

ский раздел, ориентированный на «Вовлечение граждан в цифровую среду и обучение их цифровым технологиям» посредством цифрового образования, а также предусматривает программу цифровизации школ, которая включает, помимо доступа к цифровым ресурсам и приобретения соответствующего оборудования, специализированную профессиональную подготовку преподавательских кадров.

Цифровая школы — это образовательная среда, где процесс обучения все больше насыщается цифровыми технологиями и где люди и гаджеты взаимодействуют с целью преподавания, обучения и накопления знаний. Этот путь является целью португальской системы образования, которая использует инновационные и цифровые технологии с целью ознакомить «учащихся с инструментами производственной деятельности и подготовить их к интеграции в профессиональную среду» [3, р. 15].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Dias-Trindade, S. Integración tecnológica en la educación secundaria en Portugal desde la década de 1970 hasta la actualidad / S. Dias-Trindade, J. A. Moreira, A. G. Ferreira [Recurso eletrônico] // Obra Digital. — 2021. — N 21: Competencias digitales y alfabetización mediática para una educación abierta y en red— Modo de acceso: <http://revistesdigitals.uvic.cat/index.php/obradigital/article/view/319>. — Data de acesso: 09.09.2022. <https://doi.org/10.25029/od.2021.319.21>
2. João Costa O governo está a fazer um caminho planeado e objetivo na digitalização do ensino em Portugal [Recurso eletrônico]. — Modo de acesso: <https://www.computerworld.com.pt/2021/04/23/>. — Data de acesso: 09.09.2022.
3. Conselho de Ministros (2020). Resolução N 30/2020. Diário da República N 78/2020, Série I de 2020-04-21 [Recurso eletrônico]. — Modo de acesso: <https://dre.tretas.org/dre/4086632/resolucao-do-conselho-de-ministros-30-2020-de-21-de-abril>. — Data de acesso: 09.09.2022.

АКТУАЛЬНОСТЬ ИСПОЛЬЗОВАНИЯ КОРПУСОВ ТЕСТОВ В ЛЕКСИЧЕСКОЙ РАБОТЕ ПО ИНОСТРАННОМУ ЯЗЫКУ

С. А. Шурко¹⁾, М. В. Натуркач²⁾

¹⁾Белорусский государственный университет.

пр. Независимости, 4, 220030. г. Минск, Беларусь, e-mail: shurko@bsu.by

²⁾Белорусский государственный университет,

пр. Независимости, 4, 220030, г. Минск, Беларусь, e-mail: michellinnoff@gmail.com

В настоящее время существует огромное количество текстов в электронном формате. Создание корпусов и их многофункциональность позволили организовать и структурировать различные электронные тексты в корпуса и ис-

пользовать данные корпуса в разных сферах: в переводческой, лингвистической, лексикографической и других. В современном мире огромное количество лингвистических исследований базируется на работе с корпусами, их постоянное модернизирование позволяет получать более точные и усовершенствованные данные и предлагает новые возможности для исследований.

Ключевые слова: корпус текста; корпусная лингвистика; лексико-семантическое поле; частота словоупотреблений.

THE RELEVANCE OF USING CORPORA IN LEXICAL WORK ON A FOREIGN LANGUAGE

S. A. Shurko^a, M. V. Naturkach^b

^aBelarusian State University,
Niezaliežnasci Avenue, 4, 220030, Minsk, Belarus, e-mail: shurko@bsu.by

^bBelarusian State University,
Niezaliežnasci Avenue, 4, 220030, Minsk, Belarus, e-mail: michellinnoff@gmail.com

Nowadays, there are a huge number of texts in electronic format. The creation of corpora and their versatility made it possible to organize and structure various electronic texts into corpora and use these corpora in various fields: in translation, linguistics, lexicography and etc. In modern world, a huge amount of linguistic researches are based on working with corpora. Their constant modernization allows obtaining more accurate and improved data and offers new opportunities for research.

Keywords: corpus; corpus linguistics; lexico-semantic field; word usage frequency.

Несмотря на то, что термин корпусная лингвистика был введен в употребление в 1960-е гг. в связи с развитием практики создания корпусов, которому начиная с 1980-х гг. способствовало развитие вычислительной техники, сегодня корпусная лингвистика понимается как относительно новый подход в лингвистике, который имеет дело с изучением использования языка в «реальной жизни» с помощью компьютеров и электронных корпусов. Корпусная лингвистика — раздел компьютерной лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов (корпусов текстов) с применением компьютерных технологий. Под лингвистическим, или языковым, корпусом текстов понимается большой, представленный в машиночитаемом формате, унифицированный, структурированный, размеченный, филологически компетентный массив языковых данных, предназначенный для решения конкретных лингвистических задач. Основными чертами современного корпуса являются машиночитаемый формат, репрезентативность

(сбалансированность корпуса), наличие металингвистической информации. Репрезентативность достигается с помощью специальной процедуры отбора текстов [1].

Первым большим компьютерным корпусом считается Брауновский корпус (БК, англ. *Brown Corpus*, *BC*), который был создан в 1960-е гг. К 1990 г. Уже было зафиксировано более 600 компьютерных корпусов. В последующие года количество и многообразие создаваемых корпусов шли по нарастающей [2, с. 15].

Среди современных корпусов английского языка наиболее известны Британский национальный корпус, Международный корпус английского языка, Лингвистический банк английского языка, Корпус современного американского английского. Что касается русского языка, наиболее представительным, обладающим огромным объемом лингвистических единиц и оснащенным лингвистической разметкой и инструментами поиска является Национальный корпус русского языка [3]. Последнее обновление, примененное к данному корпусу, является изменение дизайна сайта, заключенное в отображении страниц с информацией в новом интерфейсе. Так же функционируют такие корпуса в русском языке как компьютерный корпус текстов русских газет конца XX в. (является ярким примером публицистического корпуса) и путеводитель по дискурсивным словам русского языка (пример иллюстративного корпуса).

Большинство стран могут похвастаться наличием своего собственного корпуса текстов. Основным корпусом является национальный корпус. Он представляет данный язык на определенном этапе его существования, во всем многообразии жанров, стилей, территориальных и социальных вариантов и т. п.

Существует классификация корпусов текстов базируясь на различных критериях. Например, в зависимости от целесообразности использования выделяют специализированные и многоцелевые; с учетом типа языковых данных различают письменные, устные, смешанные; акцентируя внимание на термине «литературность» дифференцируют литературные, диалектные, разговорные, терминологические, смешанные; применительно в практике — исследовательские и иллюстративные т. д. [2, с. 17]. Критериев для классификации корпусов огромное множество, но следует акцентировать внимание на еще одном важном аспекте — параллельность [4]. Корпусы делятся на одноязычные, двуязычные, многоязычные. В одноязычных корпусах противопо-

ставляются диалекты и варианты одного языка, соответственно в двуязычных — два языка, в многоязычных находят свое отражение два и более языка. Данные корпуса активно используются для разработки методов перевода, в том числе, машинного, для составления двуязычных и многоязычных терминологических словарей, а также для сравнительных исследований языков [5].

Для отслеживания изменений, происходящих в языке, можно использовать корпуса текстов. Они дают возможность проводить лексикографические, лингвистические и др. исследования [6]. При помощи корпусов текстов можно установить частотность употребления лексических единиц, построить конкорданс, провести различные виды анализов (морфологический, семантический и др.) [7].

Практическое применение корпусов возможно рассмотреть на лингвистическом исследовании, проведенном для сравнения лексических единиц, находящихся в непосредственном окружении с лексической единицей «семья». При построении лексико-семантических полей, ЛЕ «семья» составляла ядро поля, остальные ЛЕ, состоящие в каких-либо отношениях (синонимических, партонимических, гипонимических и т. д.) с ЛЕ «семья» составляли ближайшую или дальнюю периферию лексико-семантического поля. Для проведения сравнительно-сопоставительного анализа лексико-семантических полей «семья» для обоих языков использовались национальные корпуса английского и русского языков.

Лексема *family* включена в лексико-семантическое поле, имеющее сложную структуру, которая включает следующие элементы: ядро, центр и периферию. Ядерным элементом являются лексические единицы с общим содержанием и наделенные сходным семантическим признаком. Ядро составляют лексические единицы: *family* — семья; *familial* — семейный.

К центральной части поля относятся:

1) узуальные синонимы: *kin* (род / родственник), *line* (род), *clan* (клан), *group* (группа), *tribe* (племя), *dynasty* (династия), *generations* (поколения), *parentage* (отцовство);

2) партонимы: *mother*, *father*, *brother*, *siblings*, *sister*, *in-laws*, *ancestors*, *children*, *descendants*, *relatives*, *grandparents*.

Ближнюю периферию составляют лексические единицы, входящие в ассоциативное поле семья: *birth* (рождение), *origin* (происхождение), *descent* (происхождение), *inheritance* (наследство), *heir* (наследник).

Так же сюда отнесем стилистически маркированные и экспрессивно окрашенные лексические единицы:

— диминутивы: *mommy* (мамуля), *daddy* (папочка), *auntie* (тетушка), *sonny* (сын);

Количество словоупотреблений в *BNC* (*British National Corpus*) представленных диминутивов следующее: *mommy* (2305 словоупотреблений), *daddy* (2090), *auntie* (635), *sonny* (210). Данные лексические единицы в большинстве случаев встречаются в устном и газетном корпусах. Например, частота слова *mommy* в устном корпусе составляет 164 (на миллион словоупотреблений), а в газетном — только 5. Что касается других типов корпусов, частота употребления данного слова в них меньше единицы, т.е. встречается там данное слово всего несколько раз.

— просторечия: (*kiddy*, *childie*, *moppet* (ребеночек), *babykins* (деточка);

Следующие примеры просторечий также освещены в *BNC*, правда, употребляются они довольно редко. Например: слово *kiddy* встречается 20 раз во всем корпусе, частота словоупотребления составляет всего 0,20, преимущественно появляясь в устном корпусе. Слова *moppet* и *babykins* встречаются дважды по всему корпусу.

— диалектизмы (*trouble and strife — wife*; *God forbid — kid*; *skin and blister — sister*; *wet and wild — child*). Данное выражение ‘*trouble and strife*’ — *wife*, принадлежащее диалекту Кокни, употребляется 7 раз в устном и газетном корпусах, остальные выражения встречаются по разу.

Дальнюю периферию поля составляют устойчивые сочетания с компонентом *Family* / семья. Данные словосочетания были взяты из Британского корпуса, который также позволил определить частотность их употребления в контексте. В данную лексико-семантическую группу можно отнести следующие словосочетания, включающие общую лексему семья: *extended family* (199 словоупотреблений), *royal family* (735), *immediate family* (100), *wealthy family* (40), *low-income family*, *foster family* (28) / *adoptive family*, *supportive family*, *noble family* (47), *needy families*, *the ruling family*, *aristocratic family*, *host family*, *poor families*, *military families*, *church family*, *immigrant family*, *single parent families*, *middle class families*, *Christian family*, *refugee families*, *farm families*, *peasant family*, *the nuclear family*, *the average family*, *Muslim family*, *disadvantaged families*, *religious family*, *a good family*, *Orthodox family*.

Значительная часть предложенных словосочетаний обозначают различные типы семей. Данные семьи можно классифицировать в соответствии с классовым делением (*noble family, middle class family*), религиозными признаками (*religious family, orthodox family*), финансовой обеспеченностью (*needy family, low-income family*), наличием или отсутствием каких-либо родственников (*extended family, nuclear family*).

Следует отметить, что при составлении конкорданса Британского корпуса многие из вышеперечисленных словосочетаний являются часто употребляемыми в мидийом дискурсе. Например: *extended family, immediate family, wealthy family*.

Однако особое внимание привлекает к себе словосочетание *royal family*, употребленное в контексте 735 раз, которое зачастую используется не только в новостных и публицистических статьях, но и занимает почетное место в заголовках данных статей. (Например: *What happens to the Royal Family in the long term is anybody's guess*). Это связано с лингвокультурологическим аспектом, который проявляется в заинтересованности людей жизнью королевской семьи, их традициями и особыми датами, которые иногда становятся национальными праздниками.

Акцентируем внимание на паремиях, составляющих дальнюю периферию лексико-семантического поля семья. Наиболее ярко культурная специфика проявляется именно в идиоматике, ориентированной преимущественно на отражении традиционной картины мира.

Количество фразеологических единиц, сочетаний и идиом, включающих в свою структуру лексику семья, не представляет собой огромный список. В результате поиска по фразеологическому словарю английского языка были выделены следующие фразеологические сращения (идиомы): *black sheep of the family* (meaning: *a disreputable or disgraced member of a family*), *in a/the family way* (meaning: *pregnant*), *run in the/(one's) family* (meaning: *to be a hereditary trait or a characteristic frequently exhibited by members of a family*). Всего было найдено 3 идиомы, образность которых раскрывается только исторически. Все идиомы наделены своей историей происхождения и отражают культуру и традиции своей страны.

Следует отметить, что даже идиомы были найдены в BNC. Например: *black sheep of the family* встречается 15 раз, 11 из 15 употреблений освещены в устном корпусе. Идиома *'in the family way'*

употреблена в контексте 13 раз, *run in the/(one's) family* — 7 раз.

Можно встретить множество фразеологических сочетаний (коллокаций): *family fortunes* (43 словоупотребления), *family man* (85), *family matters* (33), *family ties* (67), *family tree* (68), *family values* (42), *head of the family* (50), *like one of the family*, *one big happy family*, *to start a family* и т. д.

Ядро лексико-семантического поля семья в русском языке состоит из единиц, имеющих интегральное, общее с именем поля и друг с другом значение. Например: семья, семейство, семейный. К центральной части поля относятся лексемы, имеющие общие семантические признаки с ядром. К примеру, лексические единицы, вступающие в синонимические отношения с ядром поля: дом, род, фамилия, поколение, семейство, близкие, династия. Всего было подобрано 7 синонимов. Самым широко употребляемым является слово дом (792,6), частота употребления которого в 7 раз выше, чем следующих за словом дом синонимов род (165) и фамилия (118), которые также распространены и используются довольно часто.

Ближнюю периферию составляют стилистически маркированные и экспрессивно окрашенные лексические единицы. К ним относятся просторечия, диминутивы, диалектизмы, эмоциональная (экспрессивная) лексика.

Диминутивы: семейка, семеюшка, мамуля, папочка и т. д. Так, например, только к одному слову сын было подобрано 7 диминутивов: сынок, сыночек, сыночка, сына, сынулька, сынуличка, сынулик; к слову дочь тоже 7: доченька, дочурка, дочура, доча, дочурушка, дочушка, доня. Количество употреблений лексической единицы сын — 46 082. Рассмотрим диминутивы, их распространенность и количество вхождений в русский национальный корпус: сынок (3988 словоупотребления), сыночек (509), сыночка (335), сынулька (3), сынулик (1). В большинстве случаев данные лексические единицы употребляются в газетном, региональном, диалектном, поэтическом, и устном корпусах.

Просторечия: бабуля, бабуся; вот тебе, бабушка, и Юрьев день; мамаша, папаша; муж — объелся груш;

Диалектизмы: подноготная (подробности о семье и родственниках), домовничанье (ведение домашнего хозяйства в отсутствие хозяев) [8].

Данные просторечия и диалектизмы в корпусах найдены не были, не смотря на то, что национальный корпус дает возмож-

ность поиска и по региональным, диалектным, поэтическим, и устным корпусам.

Дальнюю периферию поля составляют устойчивые сочетания с компонентом семья. В русском варианте были представлены следующие словосочетания: царская семья (156 словоупотреблений), многодетная семья (29), многочисленная семья, малообеспеченная семья, императорская семья, княжеская семья, приемная семья, неполная семья, христианская семья, знатная семья, благородная семья, дворянская семья, нуклеарная семья, неблагополучная семья, аристократическая семья, религиозная семья, верующая семья, порядочная семья, патриархальная семья, православная семья, шведская семья, патронатная семья, моногамная семья.

Проверив частоту употребления данных сочетаний в русском национальном корпусе, были полученные следующие статистические данные. Самым распространенным является словосочетание царская семья, как в английском, в свою очередь, королевская семья. Остальные сочетания употребляются в контексте от 5 до 10 раз.

Таким образом, исходя из вышеупомянутого исследования, корпуса дают возможность определить частоту употребления лексических единиц, проанализировать их окружение, подобрать устойчивые сочетания с нужной лексической единицей. Они могут быть полезны для составления лексико-семантических полей, определения границ лексических единиц, входящих в поля, пополнения уже составленных лексико-семантических полей новыми лексическими единицами, в соответствии с постоянным развитием языка и благодаря непрерывной модернизации корпусов.

Более того, огромным преимуществом для пользователей корпусов является их доступность. Свободно доступные корпуса позволяют в любое время в режиме онлайн искать по всем текстам корпуса в полном объеме. В ряде случаев свободный доступ может предоставляться к части корпусных данных и не со всеми функциональными возможностями. В работе с коммерческими корпусами нужно покупать право его использования онлайн или копию на компакт-диске. Предварительно можно ознакомиться с аннотацией к корпусу или, возможно, даже поработать с корпусом в пробном режиме, но, как правило, не с большими текстами, а только с небольшим по объему подкорпусом.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Потапова, Р. К. Новые информационные технологии и лингвистика: учебн. пособие / Р. К. Потапова. — М.: Едиториал УРСС, 2005. — 368 с.
2. Захаров, В. П. Корпусная лингвистика: учебник для студентов направления «Лингвистика» / В. П. Захаров, С. Ю. Богданова. — 2-е изд., перераб. и дополн. — СПб.: СПбГУ, 2013. — 148 с.
3. Сичинава, Д. В. Национальный корпус русского языка: очерк предыстории / Д. В. Сичинава [Электронный ресурс] // Национальный корпус русского языка. — Режим доступа: <https://ruscorpora.ru/media/uploads/2022/04/20/03sitch.pdf>. — Дата доступа: 01.09.2022.
4. Беляева, Л. Н. Лексикографический потенциал параллельного корпуса текстов / Л. Н. Беляева // Труды междунар. конф. «Корпусная лингвистика — 2004». — СПб.: 2004. — С. 55—64.
5. Гвишиани, Н. Б. Практикум по корпусной лингвистике: учебное пособие по английскому языку / Н. Б. Гвишиани. — М.: Высшая школа, 2008 г. — 191 с.
6. Герд, А. С. РНК и академическая лексикография / А. С. Герд // Труды междунар. конф. «Корпусная лингвистика — 2006». — СПб.: СПбГУ; РХГА, 2006. — С. 88—91.
7. Перцов, Н. В. О роли корпусов в лингвистических исследованиях / Н. В. Перцов // Труды междунар. конф. «Корпусная лингвистика — 2006». — СПб.: СПбГУ; РХГА, 2006. — С. 318—331.
8. Дружина, Н. Л. Термины родства в русском и немецком языках / Н. Л. Дружина // Актуальные вопросы германской филологии и методики преподавания иностранных языков: материалы XIII респ. науч.-практ. конф. Брест, 20 марта 2009 г.: в 2 ч. / Брест. гос. ун-т ; редкол.: Н. А. Тарасевич [и др.]. — Брест, 2009. — Ч. 1. — С. 116—119.

ИЗУЧЕНИЕ ИНОСТРАННОГО ЯЗЫКА С ПОМОЩЬЮ РЕАЛИСТИЧНЫХ ЗАДАНИЙ (обучение на основе задач (*Task-Based Learning*))

Н. В. Шуша

*Белорусский государственный университет,
пр. Независимости, 4, г. Минск, Беларусь, e-mail: natshusha@mail.ru*

В статье рассматриваются обучение иностранному языку и изучение иностранного языка на основе реалистичных заданий. При коммуникативно-ориентированном обучении иностранному языку язык должен осваиваться в ситуациях, максимально аутентичных и повседневных. Это означает, что участники должны изучать и пробовать язык с помощью ситуаций на занятии, максимально приближенных к реальности.

Ключевые слова: *реалистичное задание; обучение на основе задач; аутентичность обучения; деятельность с интегрированными навыками; автономное обучение.*