

Оригинальная статья / Original Article

DOI: 10.31857/S160578800021458-9

Фольклор и упрек: опыт сентимент-анализа на базе конструкций

© 2022 г. Р. И. Шмурак

Кандидат филологических наук,
доцент Московского педагогического государственного университета,
Россия, 119991, Москва, ул. Малая Пироговская, д. 1, стр. 1
roman.shmurak@gmail.com

Резюме. Настоящая статья представляет собой пилотный опыт по выявлению упрека в фольклорном материале. Теоретическим ядром работы является представление о том, что упрек, как и все языковые явления, должен иметь формальное выражение в языке. Из этого следует, что упрек можно выявить в тексте на базе его формальных характеристик, опираясь на инструментарий грамматики конструкций, а не только с помощью семантики и контекста. На конкретных примерах демонстрируется эффективность метода, его достоверность и доказательность по сравнению с подходами, опирающимися на контекст и прагматическое содержание. Таким образом, перед исследованием стояло две задачи.

Во-первых, эксперимент по проведению сентимент-анализа в условиях отсутствия на момент написания статьи академического корпуса фольклорных текстов. Хотя объем фундаментальных изданий фольклорных текстов более чем значителен, закономерное отсутствие в таких текстах специализированной лексико-грамматической разметки, позволяющей создавать неоднородные по структуре корпусные запросы, возвращает исследователя к ручной обработке фольклорного материала без надежды на автоматизацию процедуры. В то же время, существование распознанных электронных текстов позволяет делать точные запросы с помощью интерфейса программ по редактированию и чтению документов в форматах PDF и DJVU.

Во-вторых, демонстрируются сам метод выявления упрека на базе конструкций и результаты применения этого метода в виде конкретных примеров реализации упрека в фольклорных текстах. Разбираются преимущества и недостатки работы с точными запросами на примере программ PDF-Xchange и DJVU-reader. Предлагается анализ полученных результатов, а также дискутируются “узкие” места и перспективы исследований подобного рода.

Ключевые слова: упрек, фольклор, анализ тональности текста, сентимент-анализ, конструкция, грамматика конструкций, корпус.

Для цитирования: Шмурак Р.И. Фольклор и упрек: опыт сентимент-анализа на базе конструкций // Известия Российской академии наук. Серия литературы и языка. 2022. Т. 81. № 4. С. 70–80. DOI: 10.31857/S160578800021458-9

Folklore and Reproach: The Experience of Sentiment Analysis Based on Constructions

© 2022 Roman I. Shmurak

Cand. Sci. (Philol.),
Associate Professor of the Moscow Pedagogical State University,
1 bld. 1 Malaya Pirogovskaya Str., Moscow, 119991, Russia
roman.shmurak@gmail.com

Abstract. This article presents a pilot experiment to identify reproach in folklore material. The theoretical core of the work is the idea that reproach, like all linguistic phenomena, should have formal expression in language. From which it follows that reproach can be identified in the text on the basis of its formal characteristics, relying on the tools of construction grammar, and not only with the help of semantics and context. Concrete examples demonstrate the effectiveness of the method, its reliability and evidence-based nature in comparison with approaches based on context and pragmalinguistic content. Thus, the study faced two tasks.

Firstly, it presents the experiment to conduct sentiment analysis in the absence of any academic corpus of folklore texts at the time of writing the article. Although the volume of fundamental editions of folklore texts is more than significant, the natural absence of specialized lexical and grammatical tagging in such texts, which allows creating corpus queries that are heterogeneous in structure, returns a researcher to manual processing of folklore material without hope of automating the procedure. At the same time, the existence of recognized electronic texts allows you to make accurate queries using the interface of programs for editing and reading documents in PDF and DJVU formats.

Secondly, the method of identifying reproach on the basis of constructions and the results of using this method in the form of concrete examples of the reproach in folklore texts are demonstrated. The advantages and disadvantages of working with precise queries in interfaces of PDF-Xchange and DJVU-reader programs are analyzed. The analysis of the obtained results is proposed, as well as the “bottlenecks” and prospects of such research are discussed.

Key words: reproach, folklore, sentiment analysis, construction, construction grammar, corpus.

For citation: Shmurak, R.I. *Folklor i uprek: opyt sentiment-analiza na baze konstrukcij* [Folklore and Reproach: The Experience of Sentiment Analysis Based on Constructions]. *Izvestiâ Rossijskoj akademii nauk. Seriâ literatury i âzyka* [Bulletin of the Russian Academy of Sciences: Studies in Literature and Language]. 2022, Vol. 81, No. 4, pp. 70–80. (In Russ.) DOI: 10.31857/S160578800021458-9

I. Введение

Не вызывает сколь-нибудь сильного удивления, что при упоминании конструкций в контексте фольклорного материала первое, что приходит на ум, — это сказочные формулы, описанные В.Я. Проппом:

- “русским духом пахнет”, “пошел стрелок в путь-дорогу”, “стань к лесу задом” [1, с. 6, 32, 45];
- “в некотором царстве”, “ехал близко ли, далеко ли, низко ли, высоко ли, не так скоро дело делается, как сказка сказывается”, “стань передо мной, как лист перед травой” [2, с. 199, 205, 215];
- “по щучьему велению”, “в некое время я тебе пригожусь”, “будь опять девицей” [3, с. 44, 51] и т.д.

Многие из сказочных формул-клише действительно имеют в своей основе конструкции, в том числе и конструкции упрека, о котором пойдет речь в статье:

- “Ах ты, волчья сыть да травяной мешок! Али ты идти не хошь, али нести не можь?” [4, с. 45];
- “Куда тебе дураку! Знал бы сидел на печи!” [5, с. 232];
- “Эх ты, старая ведьма! Не накормила, не напоила прохожего человека, а уж вестей спрашиваешь!” [6, с. 200];
- “Поделом вору и мука!” [7, с. 162] и т.д.

Тем не менее, термин “конструкция” используется нами в его грамматическом понимании, и речь в работе пойдет не о сказочных формулах-клише.

Как явствует из названия, в фокусе внимания статьи находятся конструктивная природа языка, рассматриваемая в контексте взаимоотношений фольклора и упрека. В теоретическом аспекте статья стремится ответить на вопрос, насколько в контексте фольклорного материала применим анализ тональности текста (англ. *sentiment analysis*), то есть может ли исследователь фольклора при необходимости поставить себе задачу выявить определенный вид эмоций, а в широком смысле и любые семантические поля, и возможно ли это в принципе в условиях отсутствия на момент написания статьи специализированного корпуса русского фольклора. В практическом плане статья демонстрирует возможность решения поставленной задачи инструментарием грамматики конструкций на примере семантического поля упрека. При этом в задачу настоящей работы не входит изучение фольклора как такового, он представляет собой материал, на базе которого происходит апробация. Упрощенно говоря, мы попытаемся извлечь упрек из массива различных фольклорных текстов с помощью методов грамматики конструкций.

Фольклор как разносторонне изученный концепт не нуждается в специальном представлении, в то время как понятия “конструкция” и “упрек” в рамках работы требуют четкой дефиниции во избежание дальнейших терминологических разночтений.

II. Конструкция

Термин “конструкция” принадлежит понятийному аппарату “грамматики конструкций” (англ. — “construction grammar”), см. [8]; [9]; [10], а также на русском [11]. Мы разделяем точку зрения многих авторитетных лингвистов, рассматривающих язык не как набор отдельных лексем, а как результат сложного взаимодействия различных конструкций между собой [12, с. 48]. Вот как метафорически описывает понятие конструкции О.Н. Ляшевская: «Подобно тому как можно судить о форме и структуре поверхности объекта по тени, которую отбрасывает этот объект, компьютерные лингвисты судят о семантике лексики. Они говорят: “Мы не знаем, что такое значение, но можем предположить, что присутствие тех или иных словоформ в контексте или их конфигураций соотносится с определенным типом значения”» [13, с. 261]. Идея, что существуют некие абстрактные модели, имеющие в языке свои реальные воплощения, позволяет, в свою очередь, предположить возможность выявления в тексте определенных семантических полей на формальной основе, с той или иной степенью автоматизации, без обращения собственно к семантике и контексту. Речь притом идет не о фразеологизированных сочетаниях и их вариантах, а о полноценных конструкциях, реализуемых множеством модификаций, далеко выходящих за пределы вариативности фразеологических единиц. В качестве примера рассмотрим конструкцию упрека “куда + тебе + инфинитив + !”. При самом поверхностном поиске по фольклорным текстам мы получим следующие ее реализации:

— “*Куда тебе отнять! Мой муж дюже сильный, с двадцатью головами!*” [5, с. 296];

— “*Да куда тебе со мной взапуски бегать! Мой маленький брат Заинька — и тот тебя далеко за собой оставит!*” [5, с. 336];

— “*Куда тебе мою крепость пробовать? Ты муха передо мной!*” [5, с. 282];

— “*Куда тебе, старому черту, жениться, разве мне, молодцу!*” [5, с. 284];

— “*Куда тебе идти, — отвечал убогий, — ты еще мал и глуп!*” [6, с. 311];

— “*Куда тебе, дураку, воевать; тебе впору свиньям корм давать!*” [7, с. 241];

— “*Ну, куда тебе еще странствовать! Денег у нас и так нет, что ж ты с собой возьмешь?*” [14, с. 251].

Как можно видеть из представленного примера, концепция конструктивной природы языка действительно позволяет автоматизировать поиск необходимой семантики на начальном этапе, делает этот поиск полностью автономным от контекста, значительно повышает его эффективность. В то же время неизбежно возникает ряд вопросов, без ответа на которые описание метода представляется неполным:

1. Как была найдена сама конструкция, по которой осуществляется поиск?
2. Какие параметры лежат в основе поиска?
3. Каковы ограничения поиска?
4. Что понимается в работе под термином “упрек”?

Вопросы 1–3 соотносятся с методологией исследования и будут рассмотрены в соответствующем параграфе “Методы”, а сейчас обратимся к упреку как языковому феномену.

III. Упрек

На настоящее время в отечественной лингвистике не существует однозначного толкования упрека как языкового феномена. Причина этого кроется, во-первых, в большом количестве работ, посвященных упреку и выполненным учеными разных школ и направлений, а во-вторых, в традиционном подходе к упреку с позиций семантики, при котором игнорируются формальные характеристики явления.

Поскольку наша работа посвящена “визуализации” упрека через формальную сторону, дефиниция его семантики не является ключевой задачей статьи. Однако, отдавая дань академической традиции, диктующей четко формулировать объект исследования, мы полагаем необходимым выразить нашу точку зрения по этому вопросу.

Учитывая разнообразие и значительную конкуренцию существующих определений упрека, мы не считаем целесообразным в рамках статьи ни защищать какое-то одно из них, ни создавать еще одно авторское. В том или ином ключе каждое из этих определений ориентируется на три базовые характеристики: коммуникативную цель, контекст и субъектно-объектные отношения. Основываясь на этих характеристиках, в пределах работы мы будем считать упреком высказывание двухкомпонентной структуры, содержащее в себе обвинительную составляющую (1) и трансляцию

субъектом своей эмоциональной уязвленности (2) от вымышленных или реальных действий объекта.

Как известно, одной из самых больших слабостей семантического анализа является проблема интерпретации. Проведем в качестве примера мысленный эксперимент. Группа интерпретаторов оценивает семантику одного высказывания. Интерпретатор А говорит, что высказывание является упреком. Интерпретатор В предполагает, что высказывание может быть как упреком, так и другой формой речевой агрессии, например, обвинением. Интерпретатор С абсолютно убежден, что высказывание представляет собой ярко выраженное обвинение. И, наконец, интерпретатор D считает высказывание лишенным семантики упрека и других видов речевой агрессии. Кто же прав? Очевидно, что в условиях эксперимента ответа просто не существует. Для успешной реализации коммуникативной ситуации упрека необходимы не только желание субъекта обвинять и транслировать свою эмоциональную “уязвленность”, но и акцепция — интерпретирование высказывания объектом в качестве упрека. Подход к упреку с позиций грамматики конструкций помогает отчасти нивелировать субъективность интерпретации, но только отчасти. Конечное решение, считать ли упреком пример, полученный в результате поиска по конструкции, остается все же на откуп исследователя-интерпретатора. Однако, даже с учетом погрешностей интерпретации, выдача на конструкцию 70–100% примеров, ассоциированных с упреком, будет доказательством ее регулярности. Выдачи именно таких конструкций мы будем считать упреками в рамках статьи. Подробнее об этом см. [15]; [16]; [17].

IV. Методы

Как было упомянуто выше, ключевыми вопросами для решения поставленной задачи являются следующие:

1. Как была найдена сама конструкция, по которой осуществляется поиск?
2. Какие параметры лежат в основе поиска?
3. Каковы ограничения поиска?

Поскольку описываемый метод является ядром практической части работы, необходимо тезисно остановиться на его базовых положениях.

1.

Создание конструкции представляет собой классический пример процесса научного поиска. В начале этого пути исследователь не имеет никаких предвзятых представлений о том, как

будет выглядеть будущий список конструкций. Исследователь может опираться лишь на чувство языка, присущее ему как носителю. Погружаясь в материал (в нашем случае былины и сказки), исследователь не стремится выписывать упреки, но старается увидеть закономерности тех локусов, в которых были найдены упреки. С течением времени и увеличением опыта этот поиск сужается и становится более таргетным. В распоряжении исследователя оказываются фрагменты текста, которые с той или иной степенью достоверности указывают на наличие в исследуемом материале упрека. Эти сигнальные фрагменты — **маркеры** упрека (выделено нами. — Р.Ш.). По аналогии с медицинским пониманием маркера такой фрагмент не гарантирует 100%-ное наличие упрека, а лишь позволяет исследователю “заподозрить” это наличие. Чем надежнее (“специфичнее”) маркер, тем меньше ложноположительных результатов будет в выдаче примеров, и наоборот. Эти характеристики маркера упрека делают его субъективной единицей поиска, в отличие от конструкции, в состав которой такие маркеры обычно и входят. Так, например, в описанной ранее конструкции упрека “куда + тебе + инфинитив + !” маркером в общем виде является “вопросительно-относительное местоимение / наречие + личное местоимение 2 л. ед. ч. Д. п.”. Это не особо специфичный маркер для упрека, кроме того, вне корпуса с лексико-грамматической разметкой задать его параметры невозможно. Но даже сузив маркер до сочетания “куда + тебе”, при вводе в материал мы все равно получим вместе с упреками ряд ложноположительных результатов — примеров с иной семантикой:

— “Садись скорей на меня, — сказала жар-птица, — и полетим, **куда тебе** надобно; а то баба-яга съест тебя!” [6, с. 229];

— “Ну, сынок, вот тебе ключи от всех палат; всюду ходи — **куда тебе** захочется, только не смей заглядывать в одну палату, что заперта двумя замками, золотым да серебряным” [6, с. 252].

Следующим шагом будет обнаружение в полученной выдаче закономерности, что “куда + тебе” в сочетании с инфинитивом и восклицательным знаком передает минимум в 90% случаев семантику упрека. Таким образом, на основе выдачи по маркеру может быть спрогнозирована конструкция упрека, которая будет объективной единицей поиска, поскольку в отличие от маркера, порожденного интуицией исследователя, при вводе в лексико-грамматический поиск корпуса, например, НКРЯ [18] такая конструкция позволяет получить конкретные реализации упрека в русском языке. Например:

I. Поисковый запрос:

“куда + тебе + инфинитив + !” (куда на расстоянии 1 до 3 от ты, dat на расстоянии 1 до 3 от inf, beхcl) (дата обращения: 14.02.2022)

II. Лексико-грамматический поиск:

Основной корпус: 20 документов, 20 вхождений / 1 вне семантики упрека

Газетный корпус: 3 документа, 3 вхождения / 0 вне семантики упрека

Поэтический корпус: 1 документ, 1 вхождение / 0 вне семантики упрека

Остальные корпуса: 0 документов.

III. Пример реализации в корпусе:

“– Погляди на себя, куда тебе сейчас *пить!*” [С.Ф. Буданцев. Отчий дом (1928)]

IV. Количество вхождений / количество вне семантики упрека:

24 вхождения / 1 вне семантики упрека

V. Всего упреков:

95,8%

Представляется, что такой высокий процент упреков в выдаче корпуса доказывает, что перед нами действительно конструкция упрека.

2.

В основе поиска, представленного в практической части работы, лежат следующие параметры:

1) Список маркеров и конструкций упрека, разработанный автором в процессе написания диссертации [22];

2) Дефиниция семантики упрека и научные установки, представленные в статье;

3) Основным материалом исследования послужили русские былины [4] и сказки [5]; [6]; [7], в качестве дополнительных источников использовались сказки братьев Гримм [14]; [19] и некоторые другие издания.

3.

Серьезный отпечаток на работу накладывают следующие ограничения поиска:

1) Как было отмечено, в настоящее время не существует сколь-либо серьезного академического корпуса русского фольклора. Таким образом, в распоряжении исследователя оказывается относительно ограниченный (по сравнению, например, с массивами Национального корпуса русского языка) объем разрозненных источников.

Учитывая исключительно электронный формат поиска, за пределами этих источников оказываются материалы с нераспознанным текстом, то есть такие, в которых невозможно осуществлять внутритекстовый поиск. Определенное ограничение на поиск накладывает также авторский отказ от источников иных файловых форматов в пользу PDF и DJVU, программы для которых имеют наиболее продвинутый интерфейс и улучшенный функционал внутритекстового поиска.

2) Корпусный поиск, помимо преимуществ в объеме и следовательно в репрезентативности данных, обладает для исследований подобного рода и еще одной ключевой характеристикой – возможностью практически в 100% случаев трансформировать искомую конструкцию в поисковый запрос. Это достигается за счет корпусной разметки текстов и поисковой машины корпуса, предоставляющей в распоряжение исследователя возможность создавать неоднородные по своей структуре запросы, учитывать семантические, грамматические, синтаксические и др. характеристики конструкции. Внутритекстовый поиск в программах для форматов PDF и DJVU возможен только по точным словам, что, конечно, значительно суживает поле исследования, не позволяя использовать конструкции как таковые, а лишь их составные части – маркеры упрека.

Несмотря на перечисленные ограничения, эффективность и достоверность исследования сложно подвергнуть сомнению.

Да, объем источников по сравнению с массивами корпуса не так велик. Однако о недостатке репрезентативности можно поспорить, поскольку речь идет о многотомных академических источниках, авторитетность которых самоочевидна. Кроме того, перед исследованием не стоит цель механически пересчитать все упреки в русском фольклоре. Задача состоит в том, чтобы продемонстрировать метод, с помощью которого упреки могут быть извлечены из фольклорного текста с определенной степенью автоматизации на базе грамматики конструкций. Для решения этой задачи объем не является столь уж ключевым параметром.

Да, поиск осуществляется фактически на основе маркеров, что увеличивает количество ложноположительных результатов и должно было бы значительно затруднить работу исследователя по ручному отсеиванию зерен (упреков) от плевел (высказываний с иной семантикой). Однако, во-первых, за счет относительно небольших объемов по сравнению с корпусом выдач такой отбор в принципе осуществим, а во-вторых, в любом

случае представляется в разы более эффективным, чем ручной отбор с карандашом в руках по бумажным источникам, как это происходило до появления корпусов.

V. Представление результатов

Конструкция “где уж + инфинитив”

1. Реализация в фольклорных текстах

— “Ну, ребятушки, вот какова рысь kota да барана! Евстифейка-волка умудрились сшибить и изувечить с какого высокого пня, а нам где уж на земле устоять!” [5, с. 65];

— “Где уж нынче правдой жить? Правда-то в лаптях ходит, а кривда в сапогах!” [5, с. 202];

— “Где уж тут о работе думать? Лучшие погуляем на последки” [5, с. 236];

— “И то правда, — говорит волк, — где уж тебе, сестрица, идти, садись на меня, я тебя довезу” [20, с. 241].

2. Комментарии:

Упреки, полученные на конструкцию, делятся на 3 вида: самоупрек (1), “риторический” упрек (2) и упрек со стертым обвинительным компонентом (3).

(1) Вектор классического упрека идет от субъекта высказывания к его объекту. В случае самоупрека в качестве объекта упрека выступает сам субъект, другими словами, упрек направлен “на себя”. Формально самоупрек легко определяется по наличию в составе высказывания личных местоимений 1-го и 2-го лица (я, мы во всех формах), что хорошо демонстрирует представленный пример: “Ну, ребятушки, вот какова рысь kota да барана! Евстифейка-волка умудрились сшибить и изувечить с какого высокого пня, а нам где уж на земле устоять!”.

(2) “Риторический” упрек достаточно редкий вид упрека. Его особенностью является отсутствие очевидного объекта, это упрек “в воздух”, упрек-обращение к “высшим силам”, не предполагающий получение ответа: “За что мне все это?”, “Ну что же у меня за жизнь такая?”, “Почему они всегда в мою смену приходят?” и т.п. Еще одной чертой “риторического” упрека является генерализация — субъект как бы говорит “за всех”. “Где уж нынче правдой жить?” и “Где уж тут о работе думать?” — типичные примеры этого вида упрека, обладающие обоими из перечисленных признаков. Такой упрек, вероятно, довольно близок фольклорной речи, так как генерализация и обращение к высшим силам ей в полной мере свойственны.

(3) Стертость одного из компонентов упрека выводит его на периферию семантического поля

и затрудняет определение значения высказывания. Типичные представители такой коммуникативной ситуации “ласковые упреки”. Пример “И то правда, — говорит волк, — где уж тебе, сестрица, идти, садись на меня, я тебя довезу” может быть понят и вне семантики упрека. Явственного обвинения здесь не прослеживается, тональность высказывания определенно положительная. Считать ли такое высказывание упреком остается на усмотрение интерпретатора. Наше отнесение примера к упреку мотивировано тем фактом, что, несмотря на общую положительную коннотацию высказывания (волк собирается помочь лисе), “где уж тебе, сестрица, идти” означает “плохо, что ты не можешь идти — тебе стоило бы мочь — из-за тебя у меня могут быть трудности”, что вполне попадает под семантику упрека, даже учитывая стертый характер обвинения.

Конструкция “почто + глагол + ?”

1. Реализация в фольклорных текстах

— “Ах ты, эй, Добрынюшка Никитинец! Ты, знать, порушил свою заповедь. Зачем стоптал младых змеенышей, Почто выручал полоны да русские?” [4, с. 69];

— “Ай же ты, мое чадо милое Молодой Василий да Буслаевич! Что же ты ходишь по городу уродуешь? Почто же хватаешь за руки, Руку рвешь из плеча вон? Почто хватаешь за ноги, Ног у рвешь из ходило вон?” [4, с. 114];

— “Почто вы губите себя? Разве вы можете перестоять нас? Если будете стоять и десять лет, то что сможете сделать нам?” [4, с. 383];

— “О милое мое чадо, петел! Почто ты такую речь говоришь? Почто я учиню так? Читывал ли ты притчу про мытаря и фарисея, как мытарь спасся, а фарисей погиб за гордость? Ты, мое милое чадо, без покаяния на высоком древе погибнешь” [5, с. 26];

— “Детоньки, миленьки! Не вы ли скота-то пустили? Почто же, детоньки, пакостите? Не надо, миленьки!” [7, с. 181];

— “Ох ты гой еси, ладо немилое, Немилый и постылый! Ты почто убил моего брата любимого? Своего шурина постылого?” [21, с. 570];

— “Ох ты ой есть, пресвята мать Богородица! Ты почто это меня нынче повыдала? Я за веру стоял да Христовую, Я за церкви стоял да за соборные” [21, с. 88].

2. Комментарии

Упреки, маскирующиеся под причинный вопрос, не исчезли и в современной речи и, как показывают примеры, очень характерны для фольклорных текстов. Конечно, не всегда возможно с уверенностью сказать, что упрек присутствует

в таком вопросе. При работе с выдачей семантика упрека на фоне очевидной причинности казалась нам настолько ускользающей, что, размышляя над полученными результатами, мы уже практически приняли решение не включать “почто + глагол + ?” в качестве конструкции упрека в статье. Но возникшее в ходе анализа соображение, представленное ниже, полностью нивелировало эти сомнения. Допустим, что все представленные примеры всего лишь вопросы причины, без намека на семантику упрека. Тогда получается, что субъект высказывания действительно желает знать именно причину событий, а не выражает обвинения и свою уязвленность действиями объекта? При такой интерпретации вопросы в примерах приобретают гротескную окраску. Змея, вопрошающая Добрыню, “зачем <он> стоптал младых змеенышей” и “почто выручал полоны да русские”, просто интересуется актуальной для нее информацией. Василия Буслаевича просят указать причины, по которым он ходит по городу, “уродует” и “рвет ноги из ходилов вон”. Исключительно, чтобы узнать причины и учесть ошибки, человек, который всю жизнь “за веру стоял да Христовую, за церкви стоял да за соборные”, любопытствует у Богородицы, почему она его “повыдала”, и т.д. Прием доведения аргументации до абсурда позволяет хорошо продемонстрировать, что в представленных примерах ответ в виде перечисления причин совсем не ожидается и мы действительно имеем дело именно с семантикой упрека, а не причинностью в чистом виде.

Конструкция “не + стыдно / совестно + инфинитив + ? / !”

1. Реализация в фольклорных текстах

— “Да постыдитесь же, дрянные девчонки! Как вам не совестно бежать следом за этим молодым парнем? Разве так-то водится?” [19, с. 379];

— “Что, куманек, не стыдно ли свой грех на чужих сваливать? Отпирайся теперь, вор, а укоры-то верные” [5, с. 21];

— “Как не стыдно тебе, молодой юноша, воровать! Да кто ты таков, и которые земли, и какого отца сын, и как тебя по имени зовут?” [5, с. 417].

2. Комментарии

Категория “стыда” напрямую взаимосвязана с упреком. Лексемы, имеющие в своем составе корни “стыд-”, “совест-”, “срам-”, часто обнаруживаются в упреках, и не только в составе представленной конструкции. Любопытной деталью является, что упрек обычно встречается в высказываниях, где подобные лексемы входят в состав предложений с вопросительной, восклицательной интонацией и отрицательной формой,

и совсем не характерен для утвердительных и невосклицательных предложений.

Конструкция “что + личное местоимение 1, 2 л. Им. п. + наделал + ?”

1. Реализация в фольклорных текстах

— “Ах, Иван-царевич, что же ты наделал! Если бы ты еще только три дня подождал, я бы вечно твоей была” [4, с. 173];

— “После бедный дворянин одумался и говорит сам себе: **Что я наделал?** Взятся узнать, а сам ничего не ведаю!” [4, с. 297];

— “Отец перепугался: **“Что ты наделал!** Ведь я теперь должен буду уплатить за топор, а чем я платить буду? В этом только и весь прок от твоей работы!” [14, с. 61];

— “Ах, Катерлихен! **Что ты наделала?** Ведь это же не черенки были, а чистое золото, и в том было все наше состояние. Ты бы этого не должна была делать!” [19, с. 335];

— “Доченька, — сказала мать, — **что ты наделала?** Но теперь-то уж замолчи, чтобы никто этого не знал; ведь теперь уж этого не воротить! Давай разварим его в студень” [19, с. 263];

— “**Что ты, дурак, наделал?** Зачем медведя в сарай притащил? Он всех коров у нас порешил!” [5, с. 333];

— “Ах, братец, **что ж это мы наделали?** Старичку поклона не отдали; давай нагоним его да поклонимся” [5, с. 350];

— “**Что ты наделал, старый пес?** Уходил ты моих дочек, моих кровных деточек, моих ненаглядных семечек, моих красных ягодок! Я тебя ухватом прибую, кочергой зашибу!” [5, с. 96];

— “Ах, Лебедь-птица, красная девица! **Что я наделал:** спьяну тобой похвалился, своей жизни лишился!” [7, с. 21];

— “Ах, батюшки! **Что ж ты наделал;** ведь это дочь нашего купца — страшная ведьма! Не одного молодца свела она с белого свету” [7, с. 134];

— “**Что ты наделал, собака?** И нас упекешь тут же!” [7, с. 188];

— “Эх, дурак! **Что ты наделал?** — закричали братья. — Сам пропадешь и нас загубишь! Куда теперь мертвое тело девать?” [7, с. 199].

2. Комментарии

Безусловно, смыслообразующим ядром данной конструкции является глагол “наделал”. Однако само по себе его наличие в тексте еще не означает появления семантики упрека. Ср.: “В одном городе был вороватый человек и **наделал** уже много бед” [7, с. 270]. Важными элементами являются вопросительно-относительное местоимение *что*, личное местоимение *ты* и усилительная частица *же*. Все три являются маркерами упрека и часто сопровождают его в различных конструкциях как вместе, так и по отдельности. Любопытно, что личное местоимение 2-го лица ед. ч. “*ты*” наиболее частотно из всех личных местоимений

в упреке. Это косвенно подтверждает мысль некоторых авторов о том, что упрек, как правило, обращен к близким людям — с равным или более низким социальным статусом.

Конструкция “из-за + тебя + сказуемое”

1. Реализация в фольклорных текстах

— “Ах, — сказал отец, — какие мы тревоги **из-за тебя вынесли!**” [19, с. 222];

— “Вышел солдат из дворца, забрел в дремучий лес, сел под кустик, задумался-закручинился: “Ах, дух нечистый! Все **из-за тебя пропадаю**”» [6, с. 250];

— “Ты как сюда попал? Ведь говорил тебе: берегись, так нет! Не жалко мне, что тебя прутьями били, а жалко мне, что **из-за тебя** триста рублей даром **пропали!**” [7, с. 165].

2. Комментарии

Ядро конструкции “из-за тебя”, как и в случае с глаголом “наделать”, может встречаться в высказываниях и без семантики упрека. Однако появление в предложении глагола (особенно глагола в прошедшем времени) делает вероятность упрека очень высокой.

Конструкция “напрасно + личное местоимение + сказуемое, пр. вр.”

1. Реализация в фольклорных текстах

— “Совсем **напрасно ты**, Феодул Змеуланович, на свою силу **понадеялся!** Хвастал, что Русь повоюешь, богатырей русских полонишь. А вышло не по-твоему. Так не лучше ли нам замирились подобру-поздорову?” [4, с. 237];

— “Эх, совсем **напрасно я** тогда Данилу Белого в живых **оставил!** Землю он ел, клялся не воевать с королевством Картауса, я и поверил! А он вон чего натворил! Отплатил за мою доброту. Ну, погоди, рассчитайся с тобой! Теперь пощады не будет!” [4, с. 243];

— “**Напрасно вы**, братцы, **ворчали** на меня! Ведь и колодец, и яблоня, и ковер — все это чудо-юдовы жены были” [4, с. 279].

2. Комментарии

Синонимичные лексемы “зря” и “напрасно” высокоспецифичны для упрека. Представленные примеры не отражают все многообразие вариантов выражений с семантикой упрека (или генетически родственной семантикой), полученные при работе с текстами.

За пределами конструкции остаются:

1) Примеры с тем же составом и инверсией:

— “Экой я дурак! Что ж **я напрасно съел** столько калачей. Мне бы надо сначала съесть одну баранку” [4, с. 346];

— “Нет, видно **мы напрасно смеялись** над женой Ивана-царевича, она не лягуша, а кака-нибудь хитра чародейка!” [6, с. 329].

В этих двух примерах мы не имеем дело с какой-то новой сущностью — это всего лишь инверсивный вариант изучаемой конструкции. Оба примера реализуются как ярко выраженные самоупреки, что только подтверждает их родство с конструкцией.

2) Примеры, где в составе конструкции отсутствует личное местоимение:

— «Журавль раздумался и сказал: “**Напрасно не взял** за себя цаплю: ведь одному-то скучно. Пойду теперь и возьму ее замуж”» [4, с. 145];

— “Но только **напрасно считают** нас первыми красавицами. Вот в Вахрамеевом царстве царица, дочь царя Вахрамея, — та настоящая красавица” [4, с. 238].

Здесь мы также имеем дело с вариантами все той же конструкции “напрасно + личное местоимение + сказуемое, пр. вр.”. Отсутствие местоимения (легко, впрочем, восстанавливаемого) объясняется синтаксической структурой этих примеров, это односоставные предложения определенно-личного и неопределенно-личного типа соответственно.

3) Примеры, сказуемое которых глагол не в прошедшем времени:

— “Это похвально, Еруслан Лазаревич, что ты своих близких в горе-беде не покинул, выручил, а теперь поехал живую воду для них доставать. Но скажу тебе: **напрасно ты туда поедешь**” [4, с. 246].

То, что данный пример, несмотря на явную родственность, является вариантом конструкции, вопрос дискуссионный, поскольку семантика упрека за счет еще не совершенного действия здесь не так уж и очевидна.

Почему подобные случаи вынесены за границы конструкции, хотя как минимум два из них бесспорные ее варианты? Причина этому лежит в необходимости соблюдения баланса между “количеством” и “качеством” конструкции. Излишнее расширение границ конструкции приведет не только к взрывному росту количества примеров, но и к росту ложноположительных результатов, а сверхстрогое регламентирование, хотя и уменьшит риск примеров, не попадающих под семантику упрека, приведет также к серьезному уменьшению выдачи в целом и потере интересных примеров в частности. Поскольку статья представляет собой пилотный опыт работы с фольклорным материалом, в целях большей наглядности и сохранения “качества” мы жертвуем “количеством” выдачи и рядом ценных примеров. Сказанное верно не только для данной конструкции, но и остальных конструкций, представленных в работе.

Конструкция “ты + небось + сказуемое, пр. вр.”

1. Реализация в фольклорных текстах

— “А **ты небось** думал, что у меня ноги-то лошадиные, как у тебя? — возразил ему солдат. — А с коих пор ты стал таким скрягой?” [14, с. 422];

— “Мало ли что! — отвечал ему малютка. — **Ты небось** наелся досыта, а я вот хочу повеселиться! И опять стал кричать во все горло” [19, с. 222];

— “Э-э, господин лис! — крикнула опять кошка. — Где же твоё уменье и хитрости? Кабы **ты** умел так лазать по деревьям, как я, так **небось** не **поплатился** бы шкурою” [19, с. 420].

2. Комментарии

“Небось” еще одна лексема-маркер упрека. В реальности при расширении конструкции можно получить гораздо больше интересных примеров с неизбежным при этом ростом ложноположительных срабатываний:

— “*“Люди добрые землю пашут, а ты без просыпу спишь!” — “А **ты небось** хочешь, чтоб я на тебя нахал? И не думай!”*» [5, с. 437].

Пример с настолько очевидной семантикой упрека, что кажется невозможным не включить его в состав работы. Однако включение в конструкцию помимо прошедшего настоящего времени немедленно даст и результаты вне семантики упрека или с очень смазанной семантикой:

— “*“**Ты небось** сидишь над кладом?” — сказал мужичонка. “Конечно, над кладом, — отвечал чертенок, — и в том кладе больше найдется серебра и золота, чем тебе на всем твоём веку видеть пришлось”*» [14, с. 340];

— “*Вот он однажды вечером и говорит королю: “**Ты небось** думаешь, что к тебе поступило на службу двенадцать егерей?” — “Ну, да! Конечно, двенадцать егерей!” — отвечал король*” [19, с. 392].

Примеры хорошо демонстрируют, как сказанное выше о балансе “количества” и “качества”, так и некоторую условность создаваемых конструкций. Условность состоит в том, что структура конструкций будет зависеть от опыта и интуиции исследователя, от инструментов поиска, которые исследователь может использовать, от задач, которые стоят перед исследованием, и т.д.

VI. Выводы

Как заявлено выше, задача исследования подразделялась на два аспекта:

— эксперимент по выделению упрека в фольклорном материале на базе конструкций;

— доказательство эффективности используемого метода.

Можно сказать, что в той или иной мере обе цели были достигнуты. Прежде всего, были получены конкретные реализации конструкций упрека в фольклорных текстах. Несмотря на технические ограничения и невозможность использовать лексико-грамматическую разметку, данные, полученные исключительно с помощью внутритекстового поиска программ Windjview и PDF-Xchange по разным источникам, представляются в целом (с поправками на неизбежную погрешность и техническое несовершенство) достоверными и заслуживающими внимания.

Нет сомнения, что поиск, поэтапно осуществляемый по маркерам и конструкциям упрека, в состоянии обеспечить определенный уровень автоматизации при работе с фольклорными источниками, сделать работу исследователя с большими массивами данных более продуктивной и системной. Основным преимуществом такого сентимент-анализа является его более высокая (по сравнению с классической ручной обработкой данных) объективность, достигаемая алгоритмизацией действий исследователя, автоматизацией процедуры поиска и регулярной природой конструкций. Впрочем, и продемонстрированный уровень объективности не может быть признан полностью удовлетворяющим. Следующим этапом, который мог бы поднять объективность подобных исследований на недосягаемую в настоящее время высоту, является запуск специализированного корпуса русского фольклора.

Следует также добавить, что использованный в статье метод поиска по конструкциям не привязан именно к упреку. Это дарит надежду на его применение при работе с широким спектром семантических полей, помимо собственно упрека, на базе которого был проведен описанный опыт.

СПИСОК ЛИТЕРАТУРЫ

1. *Пропп В.Я.* Исторические корни волшебной сказки / Научная редакция, текстологический комментарий И.В. Пешкова. М.: Лабиринт, 2000. 336 с.
2. *Пропп В.Я.* Русская сказка (Собрание трудов В.Я. Проппа) / Научная редакция, комментарии Ю.С. Рассказова. М.: Лабиринт, 2000. 416 с.
3. *Пропп В.Я.* Морфология волшебной сказки / Научная редакция, текстологический комментарий И.В. Пешкова. М.: Лабиринт, 2001. 192 с.
4. Былины. Русские народные сказки. Древнерусские повести / Авторы вступит. статей,

- сост. и коммент. В.П. Аникин, Д.С. Лихачев и Т.Н. Михельсон; Переизд. М.: Дет. лит., 1989. 638 с.
5. Народные русские сказки А.Н. Афанасьева / Подготовка текста, предисловие и примечания В.Я. Проппа. В 3 т. М.: Государственное издательство художественной литературы (Гослитиздат), 1957. Т. 1. 516 с.
 6. Народные русские сказки А.Н. Афанасьева / Подготовка текста, предисловие и примечания В.Я. Проппа. В 3 т. М.: Государственное издательство художественной литературы (Гослитиздат), 1957. Т. 2. 511 с.
 7. Народные русские сказки А.Н. Афанасьева / Подготовка текста, предисловие и примечания В.Я. Проппа. В 3 т. М.: Государственное издательство художественной литературы (Гослитиздат), 1957. Т. 3. 573 с.
 8. Fillmore Ch.J., Kay P. Construction grammar course book. Berkeley: University of California, 1992. 113 p.
 9. Goldberg A.E. Constructions: A Construction Grammar Approach to Argument Structure. Chicago: University of Chicago Press, 1995. 265 p.
 10. Goldberg A.E. Constructions at work: The nature of generalization in language. Oxford: Oxford Univ. Press, 2006. 280 p.
 11. Рахилина Е.В. Лингвистика конструкций / Отв. ред. Е.В. Рахилина. М.: Издательский центр “Азбуковник”, 2010. 584 с.
 12. Рахилина Е.В., Плунгян В.А. Конструкция анекдота с точки зрения грамматики конструкций // Известия РАН. Серия литературы и языка. 2009. Т. 68. № 5. С. 47–54.
 13. Ляшевская О.Н. Корпусные инструменты в грамматических исследованиях русского языка. М.: Издательский дом ЯСК, Рукописные памятники Древней Руси, 2016. 520 с.
 14. Братья Гримм. Собрание сказок / Пер. под ред. П. Полевого. В 2 т. М.: Литература, Престиж Бук, 2007. Т. 2: Сказки; Детские легенды. 464 с.
 15. Шмурак Р.И. К уточнению понятия упрека с помощью корпусных инструментов // Известия Российской академии наук. Серия литературы и языка. 2020. Т. 79. № 3. С. 24–48.
 16. Шмурак Р.И. Формальные средства выражения упрека в русском дискурсе // Слово.ру: Балтийский акцент. 2020. Т. 11. № 4. С. 65–79.
 17. Шмурак Р.И. Формальные средства выражения упрека в русском языке: к процедуре корпусного поиска (на материале конструкции “сам + небось + V”) // Вестник Московского университета. Серия 9. Филология. 2021. № 2. С. 83–92.
 18. НКРЯ [Национальный корпус русского языка: [сайт]]. URL: <http://www.ruscorpora.ru>
 19. Братья Гримм. Собрание сказок / Пер. под ред. П. Полевого; Вступ. ст. В. Неустроева. В 2 т. М.: Литература, Престиж Бук, 2007. Т. 1: Сказки. 480 с.
 20. Русские народные сказки и былины. М.: ТЕРРА—Книжный клуб, 1999. 368 с.
 21. Былины; Исторические песни; Баллады / сост. А. Калугиной, В. Ковпика. М.: Эксмо, 2008. 797 с.
 22. Шмурак Р.И. Формальные средства выражения упрека в русском языке (корпусное исследование): диссертация на соискание ученой степени кандидата филологических наук: 10.02.01 / Шмурак Роман Ильич. М., 2021. 241 с.

REFERENCES

1. Propp, V.Ya. *Istoricheskie korni volshebnoj skazki. Nauchnaya redakciya, tekstologicheskij kommentarij I.V. Peshkova* [The Historical Roots of a Fairy Tale. Scientific Editorial, Textual Commentary by I.V. Peshkov]. Moscow, Labirint Publ., 2000. 336 p. (In Russ.)
2. Propp, V.Ya. *Russkaya skazka (Sobranie trudov V.Ya. Proppa). Nauchnaya redakciya, kommentarii Yu.S. Rasskazova* [Russian Fairy Tale (Collected Works of V.Ya. Propp). Scientific Editorial, Textual Commentary by Yu.S. Rasskazov]. Moscow, Labirint Publ., 2000. 416 p. (In Russ.)
3. Propp, V.Ya. *Morfologiya volshebnoj skazki. Nauchnaya redakciya, tekstologicheskij kommentarij I.V. Peshkova* [Morphology of a Fairy Tale. Scientific Editorial, Textual Commentary by I.V. Peshkov]. Moscow, Labirint Publ., 2001. 192 p. (In Russ.)
4. *Byliny. Russkie narodnye skazki. Drevnerusskie povesti. Avtory vstupit, statej, sost. i komment. V.P. Anikin, D.S. Lixachev i T.N. Mixelson; Pereizd.* [Epics. Russian Folk Tales. Old Russian Stories. Authors of Introductory Articles, Compilers and Commentators V.P. Anikin, D.S. Likhachev and T.N. Mikhelson; Reprint]. Moscow, Det. lit. Publ., 1989. 638 p. (In Russ.)
5. *Narodnye russkie skazki A.N. Afanasjeva. Podgotovka teksta, predislovie i primechaniya V.Ya. Proppa. V 3 t.* [Folk Russian Fairy Tales by A.N. Afanasyev. Preparation of the Text, Preface and Notes by V.Y. Propp. In 3 Vols.]. Moscow, Gosudarstvennoe Izdatelstvo Khudozhestvennoj literatury (Goslitizdat) Publ., 1957. Vol. 1. 516 p. (In Russ.)
6. *Narodnye russkie skazki A.N. Afanasjeva. Podgotovka teksta, predislovie i primechaniya V.Ya. Proppa. V 3 t.* [Folk Russian Fairy Tales by A.N. Afanasyev. Preparation of the Text, Preface and Notes by V.Y. Propp. In 3 Vols.]. Moscow, Gosudarstvennoe Izdatelstvo Khudozhestvennoj literatury (Goslitizdat) Publ., 1957. Vol. 2. 511 p. (In Russ.)

7. *Narodnye russkie skazki* A.N. Afanasjeva. *Podgotovka teksta, predislovie i primechaniya* V.Ya. Proppa. V 3 t. [Folk Russian Fairy Tales by A.N. Afanasyev. Preparation of the Text, Preface and Notes by V.Y. Propp. In 3 Vols.]. Moscow, Gosudarstvennoe Izdatelstvo Khudozhestvennoj literatury (Goslitizdat) Publ., 1957. Vol. 3. 573 p. (In Russ.)
8. Fillmore, Ch.J., Kay, P. Construction grammar course book. Berkeley: University of California, 1992. 113 p.
9. Goldberg, A.E. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press, 1995. 265 p.
10. Goldberg, A.E. *Constructions at work: The nature of generalization in language*. Oxford: Oxford Univ. Press, 2006. 280 p.
11. Rakhilina, E.V. *Lingvistika konstrukcij* [Linguistics of Constructions]. Moscow, Izdatelskij centr "Azbu-kovnik" Publ., 2010. 584 p. (In Russ.)
12. Rakhilina, E.V., Plungyan, V.A. *Konstrukciya anekdota s točki zreniya grammatiki konstrukcij* [The Construction of a Joke from the Point of View of the Grammar of Constructions]. *Izvestiâ Rossijskoj akademii nauk. Seriâ literatury i âzyka* [Bulletin of the Russian Academy of Sciences: Studies in Literature and Language]. 2009, Vol. 68, No. 5, pp. 47–54. (In Russ.)
13. Lyashevskaya, O.N. *Korpusnye instrumenty v grammaticheskikh issledovaniyakh russkogo yazyka* [Corpus Instruments in Grammatical Studies of the Russian Language]. Moscow, Izdatelskij dom YaSK, Rukopisnye pamyatnik Drevnej Rusi Publ., 2016. 520 p. (In Russ.)
14. *Bratya Grimm. Sobranie skazok. Per. ed. P. Polevoy* [Brothers Grimm. Collection of Fairy Tales. In 2 Vols. Translated by P. Polevoy]. Moscow, Literatura, Prestizh Buk Publ., 2007. T. 2: *Skazki; Detskie legendy* [Vol. 2: Fairy Tales; Children's Legends]. 464 p. (In Russ.)
15. Shmurak, R.I. *K utochneniyu ponyatiya upreka s pomosh'yu korpusnykh instrumentov* [Clarifying the Concept of Reproach Using Corpus Tools]. *Izvestiâ Rossijskoj akademii nauk. Seriâ literatury i âzyka* [Bulletin of the Russian Academy of Sciences: Studies in Literature and Language]. 2020, Vol. 79, No. 3, pp. 24–48. (In Russ.)
16. Shmurak, R.I. *Formalnye sredstva vyrazheniya upreka v russkom diskurse* [Formal Means of Reproach in Russian Discourse]. *Slovo.ru: Baltijskij accent* [Slovo.ru: Baltic Accent]. 2020, Vol. 11, No. 4, pp. 65–79. (In Russ.)
17. Shmurak, R.I. *Formalnye sredstva vyrazheniya upreka v russkom yazyke: k procedure korpusnogo poiska (na materiale konstrukcii "sam + nebos' + V")* [Formal Means of Reproach in Russian: to the Procedure of Corpus Search (Based on the Construction "sam + nebos' + v")]. *Vestnik Moskovskogo universiteta. Seriya 9. Filologiya* [Moscow University Bulletin. Series 9. Philology]. 2021, No. 2, pp. 83–92. (In Russ.)
18. *NKRYA* [Russian National Corpus: [website]]. URL: <http://www.ruscorpora.ru>
19. *Bratya Grimm. Sobranie skazok. Per. ed. P. Polevoy* [Brothers Grimm. Collection of Fairy Tales. In 2 Vols. Translated by P. Polevoy]. Moscow, Literatura, Prestizh Buk Publ., 2007. T. 1: *Skazki* [Vol. 1: Fairy Tales]. 480 p. (In Russ.)
20. *Russkie narodnye skazki i byliny* [Russian Folk Tales and Epics]. Moscow, TERRA–Knizhnyj klub Publ., 1999. 368 p. (In Russ.)
21. *Byliny; Istoricheskie pesni; Ballady. Sost. A. Kaluginoj, V. Kovpika* [Epics; Historical songs; Ballads. Compilers A. Kalugina, V. Kovpik]. Moscow, Eksmo Publ., 2008. 797 p. (In Russ.)
22. Shmurak, R.I. *Formalnye sredstva vyrazheniya upreka v russkom yazyke (korpusnoe issledovanie)* [Formal Means of Reproach in the Russian Language (Corpus Study)]. *Dissertaciya na soiskanie uchenoj stepeni kandidata filologicheskix nauk: 10.02.01. Shmurak Roman Iljich* [Dissertation for the Degree of Candidate of Philological Sciences: 10.02.01. Shmurak Roman Ilyich]. Moscow, 2021. 241 p. (In Russ.)

Дата поступления материала в редакцию: 28 апреля 2022 г.
 Статья поступила после рецензирования и доработки: 31 мая 2022 г.
 Статья принята к публикации: 29 июня 2022 г.
 Дата публикации: 31 августа 2022 г.

Received by Editor on April 28, 2022
 Revised on May 31, 2022
 Accepted on June 29, 2022
 Date of publication: August 31, 2022