnom yazyke], Academy of Sciences of the USSR Publ., Moscow; Leningrad.
6. *Norde, M.* (2009), Degrammaticalization. Oxford University Press, Oxford.
7. *Plungian, V. A.* (2011), Introduction to grammatical semantics: Grammatical meanings and grammatical systems of the world's languages [Vvedenie v grammaticheskuyu semantiku: grammaticheskie znacheniya i grammaticheskie sistemy yazykov mira], Russian State University for the Humanities, Moscow.
8. *Zalizniak Anna A., Mikaelian I. L., Shmelev A. D.* (2010), Aspectual correlation in Russian: Supporting the notion of «aspectual pair» [Vidovaja korreliativnost' v russskom: v zashchitu vidovoi pary], *Voprosy yazykoznanija*. No. 1, pp. 3–23.