

**Иванова Елизавета Васильевна**  
доктор филологических наук, профессор  
Санкт-Петербургский государственный  
университет  
г. Санкт-Петербург, Россия

**Ivanova Elizaveta**  
Doctor of Philology, Professor  
St.Petersburg State University  
St.Petersburg, Russia  
e.ivanova@spbu.ru

## О КОРПУСНОМ ПОДХОДЕ К ФРАЗЕОЛОГИИ

Корпусная лингвистика открывает большие возможности для исследования языковых единиц в их реальном функционировании в текстах. Эти возможности самым непосредственным образом касаются и изучения фразеологии. В статье рассматриваются

некоторые сложности и результаты анализа фразеологических единиц американского варианта английского языка на фоне корпусных данных. На материале примеров демонстрируется необходимость ручной обработки корпусных данных при омонимичности фразеологических единиц и свободных словосочетаний. Устанавливается продуктивность исследования частотности вариативных притяжательных местоимений в составе фразеологических единиц, и очерчиваются некоторые тенденции этой вариативности. Производится сравнение количественных данных о вхождении ряда фразеологических единиц в корпус, делается вывод о важности количественной характеристики для исследований фразеологических единиц в теоретических и практических целях. Обосновывается целесообразность включения количественной характеристики в лексикографические источники. В заключение определяются некоторые направления в изучении фразеологических единиц на основе корпуса.

*К л ю ч е в ы е с л о в а: корпусный подход; фразеологическая единица; идиоматичность; частотность; n-грамма.*

## ON THE CORPUS APPROACH TO PHRASEOLOGY

Corpus linguistics opens great possibilities for the studies of language units in their real functioning in texts. These possibilities directly concern phraseological studies. The article considers some difficulties and results of the analysis of American English phraseological units against the background of corpus data. The necessity of manual processing of corpus data involving phraseological units homonymous with flexible combinations of words is demonstrated, based on a number of corpus entries. The efficiency of studying the frequency of variable possessive pronouns in the structure of phraseological units is stated, and some tendencies in the variability of this frequency are delineated. The comparison of quantitative data regarding corpus entries of several phraseological units is made, and a conclusion about the importance of quantitative characteristics for theoretical and practical phraseological studies is reached. The advisability of including quantitative characteristics into lexicographical resources is stated. In conclusion some prospects of phraseological studies based on corpus data are outlined.

*Key words: corpus approach; phraseological unit; idiomaticity, frequency, n-gramma.*

Корпусная лингвистика весьма интенсивно и широко применяется в настоящий момент во многих областях языкознания, позволяя получить неоценимые данные о функционировании языковых знаков в речи. Корпусный подход не обошел стороной и фразеологию. Данные корпуса позволяют внести коррективы в актуальное значение фразеологических единиц [1], выявить их модификации в речи, уточнить свойственное им окружение [2]. Обращение к корпусу также дает возможность установить наиболее частотные фразеологические единицы в целях оптимального отбора учебного материала при преподавании иностранного языка [3], а также значительно улучшить качество перевода [4], решая тем самым задачи прикладного характера.

Цель статьи заключается в том, чтобы на материале английских фразеологических единиц с компонентом *heart* 'сердце' рассмотреть некоторые преимущества, сложности и перспективы использования корпусного подхода при исследовании фразеологии.

Определение частотности фразеологических единиц производится на основе данных Корпуса современного американского варианта английского языка [5], который считается наиболее представительным и обширным англоязычным корпусом [6, p.17].

Фразеологический материал собран по Оксфордскому словарю идиом [7], при этом принципом отбора является учет только тех фразеологических единиц, которые входят в пласт английского языка, используемый во всех его ареальных вариантах, и в пласт американского варианта английского языка. Отметим, что среди фразеологических единиц с компонентом *heart* 'сердце' присутствуют только первые. Общее количество фразеологических единиц (ФЕ) составляет 17.

Слово *heart* 'сердце' относится к соматизмам, широко употребительным лексическим единицам в силу неизменного интереса человека к самому себе и к своему телу. По этой же причине данные слова оказываются задействованными в образовании внушительного количества ФЕ.

Что касается соматизма *heart* 'сердце', то его частотность в корпусе составляет 193144 вхождения и не является такой высокой, как, например, частотность *head* 'голова' – 340401 вхождение. Тем не менее это значительно выше, чем, например, у соматизма *ears/ear* 'уши, ухо' (60192 вхождения) и свидетельствует о достаточной востребованности слова и значимости для социума данного органа.

В Оксфордском словаре идиом все ФЕ снабжены иллюстративными примерами их употребления в текстах, но указания на их частотность отсутствуют, и можно предположить, что лексикографы в значительной степени полагались на свою интуицию при отборе материала для включения в словарь.

Отличительной особенностью всех 17 ФЕ с компонентом *heart* 'сердце' является отсутствие у них омонимичности свободным словосочетаниям. Например:

*take something to heart* 'букв. взять что-то к сердцу.', «принять близко к сердцу, относиться серьезно к чему-л».

По-иному обстоит дело со многими ФЕ с компонентом *eye* 'глаз', которые необходимо обрабатывать вручную, отделяя идиоматичные по своей природе ФЕ от омонимичных им свободных словосочетаний.

*When she awakens she does not open her eyes right away.* 'Когда она просыпается, она не сразу открывает глаза.'

*... he could be the love of her life if she would just open her eyes to the possibilities.* '... он мог бы быть любовью всей ее жизни, если бы она обратила внимание на такую возможность'.

В первом предложении *n*-грамма *open her eyes* 'открыть глаза' представляет собой свободное сочетание слов. Во втором предложении та же *n*-грамма является ФЕ со значением «принимать правду, не игнорировать реальность».

Данное обстоятельство приводит к необходимости сортировать весь собранный по корпусу материал. Для рассматриваемых здесь ФЕ с компонентом *heart* 'сердце' такой проблемы нет.

Рассмотрим вхождения ряда ФЕ в корпус.

ФЕ *wear your heart on your sleeve* 'букв. носить сердце на рукаве', «проявлять свои чувства в открытую» имеет 49 вхождений, из них преобладает употребление с притяжательными местоимениями *my* 'мой' (19) и *your* 'твой, ваш' (14).

*I am more like Andrew, I wear my heart on my sleeve.*

'Я больше похож на Эндрю, я не скрываю свои чувства'

*It's easy to wear your heart on your sleeve when you're not looking in his eyes.*

'Легко проявлять свои чувства, когда не смотришь в его глаза'

Примеры с местоимением *our* 'наш' с данной ФЕ не обнаружены, с местоимением *their* 'их' всего 3. Местоимения *his* 'его' и *her* 'ее' имеют низкое число вхождений: 7 и 6 соответственно.

*Zim doesn't really wear his heart on his sleeve.*

'Зим на самом деле не показывает своих чувств.'

*Linda doesn't wear her heart on her sleeve.*

'Линда не проявляет свои чувства.'

Частотность использования вариативных местоимений в составе ФЕ представляется достаточно показательным дополнительным фактором для характеристики этих ФЕ. В конкретном случае можно заключить, что анализируемая ФЕ преимущественно употребляется с местоимениями первого и второго лица единственного числа.

Более равномерное варьирование вхождений притяжательных местоимений обнаруживается при корпусном анализе ФЕ *in your heart of hearts* 'букв. в сердце твоего сердца', «в глубине души»: *my* 'мой' (131), *your* 'твой, ваш' (101), *their* 'их' (76), *his* 'его' (68), *her* 'ее' (36), *our* 'наш' (21).

*In my heart of hearts I just don't want to believe that he did it.*

'В глубине души я просто не хочу верить, что он сделал это.'

*In your heart of hearts, you absolutely don't believe it.*

'В глубине души ты в это абсолютно не веришь.'

*But in their heart of hearts, these boys wish these girls would wait.*

'Но в глубине души эти мальчики хотят, чтобы девочки подождали.'

*In his heart of hearts I think he always saw himself as some kind of educational reformer.*

'В глубине души, я думаю, он всегда смотрел на себя как на своего рода реформатора образования.'

*But in her heart of hearts she knew that it did not really matter to them.*

'В глубине души она знала, что для них это действительно неважно.'

*A lot of us know in our heart of hearts there's something fishy going on.*

'Многие из нас знают в глубине души, что происходит что-то подозрительное.'

Как видим, на первом месте вновь представлены местоимения *my* 'мой' (131) и *your* 'твой, ваш' (101), местоимение *our* 'наш', хотя и находится на последнем месте по количеству вхождений (21), но все-таки используется.

Более значительный контраст по сочетаемости с притяжательными местоимениями обнаруживает ФЕ *from the bottom of your heart* 'букв. из глубины своего сердца', «от всей души; в глубине души»: *my* 'мой' (487), *his* 'его' (22), *your* 'твой, ваш' (12), *her* 'ее' (6), *their* 'их' (2), *our* 'наш' (2). Частотность местоимения первого лица единственного числа в составе данной ФЕ в разы превышает частотность всех остальных местоимений.

*I hope from the bottom of my heart, that he comes to the same conclusions you did.*

'Я от всей души надеюсь, что он придет к тем же выводам, что и ты.'

Можно предположить, что отмеченная в вышерассмотренных ФЕ тенденция преимущественного использования местоимений первого лица единственного числа и второго лица единственного/множественного числа является показательной характеристикой, иллюстрирующей особенности функционирования этих ФЕ в тексте. Дополнительное изучение контекстов использования данной ФЕ может помочь найти объяснение выявленной тенденции.

Рассмотрим ФЕ без вариативных местоимений: *heart of gold* 'букв. сердце из золота', «о добром, щедром человеке» и *heart of stone* 'букв. сердце из камня', «о суровом или жестоком человеке». Обе ФЕ строятся на одинаковом образе сердца, сделанного из какого-то материала. Число вхождений первой приблизительно в четыре раза превышает число вхождений второй (442 против 93).

*He has a heart of gold, but he is a little blinkered and he never learns.*

'У него золотое сердце, но он немного ограниченный и не учится ни на чем.'

*She's got a heart of stone. Nothing ever worked with her, and I stopped trying a long time ago.*

'У нее каменное сердце. С этим ничего не сделать, и я перестал пытаться давным-давно.'

В целом рассматриваемые ФЕ обладают разной частотностью. Так, ФЕ *wear your heart on your sleeve* 'букв. носить сердце на рукаве', «проявлять свои чувства в открытую» имеет, как уже отмечалось, 49 вхождений, *in your heart of hearts* 'букв. в сердце твоего сердца', «в глубине души» в сочетании со всеми притяжательными местоимениями насчитывает 433 вхождения, *from the bottom of your heart* 'букв. из глубины своего сердца', «от всей души; в глубине души» – 531 вхождение. А ФЕ *have your heart in your mouth* 'букв. иметь свое сердце во рту', «быть испуганным» не отмечено вовсе с различными притяжательными местоимениями. Причиной этому может быть как реальная неупотребительность ФЕ, так и проблемы в программе при поиске протяженных n-грамм.

Количество вхождений ФЕ в корпус является ее важной характеристикой как в плане теоретических изысканий, так и в отношении практических исследований, направленных на поиск наиболее подходящего для использования в преподавании фразеологического материала. Представляется, что эту характеристику было бы целесообразно включать в лексикографические источники, дополняя тем самым словарные статьи.

В настоящей статье были рассмотрены следующие параметры и результаты корпусного анализа фразеологии:

- Проблема омонимичности фразеологических единиц свободным словосочетаниям и необходимость прибегать в результате к ручной обработке материала. Эта проблема касается не всех корпусов, но имеет непосредственное отношение к корпусу, задействованному для анализа в данной статье.

- Функционирование вариативных притяжательных местоимений в составе фразеологических единиц. Возможность обнаружить тенденции в использовании притяжательных местоимений для некоторых фразеологических единиц.

- Частотность фразеологических единиц в корпусе. Важность количественной характеристики фразеологических единиц для теоретических и практических исследований.

В качестве одной из перспектив исследований фразеологии видится анализ функций фразеологических единиц в конкордансе корпуса, что поможет выявить некоторые закономерности между частотностью и выполняемыми функциями, а также создание программ, обнаруживающих идиоматичность, что, безусловно, не является задачей, непосредственно возлагаемой на лингвистов.

## ЛИТЕРАТУРА

1. Баранов А. Н., Добровольский Д. О. Аспекты теории фразеологии. М.: Знак, 2008. 656 с.
2. Добровольский Д. О. Корпусы текстов и двуязычная фразеография. Вестник Новосибирского государственного педагогического университета. 2015. № 5 (27). С. 23–37.
3. Комарова И. А., Коган М. С. Исследование английской фразеологии с помощью подходов корпусной лингвистики Компьютерная лингвистика и вычислительные онтологии. 2019. Вып. 3. С.40–49.
4. Пивоварова Е. В. Корпусный анализ как инструмент для выявления семантических изменений у фразеологизмов – «ложных друзей» переводчика. Филологические науки. Вопросы теории и практики. 2020. Том 13. Выпуск 9. С. 312–320.
5. Corpus of Contemporary American English – <https://www.english-corpora.org/coca/> Дата последнего посещения: 10.04.2024.
6. Lindquist H. Corpus linguistics and the description of English. Edinburgh: Edinburgh University Press, 2009. 219 p.
7. Oxford Dictionary of Idioms. 2nd Siefiring J. (ed.). Oxford: Oxford University Press, 2004. 340 p.