

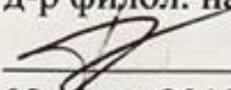
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
МОРДОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИМ. Н. П. ОГАРЁВА»

Факультет иностранных языков  
Кафедра английской филологии

УТВЕРЖДАЮ

Зав. кафедрой

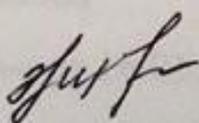
д-р филол. наук, доц.

 К. Б. Свойкин

03 июня 2019 г.

**БАКАЛАВРСКАЯ РАБОТА**

**ИСПОЛЬЗОВАНИЕ КОРПУСНОЙ ЛИНГВИСТИКИ ДЛЯ РАЗРАБОТКИ  
АНГЛОЯЗЫЧНЫХ ЛИНГВОДИДАКТИЧЕСКИХ МАТЕРИАЛОВ**

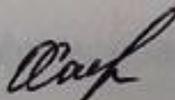
Автор бакалаврской работы  03.06.2019 К. В. Иргизова

Обозначение бакалаврской работы БР-02069964-45.03.02-16-19

Направление 45.03.02 Лингвистика

Руководитель работы

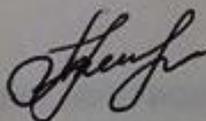
канд. филос. наук, доц.



03.06.2019 О. С. Сафонкина

Нормоконтролер

ассистент



03.06.2019 О. А. Тремаскина

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
МОРДОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИМ. Н. П. ОГАРЁВА»

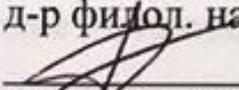
Факультет иностранных языков

Кафедра английской филологии

УТВЕРЖДАЮ

Зав. кафедрой

д-р филол. наук, доц.

 К. Б. Свойкин

16 января 2019 г.

**ЗАДАНИЕ НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ**

(в форме бакалаврской работы)

Студент Иргизова Ксения Вячеславовна

1 Тема Использование корпусной лингвистики для разработки англоязычных лингводидактических материалов

Утверждена приказом № 230-с от 16.01.2019 г.

2 Срок представления работы к защите 03.06.2019 г.

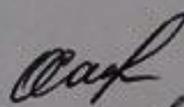
3 Исходные данные для научного исследования: научная и учебная литература интернет-источники, электронные языковые корпуса, англоязычные УМК

4 Содержание выпускной квалификационной работы

4.1 Корпусная лингвистика как новая парадигма в лингвистических исследованиях

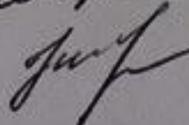
4.2 Создание англоязычных лингводидактических материалов с опорой на принципы корпусной лингвистики

Руководитель работы



16.01.2019 О. С. Сафонкина

Задание принял к исполнению



16.01.2019 К. В. Иргизова

## РЕФЕРАТ

Бакалаврская работа содержит 57 страниц, 7 рисунков, 1 таблицу, 41 использованный источник, 9 источников-примеров.

КОРПУСНАЯ ЛИНГВИСТИКА, ЯЗЫКОВОЙ КОРПУС, КОРПУСНАЯ ЛИНГВОДИДАКТИКА, АНГЛОЯЗЫЧНЫЕ ЛИНГВОДИДАКТИЧЕСКИЕ МАТЕРИАЛЫ, КОРПУСНО-ОРИЕНТИРОВАННЫЙ УЧЕБНИК

Объектом исследования является корпусная лингвистика как наука о систематизированных массивах языковых данных.

Целью бакалаврской работы является изучение применения элементов корпусной лингвистики при создании англоязычных лингводидактических материалов, а также характера функционирования лингвистического корпуса в учебнике английского языка. В работе исследование основывается на изучении принципов построения учебных пособий, в которых в качестве языкового материала используются данные, входящие в состав Кембриджского корпуса английского языка.

В ходе работы были проанализированы 172 текста из 9 учебных пособий и содержащиеся в них 2536 лексические единицы. Исследование проводилось в рамках количественного анализа учебников с использованием метода стратегической выборки.

Степень внедрения – частичная.

Область применения – лингводидактика, в частности – разработка лингводидактических материалов.

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1 Корпусная лингвистика как новая парадигма в лингвистических исследованиях	8
1.1 Формирование корпусной лингвистики как отрасли современного языкознания	8
1.2 Принципы корпусной лингвистики в контексте лингводидактики	18
2 Создание англоязычных лингводидактических материалов с опорой на принципы корпусной лингвистики	28
2.1 Разработка лингводидактических материалов по принципам корпусной лингвистики	28
2.2 Особенности лексического наполнения традиционных и корпусно-ориентированных учебных пособий	33
ЗАКЛЮЧЕНИЕ	49
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	52
СПИСОК ИСТОЧНИКОВ-ПРИМЕРОВ	56

## ВВЕДЕНИЕ

С развитием информационных технологий в настоящее время всё больше информации, которая используется нами ежедневно, существует в электронном формате. Данные из традиционных печатных источников переносятся в цифровое пространство, поскольку такая форма организации информации позволяет не только обеспечить ее долговременное хранение, но и дает возможность работать с ней на совершенно другом уровне. Одной из подобных форм организации данных является языковой корпус, который по определению В. П. Захарова представляет собой «многогранное собрание естественных случаев употребления языка в виде текстов разной жанровой и стилистической направленности и хранящееся в электронном формате». Одним из ключевых свойств языкового корпуса, которое определяет его уникальность, является репрезентативность, т.е. способность отражать все свойства проблемной области и показывать лингвистическое явление во всех возможных контекстах. Науку, которая занимается изучением принципов построения корпусов текстов, принято называть корпусной лингвистикой.

Активное развитие корпусной лингвистики привело к радикальным изменениям в различных филологических науках, таких как лексикография, теория и практика перевода и т.д. Немалое влияние корпусные технологии оказали и на лингводидактику, что привело к созданию такой междисциплинарной науки как корпусная лингводидактика. Объектом данной науки является изучение влияния использования элементов корпусной лингвистики на методику обучения иностранным языкам, в частности в области создания лингводидактических материалов.

Общеизвестен тот факт, что одной из ключевых проблем в области преподавания английского языка является нехватка релевантных учебных пособий, а также проблема их оценки, в частности лингвистического компонента. В этой связи последние исследования лингвистов посвящены анализу языкового материала в учебниках. В одном из подобных исследований,

проведенном М. Маккарти, отмечено, отмечено, что «в большинстве УМК показано не реальное употребление языка, а лишь предполагаемые варианты его использования в речи». Другими словами, тексты, которые содержатся в обычных УМК не могут полностью отразить черты, характерные для естественного разговорного дискурса, поскольку они не являются примерами живого использования языка. Следовательно, учащиеся, при работе с которыми используются искусственно созданные материалы, более ограничены в возможности совершенствовать свои языковые навыки для успешного осуществления коммуникации, выходящей за рамки аудиторных занятий. В этой связи включение в учебник данных из лингвистического корпуса, как примера живого употребления языка, может способствовать повышению его качества.

В настоящее время корпусный подход к составлению УМК активно развивается за рубежом. В нашей стране реализация подобной стратегии невозможна в связи с тем, что корпус является закрытой базой данных и доступ к нему трудноосуществим. По этой причине в настоящем исследовании был проведен сравнительный анализ традиционных и корпусно-ориентированных учебных материалов для выявления сходства и различий в технологии их построения.

**Объект настоящего исследования** является корпусная лингвистика как наука о систематизированных массивах языковых данных.

**Предмет исследования** представлен технологиями корпусных методов в процессе разработки англоязычных лингводидактических материалов.

**Цель работы** состоит в теоретическом осмыслении и практическом изучении применения элементов корпусной лингвистики при создании англоязычных лингводидактических материалов, а также характера функционирования лингвистического корпуса в учебнике английского языка.

**Задачи исследования:**

1) Собрать теоретическую базу по теме «корпусная лингвистика» и охарактеризовать ее ключевые принципы;

- 2) Изучить классификацию и структуру языковых корпусов;
- 3) Проследить использование корпусной лингвистики по различным перспективным направлениям, в частности в области лингводидактики;
- 4) Рассмотреть основные элементы технологии построения учебника по принципам корпусной лингвистики;
- 5) Определить параметры для проведения сравнительного анализа лексического наполнения традиционных и корпусно-ориентированных учебных пособий;
- 6) На основе проведенного анализа разработать ряд рекомендаций по созданию лингводидактических англоязычных материалов.

**Теоретической базой исследования** послужили работы отечественных и зарубежных ученых: Н. Б. Гвишиани, С. О. Савчук, О. В. Нагель, В. П. Захарова, В. В. Рыкова, В. А. Плунгяна, П. В. Сысоева, Д. Кристала, Р. Картера, М. Маккарти, Д. Лича, Д. Синклера, М. Льюиса, М. Скотта и других. Указанные авторы рассматривали корпусную лингвистику не только как направление языкознания, но также исследовали возможность применения корпусных инструментов в лингводидактике.

**Практическая значимость** нашей работы заключается в разработке рекомендаций по созданию англоязычных лингводидактических материалов, а также отбору и редактированию их лексического наполнения.

**Методы исследования.** Работа выполнена в русле междисциплинарного исследования, в основу которого положен принцип системности, определивший применение таких общенаучных методов, как анализ литературы и статей, изучение и обобщение отечественной и зарубежной практики, сравнение, классификация и обобщение. К эмпирическим методам исследования в первую очередь следует отнести метод стратегической выборки, с помощью которого был осуществлен количественный анализ лексического наполнения учебных пособий, что послужило фундаментальной основой настоящего исследования.

## **1. Корпусная лингвистика как новая парадигма в лингвистических исследованиях**

### **1.1 Формирование корпусной лингвистики как отрасли современного языкознания**

Корпусную лингвистику в широком понимании можно назвать одним из разделов общего языкознания, занимающимся созданием текстовых корпусов и их дальнейшим применением. Сам термин «корпусная лингвистика» был введен в употребление сравнительно недавно, в середине XX века. Однако нужно сказать, что первые языковые корпусы начали создаваться еще до возникновения корпусной лингвистики как научного направления, так как на практике, любое исследование языка, имеющее в своей основе соотнесение и анализ контекстов, является корпусным.

В отечественном и зарубежном языкознании существует несколько определений современной корпусной лингвистики. Одно из них принадлежит В.П. Захарову, автору учебного пособия по корпусной лингвистике. Он называет науку о корпусах «разделом компьютерной лингвистики, который занимается разработкой общих принципов построения и использования лингвистических корпусов (корпусов текстов) с применением компьютерных технологий» [8, с. 7]. В свою очередь, В. В. Рыков, доцент кафедры информатики МФТИ, говорит о том, что «корпусная лингвистика – это цикл исследований, связанных с правилами организации текстов в корпус, разработкой алгоритмов анализа таких текстов в рамках некоторой научной методологии» [23, с. 124]. Профессор современного английского языка в Бирмингемском университете, лексикограф Д. М. Синклер определил корпусную лингвистику (*corpus linguistics*) как «изучение языка, включающее в себя все процессы, связанные с обработкой, использованием и анализом письменных или устных машиночитаемых корпусов текстов» [40, с. 10]. Британский лингвист Д. Лич отмечает, что термин «корпусная лингвистика» относится «не столько к новой компьютерной методологии, сколько к новому

способу мышления, нестандартному методу анализа языка» [33, с. 6]. Следовательно, можно заключить, что корпусная лингвистика является одновременно наукой об использовании языковых корпусов, способом разработки алгоритмов текстовых исследований и новым подходом к анализу текста.

Ключевым компонентом в корпусной лингвистике, ее объектом является языковой корпус (*linguistic corpus/text corpus*), однако его функцию довольно трудно определить однозначно по причине его двойственной природы. Корпус одновременно является и исходным речевым материалом и непосредственным продуктом лингвистических исследований.

В современной лингвистике существует множество дефиниций термина «корпус», однако наиболее полное и многогранное определение принадлежит Д. М. Синклеру: «корпус есть собрание отрывков текстов в электронной форме, отобранных в соответствии с внешними критериями, чтобы наиболее полно представлять язык или вариацию языка; он функционирует в качестве источника данных для лингвистических исследований» [40, с. 11]. Языковой корпус способен отражать все свойства проблемной области. Данная способность называется репрезентативностью и определяется фонетическими, морфологическими, стилистическими и иными корпусными маркерами (параметрами). Другими словами, репрезентативность корпуса проявляется в тех лингвостилистических особенностях, которые присущи только ему, показывая тем самым его отличие корпуса от простого набора текстов [12, с. 4].

Нужно отметить, что корпус не является аналогом стандартного электронного библиотечного каталога, несмотря на то, что их структуры имеют похожие элементы, такие как база данных, представленная в электронном формате, поисковый интерфейс или классификация текстов по различным параметрам (жанр, автор, стиль). Однако только в корпусе представляется возможным искать отдельные фрагменты текстов по специальным параметрам и критериям, выделенным исследователем. Также следует сказать о том, что в отличие от библиотечного каталога, наполнение корпуса представлено

примерами употребления живого естественного языка. Таким образом, устройство корпуса дает возможность рассматривать язык с разных сторон, выделяя определенные закономерности и формулируя новые лингвистические законы.

Рациональность разработки корпусов и смысл их использования в лингвистических исследованиях отмечены в следующих пунктах:

- 1) внушительный объем корпуса обеспечивает типичность данных и дает возможность максимального изображения всего спектра языковых явлений;
- 2) языковые данные находятся в корпусе в своей естественной разговорной форме, что дает возможность объективно изучать их с разных сторон;
- 3) однажды разработанный массив данных может быть использован неоднократно, в зависимости от цели исследователя [19, с. 43].

Становлению корпусной лингвистики предшествовал долгий период использования корпусных методов, которые, ранее назывались иначе. Данный период называется доцифровым по причине того, что все материалы хранились без использования каких-либо электронных автоматических систем. Корпусы в то время были в основном представлены библейскими текстами и священными писаниями, а после XVIII века – еще и словарями в связи с бурным развитием лексикографии [13].

Лишь в начале XX века корпусы текстов начинают использовать при проведении языковых исследований, в том числе с целью быстрого обучения языку; в частности, языковедами использовались языковые корпусы (отобранные по специальным критериям собрания текстов) и корпусные технологии (методика подсчета частотности лексем) для создания частотных словарей.

Важной вехой в развитии корпусной лингвистики стало повсеместное использование информационных технологий в XX-XXI веках. Это позволило увеличить объемы самих корпусов, а также делать бесчисленное множество

операций, таких как разбиение фрагментов текста по нужным критериям, их маркировка и разметка, за секунду, что было невозможно при анализе текстов вручную.

В настоящее время существует большое количество лингвистических корпусов. Причиной такого многообразия является тот факт, что язык является очень разнородным и динамичным явлением, и в качестве его репрезентативной выборки трудно представить себе один единственный корпус. Для каждой конкретной исследовательской цели можно выбрать свой тип корпуса. В современном мире существует множество классификаций, из которых наиболее подробная принадлежит британскому филологу Д. Кристалу [30, с. 15]. Он выделяет:

1) Корпусы для общих и специальных целей.

Корпус для общих целей (эталонный) считается репрезентативным для конкретного языка в целом, и поэтому может использоваться для общих наблюдений относительно этого языка. Основная функция данного корпуса – это фокус на языке, который используется обычными людьми в повседневных ситуациях. Обычно он содержит письменные и устные материалы следующих типов: газеты, художественная литература, репортажи, радио- и телепередачи, дебаты и т.д. Примером данного типа корпуса может служить любой национальный корпус: Британский национальный корпус (British National Corpus, BNC), Корпус немецкого языка DWDS, созданный Берлинско-Бранденбургской Академией наук, Национальный корпус русского языка (НКРЯ), Чешский национальный корпус, созданный в Карловом университете Праги и т.д. [26, с. 105].

Напротив, корпус специального назначения проявляет себя в рамках определенного аспекта языка. Он может быть ограничен вокабуляром предметной области, типом текста по жанру, или особенностями социолекта (например, подростковый сленг). Исходя из своей специфики, такой корпус не может использоваться для наблюдения за языком в целом.

2) Корпусы письменные и устные.

Письменный корпус содержит написанные тексты, в то время как устный (разговорный) корпус состоит из стенограмм разговорной речи. Примером устного корпуса является Мичиганский корпус академического устного английского языка (MICASE), содержащий в себе более 1,5 млн. слов из разных источников (лекции, семинары, дебаты, презентации и т.д.), среди которых есть в равной степени примеры речи носителей и людей, изучающих язык. В стенограммах сохраняются все речевые явления, такие как паузация, прерывания, приемы лексической и грамматической компрессии и т.д.

Лингвистами выделяются также смешанные корпуса (В. П. Захаров). В большинстве своем это национальные корпуса, содержащие и письменные и устные тексты [8, с. 30].

### 3) Корпусы одноязычные и многоязычные.

Одноязычный корпус содержит тексты одного языка (например, инварианта и вариантов языка), в то время как наполнение многоязычных корпусов представлено текстами двух и более языков. Многоязычные корпуса в свою очередь подразделены на параллельные и сопоставленные. Параллельные корпуса включают в себя тексты на языке оригинала и тексты соответствующих переводов. По этой причине довольно часто параллельные корпуса называют переводческими (*translational corpus*). Сопоставленные корпуса, напротив, не содержат переведенных текстов. Тексты в них изначально написаны на разных языках, но во всех содержится одна и та же коммуникативная направленность и все они относятся к одной проблемной области (например, инструкция по эксплуатации оборудования) [16, с. 59].

### 4) Корпусы синхронические и диахронические.

Синхронический корпус представляет собой пример использования языка в момент определенного отрезка времени, тогда как диахронический корпус показывает динамику развития языка в течение длительного периода.

### 5) Корпусы открытые и закрытые.

Открытый (мониторный) корпус позволяет пользователю в любое время обратиться к коллекции текстов, или к их части, в ряде случаев. Открытый

корпус постоянно пополняется и в основном используется в лексикографии, поскольку создателям словарей необходимо своевременно узнавать о появлении новых слов или случаях деноминации (изменениях в значении). Напротив, закрытый корпус не подлежит каким-либо изменениям после его создания. Закрытые корпуса составляются для узкоспециализированных целей и, как правило, не предназначены для общего доступа.

Корпус не обязательно должен быть связан непосредственно с инвариантом языка, но также может проявлять себя в рамках его вариантов. В частности, всем известно, что американский английский язык отличается от британского или австралийского. В каждой из этих языковых разновидностей прослеживается разница не только в лингвистических (фонетический и лексический уровни), но также и в стилистических особенностях, например тот факт, что мы разговариваем со своими друзьями не так, как с родителями своих друзей. Исходя из этого, лингвисты начали разработку корпусов вариантов английского языка, таких как Wellington (новозеландский английский), Kolhapur (индийский английский) и т.д. [5, с. 44].

Нужно сказать, что развитие корпусной науки по-разному прослеживается в рамках отечественного и зарубежного языкознания. Наука о корпусах на Западе начала развиваться на 40 лет раньше, чем в России. Именно по этой причине западные языковеды имеют богатый опыт в создании языковых корпусов, в частности, первый известный электронный корпус – «Брауновский корпус» Г. Кучера и Н. Фрэнсиса – был создан в середине XX века. Уже через несколько лет начали издаваться первые словари на базе корпусов, как например, широко известный Collins' COBUILD (Collins Birmingham University International Language Database). Большинство современных лингвистических исследований проводятся именно на базе европейских языков, в частности английского. Возможно, причина этого заключается в том, что период расцвета британской лингвистики приходится на 60-80е гг. XX века, т.е. на момент создания первого корпуса английского языка. Ко всему прочему, активное развитие информационных технологий в

Великобритании и США могло стать катализатором прогресса создания электронных корпусов [5, с. 46].

Отечественное языкознание несколько отстает от своих зарубежных коллег в области конструирования корпусов. Первым, и на сегодняшний день единственным завершенным корпусом русскоязычных текстов, является Уппсальский корпус текстов. Он был создан в конце XX века и используется по настоящее время. Но в связи с тем, что данный корпус уже не способен полностью отвечать современным требованиям в силу ограниченного объема и отсутствия лингвистической разметки (морфологических, синтаксических, семантических маркеров), начиная с 2003 года, российские лингвисты начали работу над новым проектом, который носит название «Национального корпуса русского языка» (НКРЯ) и сейчас активно развивается [11, с. 85].

Несмотря на то, что между достижениями отечественной и зарубежной корпусной лингвистики есть различия, в подходах российских и западных лингвистов можно найти общие черты. В частности, это отношение лингвистов к корпусному подходу как самостоятельной науке. Общеизвестен тот факт, что на момент создания первых корпусов в лингвистике за рубежом активно развивался генеративный подход. Его основатель, Н. Хомский, говорил о том, что базовая информация об устройстве синтаксиса заложена в человеческом сознании с рождения, и потому, может быть применима при освоении совершенно любого языка. По его мнению, в изучении языка главным компонентом является человеческая интуиция, и, соответственно, неправильных речевых конструкций априори не существует. Данная теория подвергалась массовому обсуждению и критике, к концу XX века западные лингвисты сделали вывод о том, что создание релевантного словаря и грамматики возможно только на основе репрезентативного собрания текстов с множеством примеров актуального использования языка. Такой же концепции придерживаются и российские ученые, в частности один из руководителей проекта создания Национального корпуса русского языка В. Плунгян. Он говорит, что корпус необходим исследователям, занимающимся

систематизацией фактов об анализируемом языке, а также в академических целях, поскольку таким образом процесс освоения языковых компетенций происходит быстрее [22, с. 17]. Взгляды отечественных ученых и их зарубежных коллег относительно практики конструирования лингвистических корпусов также схожи. По своему наполнению корпусы русского и других языков (например, европейских) устроены идентично: в них присутствует информация обо всех гранях языка, например, стилистическая многожанровость, случаи употребления особых коллокаций и фразеологизмов, указаны примеры полисемии и омонимии, каждый из корпусов построен по принципу репрезентативности [5, с. 45].

Нужно сказать, что корпусная лингвистика отличается от традиционного языкознания. Данные различия можно проследить по ряду критериев. Во-первых, корпусная лингвистика – это новый взгляд на дискурс как реальный, а не фиктивный элемент коммуникации. Идеология корпусной лингвистики построена на том, что при работе используются не искусственно созданные тексты, а примеры живого использования языка, какими как раз являются языковые корпусы.

Во-вторых, это акцент на количественный анализ языка, а именно исследование элементов, наиболее часто используемых в речи, и выстраивание, согласно полученным данным, определенных принципов классификации лексического состава языка [14, с. 82].

В-третьих, это работа с языком в рамках синхронического и диахронического аспектов лингвистики. Если говорить о синхронии в языке, то она будет проявляться в одном из принципов построения языковых корпусов, а именно в жанровом и стилистическом разнообразии и различных средствах выражения одной и той же мысли в зависимости от контекста ситуации. Если же рассматривать явление диахронии в языке, то в данном случае мы можем обратиться к идеям лингвистов, описанных еще в середине XX века, в частности к работе американского лингвиста Д. Гринберга “Rethinking languages diachronically” (1979). В его статье описывается динамическое и

константное развитие языка, сосуществование новейших и устаревших элементов в дискурсе и разграничение понятий языковой нормы и языковой правильности.

Корпусная лингвистика является результатом синтеза знаний в различных областях языкознания. Так, например, в рамках сравнительно-исторического языкознания корпусная лингвистика использует технологии по реконструкции древних языков для лингвистического анализа. Также корпуса текстов могут использоваться в качестве эмпирического и иллюстративного материала для различных лексико-грамматических явлений. Социолингвистика обращается к использованию корпусных критериев для создания пособий по изучению разновидностей языка (диалект, социолект и т.д.) [19, с. 43].

В современной корпусной лингвистике выделено 6 наиболее перспективных путей использования. В первую очередь это исследования в области лексикографии, т.е. практики создания словарей. Все современные словари английского языка были созданы, в основном, на основе репрезентативных корпусов.

Во-вторых, использование корпусных методов дает возможность получить точную информацию о словарном запасе человека, исходя из частотности употребления им тех или иных слов. Эти данные можно получить используя закон Ципфа, согласно которому, если в языке все слова упорядочить по убыванию частоты их использования, то частота любого слова в таком списке окажется приблизительно обратно пропорциональной его порядковому номеру (рангу) [2, с. 215]. Основопологающей чертой данного является объективная закономерность, по которой определяется словарный запас человека: некоторые слова употребляются им постоянно, другие – крайне редко.

В-третьих, корпусная лингвистика также проявляет себя в рамках изучения лексического состава и грамматики конкретного языка. Так, например, благодаря корпусным технологиям можно проследить появление неологизмов или сочетаемость каких-либо грамматических явлений. Нужно

сказать, что в таком контексте языковые исследования становятся более объективными, так как проводятся в рамках конкретного корпуса текстов.

В-четвертых, используя корпуса, становится возможным определение функционального стиля текста через его статистические характеристики – среднюю длину слова и предложения, характерные сочетания слов и т.д. Подобные методы уже активно используются в автоматическом реферировании и тематическом поиске, причем не только в рамках письменного дискурса, но и устного.

В-пятых, корпусная лингвистика представляет особый интерес для представителей переводоведения, в частности такая ее составляющая как сопоставленный корпус, содержащий эквивалент из языка перевода каждой фразе в языке оригинала.

В-шестых, достижения корпусной лингвистики активно используются в области лингводидактики. Количественные данные о языке (частотность лексики, вероятность употребления определенной грамматической конструкции и т.д.) нужны для того, что определять содержание обучения английскому языку. Особую ценность также представляет тот факт, что корпусная лингвистика позволяет модернизировать набор примеров для преподавания языка за счет обширной эмпирической базы, которую представляют собой корпуса.

Таким образом актуальность развития корпусной лингвистики как нового направления в науке не вызывает сомнений и не оспаривается современными учеными. Благодаря своему большому потенциалу и прогрессу в рамках современных инфокоммуникационных технологий, корпусный подход к анализу языка успешно используется во всем мире, а также позволяет найти точки соприкосновения между гуманитарными и техническими науками. Создание корпусов стало революцией в области анализа дискурса, позволяя рассматривать язык с различных сторон. Но главным определяющим фактором является то, что корпус позволяет объективно рассмотреть язык таким, какой

он есть на самом деле, что представляет собой неоценимую пользу для исследователей.

## **1.2 Принципы корпусной лингвистики в контексте лингводидактики**

С целью установления связи между корпусной лингвистикой и лингводидактикой, необходимо вернуться к определению языкового корпуса, который, согласно С. Ю. Богдановой и В. П. Захарову, является «большим, представленным в машиночитаемом формате, унифицированным, структурированным, размеченным, филологически компетентным массивом языковых данных, предназначенным для решения конкретных лингвистических задач» [8, с. 7]. Исходя из данного определения, мы можем заключить, что основными характеристиками корпуса являются системность, репрезентативность и унификация языковых данных. Данные параметры корпуса дают возможность предположить, что его использование поможет организовать и упорядочить по определенным критериям многообразие языковых явлений. Для современной лингводидактики этот факт имеет особое значение, поскольку исследователи в области методики преподавания иностранных языков находятся в постоянном поиске приемов передачи сложного и многогранного языкового материала самым эффективным и доходчивым способом. Согласно Н. Б. Гвишиани, профессору кафедры английского языкознания МГУ им. М. В. Ломоносова, «корпусная лингвистика, исследуя микрокосм языкового функционирования в процессе коммуникации, пытаясь отразить большое в малом, решает задачи, сходные с теми, которые близки как общедидактическим, так и частнометодическим целям» [6, с. 57]. Также, по мнению С. О. Савчук, не последнюю роль играет тот факт, что массив языковых данных, построенный по принципам корпусной лингвистики может использоваться не один раз, что подтверждает его представительность, возможность повторных лингвистических исследований и верификации их результатов. Следовательно, мы можем утверждать, что корпусная лингвистика

и лингводидактика в своем тандеме смогут организованно и систематизировано «отразить безграничную стихию языка» [24, с. 63].

Использование элементов корпусной лингвистики и лингводидактики в своей тесной связи привело к созданию такой междисциплинарной науки как корпусная лингводидактика. Ее изучением занимаются представители различных отечественных и зарубежных школ, среди которых Н. Б. Гвишиани, С. О. Савчук, О. В. Нагель, В. П. Захаров, В. В. Рыков, В. А. Плунгян, П. В. Сысоев, Д. Кристал, Р. Картер, М. Маккарти, Д. Лич, Д. Синклер, М. Льюис, М. Скотт и многие другие.

Становлению корпусной лингводидактики предшествовали значительные изменения в современной педагогической науке. Среди них можно выделить использование ИКТ и процесс информатизации образования, что повлияло на совершенствование современных образовательных технологий в целом. Под информатизацией образования в данном случае понимается «создание информационно-коммуникативной среды, где основным ресурсом является человек, умеющий работать с информацией, усваивать и обрабатывать ее, а также целесообразно применять в зависимости от ситуации» [1, с. 119]. В работах Л. В. Московкина, Л. П. Тарнаевой, М. Скотта и других также говорится об активном процессе интеграции информационных технологий в методику преподавания иностранных языков. По их мнению, причиной подобной тенденции стала возможность новейших компьютерных средств обучения (собственно компьютер, проектное оборудование, электронные УМК, интернет-ресурсы) обеспечить практически все современные потребности учащихся и преподавателей в области иностранных языков. К числу подобных средств обучения можно также отнести электронный языковой корпус, без которого возникновение инфокоммуникационного пространства было бы невозможно. Лингвистический корпус по праву можно назвать наиболее релевантным источником языковых данных, использование которого может помочь рассмотреть наиболее употребляемые лексико-грамматические формулы, увидеть разницу в структуре устной и письменной речи, определить

необходимый словарный запас человека для общения на бытовом или профессиональном уровне и т.д. Таким образом, вышеперечисленные характеристики корпуса обуславливают возрастающую тенденцию к его использованию на занятиях по иностранному языку.

Идеи использования корпусов в лингводидактике и возрастающий интерес к вопросам применения корпусных технологий послужили отправной точкой для организации специализированных конференций в различных образовательных центрах. Ярким примером являются конференции TaLC (Conference on Teaching and Language Corpora), проводимые с 1994 года в ведущих университетах Европы. В России среди подобных мероприятий можно выделить конференции цикла “LATEUM” и «Корпусная лингвистика», проводимых на базе филологического факультета МГУ и СПбГУ.

Языковой материал, отраженный в лингвистических корпусах является одной из наилучших эмпирических баз для формирования лексических навыков учащихся. В данном контексте также играет большую роль становление так называемого лексического подхода к обучению языку, в рамках которого лексические ошибки (например, нарушение узуса и сочетаемости) считаются такими же важными, как и грамматические несоответствия. Концепция данного подхода заключается в изучении языка в формате определенных коллокаций и словосочетаний, что является альтернативой традиционному способу подстановки слов в грамматическую формулу без осознания их речевой функции. Следовательно, лексика в рамках корпуса изучается по пути «от частного к общему»: от примера реального употребления языка к его использованию в составе определенных грамматических моделей. Также применение корпусных технологий в преподавании английского языка позволяет сменить традиционный вектор «правильно/неправильно» в отношении отбора лексического материала для изучения. Вместо ориентации лишь на свое субъективное мнение педагог может обратиться к корпусу, как релевантному источнику информации, поскольку в нем слова анализируются не

хаотично, а на основе определенных количественных закономерностей [7, с. 30].

Корпусные тенденции достаточно сильно повлияли и на методику формирования грамматических навыков. Долгое время в научной парадигме преподавания английского языка господствовало мнение о том, что идеальной грамматической моделью является латинский язык, а тексты на латыни представляют собой образцы выражения этой модели. Однако в современной науке принято считать, что английский язык не нуждается в латинских шаблонах и обладает собственной, самостоятельно развивающейся грамматикой. В частности, такого мнения придерживается британский лингвист М. Маккарти [35]. В своей статье “Corpora and grammar” он анализирует результаты исследований корпуса английского языка, где описана разница в структурных элементах грамматики письменной и устной речи. Данные исследования наглядно демонстрируют, что в основном в процессе общения используется грамматика устной речи, причем данная закономерность в равной степени прослеживается в речи как носителей, так и не носителей языка. Этот факт обуславливает важность включения грамматики устной речи в курс обучения общей грамматике, поскольку без рассмотрения ключевых принципов данного аспекта английского языка полноценный процесс обучения устной речи представляется невозможным.

Еще одним аспектом, в рамках которого корпусная лингвистика может проявить себя, является определение содержания обучения английскому языку. В работе “From Corpus to Classroom: Language Use and Language Teaching” авторами которой являются М. Маккарти, А. О’Киф и Р. Картер, представлено мнение относительно отбора языкового и речевого материала для усвоения учащимися на начальном уровне (например, времена групп Present и Past в рамках аспектов Simple и Continuous). Наибольшие же разногласия, по мнению авторов, вызывают детали, касающиеся процесса обучения на продвинутом уровне. Такой же позиции придерживаются отечественные исследователи. Как отмечает Е. А. Захарова, преподаватель кафедры теории и практики

иностранных языков Института иностранных языков учебные планы и программы, посвященные обучению языку на высоком уровне, очень разнообразны, а порой их компонентное содержание не совсем понятно; иногда представляется, что включенные в них грамматические структуры иногда отбираются на основе принципа «чем сложнее и непонятнее – тем правильнее» [9, с. 44]. Однако, если выстраивать стратегию обучения грамматике на основе анализа корпуса, то содержание учебной программы может стать более четким и понятным. Так, например, благодаря использованию корпуса можно организовать систему обучения грамматике на продвинутом уровне, включающую следующие элементы:

- 1) Изучение уже известных грамматических формул, но в ракурсе новой смысловой нагрузки;
- 2) Рассмотрение уже известных речевых интенций, но выраженных новыми грамматическими структурами;
- 3) Анализ различий между грамматиками устной и письменной речи и т.д.

Классификация языковых корпусов достаточно обширна, однако не каждый корпус способен максимально эффективно проявить себя в процессе обучения. Так, например, наиболее подходящим типом корпуса для эффективного изучения английского языка принято считать корпус учащегося (*learner corpus*), который впервые был выделен как новое течение в корпусной лингвистике С. Грейнджер, сотрудником центра по изучению англоязычных корпусов текстов при Лувенском католическом университете (Бельгия). Грейнджер определяет его как «электронное собрание аутентичных текстов, написанных студентами, изучающими язык как иностранный или второй». Контент корпуса представляет собой собрание работ (примеров устной и письменной речи) обучающихся в виде электронной базы данных. Данные работы выполняются студентами в двух экземплярах на родном и на изучаемом языках. Суть использования данного корпуса учащимися заключается в том, что он позволяет увидеть выражение одной и той же мысли одним человеком

на двух разных языках, а значит, позволяет выявить наиболее типичные ошибки изучающих язык и исправить их.

Использование корпусов текстов предполагает обучение английскому языку в рамках проблемного подхода. П. В. Сысоев отмечает, что в отличие от репродуктивного подхода, характеризующегося пассивным получением информации от преподавателя, преобладание проблемного подхода на занятиях по иностранному языку способствует стимуляции речемыслительной деятельности учащихся в процессе развития речевых умений. Проблемный подход на занятиях может быть реализован посредством применения учениками онлайн-версий языковых корпусов для поиска и обработки информации или нахождения в текстах лингвистических закономерностей, на основе которых затем можно формулировать правила. Реализация проблемного подхода обеспечивает повышение самостоятельности обучающихся при выполнении заданий, развивает технологии самообучения, способствует преобладанию индукции на занятиях. Продуктом реализации данного подхода является прочное и долговременное запоминание новой информации, возможность для обучающихся самим устанавливать причинно-следственные связи между языковыми явлениями [26, с. 103].

Нельзя не сказать о том, что корпусные методы в преподавании влияют на роль учителя и учащегося как участников педагогического процесса. При исследовании языка в рамках корпуса учащийся чаще всего работает самостоятельно, становясь при этом центральным участником процесса обучения. Следовательно, преподавателю остается консультативная роль, которая сводится к экспертной оценке и контролю результатов исследования. Согласно О. С. Квашниной фокус роли учителя меняется с «непосредственного руководителя» на «наблюдателя и советчика», функция которого «научить учащихся учиться» и выявлять лингвистические закономерности при формулировке тех или иных правил [8, с. 81]. Квашнина отмечает, что педагог должен заблаговременно проанализировать все данные поиска в языковом корпусе, прежде чем предлагать его учащимся в качестве дидактического

материала. Также ему необходимо заранее определить, какие именно задачи будут решаться с помощью того или иного корпуса и понять, подходит ли подобная методика для работы в группе определенного возраста и с определенным уровнем владения языком [8, с. 82].

Использование языковых корпусов в обучении иностранному языку тесно связано с развитием коммуникативной компетенции учащихся и может проявляться на разных этапах ее формирования. Согласно О. В. Нагель, в рамках данного аспекта, корпуса могут использоваться:

1) Для отбора аутентичных образцов лексем, устойчивых выражений, грамматических моделей и правил. Учащийся может самостоятельно подвергать анализу информацию о языке, используя корпус как некий ресурс, который рассматривает его со всех сторон; таким образом, у него будет возможность наиболее полно приблизить свою речь к речи носителей.

2) Для оценки функции единиц языка, рассматривая их форму и значение. Учащийся анализирует употребление слова в определенном контексте, что дает ему возможность рассмотреть и лексические и грамматические закономерности одновременно.

3) Для более углубленного процесса изучения иностранного языка, например, используя данные из корпуса о наиболее частотных коллокациях с тем или иным словом. Польза корпуса в данном контексте может прослеживаться также при обучении грамматике, где в нем можно увидеть закономерности использования глаголов с определенными предлогами.

4) Для увеличения словарного запаса учащихся в сфере профессиональной терминологии, используя специальные корпуса (*special corpora*) [21, с. 55].

Применение электронных языковых корпусов в качестве средства обучения английскому языку предполагает использование специализированных систем компьютерной обработки данных, которые получили название «конкордансеров». Конкордансерами называют программы, позволяющие найти необходимую информацию из корпуса и использовать ее в зависимости

от своей цели. Результат работы конкордансера называется «конкордансом» и представляет собой список всех возможных употреблений заданного языкового выражения в контексте [8, с. 30].

В настоящее время наиболее известными конкордансерами являются следующие программы:

1) MicroConcord. Данная программа работает в формате KWIC (Key Word In Context). Обладает способностью осуществлять быстрый поиск. Выданный конкорданс можно просматривать в режиме прокрутки, сохранить в отдельный файл или в дальнейшем распечатать. Существуют ограничения по числу выдаваемых строк в конкордансе (1662 строки). Программа работает с европейскими языками [38].

2) Concordance. Осуществляет быстрый поиск по запросу; предлагает возможности для дальнейшей обработки найденной информации (сохранить, распечатать). Может преобразить конкорданс в html-файл. Программа также работает с европейскими языками [29].

3) MonoConc. Работает не только с KWIC, но и с полноценными конкордансами. Может создавать запросы по шаблонам (например, поиск внутри тэгов). Также способен осуществлять работу с аннотированными корпусами (формата html и xml) [37].

Такие программы «подвергают» детальному анализу крупные массивы текстов с целью нахождения каких-либо закономерностей применения в них отдельных слов или словосочетаний.

Полифункциональная природа конкордансеров ложится в основу разработки вариативных заданий по английскому языку. Они способны осуществлять поиск не только отдельных слов, но и коллокаций, категоризировать информацию согласно запросу пользователя и имеют обширную контекстную базу. Корпусы текстов, представленные в электронном виде, отражают леммы (группы слов, образованные посредством флективных изменений), морфологические характеристики слова, а в некоторых случаях даже этимологические и библиографические данные. На основе такого

широкого спектра возможностей конкордансеров преподаватель может предложить учащимся различные задания лексического/грамматического типов, например, для определения значения слова по контексту или рассмотрения полисемии отдельных лексических единиц, а также изучения управления глаголов и грамматических времен и т.д. Подобные задания также будут релевантны при составлении тестов и контрольных работ. Упражнения на базе данных корпуса считаются упражнениями на восприятие языковой формы, а значит, дают возможность учащимся увидеть те искажения, которые допускаются ими неосознанно. Следовательно, инструментарий корпусной лингвистики может найти свое отражение в случаях интуитивного нарушения учащимися правил использования лексико-грамматических структур и сочетаемости лексических единиц, несоблюдения стилевых характеристик при построении высказывания, а также интерференции родного языка.

Несмотря на тот факт, что корпусная лингвистика обладает огромным потенциалом в контексте методики преподавания иностранных языков, в настоящий момент ее использование довольно ограничено. Причина этого может заключаться в том, что работа с корпусными инструментами на занятиях требует высокого уровня подготовки, как от преподавателя, так и обучающихся. Также следует помнить о том, что корпусная лингвистика является областью компьютерной лингвистики, а значит, без определенного уровня сформированности информационной компетенции у учащихся ее использование представляется малопродуктивным, или вообще невозможным. Однако, П. В. Сысоев предлагает следующий вариант разрешения данной проблемы на начальном и среднем этапах обучения: преподаватель может предоставить учащимся данные корпусного анализа в печатном виде, заранее их проработав и удалив все неоднозначные и сложные для понимания варианты.

Кроме того, исключительность использования корпусных инструментов также является предметом обсуждений. А. Ю. Мордовин полагает, что применение языковых корпусов должно быть целесообразно и рационально,

чтобы исключить возможность возникновения прагматического дискомфорта [20, с. 279]. Соответственно, независимо от результата, полученного в ходе самостоятельного поиска и обработки информации в корпусных менеджерах (особенно при изучении грамматических структур), учащемуся все равно нужно будет обращаться к теоретическим основам, изложенным в учебных пособиях и дополнительных дидактических материалах с целью исключить неправильное запоминание информации или ее нерелевантное употребление.

В-третьих, причиной отсутствия релевантного отражения использования корпусных инструментов в образовательном пространстве является недостаточная интеграция корпусной лингвистики в методику преподавания иностранных языков. Это объясняется тем, что зачастую лингвистические корпуса создаются, наполняются и используются исследователями языка, лексикографами и переводчиками. Однако для современных педагогов, ввиду недостаточной сформированности компетенций в области работы с ИКТ, корпусная лингвистика является недостаточно изученным полем деятельности.

Таким образом, технологии применения корпусов занимают особое место в современном образовательном пространстве. Они могут создать условия для реализации проблемного подхода для занятиях по английскому языку, способствовать корректному и прочному усвоению лексико-грамматического материала, являться релевантным источником информации при определении содержания обучения. И хотя использование элементов корпусной лингвистики требует наличия специальных компетенций у педагога и учащихся для создания результативного продукта, данные аспекты несопоставимы с возможностью повышения качества иноязычного обучения благодаря применению корпусных технологий и их большому педагогическому потенциалу.

## **2. Создание англоязычных лингводидактических материалов с опорой на принципы корпусной лингвистики**

### **2.1 Разработка лингводидактических материалов по принципам корпусной лингвистики**

Корпусная лингвистика находит свое отражение в различных сферах лингводидактики, в частности в области составления учебных пособий. По мнению различных исследователей, значимость данного аспекта определяется важностью использования УМК в процессе формирования коммуникативной компетенции учащихся [25, с. 200]. Так, например, в своей работе, посвященной проблемам школьного учебника, И. Л. Бим говорит, что именно учебник считается центральным компонентом УМК, поскольку «является основным средством обучения...» [4, с. 26]. Проректор ЛГПИ им. А. И. Герцена, доктор педагогических наук И. Д. Зверев видел назначение учебного пособия в образовательном процессе в том, что в рамках урока именно оно является «структурировано разработанным «хранилищем» всей основной информации, включающим в себя необходимое количество лексико-грамматических структур для изучения на данном этапе владения языком» [10, с. 4].

Таким образом, принимая во внимание тот факт, что дидактические материалы высокого качества имеют большое значение для эффективного обучения языку, анализ подобных УМК, в первую очередь их лексического наполнения, очень важен. Однако в области современного моделирования дидактических материалов отсутствует единая классификация принципов отбора лексики, образующей базу учебника. Данный факт отражен в исследованиях зарубежных лингвистов, в частности М. Маккарти, профессора прикладной лингвистики Ноттингемского и Лимерикского университетов и одного из основоположников корпусной лингвистики. Он говорит о том, что отбор лексики для учебника довольно часто лимитирован авторским субъективным мнением и его профессиональными взглядами. При разработке

учебников следует избегать подобных тенденций, в особенности в области лексического наполнения УМК. К примеру, автор может оценить правильность использования той или иной грамматической конструкции с точки зрения собственной интуиции, однако он не сможет определить частотность употребления слова без опоры на релевантные источники информации [36, с. 75].

Несмотря на отсутствие универсальных параметров в современной лингвистике существуют различные теории относительно исследования языкового контента учебников. Одной из таких теорий является количественный анализ учебников с использованием метода стратегической выборки лексических единиц. В качестве критериев оценивания результатов данного анализа лингвист Ю. Лью впервые выделил различные параметры, такие как, повторяемость слов, повышение уровня их сложности, лексическая сочетаемость, и т.д. [34, с. 1992]. Вышеперечисленные параметры являются основополагающими для создания списков, на базе которых формируются корпусно-ориентированные учебные материалы. Примером такого списка является Academic Formulas List, организованный по принципу частотности употребления лексем. Он включает в себя повторяющиеся сочетания в корпусах устной и письменной речи, преимущественно в текстах академических жанров [39].

В рамках содержательного компонента учебных пособий параметр частотности проявляет себя следующим образом: наиболее широко в учебнике должны быть отражены часто употребляемые слова, поскольку они требуют больше времени для запоминания ввиду большого количества случаев полисемии и взаимодействия данных слов в составе некоторых коллокаций. К примеру, в классические учебники английского языка составители включают от 50 до 100 глаголов, являющиеся, по их мнению, обязательным для изучения. В отличие от традиционных УМК в корпусно-ориентированных учебниках выборка ограничивается всего 12 глаголами (say, get, go, know, think, see, make, come, take, want, give, mean), которые, встречаются в корпусе Longman Spoken

and Written English чаще, чем 1000 раз на миллион. По мнению исследователей, эти 12 глаголов составляют 45% от всех использованных глаголов в разговорной речи (кроме вспомогательных). Следовательно, можно утверждать, что составители корпусно-ориентированных УМК в первую очередь фокусируются на немногочисленной выборке слов, которая, однако, часто употребляется в речи. Таким образом, учащимся предлагается фиксированный набор лексики, которая не потребует большого количества времени для изучения и точно пригодится в процессе общения [18].

В целом, частотность лексики является не единственным параметром корпусно-ориентированных учебников, отличающим их от традиционных. В большинстве своем учебники, созданные на базе корпуса, стали достаточно популярны в последнее время, так как в отличие от классических учебных пособий они являются источником аутентичного материала в любом из стилистических регистров и отражают реальное применение языка. Однако осуществлять разработку учебников подобного типа могут лишь организации, имеющие полноценный доступ к корпусу, который, как уже было упомянуто, является закрытой базой данных. Ярким примером являются учебники издательства Cambridge University Press, которое создает учебники на базе Кембриджского корпуса английского языка. Данный корпус используется в целях моделирования учебных пособий, а также при составлении Кембриджских экзаменов [27]. Данный корпус включает в себя:

- 1) корпуса письменных текстов (периодика, художественная литература, учебники, блоги);
- 2) корпус устных текстов (лекции, телефонные разговоры, радиопередачи, речи на митингах);
- 3) специализированные корпуса (коллекции документов из области финансов, юриспруденции и т.д.);
- 4) CANCODE (Cambridge and Nottingham Corpus of Discourse in English – собрание устных текстов из различных географических районов Великобритании);

- 5) корпус американского английского;
- 6) корпус обучающегося, представляющего собой коллекцию экзаменационных эссе, написанных учениками, для которых родными являются более 100 языков.

Из вышеперечисленных составляющих Кембриджского корпуса особое значение при конструировании учебных материалов имеет корпус обучающегося. Ученические тексты в составе данного корпуса отражают самые характерные ошибки учащихся из разных стран, принадлежащих к различному социальному положению и возрастной категории. Исследование данных ошибок и результаты сопутствующего анализа очень важны для повышения эффективности дидактических материалов. Полученные данные позволяют на живых примерах показать, на какие из языковых явлений следует обратить особое внимание при изучении с целью избежать допущения новых ошибок.

Нужно сказать, что при разработке УМК корпусный инструментарий также находит свое отражение в области изучения различных стилистических регистров. В большинстве своем основоположниками корпусной лингвистики были преподаватели английского языка, разделяющие мнение о том, что между языковыми регистрами, между так называемыми хорошим и правильным языком, нет сущих различий, и что они равноценны [16, с. 95]. Такая позиция обусловлена тем, что в лингвистике нет четкого определения хорошего языка, поскольку он независим от формы выражения (устной или письменной), профессиональных (язык врачей и юристов) или социолектных особенностей (молодежный сленг) [16, с. 96]. Следовательно, речь преподавателя также невозможно считать примером хорошего языка ввиду различных социокультурных особенностей, его принадлежности к определённому регистру языка и т.д. Однако, корпус, благодаря своей аутентичной природе и структуре, может стать языковой средой, при погружении в которую учащемуся будет легче быстрее и эффективнее исследовать определенные лингвистические закономерности.

Использование корпусных принципов в качестве основополагающей базы конструирования учебников может отражать не только лексическое наполнение, но и связь лексем с грамматическими структурами. Так, например, в текстах традиционных учебников английского языка в качестве определения обычно выступают прилагательные и причастия (big house, exciting game, interested couple). Содержание корпусно-ориентированных учебников, в большинстве своем, подобно традиционным, однако в них прослеживается раздельное обучение устной и письменной речи. В качестве определения в устной речи также выступают прилагательные и причастия, но в письменной речи эту задачу чаще всего выполняют существительные. (glass windows, metal seat, tomato sauce). Следовательно, в корпусно-ориентированных УМК наблюдается четкое дифференцированное обучение устному и письменному дискурсу. Данный факт проявляется в использовании определенного типа лексико-грамматической структуры для развития конкретного вида речевой деятельности.

В качестве еще одного примера можно привести тот факт, что классические учебники характеризуются наличием длинных и трудно запоминаемых списков глаголов, которые требуют за собой употребления инфинитива, а не герундия [18]. Отсутствие четкого понимания для чего необходим тот или иной глагол, в каких категориях он себя обнаруживает, какое у него управление и т.д. снижают образовательный потенциал таких учебников. В корпусно-ориентированных УМК учебный материал выстроен в виде логичной системы. Например, наиболее частотными глаголами с инфинитивом являются всего 4 слова: want, try, seem, like. В художественной литературе к ним добавляется begin, в академической речи – tend to и appear to. Остальные достаточно частотные глаголы этой группы хорошо разбиваются на 4 категории:

- 1) “want/need” verbs: hope, like, need want
- 2) “effort” verbs: attempt, fail, manage, try
- 3) “begin/continue” verbs: begin, continue, start

#### 4) “seem” verbs: appear, seem, tend

Как правило, подобные схемы способствуют активизации мнемонических процессов и направлены на долговременное запоминание ввиду четкой категоризации и наличия смысловых связей между элементами одной системы [18].

Таким образом, нами были рассмотрены основные отличия корпусно-ориентированных учебных пособий от традиционных УМК. Можно заключить, что применение корпусно-ориентированных учебников помогает повысить эффективность обучения английскому в частности в области формирования лексико-грамматических навыков, а также изучения языковых регистров. Входящие в содержание корпусно-ориентированных УМК примеры, взятые из корпуса учащегося, позволяют увидеть типичные ошибки учащихся, для которых английский язык не является родным, и, как следствие, избежать их в дальнейшем. Однако основным критерием, отличающим корпусно-ориентированный учебник от традиционного, является его лексическое наполнение. Важность лексической составляющей УМК обусловлена тем, что именно лексика является системой знаний (набором лексических единиц), имеющих общее значение для всех членов конкретного культурно-языкового сообщества, наполняя определенным смыслом те или иные речевые интенции говорящего. Подробное описание анализа лексического наполнения корпусно-ориентированных дидактических материалов, послужившего фундаментальной основой настоящего исследования, представлено в следующей части данной работы.

## **2.2 Особенности лексического наполнения традиционных и корпусно-ориентированных учебных пособий**

В данной части настоящей исследовательской работы был проведен сравнительно-сопоставительный анализ лексического наполнения традиционных и корпусно-ориентированных учебных пособий. Исследование проводилось на базе 2 линеек зарубежных корпусно-ориентированных

учебников Кембриджского издательства “English in Mind” [42; 43; 45] и “Face2Face” [44; 46; 47], а также отечественного учебника “Spotlight” [48; 49; 50], не являющегося корпусно-ориентированным. Выбор данных УМК был обусловлен тем, что они широко используются в преподавании английского языка и направлены на интенсивное развитие коммуникативной компетенции учащихся, подготавливая их, таким образом, к реальному общению. Курс “English in Mind” отличается наличием индивидуальных проектов и лингвострановедческого материала. Программа курса “Face2Face” содержит всевозможные тексты и ситуации для обсуждения на различные темы, что одинаково хорошо подойдет как детям, так и взрослым. Однако нужно сказать, что данные учебники применяются лишь в рамках системы дополнительного образования, например в рамках учебных курсов английского языка. В отличие от “English in Mind” и “Face2Face”, курс “Spotlight” в настоящее время широко используется в российских школах, и рекомендован федеральным перечнем министерством образования.

Для проведения исследования нами были взяты учебники разных уровней. Соответствие учебников определенному уровню сложности определялось согласно Общеввропейским компетенциям владения иностранным языком (Common European Framework of Reference) [28, с. 46]. В рамках настоящей работы используется общепринятая аббревиатура данной системы (CEFR).

Учебники серий “English in Mind” и “Face2Face” являются адаптированным к определенному уровню CEFR, поскольку выпускаются издательством Cambridge University Press, которое создает учебники в соответствии с данной классификацией. В свою очередь, уровни учебников “Spotlight” не имеют четкого соответствия классификации CEFR, так как их уровни сложности определяются степенью обучения (классом). Таким образом, было необходимо провести соответствие между классификацией CEFR и степенями обучения российской школы. Так по Р. П. Мильруду, согласно системе CEFR, российские школьники изучают английский язык от уровня A1

до уровня B2, где A1 является элементарным и достигается к концу начальной школы, а B2, соответственно, является уровнем выпускника. Средним в данном случае принято считать уровень A2+/B1, которого учащиеся достигают к окончанию ступени основного общего образования, то есть, 9 класса. В соответствии с данной классификацией, для исследования были выбраны учебники “Spotlight” для 4, 9 и 11 классов. Результаты соответствия выбранных учебников уровням CEFR представлены в таблице 1.

Таблица 1 – Соответствие выбранных учебников уровням CEFR

<b>УМК</b>	<b>Пороговый уровень (A1, A1+)</b>	<b>Средний уровень (A2+, B1)</b>	<b>Уровень профессионального владения (B2, C1)</b>
<b>English in Mind</b>	English in Mind 1	English in Mind 3	English in Mind 5
<b>Face2Face</b>	Face2Face Elementary	Face2Face Intermediate	Face2Face Advanced
<b>Spotlight</b>	Spotlight 4	Spotlight 9	Spotlight 11

В настоящем исследовании был использован онлайн-сервис English Profile, который является особым проектом, созданным при поддержке Совета Европы [31]. Информация, представленная в данном сервисе, позволяет выяснить уровень компетенций учащихся на каждом из уровней CEFR. Подкатегория данного сервиса, система English Vocabulary Profile, осуществляет поиск лексем в рамках Кембриджского корпуса английского языка. Она позволяет найти информацию об употреблении слова, идиоматического или фразеологического выражения на определенном уровне владения языком, а также их релевантном использовании в зависимости от контекста. Поисковые запросы сопровождаются примерами, взятыми непосредственно из корпуса.

С целью сужения области лексических единиц для анализа нами была использована формула удобочитаемости Дейла-Челла [32]. В большинстве своем формулы удобочитаемости используются для определения сложности синтаксиса текста по таким параметрам как средняя длина слова, средняя длина

предложения и т.д. Уникальность формулы Дейла-Челла заключается в том, что она также используется для анализа сложности лексики с точки зрения ее употребления согласно уровням CEFR. В базу данных этой формулы включен словарный минимум, который должен был сформирован у человека с начальным уровнем владения языком (A0). Вся лексика, не входящая в данный минимум, является лексикой уровня A1 и выше, и таким образом, подлежит дополнительной проверке.

Исследование проводилось в рамках количественного анализа учебников с использованием метода стратегической выборки лексических единиц. В ходе данного исследования нами были проанализированы 172 текста из 9 учебных пособий и содержащиеся в них 2536 лексических единицы. Под лексическими единицами в данном случае следует понимать как отдельные слова, так и устойчивые словосочетания (фразовые глаголы, фразеологизмы).

Для проведения анализа были выделены следующие параметры:

- 1) Наличие лексических единиц, являющихся наиболее частотными, и их количество;
- 2) Соответствие уровня лексических единиц заявленному уровню учебника по CEFR;
- 3) Количество лексических единиц, объединенных в определенные семантические поля.

Параметр наличия лексических единиц, являющихся наиболее частотными, был выделен для того чтобы определить какое количество слов от необходимого словарного минимума включено в учебник. В качестве стандарта нами был выбран Oxford Wordlist, содержащий в себе 3000 наиболее часто встречающихся слов, которые рекомендованы к изучению в первую очередь. Выбор данного списка был обусловлен тем, что он составлен на основе Оксфордского корпуса [41].

Несмотря на количественную разницу в отношении данных УМК следует отметить, что каждая из представленных линеек учебников представлена построением по принципу «прототипа», впервые описанного П. В. Сыроевым

[26, с. 97]. Принцип «прототипа» является одним из содержательных принципов построения учебника и заключается в увеличении количества часто употребляемой лексики с каждым новым повышением уровня языка обучающегося. Результаты исследования по данному параметру представлены на рисунке 1.

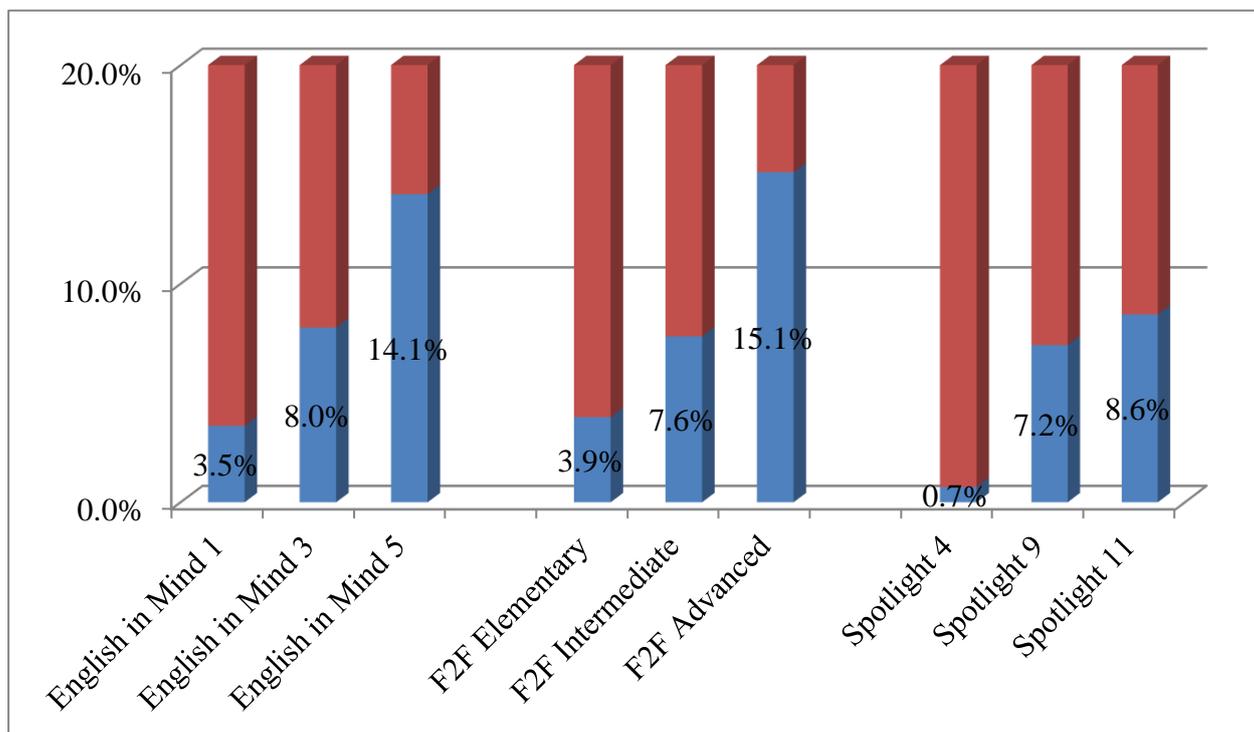


Рисунок 1 – Соотношение лексического наполнения учебника к Oxford Wordlist 3000

В ходе анализа было установлено, что количественные данные по линейкам учебников “English in Mind” и “Face2Face” являются практически идентичными (совпадают на 99,4%) в рамках всех 3 уровней сложности. Данный факт может означать, что корпусно-ориентированные учебники строятся по одной технологии и в их содержание включено примерно одинаковое количество частоупотребляемых слов. В среднем количество часто используемой лексики от Oxford Wordlist 3000 в данных сериях УМК возрастает от 3,7% до 14,5% по мере повышения уровня учебника. Учебным пособием, в котором содержится самое большое количество слов от минимального словарного запаса (15,1%), является “Face2Face Advanced”.

В отличие от корпусно-ориентированных УМК, традиционный учебник “Spotlight” характеризуется меньшим количеством часто употребляемой лексики. В среднем относительно трех уровней в учебниках данной линейки представлено на 3% меньше вышеуказанной лексики. Учебным пособием, в котором содержится наименьшее количество слов от минимального словарного запаса (0,7%), является Spotlight для 4 класса.

Таким образом, исследование показало, что в содержание корпусно-ориентированных учебники включено больше частотной лексики, чем в содержание традиционных учебников. Данный факт свидетельствует о том, что корпусно-ориентированные учебные пособия направлены на изучение учащимися фиксированного лексического минимума, необходимого для успешного осуществления коммуникации.

Параметр определения уровня лексических единиц по CEFR был выделен с целью определения степени релевантности лексического наполнения учебника его заявленному уровню. В некоторых УМК было обнаружено несоответствие между уровнем лексики, которая преобладает в учебнике и уровнем самого учебника. В качестве иллюстрации, в “Face2Face Elementary”, который, согласно CEFR, является учебником уровня A1/A1+, преобладает лексика уровня A2 (39%). В учебном пособии “Face2Face Intermediate” уровня B1/B1+, в наибольшем количестве представлена лексика уровня B1 (39%). В учебнике этой же серии уровня Advanced (C1) в основном присутствует лексика уровня B2 (31%). Таким образом, не во всех учебниках данной линейки УМК уровень лексики соответствует заявленному уровню учебника по CEFR. Учебное пособие “Face2Face Elementary” характеризуется наличием лексики с завышенным уровнем сложности, а в учебнике “Face2Face Elementary” напротив представлена лексика, которая является недостаточно сложной для данного уровня. Следовательно, лишь “Face2Face Intermediate” обладает релевантным лексическим наполнением. Результаты исследования по данному параметру представлены на рисунке 2.

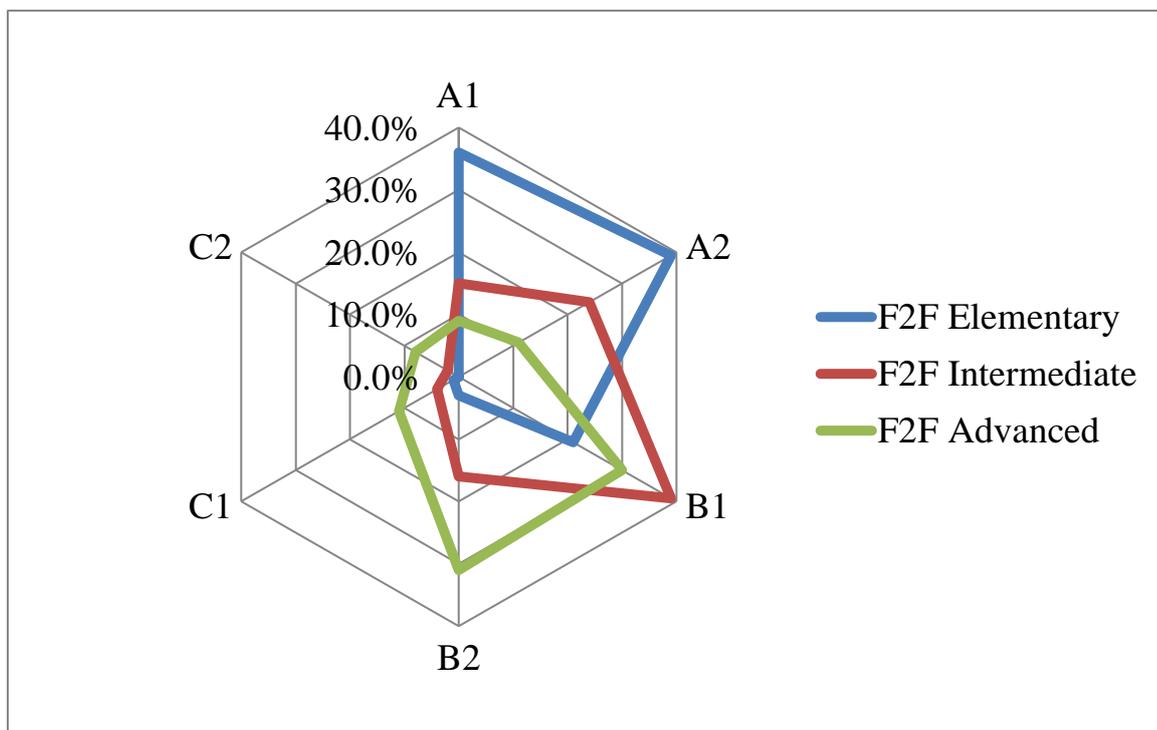


Рисунок 2 – Лексическое наполнение УМК “Face2Face” по 3 уровням

Аналогичное несоответствие между уровнем лексики, которая преобладает в учебнике и уровнем самого учебника было обнаружено при анализе УМК “English in Mind”. В “English in Mind 1”, который, согласно CEFR, является учебником уровня A1/A1+, преобладает лексика уровня B1 (37%). В учебнике “English in Mind 3”, который позиционируется как учебник уровня B1/B1+, лексика уровня B1 занимает почти половину от общего числа слов (46%). В “English in Mind 5” в основном присутствует лексика уровня B2 (31%), хотя учебник ориентирован на работу с аудиторией, владеющей языком на уровне C1. Таким образом, учебник “English in Mind 1” характеризуется наличием лексики с завышенным уровнем сложности, а в учебнике “English in Mind 5” напротив представлена лексика, которая является недостаточно сложной для данного уровня. Следовательно, лишь “English in Mind 3” обладает релевантным лексическим наполнением. Результаты исследования по данному параметру представлены на рисунке 3.

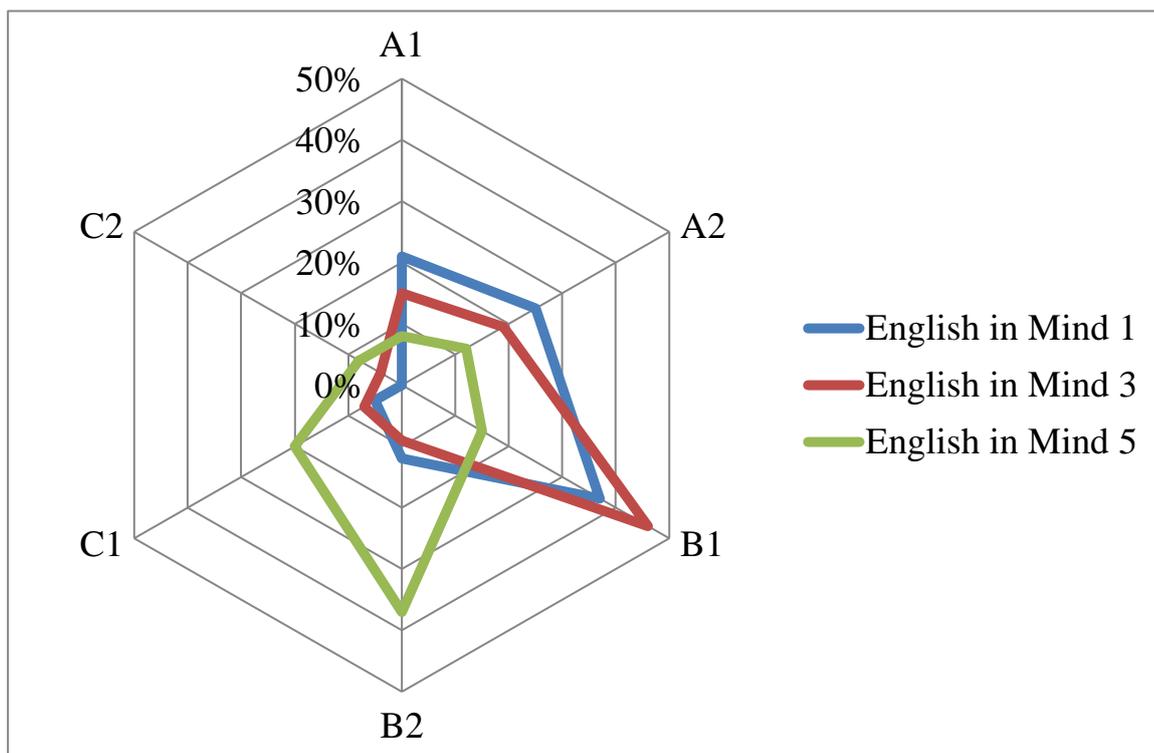


Рисунок 3 – Лексическое наполнение УМК “English in Mind”  
по 3 уровням

При анализе УМК “Spotlight”, не являющегося корпусно-ориентированным, было обнаружено, что уровень лексики, преобладающей в учебнике, полностью соответствует заявленному уровню учебника по CEFR. В “Spotlight 4” уровня A1/A1+, преобладает лексика уровня A1 (34%). В учебнике “Spotlight 9”, который позиционируется как учебник уровня B1, лексика уровня B1 занимает 35% от общего количества слов. В “Spotlight 11” уровня B2 в основном присутствует лексика соответствующего уровня (34%). Результаты исследования по данному параметру представлены на рисунке 4.

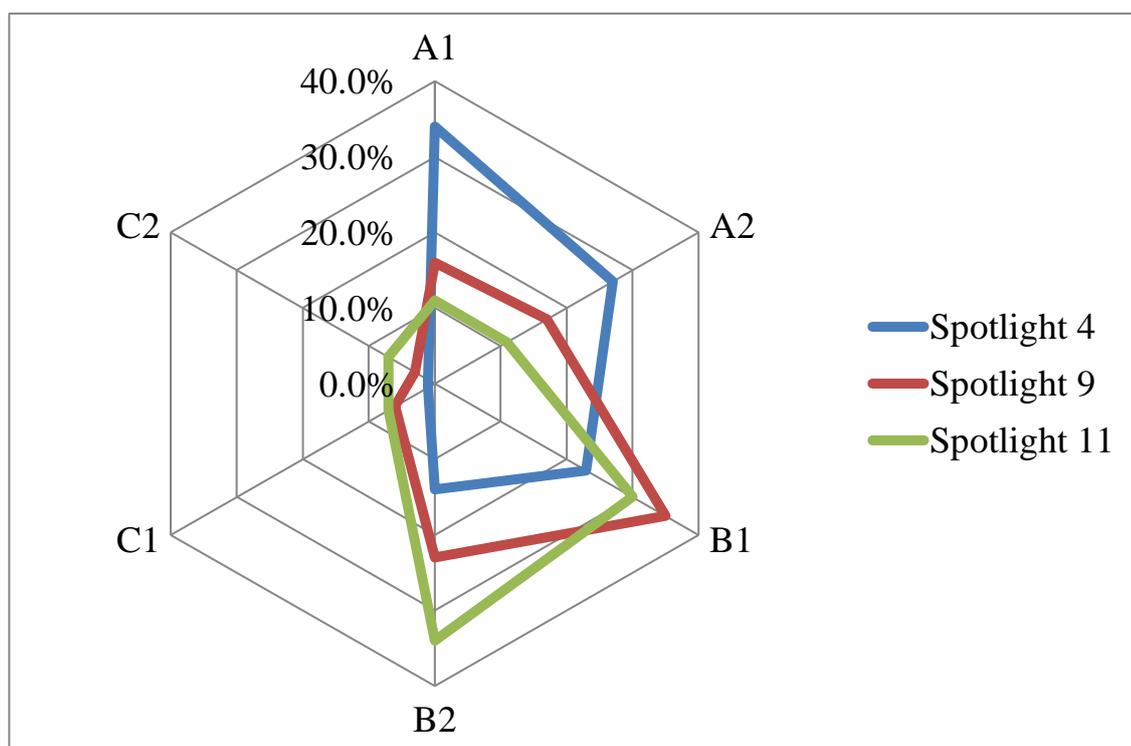


Рисунок 4 – Лексическое наполнение УМК “Spotlight”  
по 3 уровням

Во время проведения анализа корпусно-ориентированных учебных пособий нами была выявлена некоторая закономерность, при которой в учебниках уровня A1/A1+ присутствует более сложная лексика, а в учебниках высокого уровня происходит обратный процесс. С целью объяснения данной закономерности был проведен дополнительный анализ учебников на предмет повторяемости в них определенной лексики. В ходе анализа было установлено, что в рамках УМК “Face2Face” по мере повышения уровня учебника происходит в среднем повторение 22% слов. К примеру, “Face2Face Elementary” имеет в своем составе 22% лексики, которая также присутствует в учебниках следующих уровней. Аналогичный процесс происходит с УМК «English in Mind», в котором в среднем происходит повторение 20,6% слов из уровня в уровень.

Для того чтобы определить природу повторяющихся слов было решено проверить, относятся ли они к числу наиболее частотных. В качестве стандарта вновь был выбран Oxford Wordlist 3000. В ходе анализа было установлено, что в рамках УМК “Face2Face” 82% повторяющейся лексики также является

наиболее частотной. В учебниках серии “English in Mind” процент частотной лексики от числа повторяющейся составляет 86%. Полученные результаты дают основание полагать, что отклонения в лексическом наполнении корпусно-ориентированных учебников («уровень выше – лексика легче») связаны с большим процентом повторяемой лексики в рамках 3 уровней учебника. При исследовании серии “Spotlight” на предмет повторяемости лексики нами было выявлено, что лишь 6% слов повторяются из уровня в уровень; процент частотной лексики от числа повторяющейся составляет 46%. Полученный результат объясняет причину максимально точного соответствия уровня лексики заявленному уровню учебника.

Таким образом, исследование показало, что корпусно-ориентированные учебники не всегда обладают лексическим наполнением, которое соответствует уровню учебника по CEFR. Однако высокий процент повторяемости лексики в них свидетельствует о том, что данные учебные пособия направлены на закрепление частотной лексики в памяти учащихся посредством установления прочных мнемонических связей и ее постоянного повторения из уровня в уровень.

Параметр количества в учебнике лексических единиц, объединенных в семантические поля, был выделен с целью установления степени репрезентативности той или иной учебной темы в каждом из учебников. Под семантическим полем в данном случае понимается совокупность языковых единиц, объединенных общим семантическим признаком (имеющих общий нетривиальный компонент значения) [3, с. 93]. В ходе исследования был проведен анализ тематического содержания учебных пособий, в результате которого было выделено 21 семантическое поле. Для проведения количественного анализа лексических единиц каждому из семантических полей был присвоен определенный статус сложности (порогового, среднего и профессионального уровня владения) согласно дескрипторам CEFR [28, с. 97].

В ходе анализа было установлено, что некоторые из тем, например, такие как “People actions” и “Personality” прослеживались в рамках сразу нескольких

уровней. Данный факт объясняется тем, что постоянное расширение подобных тем характерно в процессе изучения английского языка в диапазоне уровней A1-C2, так как они являются наиболее общими.

При анализе учебников порогового уровня (A1, A1+) из общего количества тем было выбрано 12 обязательных к изучению на данном уровне в соответствии с CEFR. На рисунке 5 наглядно представлено, какие из тем раскрыты недостаточно репрезентативно. К примеру, в рамках УМК “English in Mind 1” это темы “Education”, “Work” и “Travel”. Аналогичная ситуация прослеживается с УМК “Face2Face Elementary” и “Spotlight 4”, где наряду с вышеупомянутыми темами также недостаточно представлены блоки “Natural world” и “Activities”. Причина этого заключается в том, что согласно CEFR развитие разговорных навыков в рамках данных тем продолжается на следующих уровнях владения языком и отличается лишь уровнем используемой лексики.

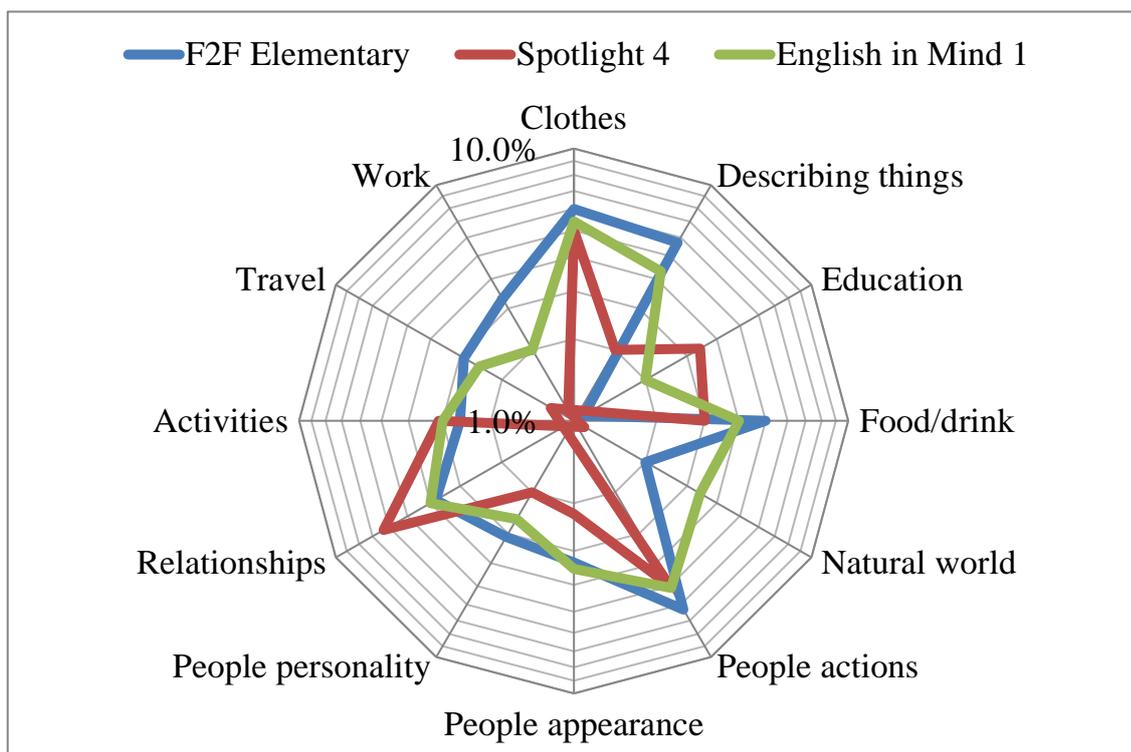


Рисунок 5 – Количество лексики в семантических полях учебников порогового уровня (A1, A1+)

Также при анализе было отмечено, что в рамках УМК “Spotlight 4” количество лексики по 8 из 12 тем (“Clothes”; “Describing things”; “Food/drink”; “Natural world”; “People actions”; “People appearance”; “Personality”; “Work”) в среднем на 1,3% меньше, чем количество лексики по этим же темам в УМК “English in Mind 1” и “Face2Face Elementary”.

При анализе учебников среднего уровня (A2+, B1) из общего количества тем было выбрано 14 обязательных к изучению на данном уровне в соответствии с CEFR. Результаты исследования представлены на рисунке 6.

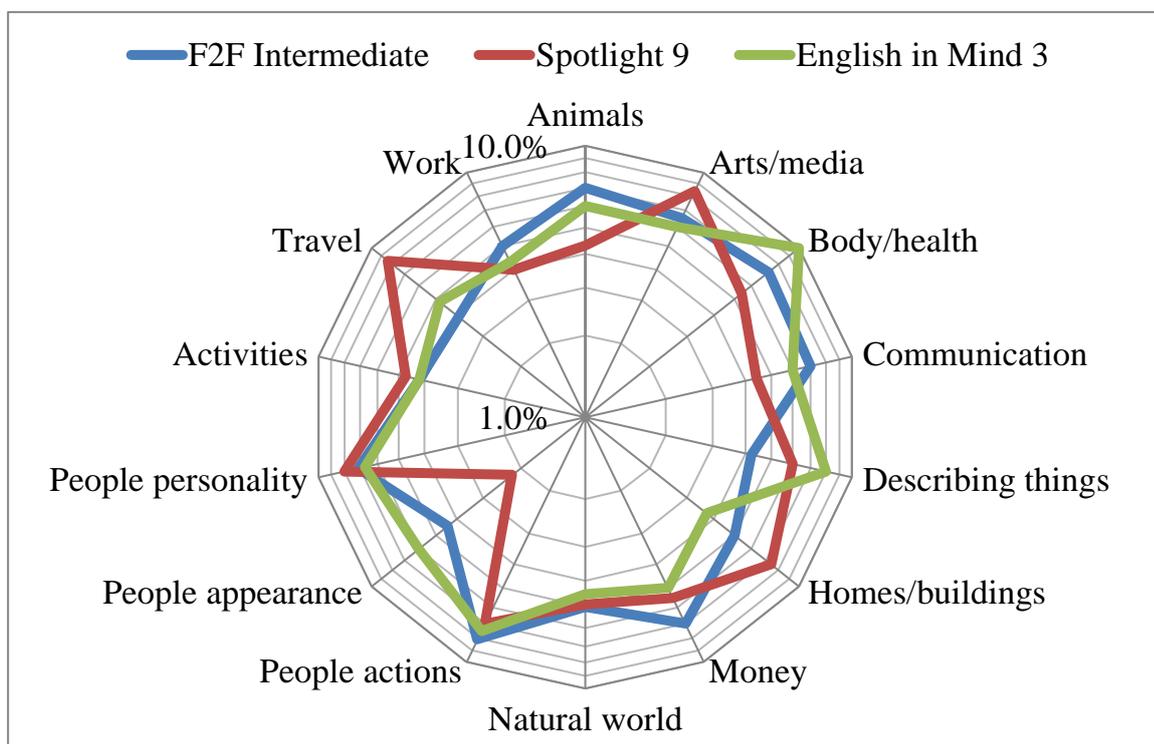


Рисунок 6 – Количество лексики в семантических полях учебников среднего уровня (A2+, B1)

В среднем по данным всех 3 УМК наблюдается рост лексики в рамках семантических полей “Activities” (1,4%), “Work” (2,3%), “Natural world” (2,7%) и “Travel” (3,6%). Как и в случае с учебниками для порогового уровня владения языком, в УМК уровня A2+/B1 наблюдается низкое количество лексики в некоторых из семантических полей. Так, в случае с учебником “Face2Face Intermediate” это темы “Travel” и “Activities”; с учебником “Spotlight 9” – “People appearance”, “Work” и “Animals”; с учебником “English in Mind 3” – “Homes/buildings”. Следует отметить, что на уровне B1 в соответствии с CEFR

должно происходить активное формирование лексических навыков в рамках данных тем. Вместе с тем факт наличия небольшого количества лексики по данным темам свидетельствует о том, о недостаточной степени репрезентативности учебных тем во всех трех УМК уровня А2+/В1.

Также при анализе было отмечено, что в рамках УМК “Spotlight 9” количество лексики по 6 из 12 тем (“Animals”; “Body/health”; “Communication”; “People actions”; “People appearance”; “Work”) в среднем на 1,9% меньше, чем количество лексики по этим же темам в УМК “English in Mind 3” и “Face2Face Intermediate”.

При анализе учебников высокого уровня (В2, С1) из общего количества тем было выбрано 8 обязательных к изучению на данном уровне в соответствии с CEFR. Результаты представлены на рисунке 7.

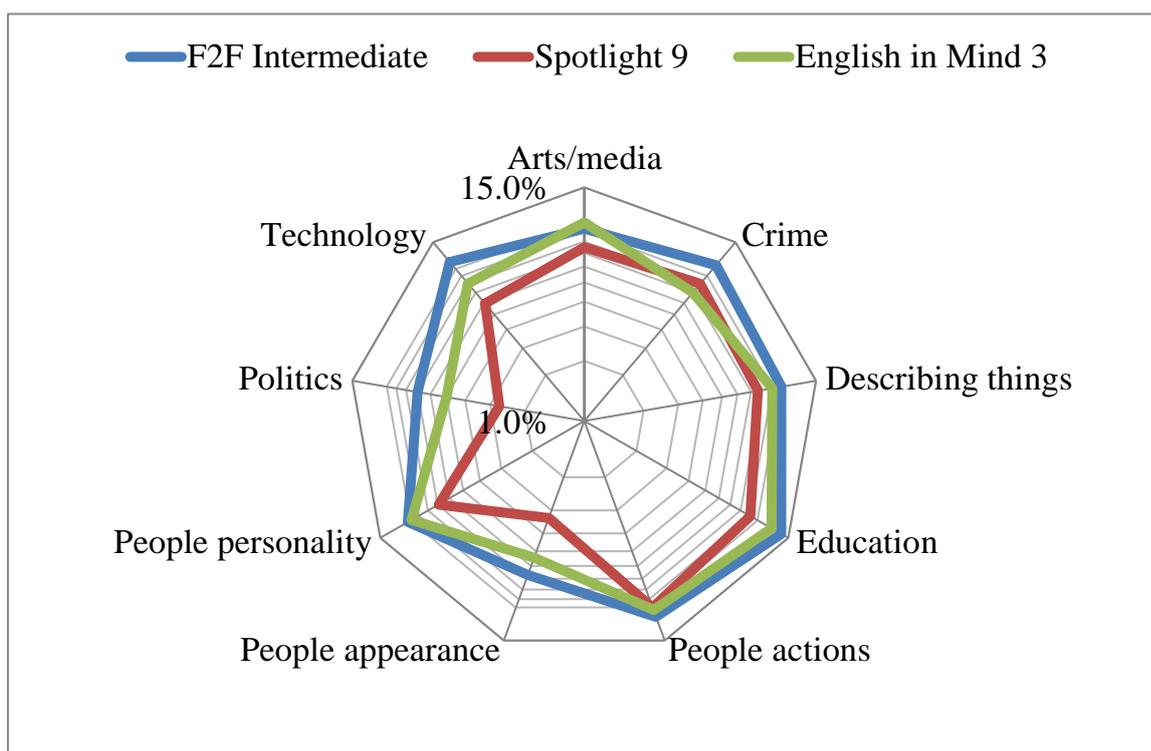


Рисунок 7 – Количество лексики в семантических полях учебников высокого уровня (В2, С1)

В среднем по данным всех 3 УМК наблюдается рост лексики в рамках семантических полей “People appearance” (0,6%), “Personality” (1,8%), “Arts/media” (1,9%) и “Describing things” (2,8%). В ходе исследования также был установлен ряд семантических полей, отличающийся наименьшей

репрезентативностью лексики: “Politics”, “Arts/media”, “Crime”, “Technology”. Было отмечено, что в рамках УМК “Spotlight 11” количество лексики по 7 из 8 тем (все, кроме “Crime”) в среднем на 2,6% меньше, чем количество лексики по этим же темам в УМК “English in Mind 5” и “Face2Face Advanced”.

Нужно сказать, что лексическое наполнение семантических полей учебников для высокого уровня владения языком представляет особый интерес в процессе подготовки учащихся к сдаче ЕГЭ по английскому языку. В качестве иллюстрации, одними из основных тем, включенных в перечень элементов содержания, проверяемых на ЕГЭ, являются «Научно-технический прогресс, его перспективы и последствия» и «Новые информационные технологии». Проанализировав содержание лексического наполнения УМК “Spotlight” для 11 класса, мы обнаружили, что семантическое поле “Technology” в нем является одним из наименее раскрытых, что означает риск недостаточной сформированности языковой компетенции по данной теме, и, соответственно, возможное снижение успешного результата при сдаче ЕГЭ.

Таким образом, степень раскрытия учебных тем в каждом из УМК варьируется согласно тематическому содержанию лексических навыков по уровням CEFR. Нужно сказать, что во всех линейках учебников в разной степени прослеживается несоответствие дескрипторам CEFR. Тем не менее, в среднем в корпусно-ориентированных учебниках наблюдается большее количество лексики в отдельных семантических полях, в сравнении с традиционными учебниками. Данный факт свидетельствует о том, что содержание корпусно-ориентированных учебных пособий направлено на более широкую репрезентацию учебного материала посредством использования тематической лексики.

Таким образом, на основе анализа традиционных и корпусно-ориентированных УМК, мы выявили следующие закономерности:

- 1) Корпусно-ориентированные учебники содержат в себе больше лексики от минимального словарного запаса в отличие от традиционных учебников. Это означает, что при создании корпусно-ориентированных

учебников авторы делают фокус на включение в его содержание самых необходимых слов, являющихся наиболее употребляемыми в речи.

2) Лексическое наполнение корпусно-ориентированных учебников в отличие от традиционных не всегда соответствует заявленному уровню учебника по CEFR. Другими словами, включение лексики в содержание данных УМК зависит не от уровня ее сложности, а от включения ее в список наиболее частотных слов. Данный факт еще раз подтверждает вывод о том, что контент корпусно-ориентированных учебников, в первую очередь, представлен частотной лексикой.

3) Семантические поля в корпусно-ориентированных учебниках раскрыты более полно, в сравнении с их репрезентацией в традиционных учебниках.

На основе вышеперечисленных положений разработан ряд рекомендаций, которые можно принимать во внимание при создании лингводидактических англоязычных материалов. К примеру, с целью повышения словарного запаса учащихся составителям УМК рекомендуется включать в учебники слова из стандартизированных списков, содержащих необходимый лексический минимум. Примерами подобных списков могут быть:

1) Oxford Wordlist, основанный на Оксфордском корпусе и включающий в себя 3000 наиболее употребляемых слов;

2) New General Service List, являющийся подкатегорией Кембриджского корпуса английского языка, и включающий в себя около 2800 часто употребляемых слов.

Следующая рекомендация применима к отбору лексического наполнения учебника. Для более прочного закрепления лексики в памяти учащихся авторам следует обращать внимание на наличие в рамках УМК блока частотных слов, повторяющегося в каждом учебнике из уровня в уровень.

Наконец, следует уделить внимание заполняемости семантических полей в соответствии с темами, предлагаемыми к изучению на определенном уровне языка. Особый акцент необходимо делать на учебники для 9-10 классов (уровни

A2+/B1+ согласно CEFR), поскольку на этой ступени обучения формируется словарный запас учащегося в рамках определенных разговорных тем.

## ЗАКЛЮЧЕНИЕ

Настоящее исследование было посвящено корпусной лингвистике и ее базисным категориям. В работе представлена широкая теоретическая база относительно данного научного направления, а именно определены ключевые понятия корпусной лингвистики, описана история становления науки о корпусах, проанализированы классификации современных языковых корпусов. Отдельно нами была подробно рассмотрена интеграция корпусной лингвистики в другие научные области. Данная отрасль языкознания находит свое отражение в различных филологических дисциплинах, таких как лексикография и переводоведение, где с помощью корпусных инструментов создаются контекстные словари и частотные списки слов.

Общеизвестно, что корпусная лингвистика является разделом компьютерной лингвистики и является продуктом развития новых информационных технологий, которые в настоящее время успешно используются в преподавании. В этой связи особого внимания заслуживает изучение синтеза корпусной лингвистики и лингводидактики, который привел к образованию такой науки как корпусная лингводидактика. В рамках данной науки особое место занимает технология разработки учебных материалов по корпусным принципам, чему и была посвящена практическая часть данной работы. Было установлено, что использование УМК, имеющих корпусную основу, способствует формированию у учащихся коммуникативной компетенции на более высоком уровне. Это связано с тем, что корпус характеризуется наличием примеров живого употребления языка и включает в себя лишь тот материал, который действительно пригодится учащимся в речи, избегая нагромождения излишней информацией.

Фундаментальной основой настоящего исследования стали результаты сравнительного анализа лексического наполнения традиционных и корпусно-ориентированных учебных пособий. Анализ проводился с помощью математического метода стратегической выборки лексических единиц.

Инструментами анализа выступили электронная система English Vocabulary Profile, построенная на основе Кембриджского корпуса английского языка и формула удобочитаемости Дейла-Челла.

Для проведения исследования были определены параметры, определяющие количество в учебнике частотной лексики, соответствие уровня сложности слов заявленному уровню учебника и процентная заполняемость семантических полей в рамках 3 линеек учебников по английскому языку.

Исследование позволило заключить, что корпусно-ориентированные учебники характеризуются наличием большего количества частотной лексики в сравнении с традиционными УМК. Также в рамках учебников, созданных по корпусным принципам, большая часть частоупотребляемых слов повторяется из уровня в уровень, что свидетельствует о том, что главной целью данных учебных пособий является постоянная тренировка и закрепление в памяти той лексики, которая в речи встречается чаще всего. Стоит также отметить, что по сравнению с традиционными учебниками тематическое содержание корпусно-ориентированных УМК раскрыто более глубоко. Семантические поля в учебниках, созданных с помощью корпуса, характеризуются наличием большого количества необходимой лексики.

На основе полученных данных был разработан ряд рекомендаций, которые можно принимать во внимание при создании лингводидактических англоязычных материалов. В число данных рекомендаций входят:

- 1) включение в учебники слов из списков с необходимым лексическим минимумом с целью повышения словарного запаса учащихся;
- 2) наличие в серии УМК блока частотных слов, повторяющегося из уровня в уровень, с целью его наилучшего запоминания;
- 3) увеличение количества лексики в семантических полях тех учебных тем, которые предлагаются к изучению в 9-10 классах общеобразовательной школы, с целью повышения качества подготовки к сдаче ЕГЭ и в целом успешного формирования словарного запаса учащегося.

В заключение, можно сделать вывод о том, что корпусный инструментарий находит свое отражение в различных областях лингводидактики, среди которых можно выделить формирование лексико-грамматических навыков учащихся, отбор лексического содержания обучения языку, реализацию проблемного подхода на занятиях и создание упражнений на основе конкорданса. Однако нужно сказать, что в первую очередь корпусные методы являются способом автоматизированной обработки данных с целью определения релевантности содержания различных источников информации. Таким образом, при решении задач в сфере оценки современных лингводидактических материалов особое значение имеют знание основополагающих принципов модели построения корпуса учебника и методологии количественного анализа его лексического наполнения.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Базарова Б. Б. Гуманитарные технологии в преподавании теории иностранного языка / Б. Б. Базарова // Отечественная и зарубежная педагогика. – М. : Изд-во Института стратегии развития образования Российской академии образования, 2013. – С. 119–121.
- 2 Баранов А. Н. Лингвистическая экспертиза текста: теория и практика : учебное пособие – 2-е изд. / А. Н. Баранов. – М. : Флинта: Наука, 2009. – 592 с.
- 3 Башарина А. К. Понятие «семантическое поле» / А. К. Башарина // Вестник Якутского государственного университета. – Якутск. : Изд-во Северо-Восточного федерального ун-та им. М. К. Аммосова, 2007. – Т. 4, № 01. – С. 93–96.
- 4 Бим И. Л. Методика обучения иностранным языкам как наука и проблемы школьного учебника / И. Л. Бим. – М. : Русский язык, 1977. – 288 с.
- 5 Волоснова Ю. А. Корпусная лингвистика: проблемы и перспективы / Ю. А. Волоснова // Вестник Московского государственного университета леса – Лестной вестник. – Москва, 2006. – № 05. – С. 43–49.
- 6 Гвишиани Н. Б. Практикум по корпусной лингвистике : учебное пособие по английскому языку / Н. Б. Гвишиани. – М. : Высшая школа, 2008. – 191 с.
- 7 Горина О. Г. Использование технологий корпусной лингвистики для развития лексических навыков студентов-регионоведов в профессионально-ориентированном общении на английском языке : дис. ... канд. пед. наук / О. Г. Горина ; Московский гос. ун-т им. М. В. Ломоносова. – Москва, 2014. – 321 с.
- 8 Захаров В. П. Корпусная лингвистика: учебник для студентов гуманитарных вузов / В. П. Захаров, С. Ю. Богданова. – Иркутск. : ИГЛУ, 2011. – 136 с.

9 Захарова Е. А. Применение результатов исследований корпусной лингвистики в обучении грамматике английского языка на продвинутом уровне / Е. А. Захарова // Вестник Российского университета дружбы народов. Серия «Русский и иностранные языки и методика их преподавания». – М. : Изд-во Российского ун-та дружбы народов (РУДН), 2016. – № 02. – С. 41–49.

10 Зверев И. Д. Школьный учебник: Проблемы и пути их развития / И. Д. Зверев // «Теория и практика создания школьных учебников»: материалы Всесоюзной конференции: вып. 20. / сост. Г. А. Молчанов. – М. : Просвещение, 1991. – С. 3–26.

11 Изотов А. И. Новые направления славянского языкознания: корпусная лингвистика / А. И. Изотов // Язык, сознание, коммуникация. – Москва, 2015. – С. 82–93.

12 Иргизова К. В. Корпусная лингвистика в отечественном и зарубежном языкознании на современном этапе [Электронный ресурс] // Огарев-online. – 2019. – №6. – Режим доступа: <http://journal.mrsu.ru/arts/korpusnaya-lingvistika-v-otechestvennom-i-zarubezhnom-yazykoznanii-na-sovremennom-etape>. – Загл. с экрана.

13 Калинина Е. История и основные понятия корпусной лингвистики [Электронный ресурс] / Е. Калинина // Изучаем Digital Humanities. – 2018. – Режим доступа: <https://dhumanities.ru/material/istoriya-i-osnovnye-ponyatiya-korpusnoj-lingvistiki/>. – Загл. с экрана.

14 Карамнов А. С. Количественная оценка повторяемости и сложности лексики в корпусе учебника английского языка / А. С. Карамнов // Филологические науки: вопросы теории и практики. – Тамбов. : Грамота, 2014. – № 06 (36). – С. 82–86.

15 Квашнина О. С. Вопросы использования корпусных данных и методов в обучении иностранным языкам / О. С. Квашнина // Филологические науки: вопросы теории и практики. – Тамбов. : Грамота, 2013. – № 08 (26). – С. 81–83.

16 Кокорева А. А. Корпус параллельных текстов в обучении иностранному языку / А. А. Кокорева // Вестник Тамбовского университета. Серия «Гуманитарные науки». – Тамбов. : Изд-во Тамбовского гос. ун-та им. Г. Р. Державина, 2013. – № 02 (118). – С. 57–62.

17 Кочетова М. Г. Социолингвистический аспект нормативности современного английского языка / М. Г. Кочетова // Филологические науки: вопросы теории и практики. – Тамбов. : Грамота, 2016. – № 10–1 (64). – С. 95–97.

18 Левинзон А. И. Введение в корпусную лингвистику [Электронный ресурс] / А. И. Левинзон // Открытое образование. – 2019. – Режим доступа: <https://openedu.ru/course/hse/CORPUS/>. – Загл. с экрана.

19 Майорова А. Д. Корпусная лингвистика: исторический и лингводидактический аспекты / А. Д. Майорова // Международный научно-исследовательский журнал. – Екатеринбург, 2017. – № 05 (59). – С. 42–46.

20 Мордовин А. Ю. Лингвистическое исследование как неотъемлемая часть применения национальных корпусов текстов в обучении иностранным языкам / А. Ю. Мордовин // Гуманитарный вектор. – Чита. : Изд-во Забайкальского гос. ун-та, 2012. – № 04 (32). – С. 274–279.

21 Нагель О. В. Корпусная лингвистика и ее использование в компьютеризованном языковом обучении / О. В. Нагель // Язык и культура. – Москва, 2018. – № 04. – С. 53–59.

22 Плунгян В. А. Корпус как инструмент и как идеология: о некоторых уроках современной корпусной лингвистики / В. А. Плунгян // Русский язык в научном освещении. – Москва, 2008. – № 02 (16). – С. 7–20.

23 Рыков В. В. Корпус текстов как реализация объектно-ориентированной парадигмы / В. В. Рыков // Труды Международного семинара Диалог-2002. – М. : Наука, 2002. – С. 124–129.

24 Савчук С. О. Национальный корпус русского языка: перспективы использования в лингвистических исследованиях и в преподавании / С. О.

Савчук // Вестник Азиатско-Тихоокеанской ассоциации преподавателей русского языка и литературы. – Владивосток, 2011. – № 02 (03). – С. 62–67.

25 Сафонкина О. С. Использование современных компьютерных технологий при обучении иностранным языкам / О. С. Сафонкина // Материалы конференции «Иностранные языки в диалоге культур: экономика, политика, образование». – Саранск, 2005. – С. 200–201.

26 Сысоев П. В. Лингвистический корпус в методике обучения иностранным языкам / П. В. Сысоев // Язык и культура. – Москва, 2010. – № 01 (9). – С. 99–111.

27 Cambridge University Press [Электронный ресурс], 2019. – Режим доступа: <https://www.cambridge.org/>. – Загл. с экрана.

28 Common European Framework of Reference for Languages: learning, teaching, assessment. – Cambridge. : Cambridge University Press: Council of Europe Publishing, 2001. – 260 p.

29 Concordance. Text analysis [Электронный ресурс], 2019. – Режим доступа: <https://concordance.apponic.com/>. – Загл. с экрана.

30 Crystal D. The Cambridge Encyclopedia of the English Language / D. Crystal. – Cambridge. : Cambridge University Press, 1995. – 491 p.

31 English Vocabulary Profile project [Электронный ресурс] : – 2015. – Режим доступа: <http://vocabulary.englishprofile.org/staticfiles/about.html>. – Загл. с экрана.

32 Free Dale-Chall Readability Formula [Электронный ресурс]. – Режим доступа: <http://www.readabilityformulas.com/free-dale-chall-test.php>. – Загл. с экрана.

33 Leech G. Computers in English language research / G. Leech, A. Beale // Cambridge language teaching surveys. – Cambridge. : Cambridge University Press, 1985. – Vol. 17(3). – P. 5–18.

34 Liu Yu. A Corpus-Based Study of College English Coursebooks / Yu. Liu // Advanced materials research, 2011. – Vol. 204–210. – P. 1990–1993.

35 McCarthy M. Corpora and grammar [Электронный ресурс] / М. McCarthy // Cambridge Assessment English / Cambridge University Press, 2019. – Режим доступа: <https://www.cambridgeenglish.org/Images/525798-corpora-and-grammar.pdf>. – Загл. с экрана.

36 McCarthy M. Language Teaching. Vocabulary / A Scheme for Teacher Education / М. McCarthy. – Oxford. : Oxford University Press, 1990. – 181 p.

37 MonoConc [Электронный ресурс]. – Режим доступа: <http://www.monoconc.com/>. – Загл. с экрана.

38 Online Qualitative Data Analysis [Электронный ресурс], 2011. – Режим доступа: [http://onlineqda.hud.ac.uk/Intro\\_QDA/index.php](http://onlineqda.hud.ac.uk/Intro_QDA/index.php). – Загл. с экрана.

39 Simpson-Vlach R. An Academic Formulas List: New Methods in Phraseology Research [Электронный ресурс] / R. Simpson-Vlach, N. C. Ellis // University of Michigan, 2009. – Режим доступа: [http://www-personal.umich.edu/~ncellis/NickEllis/Publications\\_files/AFL\\_paper\\_AppLinXPrepub.pdf](http://www-personal.umich.edu/~ncellis/NickEllis/Publications_files/AFL_paper_AppLinXPrepub.pdf). – Загл. с экрана.

40 Sinclair J. Corpus, Concordance, Collocation / J. Sinclair. – Oxford. : Oxford University Press, 1991. – 179 p.

41 The Oxford 3000 [Электронный ресурс] : Oxford Learner's Dictionaries / Oxford University Press, 2019. – Режим доступа: <https://www.oxfordlearnersdictionaries.com/wordlist/english/oxford3000/>. – Загл. с экрана.

### **СПИСОК ИСТОЧНИКОВ-ПРИМЕРОВ**

42 Английский язык. 11 класс: учеб. для общеобразоват. учреждений / О. В. Афанасьева, Дж. Дули, И. В. Михеева и др. – 2-изд., доп. и перераб. – М. : Express Publishing : Просвещение, 2009. – 244 с.

43 Английский язык. 4 класс: учеб. для общеобразоват. учреждений / Н. И. Быкова, Дж. Дули, М. Д. Поспелова, В. Эванс. – 2-изд., доп. и перераб. – М. : Express Publishing : Просвещение, 2007. – 168 с.

44    АНГЛИЙСКИЙ ЯЗЫК. 9 класс: учеб. для общеобразоват. учреждений / Ю. Е. Ваулина, В. Эванс, Дж. Дули, О. Е. Подоляко. – 2-изд., доп. и перераб. – М. : Express Publishing : Просвещение, 2010. – 216 с.

45    English in Mind 3: Student's Book / H. Puchta, J. Stranks, P. Lewis-Jones, R. Carter. – 2nd edition. – Cambridge. : Cambridge University Press, 2011. – 130 p.

46    English in Mind 5: Student's Book / H. Puchta, J. Stranks, P. Lewis-Jones. – 2nd edition. – Cambridge. : Cambridge University Press, 2015. – 131 p.

47    Face2Face Advanced: Student's Book / G. Cunningham, J. Bell, T. Clementson, C. Redston. – 2nd edition. – Cambridge. : Cambridge University Press, 2013. – 176 p.

48    Puchta H. English in Mind 1: Student's Book / H. Puchta, J. Stranks. – 2nd edition. – Cambridge. : Cambridge University Press, 2012. – 130 p.

49    Redston C. Face2Face Elementary: Student's Book / C. Redston, G. Cunningham. – 2nd edition. – Cambridge. : Cambridge University Press, 2012. – 168 p.

50    Redston C. Face2Face Intermediate: Student's Book / C. Redston, G. Cunningham. – 2nd edition. – Cambridge. : Cambridge University Press, 2013. – 176 p.

Заявление о самостоятельном характере выполнения работы

Я, Иргизова Ксения Вячеславовна, обучающийся 4 курса, направления подготовки 45.03.02 Лингвистика, заявляю, что в моей работе на тему «Использование корпусной лингвистики для разработки англоязычных лингводидактических материалов», представленной в Государственную экзаменационную комиссию для публичной защиты, не содержится элементов неправомерных заимствований.

Все прямые заимствования из печатных и электронных источников, а также ранее защищенных письменных работ, кандидатских и докторских диссертаций имеют соответствующие ссылки.

Я ознакомлена с действующим в Университете Положением о проверке работ обучающихся ФГБОУ ВО «МГУ им. Н. П. Огарёва» на наличие заимствований, в соответствии с которым обнаружение неправомерных заимствований является для отрицательного отзыва руководителя работы.

Подпись обучающегося

Дата

03.06.2019г.

Работа представлена для проверки в Системе 03.06.2019г.

Дата представления работы

подпись руководителя



## СПРАВКА

о результатах проверки текстового документа  
на наличие заимствований

Проверка выполнена в системе  
Антиплагиат.ВУЗ

Автор работы	Иргизова Ксения Вячеславовна
Факультет, кафедра, номер группы	
Тип работы	Выпускная квалификационная работа
Название работы	Использование корпусной лингвистики для разработки англоязычных лингводидактических материалов
Название файла	Диплом Иргизова.docx
Процент заимствования	4,55%
Процент цитирования	1,84%
Процент оригинальности	93,61%
Дата проверки	14:37:46 15 июня 2019г.
Модули поиска	Сводная коллекция ЭБС; Коллекция РГБ; Цитирования; Переводные заимствования; Коллекция Гарант; Модуль поиска Интернет; Модуль поиска "МГУ им. Н. П. Огарева"; Модуль поиска перефразирований Интернет; Модуль поиска общеупотребительных выражений; Кольцо Вузов
Работу проверил	САФОНКИНА ОЛЬГА СЕРГЕЕВНА ФИО проверяющего
Дата подписи	03.06.2019 Подпись проверяющего

Чтобы убедиться  
в подлинности справки,  
используйте QR-код, который  
содержит ссылку на отчет.



Ответ на вопрос, является ли обнаруженное заимствование корректным, система оставляет на усмотрение проверяющего. Предоставленная информация не подлежит использованию в коммерческих целях.

ОТЧЕТ  
о результатах проверки работы обучающегося  
на наличие заимствований

Ф. И. О. автора работы Иргизова Ксения Вячеславовна

Тема работы	Использование корпусной лингвистики для разработки англоязычных лингводидактических материалов
Руководитель	доцент кафедры английской филологии, канд. филос. наук, доцент <u>О. С. Сафонкина</u>

Представленная работа прошла проверку на наличие заимствований в системе «Антиплагиат.ВУЗ».

Результаты автоматической проверки:	оригинальность	93,61	%
	цитирования	1,84	%
	заимствования	4,55	%

Результаты анализа полного отчета на наличие заимствований:

правомерные заимствования: да, 4,55 %, наличие заимствований обосновано

корректные цитирования: да, 1,84 %, наличие цитирований обосновано

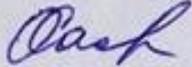
неправомерные заимствования: нет

признаки обхода системы: нет

Общее заключение об итоговой оригинальности работы и возможности её допуска к защите:

Анализ полного отчета о проверке ВКР обучающегося на наличие заимствований свидетельствует о том, что показатель оригинальности текста соответствует Положению о проверке работ обучающихся ФГБОУ ВО «МГУ им. Н. П. Огарёва» на наличие заимствований, утвержденному ученым советом университета (протокол №4 от 23.04.2018 г.), поэтому студент может быть допущен к предварительной защите и защите ВКР в ГЭК.

Руководитель  
доцент каф. английской филологии

  
«13» июня 2018 г.

О. С. Сафонкина

## ОТЗЫВ

о выпускной квалификационной работе «Использование корпусной лингвистики для разработки англоязычных лингводидактических материалов» студентки 4-го курса Иргизовой Ксения Вячеславовна

Выпускная квалификационная работа К.В. Иргизовой является завершенной работой, излагающей результаты исследования корпусной лингвистики для разработки англоязычных лингводидактических материалов. Материалом исследования послужили учебники английского языка, построенные на основе корпуса английского языка и традиционные учебники.

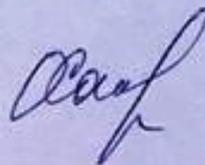
К.В. Иргизова проанализировала солидный теоретический материал. Подробно изучив принципы построения корпуса английского языка, ею была предпринята попытка разработать рекомендации для построения учебников по английскому языку, которые создаются без доступа к корпусу. В ходе работы ею были проанализированы 172 текста из 9 учебных пособий и содержащиеся в них 2536 лексические единицы.

Результаты исследования были представлены в рамках Огаревских чтений (декабрь, 2018 г.), Всероссийской студенческой научно-практической конференции «Актуальные исследования языка и культуры» (Ленинградский государственный университет им. А. С. Пушкина, Санкт-Петербург, г. Пушкин), а также опубликованы в журнале Огарёв-online (май 2019 г.) и в журнале «Нижегородское образование» (ВАК) Нижегородского института развития образования, (2019 г.). Всего автором написано 4 публикации.

Выпускная квалификационная работа состоит из введения, двух глав, заключения и списка использованных источников. Работа соответствует предъявляемым требованиям и рекомендована к защите.

Научный руководитель

к.филос.н., доцент



О.С. Сафонкина

Заведующему кафедрой английской филологии  
К. Б. Свойкину  
студента 4 курса очной формы обучения  
(на бесплатной основе)  
направления подготовки 45.03.02 Лингвистика  
факультета иностранных языков  
Иргизовой Ксении Вячеславовны

заявление

Прошу разместить мою выпускную квалификационную работу на тему «Использование корпусной лингвистики для разработки англоязычных лингводидактических материалов» в электронной библиотечной системе университета в полном объеме.

03.06.2019г.

