***С. Ю. Пужаева, Н. А. Зевахина, С. С. Джакупова***

***S. Yu. Puzhaeva, N. A. Zevakhina, S. S. Dzhakupovа***

**КОНТАМИНАЦИЯ КОНСТРУКЦИЙ В РЕЧИ НЕСТАНДАРТНЫХ РУССКОГОВОРЯЩИХ НА МАТЕРИАЛE КОРПУСА РУССКИХ УЧЕБНЫХ ТЕКСТОВ[[1]](#footnote-1)**

**CONSTRUCTION BLENDING IN NON-STANDARD VARIANTS OF RUSSIAN IN THE CORPUS OF RUSSIAN STUDENT TEXTS**

**Аннотация**. В докладе рассматривается контаминация конструкций как наиболее весомая причина ошибок в конструкциях, совершаемых студентами в учебных текстах. Исследование выполнено в рамках Грамматики конструкций [Fillmore and Kay 1992; Goldberg 1995, 2006] и грамматики ошибок [Выренкова и др. 2014]. Материалом исследования служат примеры, взятые из Корпуса русских учебных текстов, содержащего помимо метатекстовой и морфологической разметки разметку по ошибкам.

**Ключевые слова:** корпус, учебный корпус, ошибки, разметка по ошибкам, конструкции, контаминация конструкций.

**Abstract.** The paper examines construction blending as an important cause of errors in written students’ texts. The study is conducted within the framework of Construction Grammar [Fillmore and Kay 1992; Goldberg 1995, 2006] and grammar of errors [Vyrenkova et al. 2014]. It is based on the data of the Corpus of Russian Student Texts supplied with metatextual, morphological and error annotation.

**Keywords.** corpus, learner corpus, errors, error annotation, constructions, construction blending.

**1. Введение**

На протяжении долгого времени основным подходом к изучению русского языка был традиционный, при котором грамматика рассматривается как канонический свод правил, в котором закреплены нормативные варианты тех или иных грамматических явлений. Однако в противовес сложившейся традиции существует принципиально иной подход – идея для каждой грамматической зоны описывать наиболее типичные отклонения от нормы, фиксировать не-норму, тем самым создавая грамматику ошибок [Frei 1929]. В связи с отсутствием обширного корпуса материала для подобных исследований возникла необходимость создания особой базы данных, что и привело к появлению учебных корпусов.

**2. Корпус русских учебных текстов**

***2.1. Состав и особенности***

Традиционно под учебным корпусом понимается коллекция текстов, написанных людьми, изучающими язык как иностранный (L2), тексты снабжены разного рода разметкой – метатекстовой, морфологической и разметкой по ошибкам.

Корпус русских учебных текстов (КРУТ)[[2]](#footnote-2) существует с 2013 г. как один из проектов Лингвистической лаборатории по корпусным технологиям на факультете филологии в НИУ ВШЭ под руководством Е.В. Рахилиной. КРУТ – открытый бесплатный интернет-ресурс, содержащий учебные тексты студентов НИУ ВШЭ и других вузов, носителей русского языка. КРУТ интересен тем, что содержит тексты носителей языка на родном языке (L1) вместо ожидаемых в учебных корпусах текстов на изучаемом языке (L2). Объем корпуса составляет более 2,5 млн словоупотреблений, часть корпуса – 500 тыс. словоупотреблений – размечена на предмет ошибок. Всего на данный момент выделено более 10 тыс. ошибок, разметка по ошибкам производится вручную.

***2.2. Разметка***

Тексты в корпусе снабжены метатекстовой, морфологической разметкой и разметкой по ошибкам.

Разметка по ошибкам производится вручную с помощью программы Crocodiles 2.6. Классификация ошибок включает тип ошибки (лексический, грамматический, дискурсивный) и указание на возможную причину ошибки (опечатка, контаминация конструкций). Классификация ошибок содержит 39 помет, 20 верхнего и 19 нижнего уровней[[3]](#footnote-3). Главной проблемой разметки по ошибкам является субъективность – зачастую одно и то же явление разными экспертами классифицируется по-разному, более того, в некоторых случаях возникают разногласия, выделять ли ошибку как таковую. Снизить эффект субъективности помогает многомерная классификация – возможность отнесения одной и той же ошибки к разным типам.

**3. Корпус русских учебных текстов как материал исследований**

***3.1. КРУТ: практическая значимость***

Создание учебных корпусов преследует как педагогические, так и исследовательские цели. С одной стороны, на базе таких корпусов могут создаваться лингвистические тренажеры, тестовые и контрольные задания (с учетом самых распространенных ошибок), а кроме того – индивидуальные учебные подкорпусы студентов. С другой стороны, ошибки служат маркерами языковых изменений.

Целью нашего исследования было выявить наиболее распространенные типы ошибок в конструкциях, в первую очередь помеченные тегами cause, contam и lex, phrase[[4]](#footnote-4). Исследование проведено в рамках двух направлений современной лингвистики – грамматики конструкций [Fillmore and Kay 1992; Goldberg 1995, 2006] и грамматики ошибок [Выренкова и др. 2014]. В связи с этим в рамках настоящего исследования под конструкциями понимаются «языковые единицы любого уровня, если они обладают формой и содержанием» [Рахилина 2010: 19].

***3.2 Ход и результаты исследования***

На момент обращения к корпусу нашлось 130 примеров на запрос cause, contam и 700 примеров на запрос lex, phrase. В ходе анализа материала выяснилось, что контаминация (тег cause, contam) не всегда означает контаминацию конструкций [Рахилина 2010] на синтаксическом уровне[[5]](#footnote-5). Нам также удалось выявить несколько наиболее характерных явлений, лежащих в основе выделенных ошибок.

Класс 1. Существуют две семантически схожие устойчивые конструкции с двумя слотами. Один из слотов первой конструкции заполняется элементом, которым обычно заполняется соответствующий слот второй конструкции, при этом нарушается лексическая сочетаемость.

1. *Особое* ***место*** *в европеизации Российской империи* ***сыграла*** *деятельность Петра И. В данный период вся Европа находилась под влиянием идей Просвещения, что способствовало демократизации<…>*

ИГРАТЬ РОЛЬ – ЗАНИМАТЬ МЕСТО

В примере (1) обе конструкции являются устойчивыми, то есть оба слота каждой из них имеют наиболее ожидаемый вариант заполнения – конкретную лексему. Но бывает и такое, что второй элемент конструкции задается преимущественно грамматическими характеристиками, так, в примере (2) основным ограничивающим фактором на заполнение второго слота конструкции служит категория одушевленности/неодушевленности:

1. *<…> по моему мнению, честность <…> должны быть ведущими принципами среди журналистов, колумнистов и всех тех, кто взял на себя* ***роль донесения*** *фактов до массового сознания.*

РОЛЬ КОГО-ЛИБО – ЗАДАЧА ЧЕГО-ЛИБО

В рамках первого класса можно выделить несколько подклассов:

a) Существует только одна устойчивая конструкция, один из слотов которой заполняется нерелевантным элементом.

1. *В современном мире* ***ходит множество мифов*** *по поводу некой дееспособности женщин.*

ХОДЯТ СЛУХИ – (существуют) мифы

b) Помимо явления контаминации важным оказывается наличие парономазии.

1. *Яркие разноцветные и, как правило, безвкусные вывески* ***зазывают посетить*** *ту или иную закусочную и магазин.*

ЗАЗЫВАТЬ КОГО-ТО (куда-то/ с целью) – ПРИЗЫВАТЬ (кого-то) СДЕЛАТЬ ЧТО-ТО

Класс 2. Один из элементов конструкций совпадает. По сути, происходит контаминация двух конструкций, в которые входит одна и та же единица.

1. *Не раз они же отмечали отсутствие национальной идеи в России, не позволяющее массово* ***бороться против*** *таких проблем как алкоголизм и наркомания*.

Бороться против кого-то – бороться с чем-то

Класс 3. Контаминация прямого и переносного значений конструкций: метафора и метонимия.

1. *Вы посетите самые известные, старинные достопримечательности,* ***оставите свой*** *след в романтических* ***городах Италии****, ну и конечно же насладитесь прохладой кристально-голубых морей и теплотой яркого солнца.*

ОСТАВИТЬ СЛЕД НА ЧЕМ-ТО – Оставить свой след В ИСТОРИИ/ памяти (метафора)

1. *Эта страна и ее нация очень* ***близки*** *к нам* ***территориально*** *и* ***исторически****, поэтому люди здесь очень похожи на нас своей простой широкой душой и чрезмерной эмоциональностью.*

БЛИЗКИ ТЕРРИТОРИАЛЬНО – БЛИЗКИ ИСТОРИЧЕСКИ (метонимия)

Класс 4. Замена компонента устойчивой конструкции нерелевантным элементом.

1. *Вы привыкните к этому великолепному городу, как только* ***переступите порог аэропорта****.*

Переступить порог дома

Приведенная классификация является полной и носит обобщающий характер. Новые примеры, появляющиеся в Корпусе, можно распределить по выделенным случаям. Наибольшее число примеров попадает в класс 1 (51% проанализированных контекстов) и класс 2 (25%), наименее продуктивными оказываются случай 3 (9%) и 4 (15%).

**4. Заключение**

На основе материала из Корпуса русских учебных текстов нами была построена классификация видов ошибок в устойчивых выражениях. Три класса ошибок из полученных четырех можно описать при помощи контаминации конструкций: в первом случае имеет место контаминация близких по значению конструкций; во втором – контаминация конструкций, содержащих общую единицу; в третьем – контаминация прямого и переносного значений конструкции. Учитывая то, что в наших данных первый класс попадает наибольшее количество примеров, а в четвертый, напротив, наименьшее, можно утверждать, что контаминация конструкций действительно позволяет описать бо´льшую часть отклонений в употреблениях устойчивых выражений.

**Литература**

1. *Выренкова А.С., Полинская М. С., Рахилина Е.В.* (2014), Грамматика ошибок и грамматика конструкций: «эритажный» («унаследованный») русский язык. Вопросы языкознания, 3, с. 3-19.
2. *Рахилина Е. В.(ред)* (2010), Лингвистика конструкций. М.
3. *Fillmore, Ch. and Kay, P.* (1992), Construction Grammar Course Book. Berkeley.
4. *Frei*. *Henri* (1929), La grammaire des fautes. Paris.
5. *Goldberg, A.* (1995),Constructions: A Construction Grammar Approach to Argument Structure. Chicago.
6. *Goldberg, A.* (2006),Constructions at Work: The Nature of Generalization in Grammar. Oxford.

**References**

1. *Vyrenkova A. S., Polinskaya M. S., Rakhilina E. V.* (2014),

Grammatika oshibok i grammatika konstruktsiy: «eritazhnyiy» («unasledovannyiy») russkiy yazyik [Grammar of Errors and Construction Grammar: Heritage Russian] Voprosyi yazyikoznaniya, 3, pp. 3-19.

1. *Rakhilina E. V. (ed.)* (2010), Lingvistika konstruktsiy [Construction Linguistics]. Moscow.
2. *Fillmore, Ch. and Kay, P.* (1992). Construction Grammar Course Book. Berkeley.
3. *Frei*. *Henri* (1929). La grammaire des fautes. [Grammar of errors] Paris.
4. *Goldberg, A. (1995).* Constructions: A Construction Grammar Approach to Argument Structure. Chicago.
5. *Goldberg, A. (2006).* Constructions at Work: The Nature of Generalization in Grammar. Oxford.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Пужаева Светлана Юрьевна**

**Зевахина Наталья Александровна**

**Джакупова Светлана Сатыбалдиевна**

Национальный исследовательский университет Высшая школа экономики (Россия)

**Puzhaeva Svetlana**, *E-mail*: *sevtell@mail.ru*

**Zevakhina Natalia**, *E-mail:* *natalia.zevakhina@gmail.com*

**Dzhakupovа Svetlana**, *E-mail*: *svetlanads@yandex.ru*

National Research University Higher School of Economics (Russia)

1. Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ в 2015 году. [↑](#footnote-ref-1)
2. Корпус доступен по адресу <http://web-corpora.net/CoRST/search/> [↑](#footnote-ref-2)
3. Пометы верхних уровней, такие как lex, discourse указывают на то, к какому разделу языка относится данная ошибка, пометы нижних уровней носят конкретизирующий характер (например, lex, word – ошибка в употреблении конкретного слова, lex, phrase – ошибка в лексической сочетаемости). [↑](#footnote-ref-3)
4. Следует заметить, что в существующий тег constr попадают те случаи, когда ошибка затрагивает всю клаузу, если ошибка затрагивает элементы внутри клаузы, то чаще всего ей приписывается тег lex, phrase. [↑](#footnote-ref-4)
5. Контексты, выданные корпусом на обозначенные выше запросы, были просмотрены на предмет соответствия тегу. [↑](#footnote-ref-5)