

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
**«БЕЛГОРОДСКИЙ ГОСУДАРСТВЕННЫЙ НАЦИОНАЛЬНЫЙ  
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»**  
(НИУ «БелГУ»)

ИНСТИТУТ МЕЖКУЛЬТУРНОЙ КОММУНИКАЦИИ И  
МЕЖДУНАРОДНЫХ ОТНОШЕНИЙ

КАФЕДРА АНГЛИЙСКОЙ ФИЛОЛОГИИ И МЕЖКУЛЬТУРНОЙ  
КОММУНИКАЦИИ

**ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ КОРПУСНОЙ ЛИНГВИСТИКИ ДЛЯ  
РАЗВИТИЯ ГРАММАТИЧЕСКИХ НАВЫКОВ У ШКОЛЬНИКОВ СРЕДНЕЙ  
СТУПЕНИ ОБУЧЕНИЯ**

Выпускная квалификационная работа

обучающегося по направлению подготовки  
45.03.04 Интеллектуальные системы в гуманитарной сфере  
очной формы обучения,  
группы 04001415  
Погореловой Марии Александровны

Научный руководитель  
к.филол. наук, доцент  
Данилова Е.С.

БЕЛГОРОД 2018

## ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ .....	3
ГЛАВА I. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ КОРПУСНОЙ ЛИНГВИСТИКИ .....	6
1.1 История возникновения корпусной лингвистики .....	6
1.2 Корпусы первого поколения .....	11
1.3 Корпусы второго поколения .....	13
1.4 Основные параметры корпуса текстов .....	16
1.4.1 Определение корпуса текстов. Основные признаки корпуса текстов .....	16
1.4.2 Типология и обзор существующих корпусов .....	19
1.5. Компьютерные программы, используемые для анализа данных корпуса ...	23
1.6. Сферы применения лингвистических корпусов .....	27
1.6.1. Использование корпусов в обучении иностранному языку .....	28
Выводы по Главе I .....	30
ГЛАВА 2. МЕТОДИКА ПРИМЕНЕНИЯ КОРПУСНЫХ ТЕХНОЛОГИЙ В ОБУЧЕНИИ ГРАММАТИКЕ АНГЛИЙСКОГО ЯЗЫКА .....	31
2.1. Особенности развития грамматических навыков на средней ступени обучения .....	31
2.1.2 Преимущества использования подхода Data-driven learning у школьников средней ступени обучения .....	34
2.2 Создание малого корпуса с помощью инструмента «SketchEngine» .....	35
2.3 Экспериментальное исследование применения корпусных методов и технологии DDL в обучении школьников средней ступени обучения .....	38
2.3.1 Процедура проведения эксперимента .....	39
2.3.2 Результаты эксперимента по применению корпусных технологий .....	41
Выводы по Главе II .....	47
ЗАКЛЮЧЕНИЕ .....	48
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ .....	49
ПРИЛОЖЕНИЕ 1 - ГРАММАТИЧЕСКИЕ УПРАЖНЕНИЯ .....	52
ПРИЛОЖЕНИЕ 2 - АНКЕТЫ .....	58

## ВВЕДЕНИЕ

На сегодняшний день английский язык является самым распространенным языком международного общения. Более того, это язык Интернета. Знание английского языка помогает человеку развиваться, преодолевать языковые барьеры, путешествовать, общаться, получать и понимать самые передовые новости из мира науки. А поскольку наша страна ведет активную внешнюю политику, государство как никогда заинтересовано в квалифицированных лингвистах, переводчиках, и просто в людях со знанием английского языка.

Как известно, человеческий мозг наиболее активно запоминает и воспринимает информацию в течение первой половины жизни. Таким образом, в случае с английским языком крайне важным является предоставление возможности освоения языка на относительно раннем этапе развития человеческой личности. Оптимальным возрастом для предоставления такой возможности является возрастной период с 10 до 16 лет, на средней ступени школьного обучения, поскольку учащийся уже владеет базовыми навыками пользователя персонального компьютера и общим представлением о языке. Введение интернет-технологий в образовательный процесс позволит преподавателю организовать различные формы учебно-познавательной деятельности и сделать более познавательной и активной самостоятельную работу учащихся.

Одной из главных проблем у учащихся является освоение грамматики английского языка ввиду отсутствия в русском языке аналогов некоторых грамматических явлений, характерных для английского языка. При этом неправильное употребление грамматических конструкций ведет к полному искажению информации.

В таком случае, на помощь может прийти одна перспективная область – это обучение на основе данных, известное как data-driven learning (далее – DDL), технология корпусной лингвистики, при котором обучающийся может использовать «сырые» данные из корпуса текстов в своей учебной деятельности, делая при этом выводы самостоятельно о значении слов и грамматических правилах.

**Актуальность** данной работы заключается в том, что грамматика английского языка остается проблемным аспектом преподавания в российских школах. Технологии корпусной лингвистики вкуче с традиционными методами обучения могут дать эффективный результат, при котором учащиеся будут заняты самостоятельной и творческой работой.

**Объектом исследования** выступают технологии корпусной лингвистики.

**Предметом исследования** является использование корпусных технологий для развития грамматических навыков у школьников средней ступени обучения.

**Цель исследования** состоит в разработке корпуса грамматических упражнений с применением методики DDL на примере учащихся девятого класса средней ступени обучения.

Цель исследования определяет основные **задачи** работы:

- Проанализировать предпосылки создания лингвистических корпусов;
- рассмотреть особенности применения технологий корпусной лингвистики в процессе обучения английскому языку;
- создать модель обучения и разработать учебные материалы, направленные на развитие грамматических навыков;
- провести педагогический эксперимент по апробации эффективности метода DDL на учащихся средней школы.

При решении вышеизложенных задач были использованы следующие **методы**: анализ и синтез теоретической информации, метод DDL, метод педагогического эксперимента.

**Теоретическую основу** изучения составили работы МакЭнери Т., Вильсона А., Захарова В.П., Плунгяна В.А., Шарова С.А., Филлмора Ч., Синклэра Дж., Тогнини–Бонелли Е.

**Теоретическая значимость** исследования заключается в том, что в работе наравне с теоретическим обоснованием необходимости внедрения корпусных технологий в процесс обучения, разработаны лингводидактические приемы работы с корпусом.

**Практическая ценность** заключается в возможности использования результатов исследования в процессе обучения грамматике английского языка в общеобразовательной школе.

# ГЛАВА I. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ КОРПУСНОЙ ЛИНГВИСТИКИ

## 1.1 История возникновения корпусной лингвистики

Традиционно считается, что корпусная лингвистика начала активно формироваться как самостоятельная отрасль в 60–х годах XX века, когда появились первые корпуса в компьютерном формате. Например, исследователь Е. Тогнини–Бонелли корпусной лингвистики связывает историю с непосредственным развитием компьютерных технологий и выделяет следующие этапы:

Этап 1 (середина 60–х – начало 80–х годов XX века) – период получения знаний об организации и создание корпусов до 1 млн. слов, характеризуется отсутствием материалов в электронном формате и необходимостью набора текстов вручную.

Этап 2 (1980–2000 гг.) делится на два периода по десять лет:

а) 1980–е годы отличились появлением сканеров, когда даже с примитивным сканером создавались корпуса в 20 млн. словоупотреблений;

б) 1990–е годы ознаменованы расширением возможностей компьютерного набора, что облегчило доступ к большим по объемам текстовым материалам в электронном формате и способствовало значительному увеличению размеров корпусов.

Этап 3 (с начала 2000 года и по сегодняшний день) – это период электронных (виртуальных) текстов, которые никогда не имели материальной формы, предоставляет огромные возможности для создания корпусов любого неограниченного размера (Тогнини–Бонелли, 2010).

В. МакЭнери и А. Вильсон значительно отдалают во времени дату зарождения корпусной лингвистики. Ученые выделяют два периода в корпусных исследованиях.

Первый период – это стадия ранней корпусной лингвистики (1910–1960), когда происходит формирование теоретической основы и прагматических предпосылок возникновения направления, создание текстовых собраний для лингвистического исследования преимущественно на бумажных носителях. Второй период (начинается с 1960) характеризуется интенсивным развитием корпусных исследований и непосредственно связан со значительным развитием компьютерных технологий.

Массово компьютеризированные корпуса стали доступными только с середины 90-х годов XX века, когда лингвистика вдруг перешла от состояния «голода по адекватным данным» (Синклер, 1991: 1) до состояния заполненной данными. Именно в этот период корпусная лингвистика окончательно сформировалась как отдельный раздел науки о языке.

Однако, если корпусная лингвистика – это изучение языка, базирующееся на примерах реального употребления (МакЭнери, Вильсон 2001: 1), то необходимо констатировать, что масштабные эмпирические исследования языка с использованием большого объема реальных языковых данных осуществлялись и в более ранние периоды развития языкознания. Наука о языке постепенно приближалась к идее текстового корпуса.

О корпусной лингвистике высказался британский писатель и лингвист Я.Свартвик, в истории корпусов сначала был так называемый «каменный век» (Свартвик, 2007: 12), или докомпьютерный период, когда корпуса создавались вручную на бумаге. Эти первые бумажные корпуса по сути были конкордансами, то есть алфавитным списком слов в их контекстном окружении. Создание таких бумажных корпусов–конкордансов занимало много времени и требовало напряженного анализа, который осуществлялся вручную. Весомую роль бумажные корпуса сыграли в таких лингвистических проектах, как создание конкорданса Библии и литературных произведений, а также написание грамматик и словарей (Мейер, 2008).

Конкорданс Библии был первым серьезным корпусным исследованием лингвистического материала. Исследователи Библии и их помощники внимательно просматривали страницу за страницей, строка за строкой и вручную индексировали все слова. Этимология латинского слова *concordantia*—согласие, соответствие (*con*— от предлога *cum* означает «с»; *cor*—«сердце») связана с идеологическим подтекстом этой трудоемкой деятельности, а именно: подтверждает, что Библия — это гармоничное божественное послание, а не сборник текстов из разрозненных источников (Маккарти, О'Киф, 2010).

Имя францисканца Антония Падуанского (1195–1231 гг.) связывается с первым известным конкордансом Библии (*the Concordantiae Morales*), созданного на основе Вульгаты (латинского перевода Библии 5 в.). Известной работой того же периода является конкорданс кардинала Хьюго де С. Каро, который в 1230 году с помощью 500 монахов—доминиканцев в монастыре Сент—Джеймс в Париже создал указатель слов Вульгаты.

Несколько позже вслед за конкордансами Библии начали появляться конкордансы литературных произведений. К примеру, конкорданс работ В. Шекспира Эндрю Беккета (*A Concordance of Shakespeare*) (1787), конкорданс работ Дж. Чосера, созданный в 1871 году, однако опубликован лишь в 1927 году.

Кроме конкордансов значительные по объему выборки текстов использовались еще для создания ранних грамматик. Как отмечает Ч. Мейер (Мейер, 2008), даже первые известные грамматики классических языков можно считать написанными на основе корпусов текстов, например, в IV веке до н.э. грамматика Панини описывала санскрит, который сохранился только в корпусе ведических текстов. Аристон Александрийский, греческий ученый I века, в своей работе описывал нерегулярные грамматические конструкции в корпусе текстов Гомера.

Ранние грамматики английского языка также базировались на классической традиции использовать в качестве примеров цитаты из реальных



текстов, особенно из произведений, как, например, грамматика Роберт Лоуса «A Short Introduction to English Grammar» (1762 г.). Одна из самых известных грамматик этого периода – семитомный труд Есперсена (1909–1949 гг.) «A Modern English Grammar on Historical Principles» также построена исключительно на примерах, отобранных из огромного количества текстов. Отто Есперсен принадлежал к той плеяде лингвистов, которые были убеждены, что лингвистическое описание должно базироваться не на вымышленных, а на реальных примерах из текстов реальной речи. Тенденция к цитированию литературных произведений как примеров с грамматическими правилами продолжилась и в грамматиках конца XIX – середины XX века такими авторами, как Дж. Руль, Н. Поутсма, Ч. Фриз. Однако не все грамматисты придерживались этой традиции.

Например, в известной грамматике Генри Суита «A New English Grammar» (1892) для иллюстрации рассматриваемых грамматических категорий используются вымышленные исследователем примеры.

Исследования больших массивов текстов также осуществлялись с целью создания словарей. Начиная со словаря Самуэля Джонсона (1755), лексикографы использовали цитаты из текстов известных писателей, чтобы проиллюстрировать значения слов и особенности их употребления.

Лексикограф собрал 150 000 иллюстративных цитат для 40000 слов словаря. Основой Oxford English Dictionary (OED), который создавался под руководством Джеймса Мюррея (1880 г.), стал пяти миллионный корпус цитат на карточках (Маккарти, О'Киф 2010; Мейер, 2008).

Многие словари мертвых языков использовали цитаты из текстов, представлявших слово в контексте, в качестве иллюстративных примеров. В современной корпусной лингвистике это называется Key Word In Context (далее – KWIC). Несмотря на то, что современные компьютеры существенно упростили поиск и классификацию примеров употребления конкретного слова в текстах разных стилей и жанров, идеи использования текстов из корпуса все

еще очень схожи с теми, что использовались ранними лексикографами и филологами, которые не имели доступа к компьютерным технологиям.

Важнейшим доэлектронным корпусом считается The Survey of English Usage, созданный Рэндольф Квирком в 1959 году.

Корпус представлял собой большую базу данных на картонных карточках, содержащих образцы речи (как письменной, так и устной) обычных граждан. Этот проект стал переходным этапом в становлении корпусной лингвистики.

Подавляющее количество исследований в области современной корпусной лингвистики начиналось на английском языке. Как отмечает С.А. Шаров (Шаров, 2003), причиной этому было не столько активное развитие компьютерных технологий в США, сколько интеллектуальный климат в британской лингвистике в 60–80–х годах XX века. В США период формирования корпусной лингвистики (60–е годы XXв.) был ознаменован расцветом трансформационной порождающей грамматики. Корпус был неодобрительно воспринят среди научной общественности и подвергся существенной критике от основателя генеративизма Н. Хомского. Исследователь назвал корпусный способ накопления языковых данных неадекватным и ложным (МакЭнери, Вильсон, 2001; Селиванова, 2008).

Ведущие лингвисты–теоретики того времени практически не поддерживали и не проявляли интереса к корпусам. В 1960–80–х годах корпусная лингвистика была вытеснена на периферию лингвистической исследовательской парадигмы США. В отличие от США, в британской лингвистике развивались эмпирические тенденции, которые предусматривали использование реальных примеров для проверки лингвистических гипотез (в первую очередь исследования Джона Фьорса и его учеников – Грегори, Синклера, Халидея и др.) (Шаров, 2003). Это и привело к созданию корпусов и появлению корпусных исследований на материале английского языка. Однако, несмотря на непопулярность корпусных исследований в США, все–таки

первый компьютеризированный корпус, ставший стандартом для многих других корпусов, был создан именно в этой стране.

## 1.2 Корпусы первого поколения

Идея создания корпуса (уже в современном его понимании) зародилась в 60–х годах XX века под значительным влиянием эмпирических исследований. К концу 1960–х существовало несколько небольших корпусов, заключенных на разных принципах. Именно достижения в сфере компьютерных технологий, а не в лингвистике дали толчок к созданию первых электронных корпусов. В середине XX века компьютеры только входили во всеобщее употребление. Это были первые примитивные машины, с которыми было трудно работать, но их огромный потенциал был сразу распознан и привлечен к лингвистическим исследованиям. Компьютеризация текстов началась с *Father Busa's Index Thomisticus* еще до 1950 года (завершена в 1978 году). Первые лингвистические корпуса текстов на машинных носителях появились в 60–х годах XX столетия. По современным стандартам они были очень маленькими, но отличались тщательно продуманной организацией.

В начале 60–х годов появились два проекта по обе стороны Атлантики по созданию корпусов в электронном формате. В университете Эдинбурга в Шотландии формировался корпус устной речи, включавший транскрибированные версии повседневных разговоров носителей британского варианта английского языка. Этот корпус небольшого размера в 300 000 слов. Причиной этому стали трудоемкий процесс сбора и транскрибирования устной речи и отсутствие компьютера в университете. В это же время в Брауновском университете (США) Нельсон Фрэнсис и Генри Кучера начали создание одномиллионного корпуса, который был назван по месту его создания Брауновский корпусом (*the Brown Corpus*). Создание корпуса преследовало цель исследовать лингвистические особенности американского варианта

английского языка. Он содержал 500 текстовых отрывков объемом по 2000 словоупотреблений, общим объемом около 1 млн. слов.

Тексты были отобраны из пятнадцати самых крупных жанров: газетные статьи (репортажи, передовицы), религиозная литература, профессиональная литература, научно–популярная литература, беллетристика, образцы деловой прозы (в том числе правительственные документы), научная литература, художественная проза, детективы и научная фантастика, приключенческая литература и вестерны, романтическая литература, юмористические рассказы и новеллы.

Составителями учитывались такие характеристики, как:

1. Происхождение и состав текста (автор должен был быть носителем американского варианта английского языка);
2. Временная принадлежность (все отобранные к корпусу тексты были впервые опубликованы в 1961 году);
3. Сбалансированное представление различных жанров;
4. Доступность для компьютерной обработки (специальные пометки для передачи графических особенностей текста).

Появление Брауновского корпуса вызвало интерес в научном сообществе. Корпус быстро превратился в популярный объект лингвистических исследований. Постепенно в процессе его использования ученые пришли к пониманию того, что провести определенные сравнения и выявить конкретные закономерности возможно только путем анализа значительных по размеру массивов текстов, организованных по определенным правилам.

Так начали проводиться новые исследования языка уже на более высоком и надежном уровне в рамках нового направления в лингвистике, которым стала корпусная лингвистика.

Брауновский корпус превратился в некий стандарт для создания других аналогичных корпусов.

70–е годы XX века были периодом упадка корпусных исследований. Кроме неприятия идей корпусной лингвистики представителями господствующего в то время лингвистического направления: порождающей грамматики, и как следствие непопулярности корпусных исследований среди большинства лингвистов, не менее важной причиной такого состояния были ограниченные возможности компьютерной техники. Компьютеры все еще были огромными. Компьютерная обработка 60000 слов в 1950 годах занимала более 24 часов.

Несмотря на эти объективные причины, в корпусной лингвистике все же были достигнуты определенные результаты. Прежде всего, создавались корпуса свыше 1 млн. словоупотреблений. Корпусным исследователям стало очевидно, что корпус в один миллион слов недостаточно для адекватного отражения лексических и грамматических свойств языка.

### **1.3. Корпусы второго поколения**

Корпусы второго поколения – это продукты Интернета и характеризуются они значительным объемом. Так, в конце 80–х годов в Великобритании был создан первый мега–корпус, задавший новый стандарт для корпусов – Британский национальный корпус (British National Corpus). Этот корпус характеризуется объемом 100 млн. слов, использованием полных текстов, а не выборки из текстов, подкорпусом устной речи (10 млн. слов), наличием разметки и доступом через Интернет.

По заданным Британским национальным корпусом стандартам были созданы корпуса многих европейских языков. Проект Банк английского языка (the Bank of English) начал развиваться в 1980–е гг. В 1989 году его объем насчитывал 20 млн. слов, а в 2012 – 650 млн. Банк английского языка – это корпус, который призван отслеживать языковые изменения путем регулярного пополнения новыми текстами и сравнениями частотных параметров, например,

таких, как изменение частоты слов и грамматических конструкций, появление новых слов и т.д.

Затем появился Collins Birmingham University International Language Database (далее– COBUILD). Этот корпус стал основой для словаря Collins COBUILD English Dictionary и ряда английских грамматик. Еще один вид корпусов представлен Международным корпусом английского языка (theInternational Corpus of English). Этот проект начался в 1990 году с целью осуществления сравнительного исследования существующих вариантов английского языка. Двадцать центров во всех уголках мира работают над созданием подкорпусов собственного национального или территориального варианта английского языка.

Успех Британского национального корпуса в исследовании британского варианта английского языка дал толчок для создания американского аналога. В 1999 году был основан консорциум Американского национального корпуса (the American National Corpus).

В 2003 году была выпущена первая часть (около 10 млн. слов) корпуса, а через несколько лет выпущена и вторая часть (около 20 млн. слов). Особой приметой Американского национального корпуса является то, что он содержит жанры, которые не появились во время создания Британского национального корпуса. Это личные электронные тексты, Интернет–чаты и веб–страницы.

В СССР таким мега–проектом стал Машинный фонд русского языка, созданный под руководством А.П. Ершова и В.М. Андрющенко (Ершов, 1986). Главными целями этого фонда были создание корпуса и подкорпусов различных жанров и разработка компьютерных программ обработки естественного языка, с целью содействия информатизации лингвистических исследований, включая создание грамматик и словарей. К сожалению, этот проект не завершился созданием собственно корпуса, хотя были собраны коллекции текстов разных типов. Сегодня из машинного Фонда доступны некоторые тексты XIX века и газетный корпус за 1997 год.

Развитие корпусной лингвистики начала XIX в. было очень бурным, что подтверждается огромным количеством новых исследований в области. О популярности и интенсивности развития корпусной лингвистики свидетельствует как большое число теоретических и практических работ, так и самих корпусов. Речь идет об исследованиях в области лексической грамматики (Стаббс, 1996; Ханстон, Францис, 2000), лексикографии и лингводидактики (МакЭнери, О'Киф 2002, Максимов, 2008), когнитивной лингвистики (Corpora in Cognitive Linguistics, 2006; Стефанович, 2004; Шмидт, 2000), прагматики и анализа дискурса (Эймер, 2004; Арчер, 2005; Бэйкер, 2005; Орпин, 2005; Вэнь, 2005), стилистики (Чартерис–Блэк, 2004; Corpus–Based Approaches to Metaphor and Metonymy, 2006; Стаббс, 2005), переводоведения (Incorporating Corpora. The Linguist and the Translator, 2008). Исследования, основанные на данных корпусов, проходят и для изучения значения слова (Партингтон, 2004), фразеологии (Лозанская, 2009), синтаксических свойств грамматических структур (Даффли, 2003). В современной России корпусную лингвистику развивают как науку такие лингвисты, как В.П. Захаров, Н.Б. Гвишиани, Е.В. Грудева, П.В. Сысоев, С.Ю. Богданова.

Таким образом, подводя итог под историей становления корпусной лингвистики, отметим, что со времени своего становления до сих пор, корпусная лингвистика прошла сложный путь от полного неприятия в 60–е годы XX века до направления современного языкознания.

Однако необходимо помнить, что несмотря на революционные возможности корпусной лингвистики, она является лишь частью многочисленного методологического инструментария современной лингвистической науки. И как отметил Ч. Филмор (Филмор, 1992), «даже значительные по объему корпуса не в состоянии отразить все возможное в языке, но при этом небольшие по объему корпуса могут предоставить информацию, которую невозможно было бы получить, не обращаясь к корпусной лингвистике.»

## **1.4 Основные параметры корпуса текстов**

### **1.4.1 Определение корпуса текстов. Основные признаки корпуса текстов**

Центральным для корпусной лингвистики является понятие «корпус текстов». Ведь с одной стороны, он выступает главной целью, которую ставит перед собой корпусная лингвистика, а с другой, предстает объектом исследования.

Целесообразность создания и использования корпусов определяется следующими предпосылками: 1) достаточно большой (репрезентативный) объем корпуса гарантирует типичность данных и обеспечивает полноту представления всего спектра языковых явлений; 2) данные разного типа находятся в корпусе в своей естественной контекстной форме, создается возможность их всестороннего и объективного изучения; 3) однажды созданный и подготовленный массив данных может использоваться многократно, многими исследователями и в различных целях (Захаров, Богданова, 2011).

Слово «корпус» широко использовалось для обозначения коллекции определенных текстов. Лишь в 1984 году Дж. Аартс и В. Мейс в статье «Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research» (Аартс, Мейс, 1984) впервые обозначили его как часть корпусной лингвистики. К концу 90-х годов термин «корпус» прочно вошел в обиход лингвистического потребления уже в его традиционном смысле как совокупность организованных электронных текстов, предназначенных для проведения конкретных лингвистических исследований. На сегодня в



корпусной лингвистике имеются несколько трактовок лингвистического корпуса текстов (англ. Linguistic corpus или text corpus).

С целью выделения основных параметров корпуса текстов приведем несколько примеров определения термина «корпус» отечественными и зарубежными лингвистами:

- «корпус – это организованное определенным образом словесное единство, элементами которого являются целые тексты или специальным образом отобранные отрывки из текстов, доступные для лингвистического анализа» (Мейер, 2004);

- «корпус – это собрание текстов, которое считается репрезентативным по отношению к языку, диалекту или иному участку языка и предназначено для использования в лингвистических исследованиях» (Францис, 1991);

- «корпус – это определенное собрание текстов, в основе которых лежит логический замысел, логическая идея, объединяющая эти тексты. Логическая идея воплощается в правилах организации текстов в корпус, алгоритме и программе анализа корпуса текстов и в связанных с этим идеологией и методологией» (Рыков, 2007);

- «корпус – это крупное по объему цифровое собрание текстов и текстовых отрывков, служит репрезентативной выборке для определенного, ограниченного на основе различных параметров использования языка, а в общелитературном контексте выступает целевой выборкой из языка» (Асмуссен, 2007).

Таким образом, анализ приведенных выше дефиниций позволяет выделить ряд признаков, отличающих современный корпус текстов от обычных коллекций текстов в электронной форме (электронных библиотек, архивов), в том числе репрезентативность, аутентичность, сбалансированность, машиночитаемость.:

1. Репрезентативность состоит в способности корпуса отражать все свойства предметной области. Предметная область для корпуса может быть

сколь угодно большой или сколь угодно малой. Так, если речь идет об авторском корпусе и лингвистическое описание касается авторской речи, то предметная область – авторская речь – не будет слишком большой, например, по сравнению с предметной областью общезыковой корпуса. «Но если задачей является построение национального корпуса, то, соответственно, предметная область будет значительно больше» (Демская–Кульчицкая, 2005).

2. Подлинность предполагает отбор реально созданного носителем языка письменного или устного текста, отрывка текста в процессе реальной коммуникации. Соблюдение требования подлинности является одной из составляющих эмпиризации фактического корпусного материала.

3. Сбалансированность заключается во введении к корпусу пропорционального количества текстовых ресурсов. На практике, где традиционно используют различные методики отбора текстового материала, одной из довольно сложных задач является достижение сбалансированности. Для достижения сбалансированности корпуса необходимы минимальные критерии отбора текстов, должны включаться различия между художественной литературой и нехудожественной литературой; книгой, журналом или газетой; нормативным и ненормативным вариантом языка; с контролем возраста, пола и происхождения авторов.

4. Машиночитаемость является определяющим признаком к современному электронному корпусу естественного языка. Кроме электронной формы, это требование предполагает наличие кодирования первичных корпусной данных и лингвистическую аннотацию, хотя сегодня это уже параметр «по умолчанию», то есть другим современный корпус не должен быть. Во-первых, к самым большим преимуществам компьютеризации корпуса относится скорость обработки и легкость в обращении с большими массивами языковых данных (поиск, отбор, сортировка, форматирование и т. д.). Во-вторых, компьютеры предоставляют точные и объективные результаты обработки машиночитаемых данных. В-третьих, компьютеры исключают

возможность влияния человеческого фактора на анализ данных, способствуют получению объективных результатов. Наконец, машиночитаемость делает дальнейшую автоматическую обработку данных корпуса, обогащает корпус разнообразными метаданными (разными типами разметок).

Задачей авторов корпуса является собрать как можно большее количество текстов, относящихся к тому языку или подмножеству языка, для изучения которого корпус создается.

Таким образом, можно утверждать, что корпус – это уменьшенная модель языка или подязыка. Одной из важных проблем современной корпусной лингвистики является определение объема корпуса, достаточного для получения надежных выводов. Так, общепринятым минимальным объемом корпуса считается количество в 1000000 словоупотреблений. Однако некоторые корпусные исследователи отмечают, что любой текст или собрание текстов теоретически может быть признано корпусом (от лат. *Corpus*– тело), а потому корпусный анализ можно осуществлять даже на очень малых по объему текстах (например, исследование М. Стаббса (Стаббс, 1996: 81–100) сосредоточено на изучении лингвистических особенностей только двух писем объемом в несколько сотен слов).

Таким образом, корпус текстов – это машиночитаемое, сбалансированное, репрезентативное собрание размеченных (аннотированных) текстов, отобранных согласно фиксированным параметрам для достижения определенной лингвистической цели.

#### **1.4.2 Типология и обзор существующих корпусов**

Современная корпусная лингвистика имеет в своем исследовательском арсенале значительное количество корпусов различных типов и размеров. Такое разнообразие существующих корпусов определяется широким спектром

исследовательских и прикладных задач, для решения которых они создаются, и особенностями языкового материала, положенного в основу.

Однако стоит отметить, что на данный момент не существует единой общепринятой типологии корпусов, а имеющиеся систематизации отличаются основополагающими классификационными критериями. Так, например, Т. МакЭнери, Р. Сяо и И. Тоно классифицируют корпуса в зависимости от исследовательской цели и задачи, и выделяются такие типы корпусов, как общие/специализированные корпуса, корпуса письменного/устного вещания, синхронические/диахронические корпуса, ученические корпуса и мониторинговые корпуса (МакЭнери, Сяо, Тоно 2006).

В типологии корпусов В.В. Рыкова выделяются такие типы:

1. По степени организации и структурированности: электронный архив; электронная библиотека; корпус; подкорпус.
2. По хронологическому признаку: синхронический; мониторинговый; диахронический.
3. По типу разметки: простой; аннотированный.
4. По языковому признаку: одноязычный; двуязычный; многоязычный.
5. По способу применения и использования корпуса: исследовательский; иллюстративный; параллельный.

Автор российского руководства «Корпусная лингвистика» В.П. Захаров предлагает свое видение классификации типов корпусов на основе набора классификационных признаков (Захаров, 2005):

Национальный корпус – это большой по объему корпус, стремится к объединению в своей структуре текстов самых разных жанров и типов (современные технологии позволяют встраивать в национальные корпуса аудио– и видеоматериалы). Чем больше размер такого корпуса, чем разнообразнее и точнее разметка, тем эффективнее программное обеспечение такого корпуса, и тем выше его ценность как лингвистического ресурса. Именно поэтому лингвисты и программисты, занятые созданием национальных

корпусов, главным образом решают вопросы материального и технического оснащения, для того чтобы материалы корпуса были доступны для работы и удовлетворения запросов пользователей.

Среди известных национальных корпусов также назовем the British National Corpus (объем 100 млн. словоупотреблений), the American National Corpus (22 млн.), Национальный корпус русского языка.

Национальным корпусам противопоставляются специальные, которые создаются для решения конкретных лингвистических задач. Специализированный корпус – это жанрово специфический корпус, имеющий целью отразить определенный подъязык. Например, the Corpus of Professional Spoken American English (CPSA) состоит из транскриптов коммуникативных ситуаций с академической и политической профессиональных областей. The Michigan Corpus of Academic Spoken English (MICASE) содержит около 1700000 словоупотреблений (около 200 часов записей) современной устной университетской речи.

Внутри этого типа корпусов можно выделить корпуса одного автора или одного жанра (например, the Epistolary Corpus of Victorian Women Writers Letters, the Shakespeare Corpus, Корпус словаря языка Достоевского).

Очень эффективными для определения уровня адекватности перевода, совершенствования систем машинного перевода и гармонизации терминологии есть параллельные корпуса, состоящие из оригинальных и переводных текстов, выровненные по предложениям.

Особый тип специальных корпусов – это так называемые корпуса ученического вещания или ученические корпуса (learnercorpora), которые создаются из устных и / или письменных текстов, спродуцированных лицами, изучающих язык как иностранный. Здесь английский термин learner, переводится лексемой «ученический», производной от существительного ученик со значением «тот, кто учится, изучает то», то есть под этим

прилагательным следует понимать человека, который учится, безотносительно к возрасту.

Такие корпуса начали создаваться еще в конце 80–х в начале 90–х годов XX века. Выделяют следующие типы корпусов ученической речи:

1) коммерческие, инициируемые издательскими компаниями, по сравнению с академическими корпусами, заключаемыми учебными заведениями;

2) большие по сравнению с малыми по объему корпусами ученической речи;

3) ученические корпуса на материале английского по сравнению с корпусами на материале других языков.

Отмечается доминирующая роль английского языка над другими языками.

4) ученические корпуса письменной по сравнению с корпусами устной речи;

По типу языковых данных корпуса делятся на корпуса письменной, устной речи и смешанного типа. Первым современным корпусом письменной английской речи был корпус американского варианта английского языка the Brown University Standard Corpus of Present Day American English. The Australian Corpus of English представляет австралийский письменный вариант английского языка.

Корпусы устной речи – это специальные коллекции тщательно отобранных текстовых отрывков (слов, фраз, предложений), «произнесенных многочисленными говорящими при различных акустических условиях» (Тьюберт, 2007). Несмотря на трудоемкость и напряженность сбора устных данных по сравнению с письменными, эти корпуса значительно меньше по объему. Среди корпусов устной речи есть the London Lund Corpus (LLC), the Lancaster / IBM Spoken English Corpus (SEC), the Cambridge and Nottingham Corpus of Discourse in English (CANCO DE), the Santa Barbara Corpus of Spoken

American English (SBCSAE) и the Wellington Corpus of Spoken New Zealand English (WSC).

Подводя итог обзору существующих корпусов, хочется отметить, что многообразие имеющихся на сегодняшний день корпусов никоим образом не ограничивается приведенными выше данными. Учитывая расширение спектра лингвистических задач, которые ставят перед собой исследователи языка, арсенал корпусной лингвистики постоянно пополняется и будет пополняться новыми, более совершенными разработками.

### **1.5 Компьютерные программы, используемые для анализа данных корпуса**

Поскольку корпусная лингвистика пользуется в своих исследованиях большими по объему и репрезентативными выборками текстов естественного языка, существует несколько типов программного обеспечения, которые могут использоваться при исследовании: коммерческие компьютерные программы – конкордансеры (LEXA, MonoConc, MicroConcord, TACT, WordSmith, WordCruncher, Manatee (Bonito), IMS CorpusWorkbench (CQP), XAIRA, LEXA, Virtual Corpus Manager (VMC), EXMARaLDA Corpus–Manager (Co–Ma)), а также компьютерные программы, разработанные для специфических процедур анализа.

Эти типы программ применяются для сложных грамматических моделей. Работа с корпусом осуществляется с помощью специальных программных средств – конкордансеров и корпус–менеджеров, предоставляющих различные возможности по получению необходимой информации из корпуса.

Корпусный менеджер – это специальная поисковая система, включающая программные средства для поиска данных в корпусе, получения статистической

информации и предоставления результатов пользователю в удобной форме. Результаты этой процедуры представляются в виде горизонтальных строк с поисковым словом посередине. Эта процедура называется KWIC. Поисковые возможности корпус–менеджеров включают в себя поиск конкретных словоформ, поиск словоформ по леммам, поиск группы словоформ в виде разрывной или неразрывной синтагмы, поиск словоформ по набору морфологических признаков и т.д., вывод результатов поиска с указанием контекста заданной длины, сохранение отобранных материалов в отдельном файле.

Использование корпусов позволяет не только изучать лексические единицы в контекстах, но и получать данные о частоте словоформ, частоту лексем, грамматических категорий, о совместной встречаемости лексических единиц, особенности их сочетаемости, управления и т.д.

Результаты поиска выдаются в виде конкорданса, где рассматриваемая единица представлена в ее контекстном окружении и сопровождается статистической информацией. Последние могут фиксировать частотные характеристики отдельных языковых единиц, или грамм, или могут характеризовать сочетаемость нескольких лексических единиц.

Для повышения эффективности исследований к корпус–менеджерам выдвигается ряд общих требований. Корпус–менеджер должен строить полные конкордансные списки; искать не только отдельные слова, но и словосочетания; осуществлять поиск по шаблонам (сложные запросы), сортировать списки по нескольким критериям; иметь возможность отображать найденные словоформы в расширенном контексте; отображать леммы, морфологические характеристики словоформ и метаданные (библиографические, типологические), зависящие от разметки корпуса; сохранять и распечатывать результаты; работать как с отдельными файлами, так и с корпусами, неограниченными по размеру; быстро обрабатывать запросы и выдавать результаты; поддерживать различные форматы текстовых данных



(txt, doc, rtf, html, xml др.), быть легким (интуитивно понятным) в использовании как для опытного, так и для начинающего; и т.д. (Захаров, Богданова 2011: 55–56).

Как утверждают авторы учебника «Corpus Linguistics: Method, Theory and Practice» Т. МакЭнери и Е. Харди (МакЭнери, Харди, 2012), «история разработки конкордансеров (т.е. корпус–менеджеров) имеет длительную историю, которая началась с деятельности Роберто Бузы . В 1951 году он впервые создал машиночитаемый корпус и получил на его основе первый автоматический конкорданс. Хотя Буза и не изобрел конкорданс, поскольку они создавались в течение долгого времени и до него, тем не менее исследователь показал, что с незначительными усилиями процедура создания конкорданса может быть быстро и качественно применена к электронным текстам.

Это стало переломным моментом, когда создание конкорданса превратилось из трудоемкого занятия в занятие, которое можно проделать с любым текстом.

Работа Р. Буза привела к появлению конкордансеров первого поколения. Разработанные программы работали только на компьютере разработчиков, поэтому применять их на других компьютерах было практически невозможно. Поэтому каждая исследовательская группа разрабатывала свою собственную систему и применяла ее к данным, к которым имела местный доступ. Конкордансеры первого поколения были функционально ограниченными. Они могли заключать только KWIC конкордансы, а другие манипуляции с данными, например создание списка всех принятых в корпус слов, выполнялось отдельными программами.

Кроме того, первые конкордансеры не различали некоторые символы, в частности диакритические знаки. С появлением первых конкордансеров как раз и возникла необходимость создания стандартных правил аннотирования корпусных данных. Опыт использования первых конкордансеров показал, что вместо однофункциональных программ, вроде генераторов частотных списков,

лучше разрабатывать программы–конкордансеры, объединяющие различные операции и позволяющие совершать манипуляции с корпусными данными.

Конкордансеры второго поколения появились с появлением машин совместимых с IBM–продукцией. Стало возможным разрабатывать конкордансеры, которые могли без проблем работать на разных компьютерах, например, the KAYE (1990), the Longman Mini–Concordance (1989), Micro–OCP (1988). Конечно, эти программы намного упростили работу корпусных лингвистов. Благодаря этому, с конца 1980–х годов корпусная лингвистика начала стремительно развиваться. Однако следует заметить, что конкордансеры второго поколения тоже имели много недостатков и по некоторым признакам были ничем не лучше конкордансеров предыдущего поколения. Кроме того, были несогласованности относительно кодирования некоторых символов, а также они не могли оперировать большими объемами данных.

Вместе с тем, растущие мощности компьютеров способствовали появлению конкордансеров третьего поколения. Конкордансеры третьего поколения включают такие хорошо известные системы, как Wordsmith, MonoConc, AntConc, Xlira. Эти конкордансеры могут обрабатывать огромные массивы данных на ПК (например, стомиллионный Британский Национальный Корпус оперируется конкордансером Xlira).

Кроме того, эти программы предлагают больше функций и делают более сложные статистические операции. Примечательно, что конкордансеры третьего поколения очень похожи друг на друга, особенно в аспекте их базовых функций: конкордансы, частотные списки, сочетаемость (коллокации), анализ ключевых слов.

Что же касается конкордансеров четвертого поколения, то они поразительно похожи в функциональном аспекте с конкордансерами третьего поколения. Однако конкордансеры этого поколения разрабатывались не с целью расширения способов анализа данных, а для решения таких задач, как ограниченная мощность стационарных ПК; проблемы, возникающие с

несовместимыми операционными системами; юридические / правовые ограничения на распространение корпусов.

### **1.6 Сферы применения лингвистических корпусов**

Репрезентативные, сбалансированные, крупные по объему корпуса предоставляют информацию о динамике процессов изменения лексического состава языка, функционирования грамматических и синтаксических единиц, позволяют осуществлять анализ употребления языковых единиц в разных стилях, жанрах, у разных авторов и т.д. Если ранние корпусные исследования сосредоточены главным образом лишь на применимости языковых единиц (частота употребления определенных существительных, глаголов и прилагательных), то с появлением мощных компьютерных технологий исследователи получили возможность определять и анализировать более сложные лексические, грамматические и синтаксические явления. На сегодняшнем этапе своего существования корпуса масштабно используются как лингвистическими, так и нелингвистическими отраслями.

Так, лингвисты–теоретики используют корпуса в качестве экспериментальной базы для проверки гипотез и своих теорий.

Прикладные лингвисты (преподаватели, переводчики и т.п.) применяют компьютерные корпуса для обучения иностранному языку и решения своих профессиональных задач.

Компьютерные лингвисты пользуются статистическими и лингвистическими закономерностями, которые обнаружены на материале корпуса, для создания компьютерных моделей языка.

В социолингвистических исследованиях корпуса используются для изучения языкового разнообразия (например, социолектов, регистров).

Другие специалисты по языку (литературоведы, редакторы) также в ряде случаев обращаются к данным корпуса. Кроме того, корпуса используются для

анализа дискурса, в литературоведении, переводоведении, судебной лингвистике. Специалисты по общественным наукам (историки, социологи) тоже могут изучать свои объекты через язык, используя такие параметры текстов, период, автор или жанр. Наконец, корпуса используются для разработки и настройки различных автоматизированных систем (машинный перевод, распознавание речи, информационный поиск).

### **1.6.1 Использование корпусов в обучении иностранному языку**

Использование корпусов становится все более популярным в обучении и изучении языка. Учителя и ученики все чаще обращаются к корпусу для осуществления быстрого разноуровневого поиска и анализа разнообразной лингвистической информации, а конкорданс признается удобным инструментом для эффективного освоения языка.

Корпуса текстов служат неоценимым ресурсом реальных примеров употребления тех или иных лексических и грамматических единиц, при условии, что ученики и учителя имеют доступ к корпусам и знакомы с техникой работы с корпусом. Обучение на основе данных (*data-driven learning*, далее –*DLL*), когда корпус становится главным источником и способом получения лингвистических знаний и умений, получает все больше сторонников среди современных студентов и учителей. При этом подходе, базирующееся на корпусе программное обеспечение, используется для поддержки интерактивной учебной деятельности, а привлечение «живого» корпуса способствует пробуждению у студентов интереса к процессу изучения языка и поощрению их к самостоятельному изучению аутентичного языкового материала. Студент выполняет активную роль исследователя языка, а учитель лишь направляет и способствует обучению (Жуковская, 2012).

Среди наиболее популярных корпусных программ и образовательных пакетов стоит отметить *The English Collocations CD-ROM* компании *Cobuild*,

WordPilot, ClozeMaker. Корпуса в качестве источника эмпирических данных играют важную роль в лингводидактике.

Признано, что традиционные школьные грамматики и учебники довольно часто проиллюстрированы искусственно созданными или отредактированными примерами употребления лингвистической единицы или явления. Такие примеры малоэффективны для успешного обучения учащихся использованию языка как средства коммуникации, поскольку изученное ими может кардинально различаться от реального функционирования языка.

В этом отношении исследования корпусов естественного языка предоставляют более точное и детализированное описание языка, что в свою очередь отражается на содержании педагогических материалов, грамматик и словарей, а также разработке учебных программ и учебников.

Другой инновационной сферой, которая непосредственно связана с корпусной лингвистикой является создание и анализ ученических корпусов (learner corpora), которые предоставляют важную для учителя информацию о том, как ученики изучают и осваивают язык, о типичных ошибках в употреблении лексических, грамматических и синтаксических единиц. Весомым аспектом для методики обучения языку является также использование параллельных корпусов, позволяющих получить доступ к непосредственным переводным эквивалентам слова, фразы, синтаксической конструкции. Такой инструмент не только облегчает создание переводных словарей, основанных на реальных примерах переводческой практики, а и способствует легкому изучению и обучению иностранного языка в соответствии с коммуникативной стратегией обучения.

Вышеперечисленные возможности использования корпусов далеко не являются исчерпывающими. С увеличением количества корпусов вклад в различные сферы и исследовательский потенциал корпусов постоянно растет, что делает их ценным лингвистическим ресурсом.

## Выводы по Главе I

В первой главе выпускной квалификационной работы рассмотрены существующие подходы к изучению корпусной лингвистики, предпосылки появления данной области научного знания, типология лингвистических корпусов. Анализ источников теоретического материала позволил выявить основные характеристики корпуса. Так, корпус представляет собой определенное собрание текстов, в основе которых лежит логический замысел, логическая идея, объединяющая эти тексты. При этом корпус должен обладать рядом признаков, в том числе репрезентативность, аутентичностью, сбалансированностью, машиночитаемостью.

Для анализа данных лингвистического корпуса необходимо использование программного обеспечения. Под Корпусным менеджером понимается специальная поисковая система, включающая программные средства для поиска данных в корпусе, получения статистической информации и предоставления результатов пользователю в удобной форме.

Результаты поиска выдаются в виде конкорданса, где рассматриваемая единица представлена в ее контекстном окружении и сопровождается статистической информацией.

На сегодняшнем этапе своего существования лингвистические корпуса масштабно используются как лингвистическими, так и нелингвистическими отраслями.

## ГЛАВА 2. МЕТОДИКА ПРИМЕНЕНИЯ КОРПУСНЫХ ТЕХНОЛОГИЙ В ОБУЧЕНИИ ГРАММАТИКЕ АНГЛИЙСКОГО ЯЗЫКА

### 2.1 Особенности развития грамматических навыков на средней ступени обучения

Согласно Федеральному Государственному Образовательному Стандарту средняя ступень обучения или основное общее образование (5–9 классы) — вторая ступень общего образования в России и в большинстве других стран, целями которого является создание условий для становления и формирования личности обучающегося, развитие его склонностей и интересов.

Развитие грамматических навыков английского языка учащегося является актуальной проблемой, поскольку требует комплексного подхода.

Российский лингвист Е.И. Пассов определяет **грамматический навык** как совокупность автоматизированных действий, обеспечивающих адекватное морфолого–синтаксическое оформление единицы речи (Пассов, 1991).

Существуют два типа грамматических навыков: речевые и синтаксические. Под речевыми морфологическими навыками подразумевают грамматические навыки, обеспечивающие правильное и автоматизированное формообразование и формоупотребление слов в устной речи. К ним относятся навыки правильного употребления в устной речи падежных окончаний существительных и т.д. (Пассов, 1991).

Синтаксические речевые навыки – это Речевые грамматические навыки, обеспечивающие стабильно правильное и автоматизированное расположение слов (порядок слов) во всех типах предложений в аналитических (английском) и флективно–аналитических языках (немецком, французском) в устной речи, в соответствии с языковыми направлениями, то есть, навыки владения основными синтаксическими схемами (стереотипами) предложений.

Процесс формирования грамматических навыков может организовываться с точки зрения различных методических подходов. В настоящее время в методике преподавания иностранных языков можно выделить следующие подходы:

1. Эксплицитный подход;
2. Имплицитный подход;
3. Дифференцированный подход (Кузнецова, 2012).

В рамках эксплицитного подхода к формированию грамматических навыков и умений можно выделить два метода: дедуктивный и индуктивный.

Название дедуктивный метод происходит от слова “дедукция”, что означает умозаключение от общего к частному. При дедуктивном методе первый этап формирования навыков и умений – ознакомление – реализуется в процессе знакомства с правилом и примерами, второй этап – тренировка – включает отработку изолированных формальных операций, третий этап – речевая практика – организуется на базе переводных упражнений.

Другим методом эксплицитного подхода является индуктивный метод. Индуктивный метод исходит из такой формы умозаключения, как индукция, предусматривающая переход от единичных фактов к общим положениям. Индуктивный метод представляет возможность самим учащимся сформулировать правило на основе явлений, с которыми они сталкиваются при изучении иностранного языка. При индуктивном методе учащиеся находят в тексте незнакомые грамматические формы и пытаются осознать их значение через контекст. Дальнейший анализ нового явления происходит путем сравнения иностранного текста с его переводом на родной язык, после чего формулируется правило. При этом если нужно, используются подсказки преподавателя или учебника. Потом следует серия упражнений на идентификацию и объяснение нового грамматического явления на актуализацию его форм.



Имплицитный подход к формированию грамматических навыков включает два метода с различными модификациями, а именно – структурный и коммуникативный. Структурными методами можно назвать ряд методов формирования грамматических навыков, разработанных различными авторами в рамках методов, которые они называли устными, активными, структурно–функциональными и т. п.

Другой разновидностью имплицитного подхода к формированию грамматических навыков являются коммуникативные методы. К коммуникативным методам относят различные варианты интенсивных методов, так называемый метод гувернантки и просто естественное погружение в языковую среду, вынуждающее к общению в коммуникативных целях. Коммуникативный метод обучения иностранным языкам, и в том числе формирование грамматических навыков, разработан Е.И. Пассовым.

Дифференцированный подход предполагает использование различных методов и приемов обучения, различных упражнений в зависимости от целей обучения, видов речевой деятельности, этапа обучения, языкового материала и особенностей учащихся.

Согласно В.П. Натальину и Д.А. Евгененко, исследовавших типичные грамматические ошибки в устной английской речи учащихся средних школ, для школьников самой трудной оказалась глагольная система английского языка, на которую приходится 41% всех ошибок. Далее по убыванию идут ошибки на употребление артиклей, соблюдения порядка слов, употребление предлогов, местоимений, образование множественного числа существительных и другие ошибки, удельный вес которых составил 14%.

Наиболее типичные ошибки, как правило, наблюдаются в простейших грамматических явлениях. Так, например, количество ошибок в употреблении временных форм, в особенности Present Indefinite и Past Indefinite, глагольной связки *to be*, достигает почти 1/3 всех зарегистрированных грамматических

ошибок. Примитивный характер ошибок очевиден в случаях неправильного употребления артиклей, предлогов, местоимений (Натальин, 2012).

Средняя ступень обучения является завершающей на этапе создания базы для активного владения английским языком. Именно на этой ступени учащиеся овладевают практически всем грамматическим материалом, необходимым для устной речи. На этой же ступени начинается формирование дальнейшего словарного потенциала, которое продолжается и на старшей ступени. Из этого следует, что возрастные особенности данной категории учащихся мотивируют использовать в процессе обучения технологии корпусной лингвистики, а именно подхода DDL, для развития грамматических навыков. Основной упор делается на профилактику распространенных ошибок носителей русского языка.

### **2.1.1 Преимущества использования подхода Data–driven learning у школьников средней ступени обучения**

Обучение на основе данных/DDL/Data–driven learning – это учебный подход, который подразумевает использование большого количества аутентичных языковых данных, извлеченных из лингвистического корпуса, для дальнейшего исследования языковых «паттернов». Впервые подход был представлен Тимом Джонсом. Его ранние работы, основанные на применении DDL в языковом классе, вдохновили исследователей использовать различные подходы к обучению, в том числе и к обучению грамматики (Браун, 2007).

Подход DDL кратко определяется лингвистом Одлином как подход, который развивает в детях любовь к самостоятельному обучению и подчеркивает их роль как исследователей, а не просто получателей знаний (Одлин, 1994).

Кроме того, лингвист Дж. Чэмберс утверждал (2005), что подход DDL показывает, что язык проще анализировать и делать выводы самостоятельно.

В результате применения DDL–подхода как дополнения к традиционным методам обучения, в классе будет царить атмосфера веселья и приключений для учеников.

Поскольку возрастные особенности школьников средней ступени говорят о том, что у учащегося наступает макрофаза индивидуализации, то подход DDL лучше впишется именно в эту возрастную категорию. Дополнительно появится мотивация и повысится самооценка.

## **2.2 Создание малого корпуса с помощью инструмента «SketchEngine»**

Система «SketchEngine» была разработана представителями компании «Lexical computing», город Брно (Чехия), и предназначена она, в первую очередь, для лингвистов, занимающихся составлением словарей. Это программное обеспечение может быть применимо для извлечения информации, необходимой при построении словарей. Иными словами, SketchEngine выполняет такие функции, как извлечение лексикографических данных, создание и сравнение корпусов.

Важно отметить, что данная программа используется при составлении словарей лексикографами многих известных издательств, в том числе «Oxford University Press», «Cambridge University Press».

Важной особенностью совокупности языковых данных, организованной в корпус, является идея, в соответствии с которой собираются эти данные. Так, в названии «Британский национальный корпус» слово «национальный» означает в первую очередь «характеризующий британский национальный вариант английского языка», т.е. идея составления корпуса заключалась в максимально полной репрезентации этого варианта языка (Плунгян, 2005: 67).

Для создания корпуса текстов, на основе которого разработана часть упражнений для учащихся 9–х классов, было отобрано 10 самых популярных книг среди подростковой аудитории (S. Collins «The Hunger Games», A. Frank

«The Diary of Anne Frank», L.M. Montgonyery «Anne of Green Gables», A.Bouen «Street Cat», J.K.Rowling« Harry Potter: Part 1», H.Lee« To Kill a Mockingbird», J.D.Salinger« The Catcher in the Rye», J.Green« The Fault in our Stars», P.Beagle «The last Unicorn», M.L'Engle «A Wrinkle in Time»). Причиной, по которой именно эти книги были выбраны в качестве основы для корпуса, стала возможность подтолкнуть учащихся к чтению литературы на английском языке для большего понимания языка и заинтересованности в процессе обучения.

Рис 2.1 ниже показывает интерфейс и возможности настройки перед компиляцией.

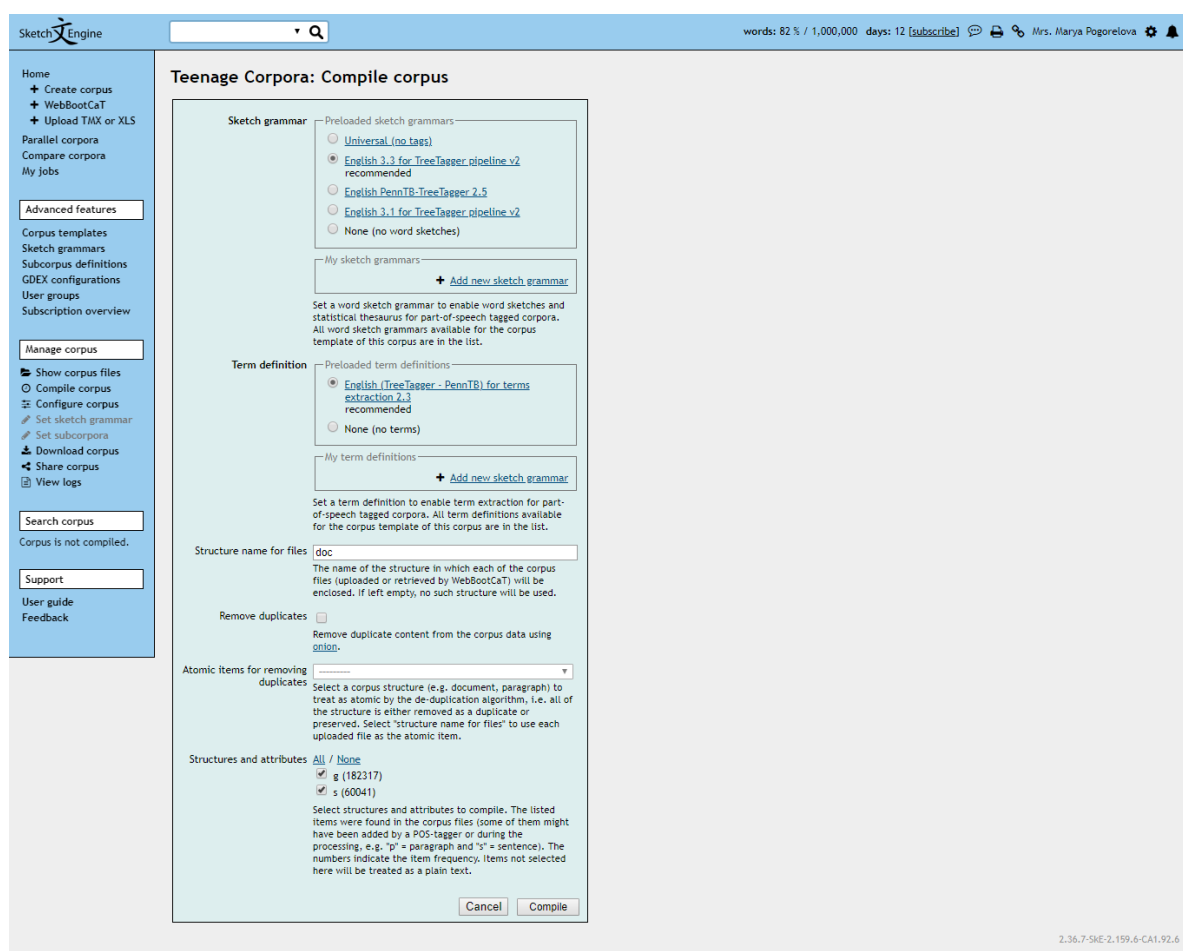


Рис. 2.1 – Интерфейс

На следующем Рис. 2.2 можно увидеть интерфейс для загрузки текстов. К сожалению, данная корпусная технология имеет ограничения в бесплатной версии.

**Teenage Corpora**  
teenage\_corpora

+ Add new file | + Add data from web using WebBootCaT | ⌚ Compile corpus | 🔍 Search corpus



































#	Original file	Plain text	Vertical	Words	Owner	
1	<a href="#">AnneFrank.txt</a>	✓	✓	102,791	Mrs. Marya Pogorelova	  
2	<a href="#">AnneGreenGamble.txt</a>	✓	✓	106,704	Mrs. Marya Pogorelova	  
3	<a href="#">Bouen A-Street-C...ob RuLit Net.txt</a>	✓	✓	65,047	Mrs. Marya Pogorelova	  
4	<a href="#">Harry1.txt</a>	✓	✓	78,447	Mrs. Marya Pogorelova	  
5	<a href="#">lee nelle to kill a mockingbird.txt</a>	✓	✓	103,588	Mrs. Marya Pogorelova	  
6	<a href="#">TheCatcherIntheRye.txt</a>	✓	✓	75,238	Mrs. Marya Pogorelova	  
7	<a href="#">TheFaultInOurStars.txt</a>	✓	✓	67,282	Mrs. Marya Pogorelova	  
8	<a href="#">TheLastUnicorn.txt</a>	✓	✓	68,631	Mrs. Marya Pogorelova	  
9	<a href="#">WrinkleinTime.txt</a>	✓	✓	49,100	Mrs. Marya Pogorelova	  
10	<a href="#">(1) The Hunger Games.txt</a>	✓	✓	99,554	Mrs. Marya Pogorelova	  

Рис. 2.2 – Интерфейс

На Рис. 2.3 показан процесс обработки и разметки пользовательского корпуса.

Sketch Engine

words: 82% / 1,000,000 days: 12 [subscribe](#)   / Mrs. Marya Pogorelova  

Home

- + Create corpus
- + WebBootCaT
- + Upload TMX or XLS

Parallel corpora

- Compare corpora
- My jobs

Advanced features

- Corpus templates
- Sketch grammars
- Subcorpus definitions
- GDEX configurations
- User groups
- Subscription overview

Manage corpus

- Show corpus files
- Compile corpus
- Configure corpus
- Set sketch grammar
- Set subcorpora
- Download corpus
- Share corpus
- View logs

Search corpus

Support

- User guide
- Feedback

**Teenage Corpora: Processing done!**

Compile corpus  100%

Word sketches  100%

Thesaurus  100%

Terms  100%

Subcorpora

✓ Compiled successfully, now you can [search the compiled corpus](#).

```
[20180618-11:05:04] >>>2:[tag=="NN" | tag=="JJ" | tag=="VVG"] I:[tag=="NN"] 3:[ic=="of"]
[20180618-11:05:04] estimated: 0 - 14173
[20180618-11:05:04] matched: 69
[20180618-11:05:04] >>>1:[tag=="NN"] 2:[ic=="of"] 3:[tag=="NN" | tag=="JJ" | tag=="VVG"]
[20180618-11:05:04] estimated: 0 - 13620
[20180618-11:05:04] matched: 38
[20180618-11:05:04] lexicon (/corpora/ca/user_data/witchlike/manatee/teenage_corpora/ter
[20180618-11:05:04] lexicon (/corpora/ca/user_data/witchlike/manatee/teenage_corpora/ter
[20180618-11:05:04] mkmap: stream sorted, #parts: 1
[20180618-11:05:04] mkmap: temporary files renamed
compiling term hashes...
24430
8913
term hashes compiled.
word sketch hashes are compiled
compiling thesaurus disabled; skipping...
compiling histograms disabled; skipping...
No parallel corpora specified in ALIGNED; skipping alignment...
Sides are compiled
compiling bilingual dictionaries disabled; skipping...
No parallel corpora specified in ALIGNED; skipping bilingual terminology...
No diachronic structure attributes specified in DIACHRONIC; skipping trends...
Checking corpus sanity disabled; skipping...
writing log to /corpora/ca/user_data/witchlike/manatee/teenage_corpora/log/compilecorp_
Term sketches compiled successfully!
```

Auto-update log

2.36.7-SKE-2.159.6-CA1.92.6

Рис. 2.3 – Процесс обработки

Таким образом, Sketch Engine представляет из себя полноценный инструмент для дальнейших исследований и может быть использован в лингводидактических целях.

### **2.3 Экспериментальное исследование применения корпусных методов и технологии DDL в обучении школьников средней ступени обучения**

Одной из главных проблем в обучении английскому языку, с которой могут столкнуться учителя, является недостаточная мотивация учащихся к изучению грамматики. Многие школьники находят освоение грамматического модуля скучным и сложным. Преподавание английского языка с годами становится все более ориентированным на решение конкретных речевых проблем, но это не мотивирует учащихся. Более того, многие студенты испытывают трудности с применением грамматических правил на письме несмотря на то, что обучение английскому языку начинается с начальной ступени школьного образования.

В данном исследовании участвовала группа учащихся 9-х классов Муниципального общеобразовательного учреждения «Дубовская средняя общеобразовательная школа Белгородского района Белгородской области с углубленным изучением отдельных предметов», обучающихся по учебно-методическому комплекту (далее УМК), для 9 класса «English» под редакцией Кузовлева В.П., который создан на основе «Программы-концепции коммуникативного иноязычного образования: Концепции развития индивидуальности в диалоге культур», разработанной профессором Е.И.Пассовым. За основу были взяты грамматические аспекты, а именно временные формы глагола, поскольку именно эта сфера вызывает наибольшие трудности в овладении языком.

При подготовке материала, перед нами возникла проблема представления данных. при этом учитывались такие аспекты как уровень общеязыковой подготовки испытуемых, отсутствие опыта использования корпусных и интернет-технологий в обучении английскому языку, возраст испытуемых, новизна подхода, а также отсутствие выхода в сеть Интернет. Поэтому в качестве подачи на первое занятие был выбран раздаточный печатный материал (Приложения 1–6) с заданиями на основе конкорданса из мини-корпуса. Второе занятие предполагает работу в компьютерном классе с корпусом British National Corpus (далее - BNC). Длительность одного занятия – 40 минут.

### **2.3.1 Процедура проведения эксперимента**

Для эксперимента был применен подход с созданием контрольной и экспериментальной группы. Однако такой подход не всегда рационален. Паусон и Тайли (1997) отмечают существование проблемных аспектов (например, эффект Хоторна – условия, в которых новизна, интерес к эксперименту или повышенное внимание к нововведению приводят к искажённому, зачастую благоприятному результату). В эксперименте ученики оценивались с двух разных точек зрения: изменения их отношения к изучению грамматики и изменения в грамматических навыках. Отношение и восприятие учениками исследовалось с помощью анкет со списком вопросов. Уровень общеязыковой подготовки учащихся варьировался от А2 до В1.

На Рис. 2.4 можно увидеть наглядно результаты опроса учащихся касательно их отношения к школьному курсу грамматики английского языка.

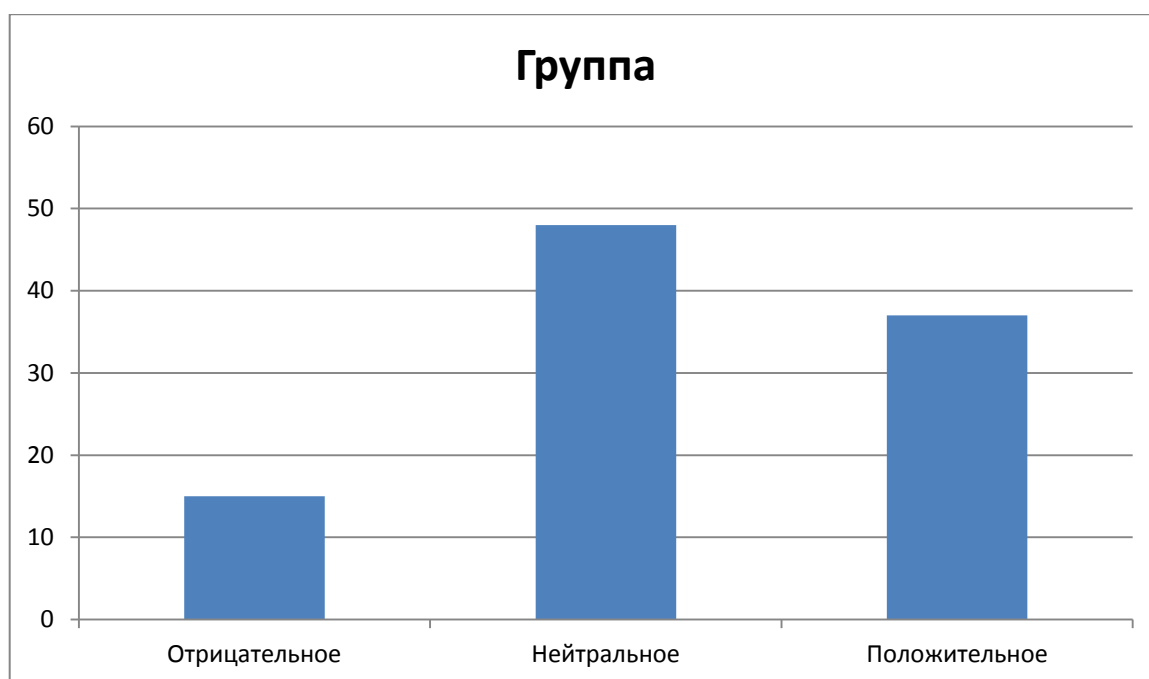


Рис. 2.4 – Отношение учащихся к школьному курсу грамматики

Вопросы в Анкете №1 (Приложение 2) (выдается перед экспериментом) акцентируют внимание на опыте, отношении, ожиданиях к курсу английской грамматики в школе. Вопросы в Анкете №2 (Приложение 2) (выдается после эксперимента) посвящены отношению, которое сформировалось у учащихся после эксперимента, а также тому насколько оправдались их ожидания после работы с корпусом текстов.

Кроме того, по прошествии двух занятий, с каждым учащимся дополнительно был проведен опрос на тему полученного опыта работы с корпусом текстов. В опросе внимание сфокусировано на вопросах, которые не были достаточно освещены в анкетах. Например, были заданы вопросы о том, насколько учащимся были понятны инструкции и пояснения для работы с упражнениями и корпусом текстов.

Копии Анкеты №1 были выданы участникам обеих групп.



## 2.3.2 Результаты эксперимента по применению корпусных технологий

### Занятие №1

На первом занятии учащимся обеих групп были рассказаны основы работы с корпусом текстов и его роль в процессе обучения английскому языку. Однако краткий устный опрос подтвердил, что у многих учащихся проблемы с пониманием этой темы. Кроме того, на эффективность занятия повлияли внешние факторы: затянувшееся введение в тему занятия и ответы на многочисленные вопросы, повышенная нагрузка у учащихся в связи с ОГЭ (из-за чего приоритетность уроков английского языка падала), низкий уровень знаний.

### Опрос №1

Была выдвинута гипотеза, что учащиеся в экспериментальной группе после ознакомления и практики с раздаточными материалами (Приложение 1) на базе корпусных технологий будут положительнее воспринимать курс грамматики английского языка. Процентное соотношение учащихся, оценивающих урок грамматики положительно к учащимся, которые оценивают отрицательно, равно 67% (8 человек из 12) к 33% (4 человека из 12). В то время как в контрольной группе соотношение представляется следующим: 83% (10 человек из 12) положительно относятся, а 17% (2 из 12) негативно. В основном, после первого занятия никто из учащихся не поменял своего мнения касательно нововведений в учебный процесс. На Рис. 2.5 можно увидеть результаты наглядно.

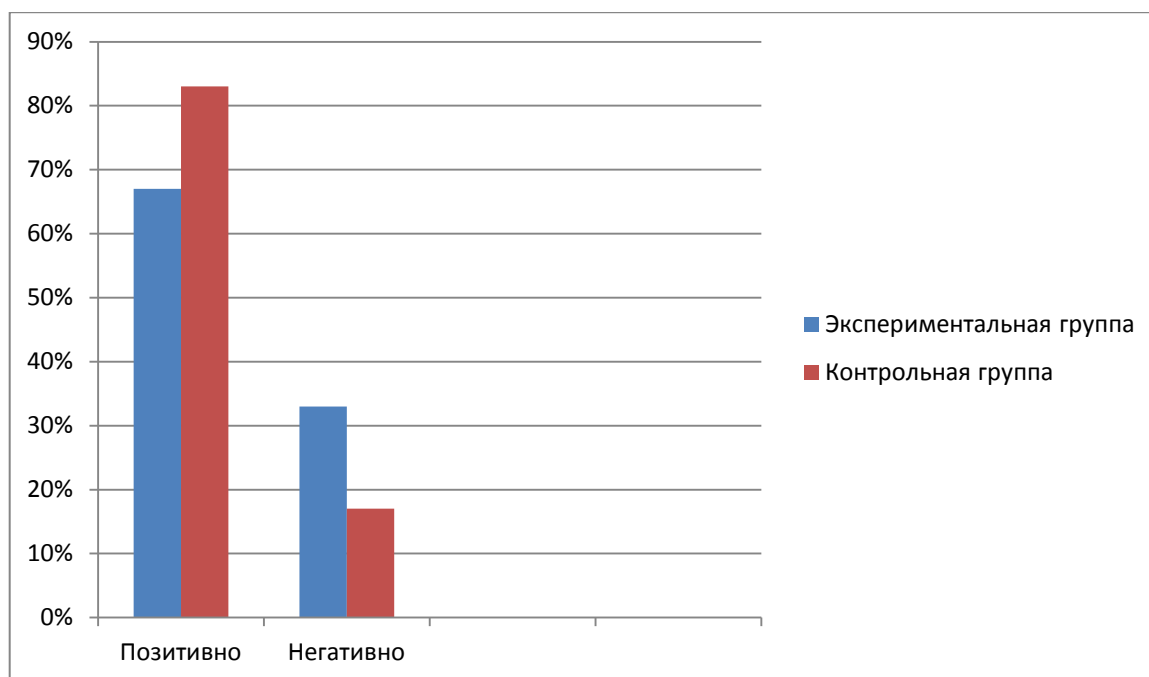


Рис. 2.5 – Отношение учащихся после первого занятия

Несмотря на то, что гипотеза не подтвердилась, учащиеся, в целом, положительно относятся к изучению грамматики английского языка. Стоит упомянуть и то, что никто из учащихся в экспериментальной группе, положительно оценивающих свое отношение, не отметил работу с раздаточным материалом и конкордансом как причину, вместо этого они акцентировали внимание на подачу материала учителем.

Один из вопросов анкеты был о том, в какой степени ожидания учащегося по улучшению навыков владения английским языком, пониманию грамматических правил, умению объяснять, были оправданы. Была выдвинута гипотеза, что ожидания учащихся в обеих группах будут выполнены. Обе группы были удовлетворены, но в контрольной группе результаты оказались лучше. В этом случае, гипотеза тоже не получила поддержки фактами, что показано на Рис.2.6.

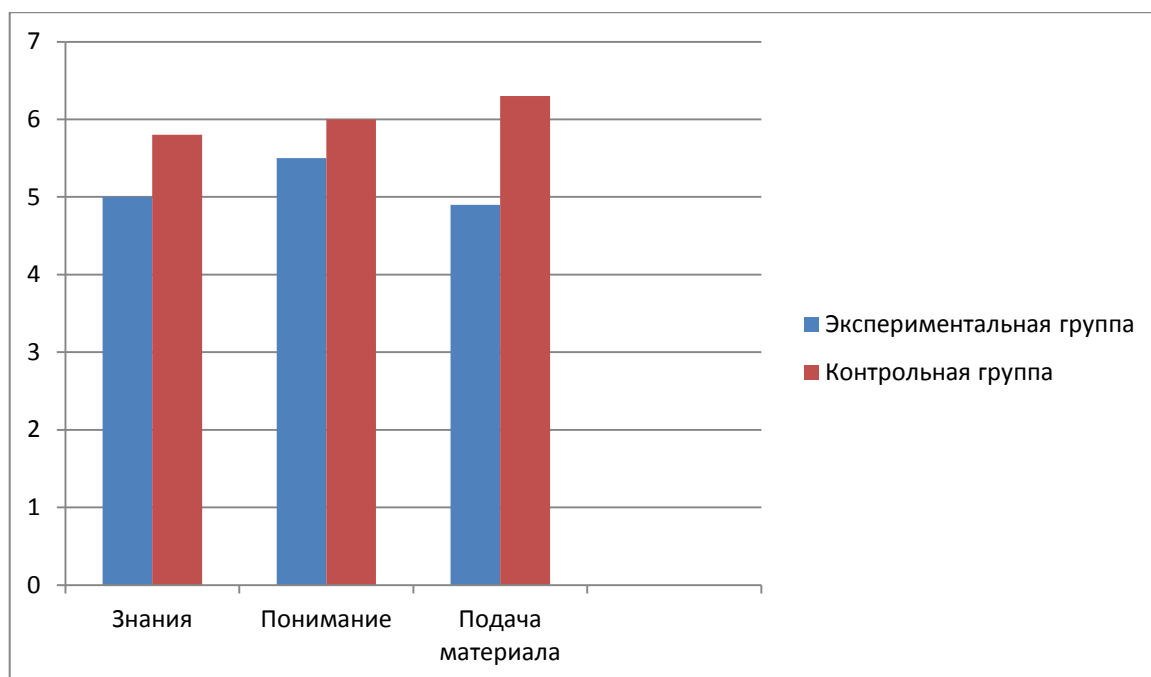


Рис. 2.6 - Степени ожидания

От учащихся экспериментальной группы были получены как положительные, так и отрицательные комментарии по окончании занятия:

*Ничего не понятно.*

*Нужно много времени, чтобы разобраться.*

*Было интересно работать с конкордансом.*

*Очень интересные задания, где нужно думать.*

*Мне понравилось.*

*Эти упражнения интереснее, чем в учебнике.*

*Тест был непростой*

Предполагалось, что грамматические навыки учащихся в экспериментальной группе после первого занятия будут выше, чем у учащихся в контрольной группе. Следует сказать, что после подсчета результатов тестирования, средний балл в экспериментальной группе составлял 63,2 из 100, а в контрольной группе 68,1 из 100.

Что касается результатов мини-сочинения на тему «Краткий сюжет любимой книги/фильма», возникало предположение, что результаты экспериментальной группы будут лучше. Мини-сочинение оценивалось по

стандартной оценочной шкале от 2 до 5. Отметки 4 и 5 – положительные (не больше 2–х ошибок, разнообразная лексика), отметка 3 – удовлетворительно (не больше 4–х ошибок, простейшая лексика), отметка 2 – неудовлетворительно (5 и больше ошибок, затруднительно понять текст). Экспериментальная группа получила средний балл – 3,2, а контрольная – 3,4. Процент отметок неудовлетворительно среди экспериментальной группы составил – 33% (4 из 12), а в контрольной – 25% (3 из 12). И снова контрольная группа показала лучший результат.

С одной стороны, негативный результат, возможно, обусловлен объективными факторами. С другой стороны, большинство учащихся положительно отнеслось к идее использования корпусов и технологий корпусной лингвистики в будущем.

### **Занятие №2**

Как было упомянуто выше, работа с корпусами проводилась во второе занятие с обеими группами. В этот раз учащиеся работали непосредственно в компьютерном классе с корпусом текстов. Перед учителем ставилась цель не объяснить материал, а помочь учащимся в освоении интерфейса корпуса. Более того, было изменено упражнение, оно стало проще, а учащимся было предложено самим формулировать запрос (Приложение 3).

### **Опрос №2**

Большинство учеников положительно оценили второе занятие. Из группы в 24 человека, 67 % (16 из 24) положительно оценили занятие, 16,5% (4 из 24) нейтрально оценили и оставшиеся 16,5% (4 из 24) оценили негативно. Хотя отношение к применению корпуса и приняло более позитивную направленность, ученики все равно отмечают, что главным фактором в развитии их грамматических навыков на занятии стал учитель. На Рис. 2.7 можно увидеть описанную ситуацию.

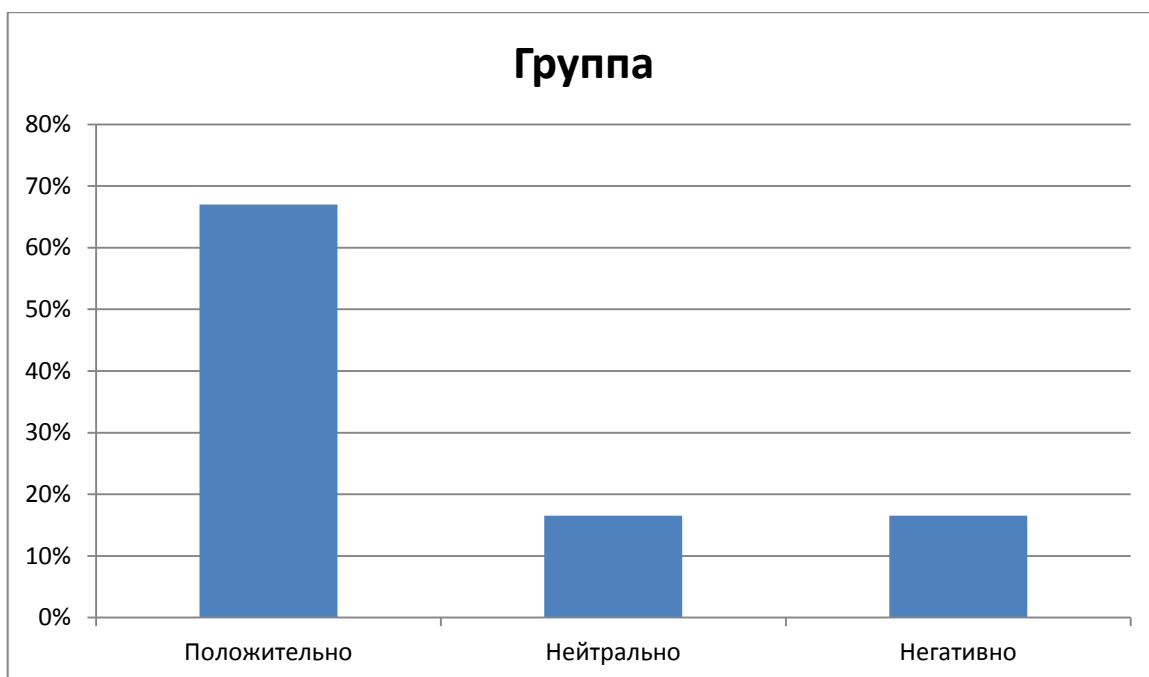


Рис. 2.7 – Оценка второго занятия

По шкале от 1 до 10 из опроса, средний бал оценки реализованных ожиданий варьируется от 6,8 до 8,3. На Рис. 2.8 наглядная диаграмма.

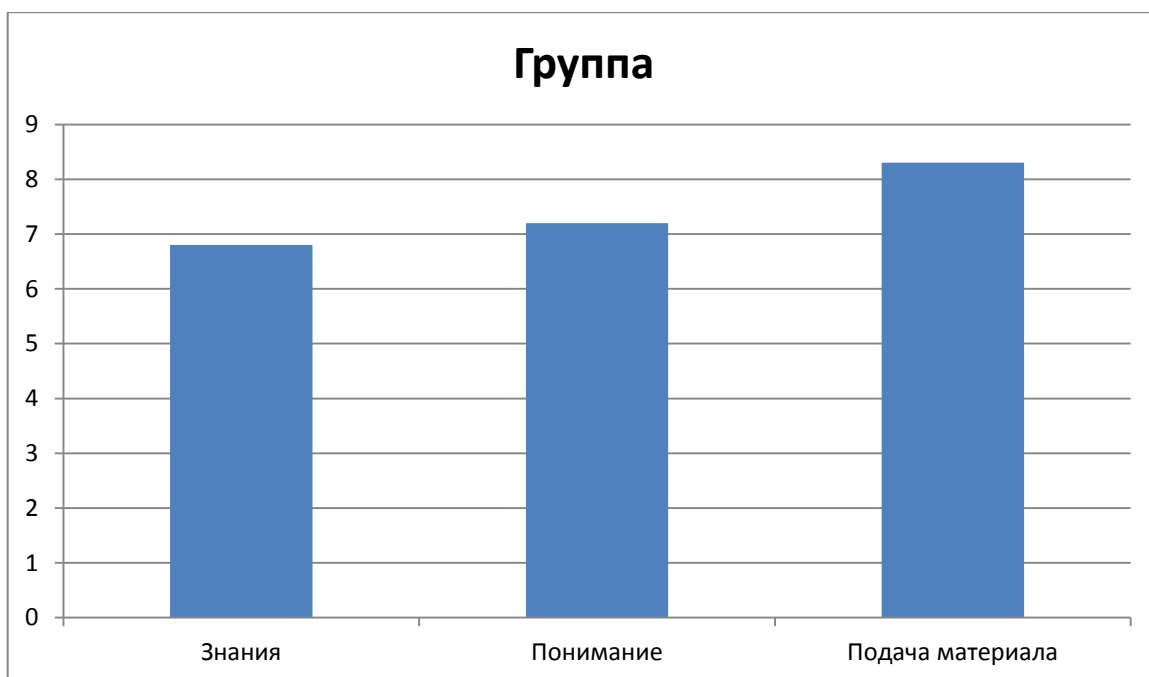


Рис. 2.8 – Оправданные ожидания после двух занятий

В заключении отметим, что опрос среди учеников на вероятность использования корпуса и технологий корпусной лингвистики в будущем показал, что большинство положительно относятся к дальнейшему использованию. На Рис. 2.9 виден положительный рост.

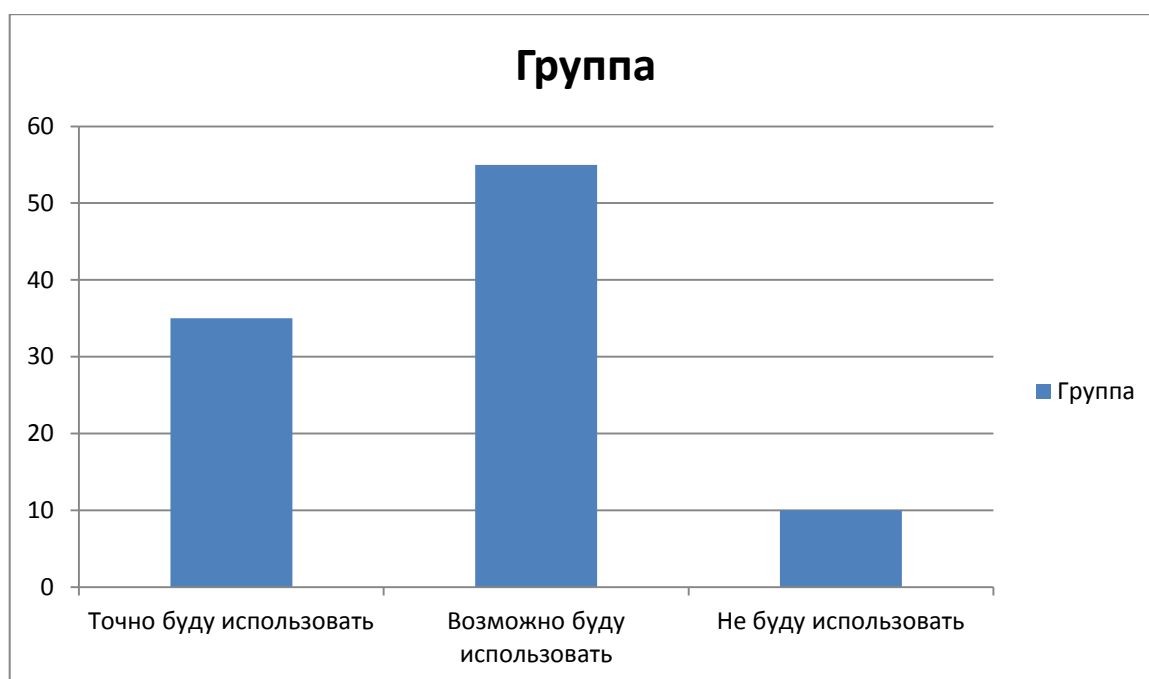


Рис. 2.9 – Оценка возможного использования

### **Интервьюирование**

Интервью, взятые у учеников после проведения эксперимента, показывают весьма неоднородное отношение к использованию корпусов в процессе обучения, но есть общие аспекты. Во-первых, корпус BNC, интерфейс и мануал по использованию полностью на английском языке. Это привело к тому, что ученики предпочитали корпусу традиционное учебное пособие. Во-вторых, для многих учеников было трудно формулировать запросы и работать со всем функционалом корпуса. Помимо двух этих аспектов, ответы кардинально различались в оценке, часть учеников крайне положительно оценила работу с корпусом и назвала этот опыт интересным, другая часть предпочла традиционный вид обучения грамматике с УМК. Некоторые ученики находили упражнения слишком простыми, некоторые наоборот слишком сложными. Некоторые ученики не понимали, как интерпретировать результаты поиска в корпусе. Они указали на то, что в корпусе представлены только примеры из текстов, нет грамматических правил. Большинство учеников сказали, что возможно будут использовать корпус, но им нужна помощь в изучении всех инструментов корпуса.

## Выводы по Главе II

Вторая глава выпускной квалификационной работы посвящена введению корпусной технологии DDL в процесс обучения грамматике английского языка учеников 9–го класса средней ступени обучения. Выявлено, что, основываясь на опыте исследователей DDL, эта технология должна благотворно влиять на учеников средней ступени обучения, а также мотивировать их на новые исследования.

Как результат, были проведены два занятия, на которых с помощью корпусных технологий пытались улучшить грамматические навыки учеников 9–го класса. Также был проведен опрос и выяснили какого мнения придерживаются ученики касательно технологий корпусной лингвистики. Несмотря на довольно осторожную оценку учебного процесса со стороны учащихся, можно сказать, что в целом DDL получила положительные отзывы.

## ЗАКЛЮЧЕНИЕ

Анализ источников теоретического материала и результатов практического эксперимента с использованием технологий корпусной лингвистики позволил всесторонне рассмотреть виды и особенности лингвистических корпусов, возможности и преимущества их применения в процессе обучения английскому языку.

А также создать грамматические упражнения на основе данных, полученных из корпуса и применить методики DDL на учащихся 9–го класса средней ступени обучения. Для достижения данной цели была изучена теоретическая литература по проблематике исследования, для углубления в проблему и лучшего понимания материала. Была рассмотрена история возникновения корпусной лингвистики, типы корпусов, их признаки и сферы применения.

Во второй главе были отмечены особенности преподавания грамматики английского языка школьникам средней ступени обучения, а также особенности подхода DDL. На базе собственного мини–корпуса, введена новая методика в учебный процесс, а также проанализирована мотивацию обучаемых и влияние применения корпусной лингвистики на результаты обучения.

Перспективы использования корпусных технологий видятся не только в преподавании и создании методического материала, но и потому, что корпус может быть полезен ряду специалистов: программистам, редакторам, лингвистам и т.д.



## СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Богданова, О. С. Лингводидактика и методика преподавания иностранных языков в схемах и комментариях: текст лекций /О.С. Богданова. . –Краснояр. гос. пед. ун-т. им. В. П. Астафьева. – Красноярск, 2011. – 176 с.
2. Богданова, С.Ю. О разработке учебного мультимодального корпуса текстов. [Текст] / С.Ю. Богданова // Материалы XLII международной филологической конференции. Секция прикладной и математической лингвистики. С. –Петербург, 2013 – С.27-33.
3. Гвишиани, Н.Б. Практикум по корпусной лингвистике: Учеб. пособие по английскому языку/ Н.Б. Гвишиани. — М.: Высшая школа, 2008. — 191 с.
4. Горчев, Ю. А. Объекты и приемы контроля в обучении грамматической стороне устной речи /Ю.А.Горчев. . –М.:, 1986. . –28 с.
5. Захаров, В.П. Корпусная лингвистика: Учебно–метод. пособие. / В.П. Захаров – СПб., 2005. – 48 с.
6. Захаров, В.П. Корпусная лингвистика: учебник для студентов гуманитарных вузов / В.П. Захаров, С.Ю. Богданова. – Иркутск: ИГЛУ, 2011. – 161 с.
7. Натальин В. П. и др. Типичные грамматические ошибки в устной английской речи учащихся средних школ /В.П. Натальин// Иностранные языки в школе. – 1981. – № 3. – С. 67-70.
8. Пассов, Е.И. Коммуникативный метод обучения иноязычному говорению/ Е.И. Пассов. - 2-е издание - М.: Просвещение, 1991. - с. 223.
9. Плунгян, В.А. Корпус как инструмент и как идеология: о некоторых уроках современной корпусной лингвистики / В.А. Плунгян // Русский язык в научном освещении. – 2008.- №2 (16).

- 10.Рогова, Г. В. Методика обучения иностранным языкам в средней школе/Г.В. Рогова, Ф. М. Рабинович, Т. Е. Сахарова. . –М.: Просвещение. – 1991. – 55 с.
- 11.Сысоев, П.В. Интегративное обучение грамматике: исследование на материале английского языка // Иностранные языки в школе.–2003.– № 6.–С. 25–31.
- 12.Сысоев, П. В. Компетенция учителя иностранного языка в области использования информационно- коммуникационных технологий: определение понятий и компонентный состав/П.В. Сысоев // Иностранные языки в школе. – 2011. – № 6. – С. 16–20.
- 13.Хэгболдт, П. Изучение иностранных языков. Некоторые размышления из опыта преподавания/П. Хэгболдт. – М. :Учпедгиз, 1963. – 160 с.
- 14.Шаров, С.А. Представительный корпус русского языка в контексте мирового опыта / С.А. Шаров // НТИ. Сер.2. – 2003. – №6. – С. 9-17.
- 15.Щерба, Л.В. Языковая система и речевая деятельность: учеб. пособие / Л.В. Щерба. - Л.: 1974. - с.58.
- 16.Aarts J., Meijs. W. Corpus Linguistics: Recent developments in the Use of Computer Corpora in English Language Research / J. Aarts, W. Meijs. – Amsterdam: Rodopi, 1984. – 425 p.
- 17.Boulton, A. Data-driven learning: reasonable fears and rational reassurance/A.Boulton//Indian Journal of Applied Linguistics. –vol. 35, no. 1. – С.22-23.
- 18.Cobb, T. Is there any measurable learning from hands-on concordancing? /T.Cobb. –System, 25 (3),2012. . – С. 301–315.
- 19.Francis W. Language Corpora B.C. / W. Francis // In: Jan Svartvik (ed.) Directions in Corpus Linguistics. Proceeding of Nobel 82, Stockholm, 4–8. August, 1991, Berlin –New York: Mouton de Gruyter. – P.17-35.

20. McEnery T., Gabrielatos C. English Corpus Linguistics / T. McEnery, C. Gabrielatos // The Handbook of English Linguistics: McMahon-Blackwell Publishing, 2006 – P. 33-72 67.
21. MacEnery T., Hardie A. Corpus Linguistics: Method, Theory and Practice / T. MacEnery, A. Hardie. – Cambridge University Press, 2012. – 294 p. 68.
22. McCarthy M., O’Keeffe A. What are corpora and how have they evolved? / M. McCarthy, A. O’Keeffe // The Routledge Handbook of Corpus Linguistics. – Routledge, 2010. – P. 3-14. 69.
23. Meyer, Ch. P. English Corpus Linguistics. An introduction / Ch.P. Meyer. – Cambridge University Press, 2004. – 168 p.
24. SearchEngine.[Электронный ресурс]. – Режим доступа: <https://the.sketchengine.co.uk/>. –(Дата обращения: 13.06.2018).
25. Sinclair, J.M. Linear Unit Grammar: Integrating Speech and Writing: Studies in Corpus Linguistics. /J.M. Sinclair. –Amsterdam/Philadelphia: John Benjamins, 2006. – 185 p.
26. Svartvik, J. Corpus linguistics 25+ years on / J.Svartvik. – Amsterdam, NY 2007. – P. 11-27
27. O’Keeffe, A. From Corpus to Classroom: language use and language teaching/A. O’Keeffe, M. McCarthy, R. Carter.– Cambridge: Cambridge Univ. Press, 2007.– 315 p.
28. O’Keeffe, A. The Routledge Handbook of Corpus Linguistics / A. O’Keeffe, M. McCarthy. – New York, The Routeledge, 2010. – 711 p.
29. Tognini-Bonelli, E. Corpus Linguistics at Work / E. TogniniBonelli. – Amsterdam: John Benjamins Publishing Company, 2001. – 224 p. 97.
30. Tognini-Bonelli, E. Theoretical overview of the evolution of corpus linguistics / E. Tognini-Bonelli // The Routledge Handbook of Corpus Linguistics / Edited by Anne O’Keeffe and Michael McCarthy. – Routledge, 2010. – 681 p.

## ПРИЛОЖЕНИЕ 1 - ГРАММАТИЧЕСКИЕ УПРАЖНЕНИЯ

### Упражнение 1

I'm not sure about is math. Anyway, all we can	xx	is wait. Until then, we keep telling each other
when Hello was at her house, she asked him, « Who	xx	you like best, Ursul or Anne?» He said, « It's
home a lot lately. There's nothing for him to	xx	at the office; it must be awful to feel you're not
« Oh, poor Margot, you have so much work to	xx	!« Every other week Mr. Kleiman brings me a
n my index finger (on my left hand), so I can't	xx	any ironing. What luck! Mr. van Daan wants me to
too lazy, so I put it away. I'm not crazy enough to	xx	it all by myself! Anne Frank PS. I forgot to
after them. CHAPTER V. Anne's History	xx	you know,« said Anne confidentially, « I've
be finished off in some way, so I put in the other.	xx	you suppose it will make any difference?«« I—I
the temptation was _irresistible_. What do you	xx	when you meet with an irresistible temptation?
longer. « Marilla,« she demanded presently, «	xx	you think that I shall ever have a bosom friend in
his goddam jaw. Only, I wouldn't have the guts to	xx	it. I'd just stand there, trying to look tough.
good time, fella? Or is it too late for you?«« How	xx	you mean?« I said. I didn't know what he was

## Отсутствующее слово:

---

**Какой тип слов чаще всего встречается после ключевого слова?**

**Какая функция ключевого слова в этих фразах?**

Упражнение 2

**В представленном ниже конкордансе подчеркни временную форму Present Progressive. Объясни ее функцию в тексте, додумав недостающий контекст.**

few, then the Badlands surfers are not complaining.' We keep meeting these people who travel down

des Arts Decoratifs. At present she is creating a line of objects and small

bags being printed today and Tony will make it today. Currently Alan is having extreme difficulty in getting hold of any dried blood. Apparently we're not killing

Wherever you are, it is alright to yell if someone is trying to hurt you. Practice yelling as loud as you can in a big,

of mind — the « continental» and the « oceanic « . There is increasing evidence that the original island peoples did not migrate eastwards

And, well, the fact is, Miss Charlotte, that... I am asking you to be my wife.' There was a long silence. I heard

There are a number of projects in their infancy. One group is developing hardware for a personal computer, sufficiently robust and inexpensive

imagine her working here?" Here?" Yes, I am looking for a receptionist. You know, somebody to answer the phone and post letters

---



---



---



---

### Упражнение 3

**Приведенные ниже примеры были взяты из корпуса текстов. Найди глагольные формы Past Simple и подчеркни их.**

1. While he drove , Uncle Vernon complained to Aunt Petunia. He liked to complain about things: people at work, Harry, the council, Harry, the bank, and Harry were just a few of his favorite subjects.

2. The potty is returned to its place, and the figure in the white nightgown (the one that causes Margot to exclaim every evening, « Oh, that indecent nighty!«) climbs back into bed.

3. My heart sank a bit when I went down again the next day and saw him back in the same position again.

4. No light for fourteen days; that's a pleasant thought , isn't it? But who knows, maybe it won't be so long!

5. I instantly turned white, my stomach churned, and my heart beat wildly – and all because I was afraid.

6. I hadn't forgotten her entirely, and yet it wasn't until I saw her before me that I thought of all her suffering.

7. « What did you do during the lunch break?«

8. I guess some people would have thought it was ironic, but there were times when I failed to see the funny side of it, I have to admit.

9. I thought and thought, and suddenly I had an idea.

**Как образована эта временная форма? Объясни ее функцию в тексте.**

### Упражнения 4

**Просмотри приведенный ниже конкорданс**

about turn, left wheel, basic foot drill you know (pause) it has been going for about ten minutes and all and we was going down towards the factory and

pound and internally at the improved cash flow and the actions that we've been taking over for the past couple of years and I can say that there's now

for the release of the Birmingham 6 and the Guildford 4, who have been protesting their innocence for the past 14 years. The two groups of demonstrators converged on

same for five years,' he said. Realistically, I have been fighting against the odds since I took over. We need three or four decent players

one Ostertag wine (the Sylvaner' VV'). The Ostertags have been making wine in Alsace for a long time — centuries rather than years. But,

**Как образуется эта временная форма?**

**Какие наречия времени употребляются?**

**Заполни пробелы, употребляя Present Perfect Progressive или Past Simple. Обоснуй свой выбор.**

was accused of severe crimes but he \_\_\_\_\_ (get) away with murder.

a typewriter table. The writing \_\_\_\_\_ (go) better since I've started

Last year the Community Initiative unit \_\_\_\_\_ (help) churches to

Over recent years the futon \_\_\_\_\_ (grow) in popularity in the West.

---



---



---



---



---



---



---



---

## Упражнение 5

**В приведенном ниже конкордансе просмотри контекст, в котором употребляется Past Perfect. В каких ситуациях оно использовано?**

survived the worst horror of the modern world	<b>had</b>	seen – and who remained triumphantly and
much for my birthday either, since the fighting	<b>had</b>	just ended in Holland. Grandma died in January
the door behind us, leaving us alone. Margot	<b>had</b>	arrived much earlier on her bike and was waiting
to cry, which made me cry too, and I already	<b>had</b>	such an awful headache. I finally told Daddy
react. Up to now very little news about the Jews	<b>had</b>	reached us here, and we thought it best to stay a
as if everything I've said and done in my dreams	<b>had</b>	never taken place! Kitty, Anne is crazy, but
asleep, but his wife told Jan that her husband	<b>had</b>	discovered the hole in the door while making his

**Дополни недостающий контекст в конкордансе.**

1. When they had driven up the further hill and
2. She had intended to
3. She had made them up herself, and they were
4. Miss Rogerson was a middle-aged lady who had taught a
5. when Anne had gone to

## Упражнение 6

**Отгадай зашифрованное слово!**

**Что это за часть речи?**

standing beside her bed with her knees knocking	xx	the boom <b>came</b> . Dussel stayed upstairs to smoke a
---	----	--



us what had happened: Peter was on the landing	xx	he <b>heard</b> two loud bangs. He went downstairs and
, too, I know she did, for she <u>was crying</u>	xx	she <b>kissed</b> me good-bye through the bookcase
became a brawl. We <u>were</u> still <u>struggling</u>	xx	Atticus <b>separated</b> us. Thats all, he said.
dirty look. « Listen,« she said. « I <u>was sleepin'</u>	xx	that crazy Maurice <b>woke</b> me up. If you think I'm-
unhinged from its meaning. I <u>was</u> still <u>saying</u> it	xx	he <b>called</b> back. « Hi,« I said. « Hazel Grace,« he

### Подчеркнутые слова

**1. Какое время использовано?**

**2. Как оно образовано?**

### **Слова, выделенные жирным шрифтом**

**1. Что их связывает с подчеркнутыми словами?**

### **Заполни пробелы в конкордансе**

1. I had also wanted to see you, sir. I \_\_\_\_\_ going to request just such

2. If only the tutors knew what they \_\_\_\_\_ doing ... If only you her

3. many from working-class homes \_\_\_\_\_ fighting to extend their ideals to

4. he realized the Japanese women \_\_\_\_\_ not going to leave and he

5. know, there was a time when people \_\_\_\_\_ fighting for voting rights for

6. that he had changed his ways as things \_\_\_\_\_ not looking good. Suddenly

## **ПРИЛОЖЕНИЕ 2 - АНКЕТЫ**

### **Анкета №1 (перед занятиями)**

1. Как ты изучал грамматику в школе? (отдельно, совместно с чтением и.т.д., не изучал?)
2. По каким учебным материалам ты изучал грамматику? (отдельный учебник, один общий учебник, не изучал)
3. Оцени отношение твоих одноклассников к изучению грамматики? (по шкале от 1 до 10)
4. Оцени отношение твоего учителя к изучению грамматики? (по шкале от 1 до 10)
5. Что ты думаешь о грамматике английского языка? (по шкале от 1 до 10): легко – сложно, интересно – скучно, полезно – бесполезно.
6. Где ты пользуешься компьютером? а) в школе, б) дома.
7. Что ты думаешь о занятиях с компьютером? (по шкале от 1 до 10): легко – трудно, интересно – скучно, полезно – бесполезно.
8. Чего ты ожидаешь от уроков грамматики английского языка? (по шкале от 1 до 10): улучшения знания английского языка, понимания как устроен язык, возможность объяснить другим людям грамматические правила, обучаться самому, другое.
9. Какой путь изучения грамматики тебе наиболее близок? (по шкале от 1 до 10): объяснения учителя – самому изучать и разбираться.

### **Анкета №2 (после занятий)**

1. Оцени занятия по шкале от 1 до 10: легко – сложно, интересно – скучно, полезно – бесполезно.
2. Как ты думаешь, изменилось ли твое отношение к изучению грамматики после этих занятий? (Изменилось в лучшую сторону, немного

изменилось в лучшую сторону, никак не изменилось, немного изменилось в худшую сторону, изменилось в худшую сторону).

3. Если твое отношение изменилось, что на него повлияло?
4. Оправдались ли твои ожидания касательно прошедших занятий?
5. По шкале от 1 до 10, что ты думаешь о а) работе с корпусом, б) о возможности исследовать все самому?
6. Как ты оцениваешь свое вовлечение в прошедшие занятия (по шкале от 1 до 10)?
7. Как тебе может помочь работа с корпусом в будущем?
8. Как ты думаешь, будешь ли ты использовать корпус в ближайшем будущем?

### **Интервью**

1. Какая реакция у тебя была на то, когда ты понял, что ты вы будете изучать английский язык с помощью компьютера?
2. Что ты думаешь о своем знакомстве с корпусом текстов? Были ли понятны тебе инструкции? Вызывали ли какие-то упражнения трудности?
3. Для каких целей ты бы использовал корпус еще? Продолжил бы ты обучение грамматике с его помощью?
4. Тебе были понятны все инструкции к упражнениям?
5. Как ты справился с составлением собственных запросов в корпусе? Тебе нужна была помощь?
6. Ты справлялся с нахождением грамматического правила под запросы?

### ПРИЛОЖЕНИЕ 3 - Упражнение с запросом

Пример упражнения, включающего запрос, для работы с корпусом

#### Упражнение *used to vs be used to*

Сформулируй запрос, чтобы выяснить разницу между глагольными формами (инфинитивом и *-ing*-формой), придерживайтесь в запросе структур *I used to...* и *I'm used to*. Совет: *I'm* в запросе – это два слова.

---

---

**Результат:**

---

---

**В чем разница?**

---

---