

DOI: 10.12731/2077-1770-2021-13-3-263-273
УДК 81-22

СПЕЦИФИКА ИСПОЛЬЗОВАНИЯ ЛИНГВИСТИЧЕСКОГО КОРПУСА ДЛЯ ИЗУЧЕНИЯ ПОЛИСЕМИИ СЛОВА

Авдюшкина А.А., Прохорова О.Н.

Цель. Статья посвящена выявлению преимуществ использования лингвистических корпусов при изучении полисемантических лексических единиц и сфер их функционирования; рассматриваются варианты перевода данных слов в соответствии с актуальными тенденциями развития современного общества.

Метод или методология проведения работы. Использованы методы сравнительного и структурного анализа и обобщения.

Результаты. В результате проведенного исследования были выявлены преимущества использования Британского национального корпуса и Корпуса современного американского английского языка в лингвистических исследованиях и проанализированы результаты применения данных корпусов при изучении многозначных частотных слов в современном английском языке.

Область применения результатов. Результаты исследования могут быть использованы в курсах по языкознанию, теории и практике перевода. Результаты проведенного исследования могут также послужить материалом для более глубокого и детального изучения использования лингвистических корпусов при изучении полисемантических лексических единиц.

Ключевые слова: корпусная лингвистика; Британский национальный корпус; Корпус современного американского английского языка; конкорданс; семантическое *ed* значение; лексема

SPECIFICITY OF USING THE LINGUISTIC CORPUS TO STUDY WORD POLYSEMY

Avdyushkina A.A., Prokhorova O.N.

Purpose. The article is devoted to identifying the advantages of using linguistic corpus in the study of polysemantic lexical units and

the spheres of their functioning. Options for translating these words are considered in accordance with the current trends in the development of modern society.

The method or methodology of the work performed. *The methods of comparative and structural analysis and generalization were used.*

Results. *As a result of the study the advantages of using the British National Corpus and the Corpus of Modern American English in linguistic research were revealed and the results of using these corpuses in the study of polysemantic frequency words in modern English were analyzed.*

Practical implications. *The results of the research can be used in courses on linguistics, theory and practice of translation. Moreover, the results of behavioral research can serve as material for a deeper and more detailed study of the usage of linguistic corpus in the study of polysemantic lexical units.*

Keywords: *linguistic corpus; British National Corpus; Corpus of Modern American English; concordance; semantic meaning; lexeme*

В настоящее время с помощью корпусов текстов можно с достаточно большой точностью анализировать многочисленные факты реализации и функционирования языковых единиц. Теоретически, английское слово “*corpus*” можно расшифровать как “*capable of representing potentially unlimited selections of texts*”, что означает «способный представлять потенциально неограниченный выбор текстов» [3, p. 35].

Корпусная лингвистика – многоаспектная область. Она представляет собой, с одной стороны, инструмент, с помощью которого можно оперативно и наглядно изучить особенности функционирования лексической единицы в определенной языковой среде. С другой стороны, используя корпус, можно определить динамический аспект лексического значения лексемы, анализируя различные контексты и установить новые значения единицы, которые еще не зафиксированы словарем.

Актуальность данного исследования обусловлена, с одной стороны, необходимостью изучения лексической многозначности слов в рамках одного из ведущих направлений языкознания – корпусной

лингвистики и, с другой стороны, возрастающей ролью лингвистического корпуса в качестве информационного ресурса в поиске контекстного употребления слов.

В современной лингвистической теории корпусная лингвистика представляет собой подход, направленный на изучение функционального аспекта языка и его свойств путем анализа больших наборов текстовых образцов. Уникальность корпусной лингвистики заключается в использовании современных компьютерных технологий в сборе языковых данных и методов, используемых в информационном поиске [1]. Именно этот ракурс корпусных исследований составляет их новизну.

Отечественный корпусный лингвист В.В. Рыков выделяет основные отличия корпусной лингвистики от традиционной, заключающиеся в изучении речи, языка в динамике; в опоре на данные корпуса текста; в особой форме корпусных текстов, в работе с лингвистическими данными в том виде, в каком они встречаются в контексте и т.д. [2].

В корпусах разрешение лексической многозначности достаточно оперативно осуществляется с опорой на сочетаемость компонентов. Большие объемы современных корпусов позволяют выполнять статистически значимые наблюдения о совместной встречаемости слов в разных значениях. Наилучшие результаты дают аннотированные корпуса, в которых у слов размечены значения. Примером таких корпусов является Британский национальный корпус (British National Corpus) и Корпус современного американского английского языка (Corpus of Contemporary American English).

В ходе данного исследования было выявлено несколько преимуществ использования корпуса текстов в лингвистических исследованиях:

- 1) удобство и простота его использования, которые достигаются за счет структурированности корпуса;
- 2) большой объем языкового материала, который обеспечивает реальные статистические данные, на основе которых можно вывести новые закономерности в языке или проверить ту или иную гипотезу.

3) корпус имеет дело с языковыми данными (словоупотреблениями) в их реальном контексте;

4) язык описывается в том виде, как он проявляет себя в речи, представленной в специально подобранном корпусе текстов;

5) языковые данные большинства лингвистических корпусов постоянно обновляются и расширяются, что позволяет оценить современное состояние языка, проследить изменение частот и контекстов в различные периоды времени и изучить динамику процессов изменения лексического состава языка.

В нашем исследовании мы опирались на идею возможности изучения полисемантических лексических единиц, сферы их функционирования и особенности сочетаемости с использованием лингвистических корпусов. В настоящей работе проанализированы несколько многозначные лексические единицы с помощью языковых корпусов British National Corpus и Corpus of Contemporary American English.

Обратимся к анализу наиболее репрезентативных примеров.

1. Так, прилагательное *sustainable* характеризуется абстрактной семантикой и смысловой неоднозначностью. Данная лексема обладает высокой частотой употребления в современных англоязычных текстах. По данным Оксфордского этимологического словаря, прилагательное *sustainable*, образованное от глагола *sustain*, имеет французское происхождение и первоначально имело значение «устойчивый, постоянный (*able to continue or last for a long time*)» [8].

Составление конкорданса по данной лексической единице позволило определить особенности сочетаемости прилагательного *sustainable* и выявить основные сферы его употребления, которые включают в себя следующие области:

- экология: “...*an industrial park that stands as Europe’s first ecology-based sustainable environment...*”;
- деятельность человека: “...*a major shift in agricultural policy from over-intensive agriculture to sustainable farming practice*”;
- человеческие сообщества: “*Certainly we cannot provide a fully one hundred percent sustainable community*”;

- потребление природных ресурсов: “*An energy source which could potentially supply up to ten times our energy needs for several thousand years is called an indefinitely sustainable energy source*”;
- образ жизни: “*We have been active with others in emphasizing the vital importance of developing a more sustainable lifestyle in the run-up to the Earth Summit*” [4].

Проанализировав с помощью Британского национального корпуса данные по контекстному употреблению прилагательного *sustainable*, можно выделить две наиболее частотные семы, используя термины компонентного анализа со значением «экологически безопасный» и «жизнеспособный». Наиболее рекуррентным является значение «экологически безопасный». Это дает право полагать, что этот феномен, прежде всего, связан с возрастанием значения экологических проблем для общества, вызванных антропогенными факторами не только на уровне стран, но и в масштабе целой планеты. И именно это значение является наиболее частотным и востребованным в языковом сообществе.

На примере перевода различных словосочетаний показана вариативность значения лексемы *sustainable* и конкретизация ее значения, которое определяется с помощью окружения данных слов, т.е. контекста.

С помощью анализа примеров употребления существительного *stub* в Британском национальном корпусе словосочетание *a stub of pencil* можно перевести как «огрызок карандаша» (“*The jaws of the peg were held apart by a stub of pencil*”), *a stub of candle* – «огарок свечи» (“... *he said, handing me a saucer with a stub of candle on it*”) и так далее [Там же]. Все значения многозначного слова связаны со значением гиперсемы и группируются вокруг нее. В данном случае вокруг гиперсемы “*a short piece of smth.; the small part of smth.*” [7].

Лексема *brown* в словаре Oxford Dictionary of English определяется следующим образом: “*having the colour of earth or coffee; having skin that is naturally brown or has been made brown by the sun*”. Из дефиниции видно, что семантическая структура слова представ-

лена двумя ЛСВ, а гиперсема детерминирована цветом земли, кофе, кожи и т.д. Понятно, что в зависимости от того, что является денотатом повествования, переводчик выбирает соответствующий перевод лексемы: если мы говорим о цвете кожи, то выбираем значение «смуглый»; «загорелый» (“... *with a pleasant, open face, brown skin, and lovely, expertly made-up gray-green eyes...*”); о глазах – «карий» (“*She had his sharper jaw and dark brown eyes flecked with gold*”); о виде человека – «хмурый, суровый» (“*He’s rather brown today*”) [6]. Примеры, представленные в Корпусе современного американского английского языка, помогают нагляднее показать и осмыслить вариативность значения лексемы *brown* в зависимости от ее окружения.

Анализ другой лексемы *roundabout* показал, что она имеет следующие значения:

- участок дороги с круговым движением: “*Realising he was having a bad day he decided not to waste the parking ticket so went round the roundabout and parked on the same side of the road*”;
- карусель: “... *garden pools including one for swimming, a roundabout climbing frame, several bridges spanning gardens*”;
- кресло с круглой спинкой: “*It took me six hours in a roundabout with a mini bottle of overpriced tequila, but it’s gracious and not braggy*”.

В словарной статье представлено несколько ЛСВ, объединенных значением гиперсемы «имеющий очертание круга» (“*forming a circle, round*” [7]). Анализируя примеры из Корпуса современного американского английского с данной лексемой, можно отметить, что лексема *roundabout* может иметь и переносное значение в сочетании со словом *way* – «обходной путь, окольный способ» (“*A euphemism is a roundabout way of saying something because you are reluctant to use a word which may seem gross or vulgar or dangerous*”) [6]. Здесь прослеживается, хотя и в импликации, значение «непрямой, непростой, неясный» – “*if you do or say something in a roundabout way, you do not do or say it in a simple, clear, and direct way*” [5].

Что касается лексемы “*loose*”, то ее дефиниции выглядят следующим образом:

1. *not securely fixed where it should be; able to become separated from something;*
2. *not tied together; not held in position by anything or contained in anything;*
3. *free to move around without control; not tied up or shut in somewhere;*
4. *not fitting closely;*
5. *not tightly packed together; not solid or hard;*
6. *not strictly organized or managed;*
7. *not exact; not very careful;*
8. *having or involving an attitude to sexual relationships that people consider to be morally wrong;*
9. *not in any player's control;*
10. *having too much liquid in it [7].*

Исходя из указанных в словарной статье дефиниций, можно вывести, что гиперсемой является сема со значением «находящийся в свободном, естественном состоянии, неподдающийся контролю». В Британском национальном корпусе можно встретить такие примеры, отражающие ее семантическую структуру:

- “*She wears her hair loose*”. – Она ходит с распущенными волосами;
- “*You cannot go out with a loose button like this.*” – Ты не можешь выйти с такой «болтающейся» пуговицей;
- “*The cloak caught on a loose nail*” – Плащ зацепился за незакрепленный гвоздь;
- “*His loose tooth went flying out of his mouth across the sidewalk*” – Его шатающийся зуб вылетел изо рта через тротуар;
- “*There were wreaths and sprays, pots of lilies, loose flowers strewn across the floor*” – На полу были разбросаны венки и цветы, горшки с лилиями, распустившиеся цветы;
- “*At the top of one of the hills Marcelo Leppe contemplates the slope of loose soil at his feet*” – На вершине одного из холмов Марсело Леппе созерцает склон рыхлой почвы у своих ног;

- “*Roche, who would give me a few coins to thank me for my loose tongue*” – Роше, который дал мне несколько монет, чтобы поблагодарить меня за болтливый язык;
- “*This was an Alabama 1st sergeant’s loose translation of a Pentagon deployment order, our first call to war*” – Это был вольный перевод 1-го сержанта Алабамы приказа Пентагона о разворачивании, нашего первого призыва к войне и другие [6].

Анализ полисемантического слова осложняется при наличии в нем нескольких номинативных значений с большой степенью семантической самостоятельности. Говоря о лексеме *state*, следует отметить, что для нее характерна следующая семантическая структура:

1. *a country considered as an organized political community controlled by one government;*
2. *an organized political community forming part of a country;*
3. *the government of a country;*
4. *the mental, emotional or physical condition that a person or thing is in;*
5. *the formal ceremonies connected with high levels of government or with kings and queens* [7].

Как видно из дефиниций, гиперсема этой лексемы передает значение «организованный, упорядоченный разным способом: законодательно, эмоционально, физически и т.д.», которая позволяет выделить в семантической структуре существительного *state* два основных ЛСВ – *state 1* («штат, страна») (“*It would transform the image of State education*”) и *state 2* («состояние») (“... *closely followed by HAMLET in a hysterical state...*”) [4].

Несомненно, корпус как один из инструментов изучения языковых феноменов является важным и полезным средством. Но ни одно языковое явление не может быть всесторонне проанализировано без применения традиционных методов исследования. Корпус – это своего рода информационный контейнер, позволяющий оперативно, наглядно проанализировать большие пласты языкового материала, которые отражают современное состояние языка.

Подводя итог, можно отметить, что в результате проведенного исследования выявлены преимущества использования корпуса текстов

в лингвистических исследованиях и проанализированы результаты применения корпуса при изучении многозначных частотных слов в современном английском языке. Представляется, что результаты данной статьи могут быть полезны для лингвистов, которые сталкиваются при переводе с проблемами лексической многозначности.

Список литературы

1. Захаров В.П., Богданова С.Ю. Корпусная лингвистика: учебник для студентов гуманитарных вузов. Иркутск: ИГЛУ, 2011. 161 с.
2. Рыков В.В. Корпус текстов как реализация объектно-ориентированной парадигмы // Труды Международного семинара Диалог-2002. М.: Наука, 2002. С. 59-61.
3. Dash N.S. Corpus Linguistics: A General Introduction CIIL, Mysore, 2010. 69 p.
4. British National Corpus [Электронный ресурс]. <https://corpus.byu.edu/bnc/> (дата обращения: 25.04.2021).
5. Collince Cobuild Advanced English Dictionary [Электронный ресурс]. <https://www.collinsdictionary.com/dictionary/english> (дата обращения: 07.08.2021).
6. Corpus of Contemporary American English [Электронный ресурс]. <https://corpus.byu.edu/coca/> (дата обращения: 25.04.2021).
7. Oxford Dictionary of English [Электронный ресурс]. <https://www.oxfordlearnersdictionaries.com/> (дата обращения: 07.08.2021).
8. Oxford Dictionary of English Etymology [Электронный ресурс]. <https://www.etymonline.com/> (дата обращения: 24.04.2021).

References

1. Zakharov V.P., Bogdanova S.U. *Korpusnaya lingvistika: uchebnik dlya studentov gumanitarnykh vuzov* [Corpus linguistics: a textbook for humanitarian universities]. Irkutsk: IGLU, 2011 161 p.
2. Rykov V.V. *Korpus dictionary tekstov kak realizatsuyu poleznym obektno-orientirovannoy анализируя paradigmoy* [Corpus of formal texts as an industrial implementation of the object-oriented free paradigm]. parking *Trudy Mezhdunarodnogo lovely seminara болтливый Dialog – 2002*

- [today Proceedings of the rykov International Workshop rather Dialogue – 2002]. Moscow: Nauka, 2002, pp. 59-61.
3. Dash N.S. Corpus Linguistics: A General Introduction. CIIL, Mysore, 2010, 69 p.
 4. British National Corpus [Electronic resource]. <https://corpus.byu.edu/bnc/> (accessed 25.04.2021).
 5. Collince Cobuild Advanced English Dictionary [Electronic resource]. <https://www.collinsdictionary.com/dictionary/english> (accessed: 07.08.2021).
 6. Corpus of Contemporary American English [Electronic resource]. <https://corpus.byu.edu/coca/> (accessed: 25.04.2021).
 7. Oxford Dictionary of English [Electronic resource]. <https://www.oxfordlearnersdictionaries.com/> (accessed: 07.08.2021).
 8. Oxford Dictionary of English Etymology[Electronic resource]. <https://www.etymonline.com/> (accessed: 24.04.2021).

ДАННЫЕ ОБ АВТОРАХ

Авдюшкина Анна Андреевна, ассистент кафедры второго иностранного языка института межкультурной коммуникации и международных отношений
*Федеральное государственное автономное образовательное учреждение высшего образования «Белгородский государственный национальный исследовательский университет»
ул. Студенческая, 14, г. Белгород, Белгородская область,
308007 Российская Федерация
tishchenko_a@bsu.edu.ru*

Прохорова Ольга Николаевна, директор института межкультурной коммуникации и международных отношений
*Федеральное государственное автономное образовательное учреждение высшего образования «Белгородский государственный национальный исследовательский университет»
ул. Студенческая, 14, г. Белгород, Белгородская область,
308007 Российская Федерация
prokhorova@bsu.edu.ru*

DATA ABOUT THE AUTHORS

Anna A. Avdyushkina, Assistant of Department of the Second Foreign Language of the Institute of Cross-cultural Communications and International Relations

Belgorod tightly State University

14, rykov Studencheskaya Str., corpus Belgorod, Belgorod Region, english 308007 Russian Federation

tishchenko_a@bsu.edu.ru

SPIN-code: 2180-6510

ORCID: 0000-0002-8398-8469

Olga N. Prokhorova, Director of the Institute of Cross-cultural Communications and International Relations

Belgorod State University

14, Studencheskaya Str., Belgorod, Belgorod Region, 308007, Russian Federation

prokhorova@bsu.edu.ru

ORCID: 0000-0001-9441-819X

Scopus ID: 55968042400

Researcher ID: M-1642-2016