

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ
УНИВЕРСИТЕТ ИМ. А. И. ГЕРЦЕНА»

На правах рукописи

Хоу Линьсюе

**«Пропедевтический учебный корпус русского языка: методы создания и
методический потенциал для обучения носителей китайского языка»**

Специальность 5.9.8. Теоретическая, прикладная и сравнительно-
сопоставительная лингвистика

Диссертация на соискание ученой степени кандидата филологических наук

Научный руководитель:
доктор филологических наук,
заслуженный деятель науки РФ
профессор Беляева Л. Н.

Санкт-Петербург

2024

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	1
ГЛАВА 1. КОРПУСНАЯ ЛИНГВИСТИКА В СОВРЕМЕННЫХ ИССЛЕДОВАНИЯХ.....	15
1.1. Корпусная лингвистика: история и развитие	15
1.2. Базовая типология лингвистических корпусов.....	21
1.3. Корпусная лингвистика в сфере образования	25
1.4. Учебные корпуса и их обучающий потенциал	36
1.5. Использование корпусов в преподавании русского языка.....	44
Выводы по Главе I.....	52
ГЛАВА 2. СОЗДАНИЕ УЧЕБНОГО КОРПУСА РУССКОГО ЯЗЫКА ДЛЯ ОБУЧЕНИЯ КИТАЙСКИХ СТУДЕНТОВ.....	55
2.1. Процедура отбора письменных текстов для создания учебного корпуса русского языка.....	56
2.2. Принципы и методы разметки в учебном корпусе русского языка	61
2.3. Идентификация и классификация ошибок при овладении русским языком в учебном корпусе русского языка	80
Выводы по Главе 2.....	90
ГЛАВА 3. КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ОШИБОК, ЗАФИКСИРОВАННЫХ В УЧЕБНОМ КОРПУСЕ РУССКОГО ЯЗЫКА, КАК БАЗА МЕТОДИЧЕСКОЙ ПОДДЕРЖКИ ОБУЧЕНИЯ	93
3.1. Типичные ошибки, совершаемые китайскими студентами под влиянием родного языка	95

3.2. Ошибки, совершаемые китайскими студентами под влиянием английского языка.....	108
3.3. Ошибки, совершаемые китайскими студентами, в сравнении с данными носителей русского языка	115
3.4. Ошибки, совершаемые китайскими студентами, обусловленные разными темами текстов.....	120
3.5. Методический потенциал учебного корпуса текстов. Предварительные оценки	130
Выводы по Главе 3	134
ЗАКЛЮЧЕНИЕ.....	138
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	144
ПРИЛОЖЕНИЕ 1. ИНФОРМАЦИЯ ОБ АВТОРАХ ТЕКСТОВ В УЧЕБНОМ КОРПУСЕ РУССКОГО ЯЗЫКА	178
ПРИЛОЖЕНИЕ 2. ПИСЬМЕННЫЕ ТЕКСТЫ КИТАЙСКИХ СТУДЕНТОВ НА РУССКОМ ЯЗЫКЕ (ФРАГМЕНТ).....	180
ПРИЛОЖЕНИЕ 3. УЧЕБНЫЙ КОРПУС РУССКОГО ЯЗЫКА ДЛЯ ОБУЧЕНИЯ КИТАЙСКИХ СТУДЕНТОВ (ФРАГМЕНТ).....	186
ПРИЛОЖЕНИЕ 4. СПИСОК КОЛИЧЕСТВА ОШИБОК В УЧЕБНОМ КОРПУСЕ РУССКОГО ЯЗЫКА	193
ПРИЛОЖЕНИЕ 5. ЧАСТОТА ОШИБОК КИТАЙСКИХ СТУДЕНТОВ С РАЗНЫМ УРОВНЕМ ВЛАДЕНИЯ АНГЛИЙСКИМ ЯЗЫКОМ	195
ПРИЛОЖЕНИЕ 6. ЧАСТОТНЫЕ СЛОВА В ТЕКСТАХ КИТАЙСКИХ СТУДЕНТОВ НА РУССКОМ ЯЗЫКЕ (ФРАГМЕНТ)	203

ПРИЛОЖЕНИЕ 7. СПИСОК ЧАСТОТНЫХ ОШИБОК В ТЕКСТАХ, НАПИСАННЫХ НА РАЗНЫЕ ТЕМЫ	205
---	-----

ВВЕДЕНИЕ

Актуальность исследования. По мере укрепления глобальных взаимосвязей в современном мире овладение одним или несколькими иностранными языками становится все более чертой и ключевой характеристикой современной коммуникации, а желание использовать иностранные языки для общения и взаимопонимания также активно растет. В соответствии с этим, успех межъязыкового общения напрямую зависит от уровня владения иностранными языками и сформированности иноязычной коммуникативной компетенции [5]. В современном мире все страны, в том числе Китай и Россия, уделяют все больше внимания подготовке специалистов, владеющих иностранными языками, как средству содействия международным экономическим, политическим, научным и культурным обменам. Благодаря успешному проведению различных экономических, культурных и научных мероприятий между Китаем и Россией, китайско-российские отношения непрерывно углубляются, особенно в последние десятилетия. В связи с этим много китайцев начали изучать русский язык, важно, что это обучение осуществляется не только в университетах, но и в средних и старших школах, где предлагаются курсы русского языка в качестве второго иностранного языка. Соответственно, процессы преподавания русского языка активизируются в КНР, при этом анализ результатов демонстрирует необходимость повышать качество и эффективность обучения русскому языку [165, с. 20].

С развитием информационных технологий появляются новые способы и подходы к анализу, синтезу и представлению текста, являющегося одним из

основных способов хранения и передачи информации, особенно это касается учебного материала. Естественно, этот процесс сопровождается разработкой новых инструментов, упрощающих процесс анализа и систематизации больших объемов информации. В результате привлечения информационных технологий к решению задач филологии и методики обучения языкам появилось новое направление – корпусная лингвистика, которая занимается обработкой большого объема информации с использованием компьютерных технологий [39; 75, с. 290]. Корпус определяется как коллекция текстов или фрагментов текстов, отобранных в соответствии с четкими критериями, корпус используется в качестве модели языка [144] и необходим для решения конкретных задач. Эти задачи не ограничиваются только сферой лингвистики, но также распространяются в методической деятельности, журналистике и других науках. Учитывая задачи, для решения которых создается тот или иной корпус, можно утверждать, что корпус всегда прагматически ориентирован [39].

Особым видом корпусов являются учебные корпуса, которые предназначены для мониторинга и анализа ошибок учащихся в процессе усвоения чужого языка [110, с. 247]. Применение корпусных технологий позволяет обнаружить наиболее распространенные орфографические ошибки и словоупотреблении [47, с. 175]. При этом следует отметить, что несмотря на растущий интерес специалистов к разработке и использованию учебных корпусов в преподавании, лексикографии, переводоведении и других практических сферах, по-прежнему мало исследований, посвященных

корпусам текстов, которые отражали бы результат усвоения русского языка носителей китайского языка в процессе его усвоения как вторым языком. Кроме того, остается неясным, каково возможное влияние изучения других иностранных языков, в частности английского, на эффективность обучения другим языкам, в данном случае русскому.

Следует также учитывать, что создание учебного корпуса осложняется необходимостью разработки и внедрения системы классификации и разметки ошибок. В связи с этим **актуальны:**

- разработка учебного корпуса текстов на русском языке, который направлен на дальнейшее исследование процесса овладения языком китайскими студентами;
- идентификация и классификация ошибок, допускаемых китайскими обучаемыми;
- количественный анализ частотных и типичных ошибок, допускаемых китайскими студентами при изучении русского языка;
- анализ влияния родного и английского языков на изучение русского языка.

Все вышеуказанное показывает, что создание пропедевтического учебного корпуса русских языка безусловно актуально, а ряд поставленных вопросов подтверждает необходимость и актуальность их решения. Актуальность исследования, посвященного проблеме анализа частотных и типичных ошибок, допускаемых китайскими студентами при изучении русского языка, в целях повышения качества и эффективности преподавания

русского языка в Китае, обусловлена также тем, что она реализуется посредством методов корпусной лингвистики.

Степень научной разработанности темы. С ростом доступности информационных технологий в последние годы, все чаще используются репрезентативные корпуса текстов в современных лингвистических исследованиях. В разное время исследованиями в сфере корпусной лингвистики занимались такие русские ученые, как А. В. Зубов, В. В. Рыков, В. А. Плунгян, В. П. Захаров, С. Ю. Богданова, Л. Н. Беляева, О. Н. Камшилова, Е. В. Рахилина и др. Особенное внимание на изучение проблемы в области корпусной лингвистики обращают такие китайские авторы, как Лян Маочэн, Сяо Чжунхуа, Сюй Цзяцзинь, Ван Кэфэй, Ван Лифи, Лю Мяо и др. В их работах рассматриваются и разрабатываются классификация корпусов по различным критериям, методы и принципы создания лингвистических корпусов русского и иностранного языков, и использование корпусов в исследованиях лингвистики.

Как в России, так и в Китае проблема создания и применения учебного корпуса привлекает огромный интерес ученых, подобным исследованиям и изучением посвящены работы А. И. Ольховской, М. К. Парамоновой, Е. В. Рахилиной, Н. Р. Добрушиной, О. Н. Камшиловой, Гуй Шичун, Вэй Найсин, Вэнь Цюфань, Ян Хуэйчжун и других исследователей.

Однако в этих исследованиях не рассматриваются проблемы создания учебного корпуса русского языка для обучения китайских учащихся, а также система классификации и разметки ошибок, допускаемых китайскими

студентами на основе учебного корпуса текстов.

Теоретическую основу исследования составили труды отечественных и зарубежных ученых, определяющие следующие аспекты темы:

- лингвистический корпусный подход нацелен на практическое изучение языка в точных ситуациях и материалах, что представляется собой важным с целью изучения и преподавания языка (В. П. Захаров, Е. П. Соснина, Л. И. Агафонова, Л. К. Раицкая, П. В. Сысоев, С. Ю. Богданова, Т. McEnergy, Y. Tono, J. Grigaliūnienė, G. Davies, H. Mortensen, M. Lamy);

- учебный корпус опирается на извлечение, классификацию и разметку ошибок в текстах, он является массивом языковых материалов, размеченный особым образом с целью решения практических и конкретных задач обучения языкам (Е. П. Соснина, О. Н. Камшилова, С. С. Замичева, С. Ю. Богданова, Е. В. Рахилина, S. Granger, J. Grigaliūnienė, Лян Маочэн, Гуй Шичун, Ян Хуэйчжун);

- ошибки являются неизбежной и неотъемлемой частью процесса изучения иностранного языка, с помощью их анализа на основе учебного корпуса можно выявлять трудности учащихся при овладении иностранными языками (Е. В. Тетерлева, Ю. В. Красиков, Ю. К. Попова, С. Н. Цейтлин, C. James, P. Corder, S. Granger);

- создание учебного корпуса русского языка требует предварительного исследования общего подхода к классификации ошибок, допускаемых китайскими студентами в процессе изучения русского языка (Е. А. Гасконь, Е. В. Жендаренко, М. О. Сорокина, Ю. И. Злобина, Ву Цзюань, Ли Гоучжин,

Лю Цянь);

- разработка принципов и методов разметки корпуса является основой для создания корпуса и увеличения ценности его применения (В. П. Захаров, Е. В. Рахилина, Е. С. Елохов, В. А. Плунгян, М. А. Куниловская, С. С. Замичева, G. Leech, M. E. Foster, Чжао Вэйнь);

- исследование изучения и преподавания иностранных языков необходимо проводить, учитывая влияние родного и второго иностранного языков на процесс овладения иностранным языком (А. А. Щелокова, Н. Д. Гальскова, Н. П. Панжиев, Р. П. Мильруд, С. П. Муродова, A. Wilson, G. Leech, P. Rayson).

Целью работы является разработка принципов и процедур количественного и качественного анализа ошибок в исследовательском учебном корпусе русского языка китайских студентов для оценки методического потенциала учебного корпуса.

Для достижения поставленной цели необходимо решить следующие **задачи**:

- 1) проанализировать опыт лингвистических исследований в области создания учебных корпусов;

- 2) определить состав, структуру, алгоритм разработки учебного корпуса и сферу его применения;

- 3) собрать и обработать текстовый материал для формирования корпуса;

- 4) определить понятие ошибки в процессе создании учебного корпуса с целью их идентификации; провести классификацию ошибок и разработать

систему их разметки;

5) разработать пилотный вариант учебного корпуса текстов, который можно рассматривать как пропедевтический корпус, в котором определены основные особенности и номенклатура ошибок, и апробировать составленный корпус с точки зрения извлечения ошибок разных типов;

6) провести количественный анализ частотных и типичных ошибок и оценить методический потенциал учебного корпуса.

Объектом исследования является учебный корпус русского языка для обучения китайских студентов как особый вариант исследовательского корпуса текстов.

Предмет исследования составляет методический потенциал учебного корпуса, созданного для выявления типичных ошибок в письменных работах китайских студентов, написанных на русском языке.

Гипотеза исследования: пропедевтический учебный корпус русского языка для обучения китайских студентов, отобранный в соответствии с требованиями корпусной лингвистики, размеченный по определенному принципу и обеспеченный специализированной поисковой программой, является базой данных, достаточной для количественного анализа ошибок и анализа влияния родного и английского языков на изучения русского языка.

Методологическая база исследования включает в себя наблюдение, описание, корпусные методы лингвистически, метод классификации, количественный анализ, сопоставительный анализ, методы компьютерной обработки данных, методы информационных технологий.

Материалом исследования послужили письменные тексты, написанные китайскими студентами, обучающимися по специальности «Русский язык», поступившими во Второй Пекинский университет иностранных языков и Хэнаньский университет науки и технологий в 2019 году. Общий объем текстов составляет 457 сочинений (82 064 слова).

На защиту выносятся следующие положения:

1) использование корпусных данных и технологий при преподавании иностранных языков становится необходимостью. Применение современных образовательных технологий в обучении русскому языку как иностранному обуславливает создание исследовательского учебного корпуса как вспомогательного инструмента для повышения эффективности и качества обучения русскому языку;

2) эффективность и репрезентативность исследовательского учебного корпуса русского языка для анализа ошибок обеспечивается корректностью установления специальных критериев отбора текстов, написанных китайскими студентами на русском языке, определением, идентификацией и классификацией ошибок при овладении русским языком и применением методов и принципов разметки на основе системы разметки *TEI*;

3) методика анализа ошибок, допускаемых китайскими студентами при создании текстов на русском языке, основывается на сочетании количественного и сопоставительного подходов. Определение оценок уровня знаний и частоты ошибок с помощью предложенного в работе коэффициента *K*, выделение частотных слов в текстах китайских студентов путем

использования корпусного менеджера *AntConc* позволяют установить влияние китайского и английского языков на изучение русского языка и выявить невидимые и типичные ошибки китайских студентов при создании текстов на разные темы;

4) созданный учебный корпус русского языка в качестве базы данных и результаты, полученные в ходе исследования, дают возможность установить методический потенциал учебного корпуса, использование которого позволяет эффективно повышать качество и улучшать методы обучения русскому языку.

Научная новизна исследования определяется тем, что в диссертации впервые:

1) разработана метод создания пропедевтического учебного корпуса русского языка для обучения китайских студентов;

2) предложены методы идентификации и классификации ошибок, допускаемых китайскими студентами в процессе овладения русским языком с учетом влияния китайского языка на изучение русского языка;

3) обобщены принципы и методы разметки для создания учебного корпуса русского языка и анализа ошибок;

4) выявлены и изучены типичные и частотные ошибки, допускаемые китайскими студентами при писании текстов на русском языке; установлено влияние китайского и английского языков на обучение русскому языку;

5) выявлен методический потенциал учебного корпуса для преподавания и изучения русского языка.

Теоретическая значимость результатов исследования заключается в

том, что разработка принципов и методов классификации и разметки ошибок, допускаемых при овладении русским языком, и методов построения учебного корпуса русских текстов для китайских учащихся вносит вклад в развитие корпусной лингвистики и использования корпусов для обучения иностранным языкам. Выводы, полученные на основе проведенного исследования, обеспечивают теоретическое исследование проблем преподавания и изучения русского языка в Китае.

Достоверность и научная обоснованность теоретических и практических результатов исследования обеспечивается использованием метода применения учебного корпуса в овладении русским языком, что позволило провести анализ на аутентичном языковом материале, написанном самими китайскими студентами. Достоверность полученных результатов подкреплена объективными количественными данными об анализе ошибок, допускаемых китайскими студентами в процессе овладения русским языком.

Практическая значимость исследования состоит в том, что

- разработанная методика в диссертационной работе может быть использован в качестве методологической основы для дальнейших исследований учебных корпусов русского языка;
- выявление типичных и частотных ошибок в письменной речи китайских учащихся предоставляет языковые данные для преподавателей при оптимизации процесса обучения русскому языку;
- разработанный пилотный вариант учебного корпуса русского языка может рассматриваться как пропедевтический корпус (учебный корпус был

загружен в систему Интернет на сайт Яндекс-Диск [<https://disk.yandex.ru/d/eNgPH51q4nOMlw>]), что позволяет его использовать как вспомогательный инструмент на занятиях по русскому языку как иностранному;

- созданный учебный корпус русского языка для обучения китайских студентов может быть использован в лексикографических исследованиях, направленных на создание русско-китайского учебного словаря.

Рекомендации по использованию результатов диссертационного исследования. Результаты исследования могут быть использованы на теоретических и практических занятиях, при обучении методам создания и применения учебного корпуса, а также при разработке специальных курсов для учителей или преподавателей, которые позволяют им приобрести навыки использования сетевых лингвистических ресурсов с целью решения прикладных лингвистических задач. Созданный исследовательский учебный корпус русского языка может быть использован преподавателями для работы по обучению русскому языку как иностранному и для создания современных дидактических материалов, а также учащимися для самостоятельной работы.

Апробация результатов исследования. Основные положения и результаты исследования были изложены в докладах на следующих научных и практических конференциях: Международная научная конференция «Язык. Коммуникация. Культура» (Гродно, Беларусь, 28–29 октября 2021 г.); IX (XXIII) Международная научно-практическая конференция молодых ученых «Актуальные проблемы лингвистики и литературоведения» (Томск, 14–16

апреля 2022 г.); VI Международная научно-практическая конференция «Российско-китайское сотрудничество: на пути к глобальному устойчивому развитию» (Новосибирск, 10–11 ноября 2022 г.); 25 Межвузовская студенческая научная конференция «Студент – Исследователь – Учитель» (Санкт-Петербург, 3–17 апреля 2023 г.). По теме диссертации опубликовано 5 работ, из них 3 – в журналах, рекомендованных ВАК РФ.

Структура и объем диссертационной работы определены поставленной целью и задачами. Диссертация состоит из введения, трех глав и выводов к ним, заключения, списка использованной литературы, включающего 182 наименования, из них 124 наименования на русском языке, 29 наименований на английском языке и 29 наименований на китайском языке, и 7 приложений. Общий объем диссертационного исследования составляет 214 страницы печатного текста, из них 177 страниц занимает основное содержание работы, включая 16 таблиц и 17 рисунков, приложения изложены на 37 страницах.

Во **введении** обосновывается выбор темы исследования и ее актуальность, степень научной разработанности, определяются объект, предмет, цель, задачи, гипотеза, теоретическая основа, методы исследования и теоретическая основа, даются сведения о языковом материале, раскрыта его научная новизна, сформулированы положения, выносимые на защиту, аргументируется теоретическая и практическая значимость, даны рекомендации по использованию результатов исследования, отражена апробация результатов работы.

Первая глава диссертации «Корпусная лингвистика в современных исследованиях» посвящена рассмотрению этапов развития корпусной лингвистики и базовой типологии корпусов; рассматриваются образцы корпусов на различных языках, особое внимание уделяется корпусам, созданным в РФ и КНР, в целях анализа места корпусной лингвистики в сфере образования; исследуются педагогические характеристики учебного корпуса и его возможности при применении в образовательной деятельности; характеризуется обучающий потенциал корпусных технологий в изучении и преподавании русского языка.

Во второй главе диссертации «Создание учебного корпуса русского языка для обучения китайских студентов» рассмотрена разработка пилотного варианта учебного корпуса русских текстов китайских студентов, обучающихся по специальности «Русский язык». В этой главе определяются процедура реализации каждого этапа разработки учебного корпуса и критерии отбора текстов, написанных китайскими студентами на русском языке для создания учебного корпуса; описываются принципы и методы проведения метаразметки и разметки ошибок в учебном корпусе русского языка; исследуется понятие ‘*ошибка*’ при обучении русскому языку для идентификации и извлечения ошибок в текстах; предлагаются теги классификации ошибок, отмеченных в текстах для создания учебного корпуса русского языка.

Третья глава диссертации «Количественный анализ ошибок, зафиксированных в учебном корпусе русского языка, как база методической

поддержки обучения» посвящена анализу ошибок, допускаемых китайскими студентами в процессе изучения русского языка. В этой главе предлагается метод оценки уровня знаний и частоты ошибок, для вычисления которого вводится коэффициент K ; рассматриваются частотные ошибки, допускаемые китайскими студентами при создании текстов на русском языке в аспекте интерференции китайского и английского языков; осуществляется сопоставление частотных слов в текстах китайских студентов и речи носителей русского языка с целью выявления невидимых ошибок китайских учащихся при изучении русского языка; описываются типичные ошибки в текстах на русском языке с разными темами, устанавливается методический потенциал созданного учебного корпуса.

В Заключении сформулированы основные выводы по полученным результатам проведенного исследования и оценен методический потенциал использования пилотного варианта учебного корпуса русского языка для обучения китайских студентов.

В Приложениях представлены: информация об авторах текстов в учебном корпусе; фрагмент оригиналов текстов, написанных китайскими студентами на русском языке; список количества ошибок в учебном корпусе русского языка; список частотных слов в текстах китайских студентов на русском языке; список частоты ошибок китайских студентов с разным уровнем владения английским языком; список частоты ошибок в текстах, написанных на разные темы; фрагмент учебного корпуса русского языка для обучения китайских студентов.

ГЛАВА 1

КОРПУСНАЯ ЛИНГВИСТИКА В СОВРЕМЕННЫХ ИССЛЕДОВАНИЯХ

Постоянное развитие глобальных информационных процессов предъявляет особые требования к работе с информацией. С развитием информационных технологий появляются новые способы синтеза, накопления и представления данных, возникают новые инструменты, позволяющие упрощать процессы анализа и систематизации больших массивов информации. Появление компьютерных технологий предоставило мощные исследовательские инструменты для различных дисциплин, особенно это касается лингвистических исследований.

Корпусная лингвистика как относительно новое направление в языкознании занимается разработкой принципов построения и использования лингвистических корпусов на основе компьютерных технологий [36, с. 5]. В рамках корпусной лингвистики, с одной стороны, решаются такие задачи как отбор, классификация, аннотирование, хранение и управление крупномасштабными текстами, а с другой стороны, созданием различных приложений на основе материалов корпусов, к ним относится весь комплекс систем обработки текстов на естественных языках [10]. Можно сказать, что корпусная лингвистика – это новая междисциплинарная наука на основе языкознания и компьютерных технологий [157].

1.1. Корпусная лингвистика: история и развитие

Корпусная лингвистика как наука зародилась в конце 1960-х гг., однако

методы, лежащие в ее основе, были известны с XIII века, то есть в эпоху создания доэлектронных корпусов, формируемых преимущественно для теологических, литературоведческих и лексикографических исследований [101, с. 132]. Одним из наиболее типичных примеров в этот период является «Обзор практического употребления английского языка» (*The Survey of English Usage*) [146, с. 389], цель создания которого заключалась в том, чтобы возможно полно описать грамматическую систему английского языка в речи носителей. Этот проект заложил прочный фундамент для современной корпусной лингвистики, определил необходимость проведения исследований на основе аутентичных, систематических языковых материалов [159, с. 2]. В рамках этого проекта были разработаны корпуса устной и письменной речи английского языка британскими лингвистами Рандольфом Квирком, Сидни Гринбаумом и Джеффри Личем и корпус устной речи Лондон-Лунд под руководством Рандольфа Квирка, этот корпус был одним из первых машиночитаемых корпусов [151].

Развитие информационных технологий во многом определило развитие корпусной лингвистики. Электронная эпоха корпусной лингвистики началась в 1960-е годы, этот период можно разделить на следующие этапы в зависимости от состава и размера корпуса [139]:

Первый этап 1960–1980 годы: характерной чертой электронных корпусов, созданных на первом этапе, является их нацеленность на изучение текстов отдельных жанров и/или речи социальных групп. Например, первый компьютеризированный Брауновский корпус (*The Brown corpus*) был

разработан Нельсоном Френсисом и Генри Кучерой в 1963 году, он включает 500 текстов из американских книг, газет и журналов, изданных в 1961 г. [104, с. 101]. Этот корпус был разработан для проведения системного исследования текстов, написанных на английском языке, принадлежащих к различным жанрам, а также для того, чтобы заинтересовать ученых этим новым научным направлением [72].

Второй этап 1980–1990 годы: исследования, проведенные на этом этапе, были ориентированы на определение терминологии корпусной лингвистики и формирование особого корпусного подхода, определяющего правила организации текстов в корпусе, а также принципы их разметки и анализа. Так, например, в начале 1980-х гг. был разработан Единый стандартный язык разметки (*Standard Generalized Markup Language*), который представляет собой набор тегов, стандартизирующий разметку текстов [127, с. 149].

Наиболее репрезентативным и прорывным проектом в конце 1980-х гг. стала Международная база данных языка, созданная в Бирмингемском университете и компанией Коллинз (*Collins Birmingham University International Language Database*), со следующими характеристиками: объём его превышал 20 миллионов словоупотреблений, источниками послужили полные тексты устной и письменной речи различных жанров [136]. Проект *COBUILD* лег в основу Словаря английского языка издательства Коллинз (*The Collins COBUILD Dictionary of English*), посвященного отдельным лингвистическим явлениям: фразеологизмам, предложениям, метафорам и т. д. [42, с. 2]. Успешное создание и использование корпуса *COBUILD* позволили лингвистам признать

огромные возможности и потенциал корпусов, что привело к буму в создании и использовании таких корпусов второго поколения, таких как Британский национальный корпус (*British National Corpus*), Международный корпус английского языка (*The International Corpus of English*), Корпус современного американского английского языка (*Corpus of Contemporary American English*) и другие [169, с. 26].

Наиболее известным из них является Британский национальный корпус (*British National Corpus – BNC*), разработанный в начале 1990-х годов специалистами из Оксфордского и Ланкастерского университетов. Тексты, представленные в корпусе, отбирались и балансировались по трем основным критериям: время, область, которую данный текст описывает, и тип издания [39, с. 88]. Общий объём корпуса составляет 100 млн слов. По мере совершенствования системы поддержки и ведения корпуса Британский национальный корпус стал международным эталоном создания корпусов, фундаментальным корпусом для обучения английскому языку, для исследований в области теории языка, социолингвистики и прикладной лингвистики.

Третий этап начался в 21 веке и продолжается до сих пор: этот этап характеризуется выдвижением корпусной лингвистики в отдельную прикладную лингвистическую науку, выделением метода корпусного анализа как отдельного метода и его широким применением во всех сферах лингвистики и смежных науках [75, с. 290]. Так, например, корпус текстов *COBUILD*, размер которого в 2000 году составлял почти 415 миллионов слов,

постепенно был преобразован в корпус, предназначенный для мониторинга и выявления тенденций развития английского языка [177, с. 45].

В первой половине 1990-х годов корпусная лингвистика окончательно сформировалась как отдельное направление науки о языке. Появление этого направления в значительной мере облегчило процедуру сбора и хранения информации и позволило использовать корпуса текстов при различных лингвистических исследованиях [63].

В России корпусная лингвистика появилась позже, чем в зарубежных странах, хотя следует отметить, что первый вариант корпуса русскоязычных текстов был создан в конце XX столетия (Упсальский корпус русского языка). Небольшой объем и другие ограничения сделали работу с инструментарием этого корпуса затруднительной [42]. Впоследствии российские исследователи начали активно создавать корпуса русского языка, такие как Компьютерный корпус текстов русских газет конца 20-го века (КГТ), Национальный корпус русского языка (НКРЯ), Хельсинкский аннотированный корпус (ХАНКО) и другие [37]. На сегодняшний день Национальный корпус русского языка является самым большим и представительным корпусом, наличие в нем текстов разных типов и исторических периодов позволяет найти примеры как современных, так и исторических текстов различных жанров [72; 87]. НКРЯ значительно увеличивает возможности изучения и применения корпусной лингвистики в русском языке для лингвистов и преподавателей, не только для носителей русского языка, но и для носителей других языков, китайского в том числе.

В Китае разработка и создание корпусов начались в конце 1970-х гг. С 1990-х гг. китайские лингвисты широко использовали корпуса в области обработки информации на естественном языке, а также при изучении китайского языка. В это время были разработаны и созданы Корпус современного китайского языка, Корпус сбалансированного современного китайского языка, Аннотированный корпус Жэньминь Жибао и другие [158, с. 9]. Следует отметить, что в настоящее время большое внимание уделяется созданию учебных корпусов, в частности, учебных корпусов английского языка, см., например, Учебный корпус английского языка, созданный в Гонконгском университете науки и технологии, Китайский учебный корпус текстов английского языка, Учебный корпус устной речи и т. д. Разработка и создание вышеуказанных корпусов способствовали развитию исследований в сфере лексикографии, переводоведения, обучения китайскому и иностранным языкам в Китае [167, с. 324].

Развитие возможностей использования больших данных (*Big Data*) в последнее время выводит корпусные исследования на совершенно новый уровень. Соответственно, реальной становится возможность использовать большие массивы систематизированного материала для научных исследований [12]. Кроме того, происходит дальнейшее раскрытие потенциала и возможностей использования корпусов в различных областях. В следующем разделе представлена основная типология корпусов для решения различных задач лингвистического и методического анализа.

1.2. Базовая типология лингвистических корпусов

Существует множество классификаций корпусов, что определяется как многообразием исследовательских и/или прикладных задач, ради которых они создаются, так и различными основаниями классификации – технологическими (касающимися технологий корпусного строительства) и лингвистическими (определяющими характер текстов в корпусе) [34, с. 20–29]. С точки зрения потенциала и целей использования принято выделять следующие виды корпусов:

- По критерию соответствия цели исследования различаются **исследовательские и иллюстративные корпусы**. Исследовательские корпусы разрабатываются для решения задач изучения различных аспектов функционирования языковой системы, они создаются до начала проведения исследования и ориентированы на решение конкретных лингвистических задач. Иллюстративные корпусы создаются после проведения исследования для подтверждения полученных результатов.

- По критерию *динамичность* корпусы подразделяются на **статические и динамические (мониторные) корпусы**. Статические корпусы предназначены для определения временного среза языковой системы, к ним относятся авторские корпусы – коллекции текстов писателей. Динамические (мониторные) корпусы фиксируют динамические изменения языка и в промежутке времени происходит обновление и/или дополнение объема корпусов [86, с. 25].

- По критерию *степень охвата данных* корпусы можно разделить на

национальные и специализированные (специальные). Национальный корпус – многоцелевая электронная коллекция текстов, объем которой может составлять миллионы и даже миллиард слов. Национальный корпус предназначен для объединения в своей структуре текстов самых разных типов и жанров: письменных и устных, литературных и разговорных, представленных сбалансированно с целью отобразить реальное соотношение их долей, характерного для современного состояния языка [9; 74].

Специализированный (специальный) корпус тесно связан с конкретными исследовательскими целями и намерениями разработчика, его размер зависит от исследовательской задачи. Построение специализированных корпусов текстов может быть активно востребовано в отдельных областях деятельности, так как они позволяют фиксировать информационно значимое и реальное положение дел в профессиональных языках [46, с. 116].

Особыми видами исследовательских корпусов являются **параллельные корпуса** и **учебные корпуса**. Параллельный корпус представляет собой коллекцию текстов на двух или более языках, в нем соединяются текст-оригинал и его перевод на другие языки. Исследование корпуса параллельных текстов (параллельных корпусов) помогает решению многих прикладных задач в сфере сравнительной лингвистики, переводоведения и обучения иностранным языкам. Разновидностью параллельных корпусов текстов являются псевдопараллельные (сопоставимые) корпуса. Эти корпуса позволяют добиться гармонизации терминологии в рамках отдельных предметных областей, особенно тех, которые отличаются постоянной

динамикой [7, с. 55; 143].

Учебный корпус является электронной текстовой базой аутентичных данных учащихся, формируемой и пополняемой в процессе овладения иностранными языками с целью оценки эффективности освоения изучаемого языка и выявления способов ее повышения. Новым направлением в создании корпусов является построение лонгитюдных корпусов, представляющих собой результат накопления текстов одного и того же автора (авторов) в течение некоторого времени, что позволяет представить процесс овладения языком в динамике [51, с. 104; 52]. В принципе, лонгитюдный подход может быть реализован и при создании учебных корпусов.

Помимо вышеперечисленных основных типов корпусов в зависимости от набора классифицирующих признаков можно определить дополнительные параметры [39, с. 57] (см. Табл. 1).

Таблица 1

Классификация корпусов по дополнительным параметрам

Признак	Типы корпусов	Признак	Типы корпусов
форма фиксации данных	письменные	разметка	аннотированные
	устные (речевые)		неразмеченные
	смешанные	характер разметки	морфологическая
свободно доступные	синтаксическая		
коммерческие	семантическая		
доступность	закрытые		анафорическая и т. д.

Следует отметить, что один и тот же корпус может быть описан по-разному в соответствии с различными критериями классификации, и в то же время эти критерии не противоречат друг другу.

В целом, различные корпусы создаются для решения конкретных задач

исследований. В письменных корпусах тексты представлены с целью анализа характеристик письменной речи, а в устных корпусах представлена только устная речь для выявления фонетических изменений. В размеченном корпусе для разных исследовательских целей словам или предложениям присваиваются теги в соответствии с характером разметки: морфологические, синтаксические, семантические, просодические и др. Свободно доступные корпуса позволяют обращаться к ним в режиме реального времени и искать информацию в полном объеме. В свою очередь закрытые корпуса предназначены для специальных целей и не предполагают общественного использования [34].

Естественно, данная классификация весьма условна и неполна. Но, к сожалению, можно утверждать, что представление о том, что корпуса создаются просто путем сбора больших объемов данных и их объединения, является распространенным заблуждением [129]. Лингвистические (языковые) корпуса обязательно должны быть достаточно репрезентативны, построены по определенным правилам и предназначены для решения конкретных задач в исследовательской и методической деятельности, лингвистике, журналистике и других науках. Корпусные технологии также используются при преподавании дисциплин, при изучении динамики развития языков, при разработке автоматических систем, таких как машинный перевод, распознавание речи, информационный поиск [55].

Следующий раздел работы посвящен исследованию места корпусной лингвистики в сфере образования.

1.3. Корпусная лингвистика в сфере образования

Корпусная лингвистика обрела широкий спектр применения, и ее методы могут использоваться практически во всех направлениях, связанных с исследованиями языка и разработкой современных методов преподавания и изучения языка. Анализ корпусных данных позволяет оценить действительное функционирование конкретного языка, изучать динамику уровня владения языком обучающимися [23].

В последние годы ведущие вузы во всем мире внедряют корпусную технологию в эмпирической части лекционных курсов и опираются на корпусные данные при разработке заданий и практических проектов, соответствующих уровню подготовки студентов. Например, в отличие от традиционных учебников, программы для изучения иностранных языков, составленные Ланкастерским университетом на основе корпуса текстов, дают студентам возможности получить реальное представление о современном состоянии языка, о разнообразии жанров, что позволяет более активно развивать их коммуникативную компетенцию [77, с. 57].

Сотни вузов в России, включая Российский государственный педагогический университет им. А. И. Герцена, Санкт-Петербургский государственный университет, Санкт-Петербургский политехнический университет Петра Великого, Московский государственный университет им. М. В. Ломоносова, Томский политехнический университет и другие, предлагают корпусную лингвистику в качестве основного курса, направленного на развитие у студентов компетенций в области использования

корпусных технологий как в изучении языка, так и в научных исследованиях [24, с. 187]. Китайские вузы также стремятся к введению в университетах дисциплины «Корпусная лингвистика» для дальнейшего развития прикладной лингвистики в Китае [177, с. 45].

Корпусные технологии используются в различных областях науки и в методических исследованиях, к ним относятся:

Лингводидактика. С помощью инструментов поиска на основе корпусов можно расширить традиционный подход к лингвистическому исследованию. Корпусы позволяют наблюдать за распределением употребления слов и их коллокаций, появлением неологизмов, а также анализировать морфологическое и лексическое развитие языка. Так, например, Корпус *TCEEE (The Tübingen Corpus of Eastern European English)* был создан для анализа разнообразия английского языка в Восточной Европе [150, с. 143]. Именно из корпусов, снабженных разными видами разметки, можно извлекать необходимые данные о грамматической структуре коллокаций для более эффективного анализа их структуры, например, осуществлять поиск именных групп – словосочетаний, включающих прилагательное и существительное.

Кроме того, опираясь на корпусы, исследователи могут получить частотное распределение различных языковых структур, анализировать связь между грамматическими структурами и различными уровнями языка, а также между языковыми и неязыковыми факторами для того, чтобы объяснить причины, по которым говорящие выбирают те или иные выражения. Результаты этих исследований широко применяют в областях обучения языку,

анализа текста и машинного перевода, к примеру, Пекинский университет языка и культуры использует корпус для подсчета частоты различных моделей предложений в современном китайском языке, предоставляя надежные данные с целью поддержки обучения китайскому языку как иностранного, составления учебных материалов и разработки стандартов тестирования [157, с. 161].

Лексикография. Словари всегда привлекают большое внимание как источник лингвистических и культурных знаний и как средство обучения. Лексикография представляет собой результат углубления исследования лексики и представления разнообразной лингвистической информации [148]. Следовательно, составление словаря должно основываться на богатом и аутентичном лингвистическом материале. Современный подход к созданию словарей предполагает формирование и использование корпуса реальных текстов, который может рассматриваться как богатая база данных для решения не только исследовательских задач, но и практических лексикографических задач [8, с. 89]. Начиная со сбора языковых материалов и заканчивая созданием конкордансов и составлением толкований или дефиниций в словарной статье, корпус оказывает активное влияние на лексикографию. По мере развития и обновления корпуса можно создавать базу данных размером в десятки и даже сотни миллионов слов для разработки учебного словаря [28].

Некоторые из наиболее влиятельных современных словарей и учебников грамматики были созданы на основе корпусов. Например, словарь современного английского языка издательства Longman основан на

Британском национальном корпусе; в Центре изучения употребления английского языка при Лондонском университете составлена Оксфордская грамматика английского языка на основе созданного там международного корпуса английского языка [69, с. 226]. В Институте русского языка им. В. В. Виноградова Российской академии наук на основе НКРЯ составлено четыре новых словаря: «Грамматический словарь новых слов русского языка», «Новый частотный словарь русской лексики», «Словарь русской идиоматики», «Словарь глагольной сочетаемости непредметных имен русского языка» [98].

С помощью масштабного корпуса можно оптимизировать иллюстративную зону в словаре и предоставить подходящие примеры для того, чтобы создать словарь, наиболее соответствующий реальным потребностям пользователей, например провести разметку самых частотных слов, добавить напоминания о частотных ошибках, возможных синонимах и антонимах, ввести распределение значения слов и коллокаций в соответствии с частотой употребления. Вышеперечисленные методы оптимизации лексикографического описания не могут быть применены без помощи корпуса.

Переводоведение. Ресурсы корпусной лингвистики являются эффективным инструментом профессиональной деятельности переводчика [105, с. 209]. По сравнению с традиционными инструментами перевода параллельные корпуса предоставляют массу аутентичных языковых материалов, в этих корпусах текст перевода соотнесен с текстом оригинала, что облегчает переводчику быстрый поиск наилучшего варианта перевода. Именно на основе этих особенностей параллельные корпуса играют важную

роль в создании двуязычных и многоязычных переводных словарей, пополнении словарей для машинного перевода, устранении полисемии лексических единиц [68]. Кроме того, параллельные корпуса позволяют переводчику повысить качество перевода. Дело в том, что качество переводной терминологии во многом определяет качество перевода в целом [3, с. 160]. Опираясь на данные специальных параллельных корпусов, можно выявить и систематизировать терминологию и заложить основу для составления переводного словаря терминов [140].

Корпусы также дают переводчику возможность реально ориентироваться в языке и оперативно решать целый ряд лингвистических и экстралингвистических проблем, возникающих в процессе перевода [120, с. 91], например, можно осуществлять сравнение языковых явлений для изучения стиля перевода, влияния языка оригинала на структуру языка перевода и т. д.

Для подготовки переводчиков подходы корпусной лингвистики позволяют развивать исследовательскую компетенцию обучающихся [65, с. 86]. Под руководством преподавателя студенты могут использовать корпус для анализа и обобщения переводных методов для совершенствования своих переводческих навыков. Опираясь на корпус, можно разработать комплекс упражнений, которые способствуют развитию языковой догадки, способности вывести значение незнакомого слова (например, фразеологизма) по полученным результатам корпусного поиска [82]. В некоторых российских и китайских вузах дисциплина «Корпусная лингвистика и перевод» реализуется

в рамках основных учебных программ подготовки переводчиков, см., например, программы Кемеровского государственного университета, Тюменского государственного университета, Хэйлунцзянского университета, Пекинского университета иностранных языков и другие.

Корпусные исследования затрагивают не только практическую, но и теоретическую области переводоведения. Первый переводный корпус английского языка был создан в 1996 году профессором Сарой Лавиоза-Брейтвейт университете Манчестерского института науки и технологий (*University of Manchester Institute of Science and Technology*) [141], а затем Дороти Кенни, Мона Бейкер [126] и другие последовательно совершенствовали переводной корпус [156, с. 62]. На основе созданного корпуса Мона Бейкер разработала свою теорию переводных корпусов и основала Центр исследований перевода [161, с. 63]. В рамках следования принципам теории переводоведения тексты переводного корпуса берутся из массивов официально опубликованных переводных текстов, используются для изучения переводческого процесса посредством анализа того, как термин одного языка передается на другом, а также путем сравнения особенностей языковых единиц и частотности их употребления в оригинальном и переводном текстах [95, с. 201].

Обучение языкам. Применение корпусов для решения практических задач служит важным направлением в корпусной лингвистике, в частности, к нему относятся задачи обучения языкам. В зависимости от того, как корпусы используются в преподавании языка, их применение можно разделить на:

1) непосредственное применение корпусов текстов, включающее использование корпусных ресурсов на занятиях и развитие навыков создания и использования корпуса для самостоятельного исследования;

Применение корпуса в обучении языкам в основном связано с изменением педагогических концепций и привело к появлению новой модели обучения с управлением от данных (*DDL – data-driven learning*), которая является одной из наиболее принятых моделей непосредственного использования корпусов на занятиях [93]. Согласно традиционному подходу к обучению языкам учащимся, когда они сталкиваются с проблемами, возникающими при овладении языком, обычно приходится полагаться на объяснения преподавателей или на учебники, которые не соответствуют современным требованиям и ограничены по содержанию. Все это препятствует развитию навыков самостоятельного обучения и исследования у учащихся. В отличие от этого корпус предлагает им новые возможности для изучения языка. Благодаря корпусу стирается пропасть между исследованием и обучением, что позволяет учащимся самим становиться исследователями, изучая особенности текстов и делая выводы на основе данных, которые они могут получить с помощью корпуса [64, с. 41].

В этом случае студенты проводят собственные уникальные исследования вместо того, чтобы просто опираться на содержание учебников [89]. Они используют данные, предоставляемые корпусом, чтобы самим искать ответы на вопросы, возникающие при овладении языком, и пытаться их решить, тем самым стимулируя мотивацию к дальнейшему изучению и исследованию.

С точки зрения преподавателей корпусы позволяют им анализировать

типичные/нетипичные языковые явления с целью оптимизации курсов и содержания обучения. Преподаватели имеют возможность сосредоточиться на типичных языковых явлениях, т. е. на частотных словах и словосочетаниях, самых частотных грамматических конструкциях и т. д., чтобы добиться дифференцированного и целенаправленного преподавания.

Кроме того, корпуса предоставляют преподавателям богатые языковые ресурсы для разработки таких упражнений как определение значения слов с помощью корпуса, изучение грамматических структур, запоминание структуры словосочетания и т. д. [104, с. 110]. Для работы с подобными упражнениями и их разработки корпуса могут стать вспомогательным инструментом для преподавателей при решении различных методических задач [55, с. 77].

2) опосредованное применение корпусов предполагает составление грамматических справочников и учебных материалов на основе корпусов, разработку мультимедийных курсов, инструментов тестирования и т. д. (см., например, [152]).

Практические задачи, возникающие в процессе обучения языку, связаны с формированием и совершенствованием у учащихся умений активного владения языком [85, с. 87]. В этом процессе корпуса могут быть чрезвычайно полезны. На подготовку и издание учебной литературы, которую используют в образовательном процессе, уходит несколько лет, используемая в ней лексика устаревает, поскольку язык в процессе коммуникации обновляется быстро. Благодаря способности корпусов обновлять текстовые базы данных,

становится возможным изучение актуальной, обновленной лексики.

Корпусы позволяют сделать учебные материалы более соответствующими реальным потребностям учащихся и близкими к ситуации их практического обучения. Анализ учебных и методических материалов, осуществляемый на основе корпусных исследований, может определить, насколько сбалансированно в учебниках представлены различные аспекты языка, соответствуют ли лексические единицы, представленные в учебном материале, действительному состоянию языка [77, с. 57], в корпусах можно найти и искать примеры лексических единиц и грамматических конструкций, используемых носителями языка, что дает возможность правильно объяснить учащимся их значение и особенности употребления. Поскольку в реальной жизни для учащихся изучение языка чрезвычайно важно, а корпусный подход дает новые возможности для эффективного им овладения, то преподавателям необходимо использовать корпусный подход, ориентированный на функционирование языка в реальных ситуациях и текстах [121, с. 103].

В последние годы в Китае активно развивается проект информатизации образования, в рамках которой множество вузов внедряют инновационные методы работы с учебным материалом, в частности с помощью корпусов. Рациональное сочетание инновационных и традиционных методов обучения не только улучшает уровень профессиональной компетентности преподавателя языка, но и способствует повышению интереса студентов к изучению [54]. На основе шести крупных китайских и зарубежных корпусов английского языка команда исследователей Шанхайского университета

иностранных языков под руководством профессора Ху Кайбао разработала интеллектуальную многоязычную платформу обучения и исследований [158]. Помимо предоставления преподавателям учебных материалов с различными жанрами для проведения обучения и исследований, платформа также позволяет преподавателям загружать и аннотировать свои собственные языковые материалы, а также их анализировать. Все это обеспечивает своевременное обновление содержания учебных материалов, разнообразие упражнений на занятиях и дальнейшее повышение качества обучения языкам.

Несмотря на множество достоинств и большой потенциал корпусов текстов, у их применения есть некоторые ограничения, мешающие широкому использованию таких технологий. Прежде всего это связано с недостаточными компетенциями преподавателей в области информационных технологий в целом, а также с психологическими трудностями преподавателей, привыкших к традиционным методам обучения. Однако в современных условиях педагогу недостаточно быть только пользователем информационных технологий, необходимо уметь разрабатывать собственные электронные учебные средства для максимально повышения эффективности обучения языкам [19, с. 5]. Поэтому важно разработать программы повышения квалификации преподавателей и системы тренингов для них.

Сами преподаватели также должны делать акцент на обучении студентов практическому применению корпусных технологий. Поэтому необходимо четко формулировать цели каждого специального упражнения для того, чтобы студенты могли успешно освоить такие новые методы как корпусные

технологии, выполнять задания и собственные исследования на их основе [1].

Вопросы разработки и использования учебных корпусов в образовательном процессе рассматриваются в следующем разделе.

1.4. Учебные корпуса и их обучающий потенциал

С развитием инновационных компьютерных технологий появились новые методы анализа, синтеза, хранения и передачи информации в области научных и методических исследований. Одним из наиболее перспективных направлений стала корпусная лингвистика, открывающая новые возможности для исследования и практики в области обучения языкам [73, с. 209]. Что касается овладения иностранными языками, большое количество иноязычных текстов, представленных в корпусе, может быть преобразовано в учебный контент, тем самым обогащая словарный запас учащихся и улучшая их навыки использования языка [53].

Особое место в таком относительно новом направлении как корпусная лингвистика занимает учебный корпус, который доказал возможность своего использования при обучении родным и иностранным языкам [49, с. 56]. Учебный корпус является особым видом специализированных корпусов, представляющим собой сбалансированную по формату, жанру и типу текста коллекцию речевых (устных и письменных) произведений на изучаемом языке [48, с. 10]. Авторами текстов учебного корпуса являются лица, изучающие иностранные языки [73, с. 209], поэтому данные, основанные на учебном корпусе, действительно выявляют наиболее реалистичную ситуацию обучающихся в процессе овладения иностранными языками, поскольку учебный корпус (как и всякий лингвистический корпус) представляет собой модель языка в процессе его усвоения носителем.

Создание учебного корпуса преследует как исследовательские, так и

педагогические цели. С одной стороны, ошибки являются важными маркерами языковых изменений, с помощью учебного корпуса можно отследить и проанализировать тенденции развития промежуточного языка (интеръязыка) учащихся. С другой стороны, педагогической целью учебного корпуса является анализ реальных текстов, создаваемых обучаемыми, на предмет выявления способов и эффективности освоения изучаемого языка [102]. На основе учебного корпуса можно создавать лингвистические тренажеры, формировать контрольные задания с учетом ошибок учащихся, а также индивидуальные учебные подкорпусы.

Учебный корпус должен отвечать следующим требованиям:

- предназначаться для изучения или преподавания второго/иностранного языка;
- предоставлять в электронном виде реальные материалы (устные или письменные) учащихся, которые изучают второй/иностраный язык;
- формироваться единообразно в соответствии с заранее заданными критериями;
- предоставлять экстралингвистическую информацию [132, с. 7].

Именно эти требования и позволяют учебному корпусу обладать рядом особых возможностей для анализа:

- поиск конкретной информации;
- сортировка результатов поиска по задаваемым критериям;
- предоставление контекстов найденных элементов;
- поддержка разнообразия жанров и стилей текстов;

- обновляемость информационных данных [99].

Все это актуализирует применение учебного корпуса в практической педагогической деятельности, что может привести к росту количества корпусов такого типа. Согласно данным Левенского католического университета до 2020 года число учебных корпусов в мире составляло около 188, включая английский, русский, французский, китайский, немецкий, венгерский, испанский, итальянский, чешский и более десятка других языков [174, с. 11]. Языковые материалы в этих корпусах также являются разнообразными, не только устными (разговорными) и письменными, но также включают мультимедийные материалы, что дает возможность максимально полно выявлять характеристики языка учащихся (см. сайт Левенского католического университета <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>, на котором в общем большом списке представлено 17 учебных корпусов, включающих тексты на русском языке).

Первым и самым известным учебным корпусом является Международный учебный корпус английского языка (*The International Corpus of Learner English*), созданный в Центре развития английской корпусной лингвистики (*Center for English Corpora Linguistics – CECL*) под руководством бельгийского лингвиста Сильвиан Грейнджер в конце 1980-х – начале 1990-х годов для изучения языковой характеристики тех, кто овладевает английским языком на фоне разных родных языков. До 2009 года этот корпус охватывал письменные сочинения изучающих английский язык с 16 различными

родными языками [131]. С целью изучения устной речи студентов в Центре *CECL* также был создан Левенский учебный корпус английской устной речи (*The Louvain International Database of Spoken English Interlanguage*), материалами которого являются устные интервью учащихся английского языка с 14 различными родными языками.

В последнее время активизировалась разработка корпусов ошибок в Китае и России [22; 27]. К созданным корпусам относятся Санкт-Петербургский учебный корпус текстов школьников, изучающих английский язык (*SPbEFL LC*) [45], Корпус русских учебных текстов (КРУТ) [59], Хельсинкский аннотированный корпус русских текстов (ХАНКО) [58], Устный учебный корпус русского языка как иностранного [17], Корпус учебных переводов (*Russian Learner Translator Corpus*) [66], Учебный корпус английского языка при Гонконгском университете науки и технологии [175, с. 20], Китайский учебный корпус текстов английского языка [179], Учебный корпус русских текстов [172] и другие.

Изначально ученый корпус использовался в основном для проверки диктантов учащихся, но позже он постепенно стал использоваться для анализа различных ошибок, допускаемых учащимися в процессе овладения иностранным языком [48, с. 11]. Можно сказать, учебный корпус представляет собой аннотированный корпус ошибок, которые допускают школьники или студенты при изучении ими второго языка [39, с. 56]. В учебном корпусе классификация и разметка ошибок, допускаемых учащимися при овладении иностранным языком, предоставляет возможности их системного выявления и

анализа. Поговорка «на ошибках учатся» точно характеризует метод применения учебного корпуса при овладении иностранным языком.

Ошибки были, есть и будут составной частью изучения иностранного языка, поскольку они действительно демонстрируют «слабые места» обучаемых [44, с. 51], и их учет и исследование способствуют повышению эффективности преподавания. По сравнению с традиционным методом анализа ошибок [43, с. 2], анализ ошибок на основе учебного корпуса опирается на аутентичные данные и позволяет извлекать контекст ошибок, одновременно анализируя ошибки так, что они становятся более объективными и понятными. Применение корпуса помогает преподавателям оптимизировать стратегии обучения и повысить качество преподавания. Для учащихся обучение на сделанных ими ошибках способствует формированию навыков самостоятельного изучения языка. Большой объем языкового материала необходимо и важен для овладения языком, и только предоставляя учащимся доступ к большому количеству ошибок, можно улучшить их способность распознавать и самостоятельно исправлять ошибки.

Кроме того, в рамках решения методических задач обучения иностранным языкам сложился новый подход к использованию учебных корпусов, а именно сопоставительное изучение языка учащихся, например, сравнение языка учащихся и речи носителей языка с целью выявления ошибок учащихся при овладении иностранным языком, таких как расширение употребления конструкции, недостаточное использование слов из-за непонимания их точного значения и другие. В педагогической сфере учебные корпуса

создаются как вспомогательный инструмент, позволяющий сравнивать характеристики изучаемых языков и результаты обучения [27, с. 65].

Сопоставительное исследование на основе учебного корпуса позволяет сравнивать язык носителей с языком учащихся с целью выявления характеристик языка учащихся, не проявляющихся в языке носителей, т. е. такие невидимые ошибки как расширение использования слов и грамматической структуры. С помощью разметки различных метаданных, проведенной в корпусе, при обучении иностранному языку можно выявить различные характеристики самих учащихся, у которых разные родные языки, чтобы установить, как различные родные языки и в какой степени влияют на овладение иностранным языком и, таким образом, целенаправленно оптимизировать содержания обучения. Сравнивая язык учащихся на разных уровнях овладения иностранным языком, можно сделать вывод о развитии языка учащихся, чтобы сделать акцент на трудностях, возникающих на разных этапах изучения, вовремя регулировать стратегии и методы преподавания.

Успешный анализ и использование данных учебного корпуса также отражается в новом подходе к созданию учебных словарей на основе распространенных ошибок учащихся [153], например, в учебном словаре Кембриджского университета (*Cambridge Advanced Learner's Dictionary*) есть зона предупреждений, которая напоминает учащимся о словах, в которых ошибки возникают наиболее часто, например, ошибки при использовании слова *information* (Рис. 1).

Common Learner Error

information

Remember you cannot make **information** plural.

Do not say 'informations'.

Could you send me some information about your courses ?

We've been able to find out several pieces of information.

For more information please contact our office.

~~*For more informations please contact our office.*~~

Рисунок 1. Зона предупреждений в учебном словаре Кембриджского университета [128]

Новое направление в учебных корпусах — создание лонгитюдных корпусов, накопление текстов одного и того же автора (авторов) в течение некоторого времени. Лонгитюдными называются исследования, при которых многократно фиксируются определенные показатели через строго установленные отрезки времени для определения динамики их изменения и взаимовлияния [99]. Лонгитюдный учебный корпус позволяет сосредоточиться на описании и проверке динамического процесса овладения языком, выявлении развития промежуточного языка, что имеет большое значение для разработки программ обучения, составления учебных материалов и т. д.

Несмотря на то, что учебный корпус может предоставить огромное количество полезной информации по многим аспектам в изучении и преподавании иностранного языка, у учащихся могут возникнуть проблемы с его использованием. Главные трудности учащихся при выполнении корпусных

заданий возникают из-за недостаточных умений пользоваться корпусами. Чтобы эффективно и правильно выполнять корпусные упражнения, студентам требуется регулярная практика [80].

В дополнение к этому корпус должен предоставлять студентам вспомогательные средства, которые помогут им лучше освоить технологию корпусов. Так, например, образовательный портал НКРЯ предлагает видеоролики и методические рекомендации по проведению практических занятий и уроков в студенческой аудитории на основе использования НКРЯ, научную и методическую литературу по корпусным методикам, информацию о конференциях и новых публикациях [9, с. 51].

Учебный корпус обеспечивает репрезентативный и объективный анализ результатов учащихся при обучении иностранным языкам. Сегодня особенно важно при подготовке будущего учителя русского языка оптимальное сочетание современных информационно-коммуникационных и педагогических технологий [4, с. 89]. Далее рассмотрим корпусы, используемые в преподавании русского языка.

1.5. Использование корпусов в преподавании русского языка

Корпусные методы зарекомендовали себя в мировой практике лингвистических исследований и преподавании иностранных языков как высокоэффективные инновационные дополнения к традиционным образовательным технологиям [77, с. 59]. В связи с этим при преподавании русского языка как иностранного также активно модернизируется учебный процесс, лингвистический корпус внедряется в педагогическую деятельность [81].

Созданный в 2004 году, регулярно обновляемый и пополняемый, Национальный корпус русского языка (НКРЯ) предоставляет практически неограниченную исследовательскую и учебную базу для специалистов по русскому языку как иностранному, а также для изучающих русский язык как родной [88, с. 111], по сути не являясь учебным корпусом, НКРЯ обладает огромным потенциалом в процессе обучения русскому языку как иностранному. Этот корпус отвечает критерию репрезентативности и другим требованиям, предъявляемым к современным корпусам, о чем свидетельствуют его характеристики:

- большой объем в сумме составляет около 2 миллиардов слов [78];
- наличие текстов, относящихся к разным периодам создания, что позволяет проследить изменения в употреблении языковых явлений;
- жанровое разнообразие текстов, которые относятся ко всем основным сферам использования русского языка [26].

Именно эти характеристики позволяют глубоко изучить лексику,

морфологию и синтаксис русского языка, и, возможно, установить динамику языковых изменений. Этот корпус позволяет искать и анализировать тексты для исследований при написании различных работ в рамках научно-исследовательской деятельности [41]. Например, «Русская корпусная грамматика», разработанная в рамках программы Министерства науки и высшего образования, предоставляет синхронное описание фрагментов русской грамматики, выполненное на материалах НКРЯ с использованием количественных методов корпусного анализа [91].

Взаимодействуя с НКРЯ, обучаемый развивает исследовательскую самостоятельность, учится организовывать индивидуальную и совместную работу с информацией, наглядно представлять полученные результаты [108]. Для преподавателей НКРЯ предоставляет обширную базу данных, будь то подготовка учебных материалов перед занятием или разработка заданий на уроках, что позволяет упростить подготовку урока, обогатить виды учебных упражнений и повысить качество преподавания. Например, в Череповецком металлургическом колледже использовали НКРЯ на занятиях по дисциплинам «Русский язык» и «Русский язык и культура речи» [22, с. 65].

Фиксируемые в корпусе ошибки приобретают особую значимость при обучении русскому языку, поскольку они обеспечивают «окно», через которое можно заглянуть в язык учащихся [149, с. 5]. Например, существует практика создания базы данных речевых ошибок детей, не являющихся носителями русского языка, особенно ошибок, допускаемых школьниками [27]. В корпусе русских учебных текстов (КРУТ) объемом около 3,1 млн слов, написанных

студентами разных специальностей, тексты снабжены несколькими типами разметки: метатекстовой, морфологической и разметкой по ошибкам [59]. Все это дает возможность предоставить аутентичные и объективные данные и повысить эффективность анализа ошибок.

Типы, характеристики и частота ошибок, возникающих на каждом этапе процесса обучения иностранному языку, зависят от принятой методики преподавания, мешающего влияния родного языка, индивидуальных способностей учащихся и т. д. Кроме того, они отражают языковые особенности разных учащихся. Таким образом, при отборе материалов и проведении разметки в учебном корпусе важно акцентировать внимание на различных параметрах для того, чтобы обеспечить эффективность использования учебного корпуса при анализе ошибок с разных точек зрения, см., например, схему на Рис. 2.

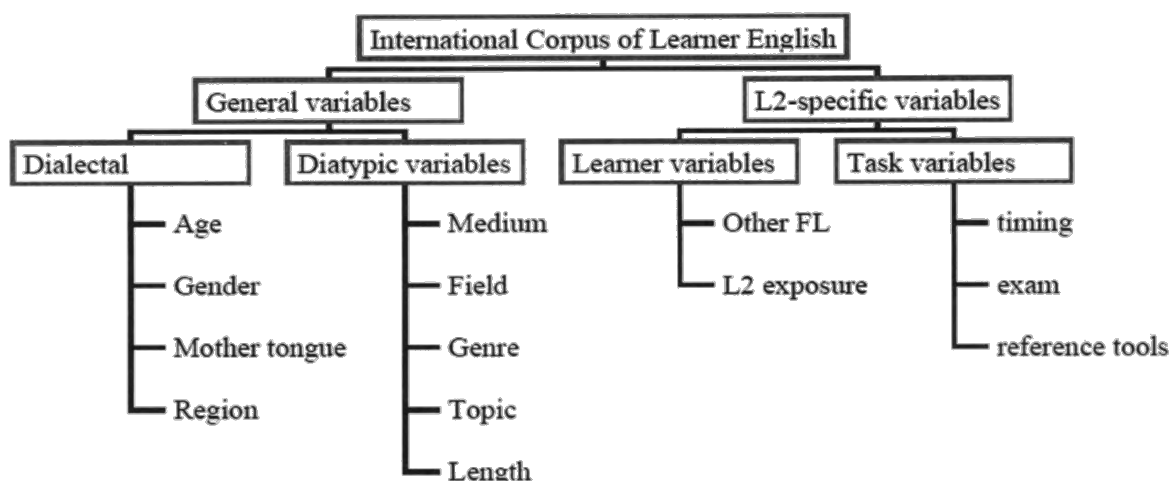


Рисунок 2. Параметры, применяемые для отбора текстов и разметки при создании учебного корпуса (цит. по [133, с. 126])

Так, например, разработанный в Лаборатории по корпусным исследованиям Национального исследовательского университета «Высшая школа экономики» Русский учебный корпус (*RLC*) содержит образцы нестандартной письменной и устной речи изучающих русский язык как иностранный [59; 30, с. 10]. Анализируя языковые ошибки разных студентов, преподаватели могут систематизировать примеры языковых явлений, возникающих в процессе усвоения русского языка иностранными студентами, и понять привычки изучения и особенности иностранных студентов. Это не только помогает преподавателям оптимизировать методы обучения в соответствии с потребностями учащихся, но и своевременно регулировать стратегии обучения [106, с. 69].

Корпус дает возможность не только увидеть те «слабые места», где носителей конкретного языка потенциально поджидают сложности, но и предоставляет богатый материал для подготовки тренировочных и тестовых заданий на типичные ошибки [30]. В настоящее время все больше и больше российских лингвистов начинает изучать практическое использование корпусных технологий, предлагать основные принципы использования корпусов на уроках и методы разработки конкретных упражнений на основе корпусов текстов (см., например, [104; 119]).

Научно-учебная группа студентов и преподавателей Национального исследовательского университета – Высшей школы экономики разработала на основе данных НКРЯ электронное учебное пособие по теме «Русский глагол. Продуктивное префиксальное словообразование», предназначенное для

обучения иностранных студентов русскому языку [96, с. 509]. В состав этого пособия включены не только определенные грамматические упражнения, но и видеопражнения, составленные на основе мультимедийного подкорпуса НКРЯ.

Кроме того, к перспективным направлениям использования корпуса текстов при составлении учебников русского языка относят: поиск общего терминологического ядра, анализ лексического состава и степени доступности содержания текстов в зависимости от возраста учеников, исследование социокультурной информации, а также установление роли иллюстративного компонента в восприятии информации и т. д. [19]. Все эти направления подтверждают большой методический потенциал использования корпусов в обучении русскому языку и, следовательно, привлекает внимание китайских ученых.

По мере всестороннего развития отношений между КНР и РФ в последние годы все больше и больше университетов в Китае открывают специальность русский язык, чтобы готовить профессиональных переводчиков и преподавателей русского языка [94, с. 85]. Более того, многие китайские студенты решили учиться в России, и для них усвоение русского языка является первым шагом на пути к обучению [113, с. 91]. В связи с этим китайские лингвисты уделяют внимание корпусной лингвистике с целью повышения качества и эффективности обучения русскому языку [112, с. 126].

В настоящее время использование корпусной лингвистики в образовании только начинает развиваться в Китае, при этом большинство исследований

сосредоточено на преподавании языков и переводе. На сайте *СНКІ* можно найти более 40 работ, выполненных в период 2013-2023 гг., использующих корпусы для изучения особенностей русской лексики и грамматических явлений, построения корпуса и т. д. Большинство этих исследований используют НКРЯ, например, с помощью НКРЯ проводили синхронное описание слова *поезд*, изучение семантических, синтаксических и прагматических различий между разными значениями предлогов; исследование лингвистических особенностей грамматической структуры *только и (делать / знать / уметь), что...* и т. д.

В то же время китайские исследователи также активно создают свои собственные корпусы русского языка и пытаются применить их в исследованиях и преподавании русского языка. Корпус, созданный профессорами Чжан Лупэном и Чжан Чаоцзином, в основном был отобран из аутентичных текстов, опубликованных основными российскими СМИ на политические, военные и социальные темы, его объем составляет 50 миллионов слов [181, с. 62]. В исследовании рассматривается использование корпусной технологии для поиска лексики и коллокаций, анализа конкретных слов, составления списков тематических слов в политической и военной сферах, а также изучаются статистические методы для определения вероятности грамматических категорий, что улучшает подготовку учебных материалов и разработку учебных программ по соответствующим темам.

Китайский учебный корпус русских текстов (*WRCCL*) был создан Ли Вэньсином и Гао Фэнланем. Аутентичный текстовый материал был собран в

30 разных университетах в период с 2015 года по 2016 год [168, с. 62]. Авторами текстов являются 3413 студентов, учащихся по специальности «Русский язык» на разных этапах обучения. Объем этого корпуса составляет более 600 тыс. слов, на его основе устанавливаются характеристики владения языком у китайских студентов, учащихся по специальности «Русский язык», при употреблении слов и формировании предложений, выявляются проблемы китайских студентов на разных уровнях обучения, возникающие при использовании лексики и синтаксиса. Это исследование используется для того, чтобы разработать методические рекомендации для преподавателей.

В переводоведении и подготовке переводчиков в Китае также внедряют корпусные технологии. На основе параллельного корпуса чеховских романов и их переводов на китайский язык, разработанного учеными Лю Мяо и Шао Цином, исследуются методы создания переводного корпуса и перевода русских произведений; профессор Юань Вэй дает представление о создании, оценке и применении параллельного корпуса русско-китайских новостей; ученые Цуй Вэй и Чжан Лань создали параллельный русско-китайский переводной корпус с целью гармонизации военной и дипломатической терминологии и формирования переводного словаря; Китайско-русский двуязычный корпус, на основе которого были собраны китайско-русские двуязычные тексты различных жанров и тематик, играет особую роль в изучении вопросов, связанных с переводом китайской специфической лексики, уникальной структурой предложений китайского языка [164, с. 53]. В то же время эти переводные корпуса способствуют подготовке учебников по

переводоведению в Китае, созданию переводческих дисциплин и оказанию помощи в практической переводческой деятельности [174, с. 174].

Несмотря на то, что в Китае использование корпусных технологий в преподавании русского языка только начинает развиваться, корпус стал одним из незаменимых инструментов для повышения качества и эффективности преподавания русского языка благодаря его аутентичности, простоте использования и разнообразию данных. Однако разработка учебного корпуса русского языка в Китае идет относительно медленно, потому что отсутствуют соответствующие стандарты построения учебного корпуса, принципы разметки ошибок, опыт его применения и т. д. [111, с. 213].

В связи с этим особенно актуальной является разработка пилотного варианта учебного корпуса текстов русского языка. Настоящее исследование посвящено разработке и анализу методического потенциала учебного корпуса текстов русского языка, написанных китайскими студентами по специальности «Русский язык».

Выводы по Главе I

Глава I посвящена описанию состояния корпусной лингвистики и её места в сфере образования, в ней:

- описаны и обобщены этапы развития корпусной лингвистики и основные типы корпусов;
- рассмотрены корпусы текстов на различных языках, особенно корпусы, созданные в РФ и КНР;
- проанализировано место корпусной лингвистики в образовании;
- охарактеризованы возможности учебных корпусов при применении в образовательной деятельности;
- рассмотрены возможности использовать корпусные технологии в изучении и преподавании русского языка.

Постепенное развитие корпусной лингвистики, начиная с конца 1950-х годов, сталкивалось с сомнениями в целесообразности этой технологии в период ее становления. Но по мере развития компьютерных технологий современные корпусы демонстрируют огромный потенциал, а корпусные технологии и методы все шире внедряются в разные области современных исследований. Проведенный анализ подтвердил, что корпусная лингвистика представляет собой область, которая привлекает интересы ученых и активно применяется в сферах, связанных с лингвистикой, особенно в сфере образования. Корпусы характеризуются репрезентативностью, сбалансированностью и ориентированностью, и именно эти делает их применение важным в области лексикографии, лингводидактики,

преподавания иностранных языков и переводоведения.

Как особый вид корпуса, учебные корпуса основываются на аутентичных данных учащихся в процессе изучения ими иностранного языка и фокусируются на ошибках учащихся. Используя учебные корпуса, преподаватели могут быстро и точно анализировать ошибки, допускаемые учащимися в процессе овладения иностранным языком, чтобы постоянно корректировать свою педагогическую деятельность и более эффективно обучать иностранному языку [27, с. 65]. Кроме того, учебные корпуса предоставляют самим учащимся богатые примеры и реальные контексты ошибок, чтобы они могли легко принимать и запоминать языковые правила иностранного языка [113, с. 91].

Хотя в Китае корпусная лингвистика появилась позже, чем в России, в последние годы все больше китайских ученых активно проводят исследования в области корпусной лингвистики, в частности, пытаются создавать учебные корпуса для исследования и практики в сфере преподавания иностранных языков. Они отмечают востребованность внедрения корпусных технологий в педагогическую деятельность и эффективность их использования при обучении. Проведенный анализ показал, что большинство исследований по русскому языку в Китае основано на использовании НКРЯ, созданного в России, однако учебные корпуса русского языка, которые фокусируются на реальной ситуации обучения китайских студентов, разрабатываются и используются крайне мало.

Эффективность использования корпусных технологий в педагогической

деятельности и отсутствие учебных корпусов русского языка, направленных на обучение китайских студентов, являются основанием, подтверждающим актуальность и перспективу дальнейшего развития корпусной лингвистики в сфере образования в Китае.

ГЛАВА 2

СОЗДАНИЕ УЧЕБНОГО КОРПУСА РУССКОГО ЯЗЫКА ДЛЯ ОБУЧЕНИЯ КИТАЙСКИХ СТУДЕНТОВ

Лингвистический корпус позволяет ускорить исследования языка и многократно повысить их эффективность, достоверность и проверяемость, а также получить результаты, требующие обработки таких массивов текстов, с которыми обычный исследователь справиться просто не в состоянии. Именно эти преимущества привели к тому, что все больше китайских и русских ученых активно экспериментируют с использованием корпусных технологий в преподавании русского языка, в частности, с учебным корпусом, который дает возможность выявить реальные ситуации обучения учащихся. Необходимо отметить, что на сегодняшний день ни в Китае, ни в России не существует учебного корпуса, специально ориентированного на китайских учащихся, у которых специальностью является русский язык. Эта ситуация является серьезным препятствием для практического преподавания русского языка как иностранного и развития этой отрасли методической науки.

Таким образом, для дальнейшей оптимизации стратегии и содержания обучения русскому языку в Китае необходимо создание специализированного учебного корпуса русского языка, предназначенного для китайских студентов. В связи с этим данная глава посвящена рассмотрению методов и принципов создания учебного корпуса русского языка. Процесс его создания можно разделить на три основных этапа:

1. Отбор письменных текстов на русском языке, написанных китайскими

студентами для создания учебного корпуса.

2. Проведение разметки ошибок и метаразметки в учебном корпусе.

3. Идентификация и классификация ошибок, допускаемых китайскими студентами в процессе овладения русским языком.

2.1. Процедура отбора письменных текстов для создания учебного корпуса русского языка

В области корпусной лингвистики лингвистические корпуса соответствуют четырем признакам:

- 1) расположение корпуса на машинном носителе;
- 2) стандартизованное представление словесного материала на этом машинном носителе, позволяющем применять стандартные программы его обработки;
- 3) репрезентативность как результат особой процедуры отбора;
- 4) конечный размер [5].

Для того чтобы обеспечить вышеуказанные характеристики корпуса, языковые материалы должны отвечать ясным и четким критериям отбора:

- заранее определенное происхождение и состав текста;
- синхронизация (материалы созданы в заранее установленный период);
- продуманная представленность различных жанров;
- доступность для компьютерной обработки [92; 118, с. 128].

На основе перечисленных выше критериев при создании пилотного варианта учебного корпуса как особого вида корпуса учитывались следующие параметры отбора текстов на русском языке:

1) **Разговорный стиль.** С целью удовлетворения специфической потребности общения в той или иной области выделяется 5 функциональных стилей (официально-деловой, научный, разговорный, религиозный и художественный) [117], каждый из них имеет свои языковые характеристики. В отличие от разговорного стиля, остальные четыре стиля текста имеют фиксированное выражение и структуру, которые не могут объективно и реально выявить настоящую ситуацию владения учащимися русским языком. Кроме того, разговорный стиль более привычен для учащихся и позволяет им более свободно выражать свои мысли, поэтому в корпус отбирались тексты, написанные разговорным стилем.

2) **Тема текстов.** Чтобы студенты не допускали ошибок из-за непонимания тем, предлагаемых для рассмотрения, в процессе создания учебного корпуса русского языка были выбраны 10 тем для сочинений, близкие к жизни учащихся: «Мое интересное путешествие», «Мой родной город/Мой любимый город», «Интернет в нашей жизни», «Мои увлечения», «Моя студенческая жизнь», «Книга в нашей жизни», «Как я провел(а) свой день рождения», «Моя мечта», «Как мы провели праздник Весны», «Самый любимый человек».

При этом учитывалось, что выбранные сочинения — это обычные домашние задания студентов, а не экзаменационные тексты. Дело в том, что для экзамена учащиеся заучивают готовые тексты и избегают использования тех грамматических правил, которыми они недостаточно хорошо владеют, что не дает возможности оценить реальное владение языком, а домашние

письменные задания точно показывают уровень усвоения учащимися русского языка. Более того, преподаватели просят студентов воздержаться от использования таких вспомогательных средств как электронные и бумажные словари при выполнении ежедневных заданий, чтобы сочинения могли дать наиболее полное представление о реальном уровне знаний студентов. Фрагмент письменных текстов представлен в Приложении 2, полные тексты были загружены в систему Интернет на сайт Яндекс-Диск [<https://disk.yandex.ru/d/FIWmfjglejynBA>].

3) Уровень лингвистических знаний и общие характеристики авторов текстов. С целью рассмотрения влияния английского языка на усвоение русского языка, необходимо убедиться, что студенты владеют английским языком на разном, но определенном уровне. Все авторы текстов в данном корпусе набрали 90 баллов и выше (из 150) на ЕГЭ Китая, что означает, эти студенты имеют в английском языке словарный запас объемом в 3000–3500 слов и могут использовать английский язык для повседневного аудирования, чтения, письма и разговора. Владение английским языком студентов было разделено на четыре уровня в соответствии с их баллами по английскому языку. Подробнее об уровнях владения английским языком и их разметке см. в разделе 2.2.

Кроме того, все авторы текстов в корпусе владеют 1200–2200 русских слов и могут достичь уровня A2 по русскому языку. Они находятся в возрастной группе от 19 до 21 года, материалы представителей этнических меньшинств не использовались. Это позволяет избежать ошибок, допускаемых учащимися в

силу уровня овладения русским языком, возраста и этнической принадлежности. В итоге для создания учебного корпуса русского языка отобраны тексты, которые были написаны 90 китайскими студентами, учащимися по специальности «Русский язык». Информация об этих студентах представлена в Приложении 1.

4) Репрезентативность корпуса. Репрезентативность корпуса определяет достоверность полученных на его материале результатов и также является одним из важных критериев отбора текстов для создания корпуса. Современная корпусная индустрия, стремясь к обеспечению репрезентативности, идет по пути наращивания объема корпусов [13, с. 163], именно объем является ключевым моментом для его достижения. А объем и размер корпусов зависит от цели исследования [39, с. 73]. В данном исследовании целью создания учебного корпуса русского языка является анализ ошибок, допускаемых китайскими студентами в процессе обучения русскому языку, поэтому объем корпуса тесно связан с идентификацией ошибок учащихся. Для того чтобы обеспечить точность идентификации ошибок в учебном корпусе, первоначальная идентификация и классификация ошибок была сопоставлена с коррекцией, выполненной преподавателями вуза. По мере увеличения объема корпуса новых типов ошибок не было выявлено, в итоге в учебный корпус русского языка были отобраны 457 сочинений, и их общий объем составляет 82 064 слова.

5) Тексты машиночитаемого формата. Собранные материалы учебного корпуса представляют собой электронные тексты, созданные за период

онлайн-обучения, которые были преобразованы из формата *.dock* в формат *.txt* (кодировка *UTF-8*) после первоначального отбора для облегчения создания учебного корпуса.

В качестве корпусного менеджера русских текстов использована специальная программа *AntConc 3.5.8*, которая работает на персональном компьютере. Эта программа характеризуется простотой использования и открытым доступом [125, 137, 138]. *AntConc* позволяет искать разметку ошибок в массиве текстов, представлять их количественные показатели, формировать конкорданс [61, с. 36]. Эта мощная функциональная система предоставляет техническую поддержку для анализа ошибок в учебном корпусе русского языка.

Перед разметкой ошибок письменные сочинения были первоначально проверены, чтобы исключить похожие тексты (плагиат), которые не могут показать настоящую ситуацию владения студентами русским языком. Следует отметить, что при обработке сочинений необходимо было соблюдать принцип аутентичности, т. е. с уважением относиться к оригиналам без внесения изменений, что обеспечивает введение текстов в компьютер с целью получения достоверных результатов.

2.2. Принципы и методы разметки в учебном корпусе русского языка

Под учебным корпусом поднимается электронный сборник аутентичных текстов, написанных лицами, изучающими иностранный язык или второй язык [134, с. 12]. Аутентичность языковых материалов, большой объем материала обучающихся и удобные средства поиска на основе корпусных и компьютерных технологий – все это делает учебный корпус важным ресурсом для изучения и преподавания иностранного языка. Разметка является ключевой частью отражения разнообразных данных о представленных в корпусе материалах [33, с. 45] и придает корпусу особую ценность.

Под разметкой (англ. *tagging, annotation*) понимается процесс или результат приписывания текстам и их компонентам специальных меток – тегов [131, с. 310]. Разметка даёт возможность идентифицировать тексты по различным параметрам, позволяя осуществлять осмысленный поиск по корпусу [29, с. 514], а также увеличивает возможность многократного использования корпуса. Кроме того, сами теги разметки обладают конкретными признаками, которые могут четко отразить семантические характеристики языковой единицы [67]. На основе количественного анализа результатов разметки можно обнаружить языковые явления, которые трудно установить, основываясь на личном опыте, что может в свою очередь содействовать развитию лингвистических и методических исследований [112, с. 127].

Разметка ошибок (англ. *error tagging*), которая выявляет «слабые места» учащегося, является не только уникальным типом разметки, характерным

именно для учебного корпуса, но и одним из наиболее важных типов разметки корпуса [132, с. 18]. Используя разметку ошибок, преподаватели могут быстро и точно анализировать ошибки, допускаемые учащимися в процессе овладения иностранным языком, чтобы постоянно корректировать свою педагогическую деятельность и более эффективно обучать иностранному языку [27, с. 65]. Кроме того, разметка ошибок в учебном корпусе предоставляет самим учащимся богатые примеры и реальные контексты ошибок, чтобы они могли легче принимать и запоминать языковые правила иностранного языка [111, с. 214]. Можно сказать, что разметка ошибок актуализирует применение учебного корпуса на практике.

А. Принципы разметки ошибок в учебном корпусе русского языка.

Основные принципы разметки корпуса были предложены Джефффри Личем [142, с. 277]:

- разметка должна быть независима от текста: должна быть возможность убрать разметку и просмотреть текст без неё;
- принципы разметки и их разработчики должны быть известны конечному пользователю;
- пользователь должен быть поставлен в известность о том, что разметка представляет собой лишь потенциально полезный инструмент;
- в основу разметки должны быть положены общепринятые и, по возможности, теоретически нейтральные лингвистические принципы;
- ни одна разметка не может априорно считаться стандартом.

Резюмируя, следует отметить, что разметка тесно связана с применением

корпуса, она должна быть максимально удобной для пользователей. Следовательно, для облегчения практического применения теги разметки ошибок должны отличаться от представления текстов в учебном корпусе русского языка. Помимо этого, при этом подходе разметка ошибок не изменяет исходный корпус, что позволяет в наибольшей степени сохранить аутентичность корпуса [111, с. 215].

Кроме следования вышеуказанным принципам разметка ошибок учебного корпуса должна учитывать такие параметры как классификация ошибок, выбранный метод разметки, характеристика ошибок учащихся и т. д. На основе изучения знаний китайских студентов по специальности «Русский язык» и с учетом перечисленных выше факторов предлагаются следующие принципы разметки ошибок в учебном корпусе русского языка:

1) Простота тегов разметки, что позволит обеспечить их практическую применимость.

Механизм разметки корпуса может быть автоматическим, ручным или автоматическим с ручной корректировкой [38, с. 48]. В процессе разметки учебного корпуса были опробованы некоторые программы автоматического аннотирования, такие как *Dossano*, но результаты этой разметки плохо представлены в корпусном менеджере *AntConc*. В связи с этим для разметки в учебном корпусе русского языка в качестве основной была выбрана ручная разметка [111, с. 215]. Ручная разметка требует много времени и сил, следовательно, теги должны быть краткими и простыми в применении. С точки зрения пользователя, чем более подробно аннотирован корпус, тем

лучше, но для практики также необходимо учитывать выполнимость и целесообразность разметки. Детальность разметки ошибок способствует информативности извлечения ошибок, но стремление к чрезмерно детальной разметке приводит к появлению слишком большого количества тегов. Это не только усложняет ручную разметку, но и не способствует ее практическому использованию [111, с. 216].

2) Иерархическое устройство разметки, что обеспечивает обзорность классов ошибок.

В соответствии с методом классификации ошибок в учебном корпусе русского языка, устройство разметки ошибок должно быть иерархическим, обеспечивающим обзорность классов ошибок учащихся [66, с. 150], например: орфографические ошибки отмечаются строчными латинскими буквами $\langle fm \rangle$, неверное словообразование и неверное использование прописной буквы, которые относятся к тому же типу ошибок, помечаются тегами $\langle fm1 \rangle$, $\langle fm2 \rangle$.

3) Возможность индивидуально настраивать разметку на основе практической цели исследования.

Как правило, разметка учебного корпуса связана с особенностями его применения, хотя, естественно, можно выделить те типы разметки (метаразметка, морфологическая разметка и др.) которые являются универсальными. В случае учебных корпусов в зависимости от планируемого применения корпуса уровень и направленность разметки различны и индивидуальны [182, с. 2]. В случае учебного корпуса русского языка для

обучения китайских студентов следует учитывать, что большинство студентов, обучающихся по специальности «Русский язык» в китайских университетах, изучало английский как второй иностранный язык в начальной, средней и старшей школе. Поэтому на начальном этапе изучения русского языка им часто мешает английский язык и возникает дополнительная интерференция. Некоторые студенты изучают русский язык со старшей школы, поэтому английский язык на них влияет меньше. Таким образом, при разметке следует учитывать наличие второго иностранного языка учащихся и степень их владения иностранным языком для дальнейшего анализа причин возникновения ошибок [111, с. 217].

Рассмотренные выше принципы разметки реализованы при решении задач создания исследовательского учебного корпуса русского языка.

Б. Использованные методы разметки при создании учебного корпуса.

С точки зрения современной корпусной лингвистики корпус должен быть большим, размеченным, репрезентативным и представленным в машиночитаемом формате [11, с. 47]. Ключ для реализации машиночитаемого формата корпуса и увеличения ценности его применения, заключается в разметке корпуса [112, с. 127].

С момента появления крупномасштабных корпусов текстов методы разметки изменились. В настоящее время существует два основных типа системы разметки: первая — справочная система *Soso*, которая является самой ранней компьютерной системой для извлечения индексных линий из машиночитаемых текстов. Вторая — система разметки *TEI* (*Text Encoding*

Initiative), на основе которой в настоящее время построены многие крупные корпуса [112, с. 127].

Особенность справочной системы *Cocoa* заключается в том, что она позволяет эффективно извлечь информацию текстов в кратчайшие сроки [147]. Например, эта система разметки используется в корпусе Лонгман-Ланкастер (*Longman-Lancaster Corpus*) и в Хельсинкском аннотированном корпусе русских текстов. Однако, справочная система *Cocoa* может аннотировать только ограниченную информацию текста и, скорее, является системой метаразметки, которая позволяет размечать автора, дату и темы, в результате эта система не может соответствовать все более сложным и диверсифицированным требованиям к современной разметке корпуса [112, с. 127].

Система разметки *TEI* разработана при участии таких международных организаций как Ассоциация по вычислительной лингвистике (*Association for Computational Linguistics*), Ассоциация по компьютерам и гуманитарным наукам (*Association for Computers and the Humanities*) и Ассоциация по компьютерным технологиям в литературе и лингвистике (*Association for Literary and Linguistic Computing*) [163, с. 55]. По сравнению со справочной системой *Cocoa* система *TEI*, которая основана на обобщенном стандартном языке разметки *SGML* (*Standard General Markup Language*) и языке *XML* (*Extensible Markup Language*), в настоящее время является самой передовой международной системой разметки. Система *TEI* дает информацию об электронных текстах (идентификация текста, представление, анализ и

интерпретация, метаязык описания и кодировки) [35, с. 203]. Эта система обладает большой гибкостью, полной комплексностью, что предоставляет возможность идентификации и кодирования разных форматов текста [18, с. 30].

Язык *XML* как технологическая основа для системы *TEI* позволяет задавать структуру любого документа с помощью набора вложенных друг в друга элементов (дерева *XML*) [97, с. 93]. Элементы отражаются с использованием открывающих и закрывающих тегов, имеющих вид <имя элемента> и </имя элемента>, например: Буква «*N*» является тегом существительного, буква «*P*» — фразы, а именная группа типа «институт русского языка» помечается как <*NP*> институт русского языка</*NP*> [112, с. 128].

Разметка в рамках системы *TEI* может быть разделена на две части: первой частью является контекстуальная информация (называемая <*teiHeader*>), которая содержит общее описание документа: заголовок, автор, время написания и другую метаинформацию (Рис. 3).

```
<teiHeader>
  <fileDesc> <!-- начало элемента <fileDesc> -->
  <titleStmt>
    <title>Повесть «Детство». Электронное издание.</title> <!-- назва-
ние нашего нового электронного документа -->
    <author>Толстой Л.Н.</author> <!-- автор текста-->
    <editor>Иванов И.И.</editor> <!--здесь может быть указан редактор
документа или составитель коллекции-->
```

Рисунок 3. Пример <*teiHeader*> повести Л. Н. Толстого «Детство»

(цит. по [97, с. 95])

Вторая часть предоставляет лингвистическую информацию о тексте (называемую *<text>*), включающую структуру текста, морфологическую разметку, синтаксическую разметку и т. д. (Рис. 4).

```
<text>
  <I>Вы хотите быть игрушечной</I>
  <I>Но испорчен Ваш завод,</I>
  <I>К Вам никто на выстрел пушечный</I>
  <I>Без стихов не подойдет.</I>
</text>
```

Рисунок 4. Пример разметки типа *<text>* художественного текста

(цит. по [97, с. 93])

В самом общем виде система разметки *TEI* представлена следующим образом, см., например:

```
<TEI>
  <teiHeader> ... </teiHeader>
  <text> ... </text>
</TEI>
```

Опираясь на систему разметки *TEI*, можно разделить виды разметки китайского учебного корпуса русского языка на две части: первая — метаразметка (экстралингвистическая разметка), которая включает контекстуальную информацию, информацию об авторе и т. д. [113, с. 92]. Вторая — собственно разметка ошибок, которые допускают китайские студенты при овладении русским языком. При создании китайского учебного

корпуса русского языка эти виды разметки были специально настроены в соответствии с исследовательскими задачами и характеристиками знаний китайских студентов [114, с. 198]. Рассмотрим эти виды разметки учебного корпуса русского языка.

1) Метаразметка (экстралингвистическая разметка)

Как правило, разметка учебного корпуса связана с особенностями его применения. Метаразметка может наглядно предоставить метаданные по тексту, позволяет осуществлять поиск по запросу пользователей, а также обеспечивает поддержку для контрастного анализа и анализа ошибок учащихся. В случае учебных корпусов в зависимости от планируемого применения корпуса уровень и направленность разметки различны и индивидуальны [182, с. 2]. Так, например, в случае учебного корпуса русского языка для обучения китайских студентов следует учитывать, что китайские студенты, обучающиеся по специальности «Русский язык» в китайских университетах, уже имеют опыт изучения английского языка [112, с. 128].

В связи с этим английский язык часто им мешает и приводит к ошибкам, особенно на начальном этапе изучения русского языка, поэтому эти характеристики авторов текстов должны быть указаны при метаразметке корпуса [111, с. 216].

Таким образом, помимо такой основной метаразметки как возраст, пол, тема текста, следует учитывать наличие второго иностранного языка учащихся и степень владения этим языком для дальнейшего анализа причин возникновения ошибок [112, с. 128].

Для метаразметки материалов корпуса все тексты были пронумерованы в соответствии с различными темами (см. Табл. 2).

Таблица 2

Теги метаразметки тем в учебном корпусе русского языка

Тема письменных текстов	Тег разметки темы
Мое интересное путешествие	1
Мой родной город/мой любимый город	2
Интернет в нашей жизни	3
Мои увлечения	4
Моя студенческая жизнь	5
Книга в нашей жизни	6
Как я провел(а) свой день рождения	7
Моя мечта	8
Как мы провели праздник Весны	9
Самый любимый человек	10

Теги, задаваемые буквами «*F*» и «*M*», приписывались в соответствии с гендерными характеристиками авторов текстов. Уровень второго иностранного языка определялся в соответствии с баллами по иностранному языку, полученными на ЕГЭ [112, с. 129]. Перечень тегов уровня второго иностранного языка представлен в Таблице 3.

Таблица 3

Теги уровня второго иностранного языка в учебном корпусе русского языка

Оценка по английскому языку на ЕГЭ	Тег английского языка
150-135	A
134-120	B
119-105	C
104-90	D

Тег A означает, что баллы по английскому языку находятся в интервале 150–135, тег B в интервале 134–120, тег C в интервале 119–105, тег D в интервале 104–90. Так, например, тег <4.23.20.F.C> означает, что тема текста

– Мои увлечения, автор - студент с кодом 23, возраст 20 лет, пол женский, баллы по английскому языку находятся в интервале 119–105 [112, с. 129]. С помощью тегов метаразметки можно позволить пользователям не только найти подходящие языковые материалы в соответствии с их собственными потребностями, но и предоставить данные для анализа причин возникновения ошибок учащихся и особенностей овладения ими русским языком [112, с. 129]. Например, в учебном корпусе русского языка можно быстро извлечь письменные тексты студента с кодом 25 (Рис. 5).

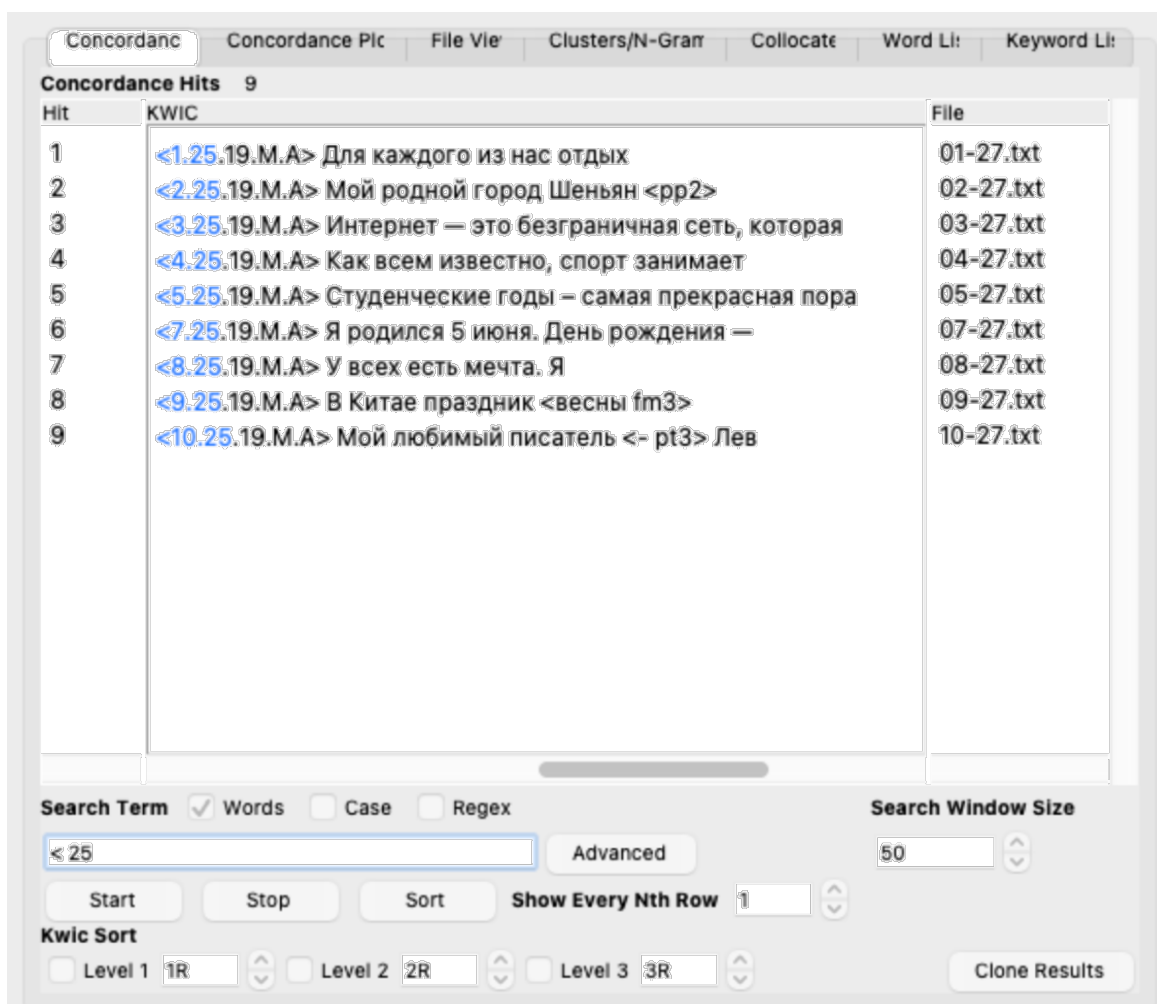


Рисунок 5. Фрагменты текстов студента с кодом 25 в учебном корпусе русского языка

В соответствии с тегами метаразметки, представленными на Рисунке 5, устанавливается, что автор – студент с кодом 25, возраст 19 лет, пол мужской, баллы по английскому языку находятся в интервале 135–150. В учебный корпус русского языка включены его 9 текстов на разные темы.

С помощью функции *File View* в инструменте *AntConc* также можно показать полные тексты, написанные этим студентом в соответствии с требованием поиска. Все это актуализирует анализ ошибок и особенностей усвоения русского языка на основе учебного корпуса русского языка.

2) Разметка ошибок

Разметка ошибок является обязательной и ключевой частью при создании учебного корпуса, а также важным предметом исследования в этой области. На основе разметки ошибок, применяя инструменты поиска, можно значительно повысить эффективность извлечения и количественного анализа ошибок, более всесторонне и систематически отразить распределение ошибок учащихся [112, с. 129].

На основе приведенных выше принципов и проведенного предварительного исследования массива текстов, а также с использованием принятой в Китае классификации ошибок на основе Учебного корпуса английского языка при Гонконгском университете науки и технологии [175], Китайского учебного корпуса текстов английского языка [179], была разработана базовая система тегов ошибок на основе классификации ошибок по аспектам языка (см. Табл. 4) [112, с. 130]. Подробная классификация ошибок приведена в разделе 2.3.

Таблица 4

Теги ошибок для разметки в учебном корпусе русского языка

Тип ошибки	Подкатегория ошибки	Название ошибки	Тег
1	2	3	4
орфографические ошибки	нарушено правописание	орфографическая ошибка за счет незнания орфограмм	fm1
		неверное словообразование	fm2
		неверное использование прописной буквы	fm3
морфологические ошибки	ошибка в использовании глаголов	неверное использование возвратной формы	vp1
		неверное использование форм несовершенного или совершенного вида	vp2
		неверное управление глагола	vp3
		неверное использование деепричастия	vp4
		неверный выбор вида причастия	vp5
	ошибка в использовании имен существительных	неверное использование форм числа имен существительных	np1
		отсутствие согласования по роду	np2
		неверное образование форм склонения	np3
		нарушение падежных форм	np4
	ошибка в использовании предлогов	неверное управление предлога	pp1
		использование неверного предлога	pp2
	ошибка в использовании имен прилагательных	неверное использование полной или краткой формы	aj1
		неверное использование степеней сравнения	aj2
	ошибка в использовании местоимений	неверный выбор местоимения	pr1
	ошибка в использовании наречий	неверное использование степеней сравнения	ad1
		неверное использование форм долженствования	ad2
	ошибка в использовании числительных	неверное использование форм числительных	nu1

Таблица 4. Продолжение

1	2	3	4
лексические ошибки	ошибка в использовании слова	употребление слова в несвойственном ему значении	wd1
		тавтология и повтор	wd2
		смешение синонимов и паронимов	wd3
	неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	cc1
		фразеологическая ошибка	cc2
синтаксические ошибки	нарушение структуры предложений	пропуск членов предложения	sn1
		неверный порядок слов	sn2
	нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	sn3
		нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях	sn4
	пунктуационная ошибка	отсутствие знака препинания	pt1
		лишний знак препинания	pt2
		неверное использование знака препинания	pt3
коммуникационные ошибки	логическое нарушение	нарушение логики при перечислении	lc1
		отсутствие формальных показателей связи между предложениями	lc2

Разметка ошибок проведена автором диссертации, ее результаты выборочно проверялись 3 экспертами — преподавателями русского языка во Втором Пекинском университете иностранных языков и Хэнаньском университете науки и технологий.

С учетом необходимости проведения разметки вручную вид разметки ошибок был определен на основе системы *TEI*. В соответствии с этим в учебном корпусе русского языка ошибки помечаются строчными латинскими буквами и цифрами, включенными в угловые скобки <...>. Например: *Я люблю*

путешествовать и <побывал много мест врз> Китая [112, с. 129].

Рассмотрим в качестве примера неверное использование прописной буквы (тег *fm3*) в учебном корпусе русского языка (Рис. 6).

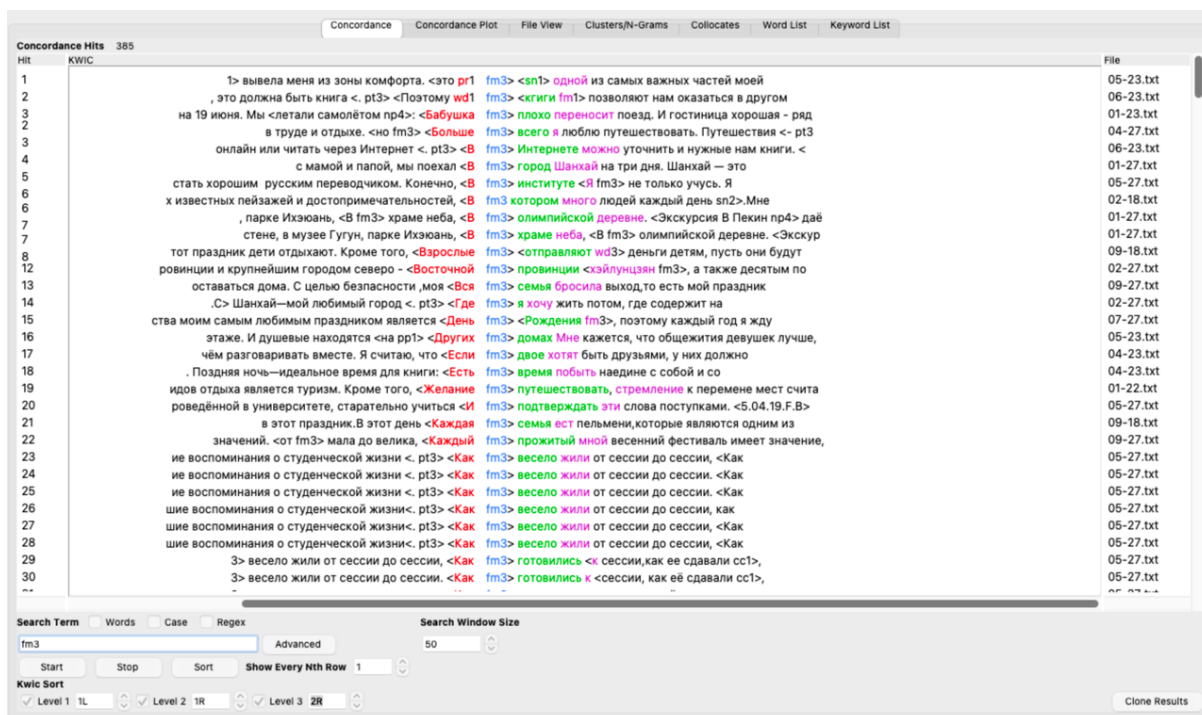


Рисунок 6. Пример разметки ошибок (тег ошибки *fm3*) в учебном корпусе русского языка

Согласно результатам, представленным на Рисунке 6, с помощью разметки ошибок учебного корпуса было обнаружено 385 ошибок использования прописной буквы в письменных текстах на русском языке, написанных китайскими студентами. Кроме того, в зоне *KWIC* показаны ошибки и их контекст, в зоне *File* — точное место ошибок, т. е. информация о том, в каком сочинении студенты допускают такие ошибки. Опираясь на метаразметку учебного корпуса, можно проанализировать, связаны ли такие ошибки с темой текста и уровнем владения студентом английским языком.

Приведем пример разметки текста китайского учебного корпуса русского языка, см., например текст [112, с. 131]:

<03.17.19.F.C>

С ростом <научной техники ss1>, всё больше и больше людей начинают использовать сеть. <Все известно ad2>, что Интернет делает нашу жизнь удобнее. <даже fm3> некоторые люди <предаются vr2 компьютером vr3>. Так, что <приносят нам Интернет sn3>? Пользу или вред?

По моему мнению, <Интернет приносят sn3> нам сегодня как позитивные, так негативные <впечатления wd3>. А для меня <Интернет приносят sn3> нам <больше пользу ad1>, чем вред.

Интернет предоставляет нам возможность искать материалы и рефераты. Просто нужно щёлкать мышью и на экране появляются многие <информайии fm1>. Так удобно! Интернет <способствует общение vr3> с разными людьми. Есть <огромное Интернет-площадок pr3> для общения, например: форум, чат, аська и так далее. <Мы sn1 на связи <без выходя pp1> из дома>. <Интернет делать sn4> нашу жизнь насыщеннее и ярче. Мы <заказаем vr2> номер гостиницы, покупаем новые вещи или <найдем vr2> партнеров по бизнесу <в интернет pr1 fm3>. Если у нас нет времени путешествовать, <Интернет дать sn4> нам возможность посмотреть фотографии онлайн.

Конечно, интернет не всемогущ. Некоторые люди обманывают других <в Интернет pp1>, и все больше людей становятся зависимыми от

Интернета. На мой взгляд, это потому, что <этих людей pp2> слабая сила воли. Они закрывают глаза на <актуальные реальные wd2> проблемы. Для них <, pt2> Интернет <pt1> это бегство от жизни. Я хочу сказать этим людям, что за пределами монитора также много всего интересного. Общение с родителями с глазу на глаз и прогулка в парке с друзьями могут приносить огромное количество эмоций. Компьютер и сеть только лишь та часть, которая <помогает fm1> нам лучше жить.

В соответствии с тегом метаразметки <03.17.19.F.C> тема текста – Интернет в нашей жизни, автору присвоен код 17, возраст 19 лет, пол женский, баллы по английскому языку находятся в интервале 119–105 [112, с. 131]. В тексте эта студентка допустила 27 ошибок, большинство ошибок связано с неверным использованием частей речи и нарушением норм формального синтаксиса, например, неверное управление глагола (<способствует общению vr3> вместо *способствует общению*), неверное управление предлога (<в Интернет pp1> вместо *в Интернете*), нарушение согласования подлежащего и сказуемого (... что <приносят нам Интернет sn3> вместо ... что *приносит нам Интернет*, нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях (<Интернет дать sn4> *нам возможность ...* вместо *Интернет дает нам возможность ...*). Именно эта информация, полученная с помощью разметки, облегчает анализ причин возникновения ошибок и факторов, влияющих на усвоение русского языка китайскими студентами.

Кроме того, как видно из вышеуказанного текста, включенного в учебный

корпус с разметкой, при этом подходе разметка ошибок не изменяет исходный текст, что позволяет в наибольшей степени сохранить аутентичность учебного корпуса русского языка и также упростить работу по ручной разметке ошибок [112, с. 131].

При этом нужно учесть, что поскольку некоторые предложения относительно длинные, на странице поиска не может отображаться полное предложение и перечень ошибок. Посредством применения функции *File View* в инструменте *AntConc* можно просмотреть полное предложение и контекст ошибки, чтобы лучше проанализировать причины возникновения ошибок учащихся при овладении русским языком [112, с. 131]. Рассмотрим в качестве примера отсутствие формальных показателей связи между предложениями (тег *lc2*) в учебном корпусе русского языка (Рис. 7).

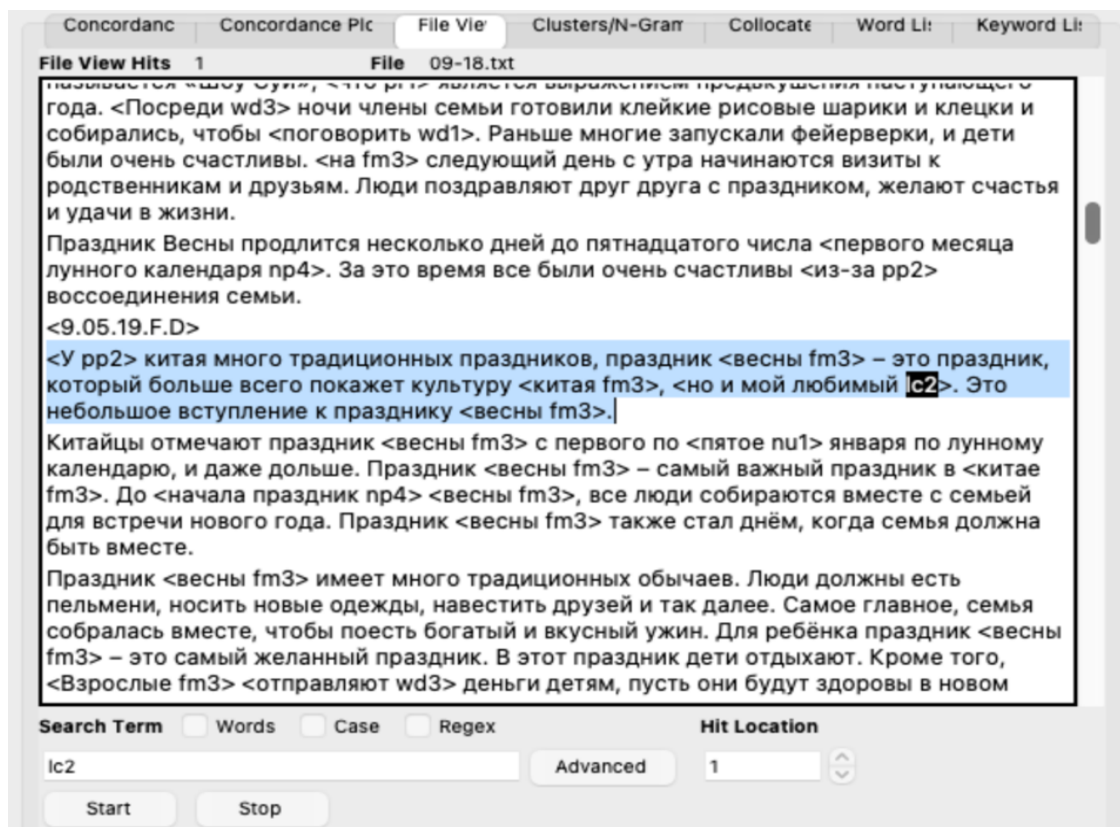


Рисунок 7. Пример результатов работы программы *AntConc - File View*

В примере можно увидеть логическое нарушение, которое необходимо анализировать с помощью контекста, например ... *это праздник, который больше всего покажет культуру Китая, <но и мой любимый Is2>* вместо ... *и мой любимый.*

Вышеописанная классификация и принципы разметки использованы при создании пилотного варианта учебного корпуса русского языка для обучения китайских студентов.

2.3. Идентификация и классификация ошибок при овладении русским языком в учебном корпусе русского языка

В современной методике обучения русскому языку как иностранному принято выделять несколько видов ошибок, допускаемых иностранными студентами и не только ими [110, с. 249]. Результаты классификации ошибок представлены в работах Ю. В. Фоменко [107], С. Н. Цейтлин [115; 116], Е. А. Гасконь [21], Е. В. Жендаренко [31], Е. В. Рахилина [90], Ю. И. Злобиной [40], Лю Лифэнь [166], Ву Цзюань [172] и др. Наиболее распространенная классификация ошибок в речи иностранных студентов основана на следующих аспектах языка: словообразовательная ошибка (неоправданное сочинительство); морфологическая ошибка (ненормативное образование слов и употребление частей речи); синтаксическая ошибка (неверное построение предложений); лексическая ошибка (нарушение лексической сочетаемости) и стилистическая ошибка (нарушение единства стиля языка) [15; 31, с. 81].

Опираясь на классификацию ошибок по аспектам языка, отдельно рассматривались ошибки в использовании каждой части речи для того, чтобы выявить конкретные трудности китайских учащихся при использовании частей речи и их языковые характеристики при овладении русским языком [110, с. 249]. Дело в том, что такая классификация в основном связана с различием между китайским и русским языками и механизмом разметки ошибки. С одной стороны, между китайским и русским языками существуют большие различия. Русский язык – синтетический, то есть в нем преобладают синтетические формы выражения грамматических значений [40, с. 131]. Грамматические

значения слова обычно передаются с помощью необходимых морфологических форм. В связи с этим русский язык сильно отличается от китайского языка, передающего грамматические отношения через отдельные служебные слова, через фиксированный порядок слов и контекст [110, с. 249]. Из-за этих различий китайские учащиеся часто встречаются с трудностями и допускают больше ошибок в использовании частей речи русского языка [109, с. 106]. И эти ошибки заслуживают особого внимания при классификации и анализе ошибок, допускаемых китайскими учащимися, они должны быть особым образом размечены в корпусе.

С другой стороны, для разметки в учебном корпусе русского языка следует учитывать следующее: нельзя сводить множество типов ошибок, выявленных в конкретном месте текста, к использованию одного идентификатора – тега [110, с. 249]. Если один и тот же тег характеризует множество ошибок, трудно выделить характеристики распространенности ошибок учащихся, это может повлиять на анализ конкретных трудностей в обучении. Кроме того, следует учитывать, что детальность классификации ошибок способствует ее информативности, но одновременно нужно принимать во внимание, что стремление к детальной классификации ошибок приводит к появлению слишком большого количества тегов [110, с. 249]. Это не только усложняет ручную разметку, но и не способствует ее практическому использованию.

Учебный корпус текстов создается на основе данных ошибок, поскольку именно идентификация и классификация ошибок являются актуальной и

ключевой проблемой для создания и разметки данного корпуса. С одной стороны, точность и адекватность классификации и идентификации ошибок определяют эффективность использования учебного корпуса, с другой стороны, они представляют особую трудность для создания корпуса такого типа [110, с. 249]. Поэтому создание учебного корпуса текстов китайских студентов на русском языке требует предварительного рассмотрения общего подхода к определению того, что понимается под этим термином, и классификации ошибок, необходимой для их разметки в учебном корпусе.

А. Идентификация ошибок в учебном корпусе русского языка

Для того чтобы классифицировать ошибки, необходимо определить, что понимается под этим термином. Ошибка появляется в любой сфере человеческой деятельности, но в разных научно-исследовательских сферах и предметных областях она определяется по-разному. В данной работе рассматриваются ошибки, возникающие при овладении иностранным языком [110, с. 247].

Широкое толкование понятия '*ошибка*' определяется в свете отклонения от стандарта и влияния межъязыковой интерференции [122, с. 390]. С точки зрения отклонения от стандарта ошибка рассматривается как нарушение правильного использования языковых единиц и их форм, т. е. результат ошибочного действия [2, с. 182]. При этом отклонением от стандартизованного использования является некорректный выбор единицы из ряда одноуровневых единиц или члена одной парадигмы для данных условий функционирования [57, с. 28]. Отсюда следует, что неверное действие учащегося приводит к

ошибке, а ошибка является проявлением речевого и мыслительного действий студентов.

Ошибка может возникать под влиянием межъязыковой интерференции, под которой понимается расхождение с нормой в системе изучаемого языка под влиянием языка родного [21, с. 81]. Явление межъязыковой интерференции является одной из основных причин возникновения ошибок при изучении иностранного языка. В этом случае ошибка определяется как явление, связанное с отрицательным влиянием знаний родного языка на знание изучаемого языка. Из-за трудности усвоения абсолютно новых, ранее в родном языке не встречавшихся явлений изучаемого языка, обучающийся допускает множество ошибок при его овладении [110, с. 248].

Анализируя ошибки учащихся, можно выявить трудности, возникающие в процессе овладения иностранным языком, и учесть их при определении методической направленности для преподавателя и улучшения качества преподавания. Однако, при таком подходе нет точного критерия для оценки отклонения учащегося от стандарта и влияния именно интерференции, в большинстве случаев установление ошибки зависит от личного опыта, поэтому является субъективным [110, с. 248].

Узкое толкование понятия '*ошибка*' связано с нарушением норм [122, с. 390]. При этом принято определять ошибку с позиции преподавателя и с позиции носителя языка. С точки зрения преподавателя ошибкой является нарушение нормы, которая не только положена в основу учебников по грамматике, но и существует в сознании преподавателя, то есть то, что он

считает неправильным, оценивается им как ошибка [70, с. 65]. Однако и здесь имеются свои недостатки: идентификация ошибки опирается на мнение преподавателя, но трудно объяснить, кто контролирует норму, на которую опирается преподаватель, и как эта норма возникла.

По сравнению с позицией преподавателя определение ошибки с позиции носителя языка более объективно и легче применяется в практических целях. В отношении классификации ошибок, допускаемых учащимися, целесообразна позиция С. Джеймса: ошибка – нарушение норм речи носителя языка [135, с. 67]. На этой основе можно определить ошибку по трем параметрам: грамматичность (правильность), приемлемость (понятность) и применимость:

- грамматичность (правильность), ошибка – нарушение требований нормы, например ошибка в словообразовании, лексическая ошибка и т. д.
- приемлемость (понятность): ошибка – это то, чего не понимает и не принимает носитель языка, например, семантическая ошибка;
- применимость: ошибка – неправильное использование стиля речи, например использование в научном тексте разговорных выражений.

Важным компонентом работы является корпусное исследование ошибок в русском языке, которые допускают китайские студенты, и создание правил их разметки в корпусе. На основе рассмотренного выше определения *ошибка* понимается как нарушение грамматичности, правильности и приемлемости речи с точки зрения носителя русского языка [110, с. 248]. Иначе говоря, нарушение в речи китайских студентов, противоречащее правилам и нормам

русского языка и непонятное носителям русского языка, воспринимается как ошибка.

Следует отметить, что идентификация ошибок в учебном корпусе русского языка в основном зависит от параметров грамматичности и правильности, остальные критерии играют решающую роль только в определенных случаях [110, с. 248]. Вопрос о том, соответствуют ли используемые студентами слова требованиям стиля речи, выходит за рамки данного исследования, поскольку определение таких ошибок более субъективно [110, с. 248–249].

Б. Классификация ошибок в учебном корпусе русского языка

В современной работе по учебным корпусам никак нельзя обойтись без внимания к разработкам группы под руководством Е. В. Рахилиной, в результате которых созданы два представительных корпуса русского языка, включающих тексты студентов-носителей самых разных языков и разных жанров (КРУТ и Русский учебный корпус). Для создания пропедевтического учебного корпуса русского языка также учитывался опыт разработки и классификации ошибок этих корпусов. Однако из-за языковых характеристик китайских учащихся и больших различий между китайским и русским языками при разработке учебного корпуса русского языка в основном наследовался опыт классификации ошибок на основе Учебного корпуса английского языка при Гонконгском университете науки и технологии [175], Китайского учебного корпуса текстов английского языка [179].

Учитывая все сказанное выше, а также проведенное предварительное

исследование массива текстов, первоначально ошибки китайских учащихся были разделены на 5 основных типов (см. Табл. 5) [110, с. 249]. По мере увеличения количества письменных текстов в учебном корпусе русского языка классификация ошибок оптимизировалась [113, с. 92], и в итоге ошибки, допускаемые китайскими студентами при овладении русским языком, разделены на 5 основных типов и 14 подкатегорий, всего выделено 34 варианта.

Таблица 5

Классификация ошибок в учебном корпусе русского языка

Тип ошибки	Подкатегория ошибки	Название ошибки	Пример
1	2	3	4
орфографические ошибки	нарушено правописание	орфографическая ошибка за счет незнания орфограмм	много; правильный; претсдавляет; английский; учят
		неверное словообразование	прийдти; весвый; суперmarket
		неверное использование прописной буквы	Я поздравляю вас с днём Победы!
морфологические ошибки	ошибка в использовании глаголов	неверное использование возвратной формы	Я уже два года не виделась друзей.
		неверное использование форм несовершенного или совершенного вида	Я хочу изучать русский язык в течение двух лет. Давайте защитить окружающую среду вместе. В университете я продолжаю изучить русский язык.
		неверное управление глагола	Родители должны уделять большое внимание своих детей.
		неверное использование деепричастия	Выйдя из метро, пошёл дождь.
		неверный выбор вида причастия	Из всех языков, изучающих в университете, студенты часто выбирают английский язык.

Таблица 5. Продолжение

1	2	3	4
морфологические ошибки	ошибка в использовании имен существительных	неверное использование форм числа имен существительных	Родители стараются создать хорошее условие , чтобы дети получили хорошее образование. Она любит купить одежды .
		отсутствие согласования по роду	Мы должны использовать каждый возможность учиться.
		неверное образование форм склонения	Студенты из деревен учатся лучше, чем студенты из городов.
		нарушение падежных форм	Здесь тоже существует несколько вопросов об использовании Интернетом .
	ошибка в использовании предлогов	неверное управление предлога	Мая мама в свободное время часто гуляет по магазинами .
		использование неверного предлога	В третьем курсе передо мной появилась куча новых вопросов: много новых слов в текстах, чтение статей.
	ошибка в использовании имен прилагательных	неверное использование полной или краткой формы	Из-за того, что мой русский язык слабый , я решила стараться учиться. Я очень согласена с этим мнением!
		неверное использование степеней сравнения	Русский язык более труднее английского.
	ошибка в использовании местоимений	неверный выбор местоимения	Кроме того, ваш друзья и родитель будут дарить вам подарки.
	ошибка в использовании наречий	неверное использование степеней сравнения	Он говорит по-русски лучше в нашей группе.
		неверное использование форм долженствования	Китайцы обычно надо есть лапшу.
	ошибка в использовании числительных	неверное использование форм числительных	В городе ещё один красивая площадь

Таблица 5. Продолжение

1	2	3	4
лексические ошибки	ошибка в использовании слова	употребление слова в несвойственном ему значении	Сейчас большинство мальчиков ценят девушку только по красивому лицу и стройной внешности .
		тавтология и повтор	Следует различать различные подходы к проблеме.
		смещение синонимов и паронимов	Я хочу вступить в аспирантуру после окончания университета.
	неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	получение больших успехов; знания возросли; взять образование; крепко понять; совсем ленивый
		фразеологическая ошибка	Тебе нельзя сидеть, сложив руки и только ожидать помощи.
синтаксические ошибки	нарушение структуры предложений	пропуск членов предложения	Речь идёт о своём дне рождения, мне тоже очень нравится .
		неверный порядок слов	Главное – подаренные цветы с любовью.
	нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	Она улыбнулась и сказал , что ей больше нравится заниматься спортом.
		нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях	По мере того, как развиваются наука и техника, рос поток информации. Однажды моя нога была ранена, я не могу пойти в школу три месяца.
	пунктуационная ошибка	отсутствие знака препинания	На зимних каникулах я заметил что сосед купил сыну велосипед.
		лишний знак препинания	В университете, она часто помогает мне.
		неверное использование знака препинания	Китайская кухня, это самая острая кухня в мире.

коммуникационные ошибки	логическое нарушение	нарушение логики при перечислении	На уроке присутствовали директор, профессора, также Анна Петровна Иванова, Чжан Цин.
		отсутствие формальных показателей связи между предложениями	Мир развивается. Страна развивается. Город развивается. Во многих областях нужна вода...

Опираясь на разработанную классификацию ошибок, можно с помощью учебного корпуса точно анализировать трудности и языковые характеристики китайских студентов при овладении русским языком [113, с. 93]. Этот корпус предоставляет базу для дальнейшего анализа ошибок, допускаемых китайскими студентами при изучении русского языка.

Выводы по Главе 2

С целью выявления и анализа ошибок, допускаемых китайскими студентами в процессе усвоения русского языка, обеспечения аутентичности текстов учебного корпуса русского языка, при создании корпуса были установлены следующие критерии отбора письменных текстов:

- обычные домашние задания с заданными темами (в учебный корпус включаются только обычные домашние задания с темами, которые близки жизни учащихся);
- определенный источник текстов (в корпус отбираются тексты, написанные китайскими студентами с разным, но определенным уровнем владения английским языком);
- репрезентативность (увеличение корпуса не ведет к увеличению числа типов и изменения типологии ошибок).

Согласно всем перечисленным выше критериям для создания учебного корпуса русского языка были использованы 457 письменных текстов по 10 темам, написанных 90 студентами из китайских университетов, обучающихся по специальности «Русский язык». Все тексты преобразованы из формата *.docx* в формат *.txt* для компьютерной обработки. Объем текстов, включенных в учебный корпус русского языка, предназначенный для анализа ошибок, допускаемых китайскими студентами в процессе обучения русскому языку, составляет 82 064 слова.

Разметка текстов в процессе создания китайского учебного корпуса русского языка для обучения китайских студентов проводилась в соответствии

со следующими принципами:

- разметка должна быть независима от текста и при необходимости должна предусматриваться возможность убрать разметку;
- теги разметки должны отличаться от представления текстов в корпусе;
- теги разметки должны быть простыми и индивидуальными для практической применимости;
- устройство разметки ошибок должно быть иерархическим, обеспечивающим обозримость классов ошибок учащихся [112, с. 129–130].

Опираясь на систему разметки (стандарт) *TEI*, разметка учебного корпуса русского языка была разделена на две части:

- первая — метаразметка (экстралингвистическая разметка), например тег *<4.23.20.F.C>* означает, что тема текста – Мои увлечения, автор - студент с кодом 23, возраст 20 лет, пол женский, баллы по английскому языку находятся в интервале 119–105 [112, с. 128]. Следует отметить, что уровень владения английским языком китайских студентов был включен в метаразметку с целью анализа влияния английского языка на обучение русскому языку.

- вторая — разметка ошибок, которые помечаются строчными латинскими буквами и цифрами, заключенными в угловые скобки *<...>*, например: *Я люблю путешествовать и <побывал много мест yr3> Китая* [112, с. 129].

Для того чтобы выявлять и классифицировать ошибки, необходимо определить, что понимается под термином *ошибка*. В учебном корпусе русского языка термин *ошибка* поднимается как нарушение грамматичности,

правильности и приемлемости речи с точки зрения носителя русского языка. Это значит, что ошибка в речи китайских студентов, противоречащая правилам и нормам русского языка, воспринимается как неправильность [110, с. 248].

Разработанная классификация ошибок в основном связана с различием между китайским и русским языками и определяется механизмом разметки ошибки. По мере увеличения количества письменных текстов в учебном корпусе русского языка классификация ошибок оптимизировалась, и в итоге ошибки, допускаемые китайскими студентами при овладении русским языком, разделены на 5 основных типов и 14 подкатегорий, всего выделено 34 варианта [110, с. 249]. Предложенная разметка позволяет создать высокоэффективный учебный корпус как поисковую систему.

Из-за большого объёма учебного корпуса русского языка в работе показан только фрагмент данного корпуса (см. Приложение 3), а полный учебный корпус русского языка был загружен в систему Интернет на сайт Яндекс-Диск [<https://disk.yandex.ru/d/eNgPH51q4nOMlw>].

ГЛАВА 3

КОЛИЧЕСТВЕННЫЙ АНАЛИЗ ОШИБОК, ЗАФИКСИРОВАННЫХ В УЧЕБНОМ КОРПУСЕ РУССКОГО ЯЗЫКА, КАК БАЗА МЕТОДИЧЕСКОЙ ПОДДЕРЖКИ ОБУЧЕНИЯ

Ошибка является неизбежной и неотъемлемой частью процесса изучения иностранного языка. Важность анализа ошибок, допускаемых при изучении иностранных языков, определяется следующими причинами: во-первых, на основе анализа ошибок можно отслеживать прогресс учащихся в процессе обучения иностранным языкам, чтобы преподаватели могли постоянно стимулировать педагогическую деятельность; во-вторых, ошибки предоставляют исследователю информацию о том, как язык изучается или приобретается, и какие стратегии или процедуры использует учащийся при овладении языком; в-третьих, ошибки могут рассматриваться как средство, которое учащиеся используют для улучшения своих методов изучения [60, с. 167].

С развитием прикладных лингвистических технологий в области преподавания и изучения иностранных языков расширяются возможности применения анализа ошибок в области овладения иностранными языками, в том числе и метода анализа ошибок на основе учебных корпусов [103, с. 39]. По сравнению с традиционными интуитивными эмпирическими методами анализа ошибок, аутентичные языковые материалы учащихся, разметка ошибок и метаразметка текстов в учебном корпусе позволяют анализировать ошибки так, что они становятся более объективными и более

соответствующими реальной ситуации учащихся, и в то же время предоставляют контекст ошибок, что значительно повышает эффективность соответствующих исследований.

Анализ ошибок связан с оценкой частоты их появления. В данном исследовании для оценки уровня знаний и частоты ошибок, допускаемых китайскими студентами при создании текстов на русском языке, предлагается использовать относительное значение числа допускаемых ошибок, для вычисления которого вводится коэффициент K [114, с. 198]:

$$K = F_f / L$$

где F_f – частота ошибки типа f в исследуемом тексте,

L – объем текста в словах.

Значения этого коэффициента вычисляются в интервале $0 < K < 1$, при этом, чем ниже значение коэффициента, тем выше уровень знаний. Коэффициент вычисляется для разных авторов, тексты которых вошли в учебный корпус, и для разных типов ошибок, определяемых построенной типологией [114, с. 198].

Поскольку значения частот определяются десятками единиц, а объемы текстов в корпусе – тысячами слов, то значения коэффициента K определяются десятичными дробями с большим количеством нулей после запятой. См., например, значение параметра $K = 3/82064$, что при оценке частоты неверного использования форм деепричастия китайского обучаемого оценивается как $K_1 = 0,0000365568$, а тот же коэффициент при неверном использовании возвратной формы глаголов $K = 16/82064$ получает значение $K_2 = 0,0001949697$. Для удобства сравнения этих значений введено сокращение десятичной дроби

до 6 знаков: $K_1 = 0,000037$ и $K_2 = 0,000195$, что создает больше удобства для сравнения данных, не нарушая их соотношение [114, с. 199].

3.1. Типичные ошибки, совершаемые китайскими студентами под влиянием родного языка

После извлечения и проверки всех ошибок в пропедевтическом учебном корпусе русского языка для обучения китайских студентов было обнаружено, что общий объем письменных текстов составляет около 82064 слова, и в этих текстах было допущено 3176 ошибок, исключая 80 ошибок, которые вызваны тем, что учащиеся недостаточно владеют компьютером, например некоторые пунктуационные ошибки из-за разных методов ввода на компьютере между китайским и русским языками: *Сюй Чжимо является моим любимым писателем.* вместо *Сюй Чжимо является моим любимым писателем.* Список количества ошибок в пропедевтическом учебном корпусе русского языка представлен в Приложении 4.

Частота типов ошибок в учебном корпусе показана на Рисунке 8 и в Таблице 6.

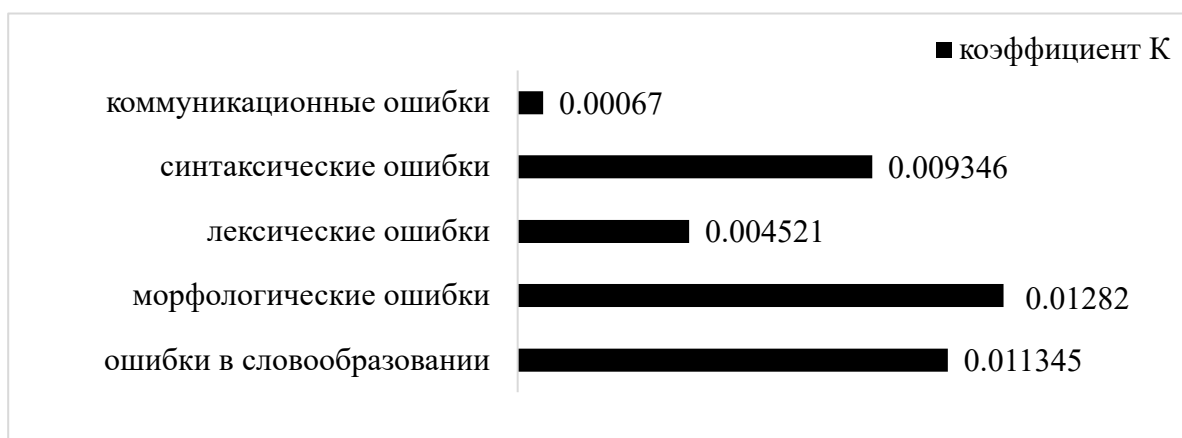


Рисунок 8. Значения коэффициента K при различных типах ошибок в учебном корпусе русского языка

Таблица 6

Значения коэффициента K для разных подкатегорий ошибок в учебном корпусе

Тип ошибки	Подкатегория ошибки	Коэффициент K	Пример
орфографические ошибки	нарушено правописание	0,011345	интересо, праздник весны, знакомства
морфологические ошибки	ошибки в использовании частей речи	0,012819	В наше жизни интернет играет важное роль . Мы с друзьями можно выбирать обед, который тебе нравится.
лексические ошибки	ошибки в использовании слова	0,002644	Нам только надо научиться правильно его пользовать . В Пекине есть великий аэропорт.
	неверное использование устойчивых словосочетаний	0,001877	Верно говорят: В здоровом теле-здоровый .
синтаксические ошибки	нарушение структуры предложений	0,001938	Я родился в небольшом городе, но много людей живут .
	нарушение норм формального синтаксиса	0,002571	Книги помогает нам в жизни. В праздник Весны вся наша семья устала, после 12 часов ночи счастлива спать .
	пунктуационная ошибка	0,004838	Для меня, жизнь в общежитии помогла мне познакомиться с друзьями. Поэтому мой любимый писатель, это Сюй Чжимо.
коммуникационные ошибки	логическое нарушение	0,000670	На площади Тяньаньмэнь находятся многие известные памятники, музей Гугун, парк Ихэюань, Великая китайская стена и другие памятники.

Согласно результатам, представленным в Таблице 6, морфологическая ошибка лидирует по частоте допущенных ошибок [114, с. 200]. При этом

студенты допускали ошибки, выбирая неверное спряжение, склонение и управление слов. Второй по частоте оказалась орфографическая ошибка, к этому типу относятся неверное использование прописной буквы и орфографическая ошибка за счет незнания орфограмм [114, с. 200]. Третье место заняла синтаксическая ошибка, включающая пропуск членов предложения, нарушение согласования членов предложения, отсутствие знаков препинания и т. д. [114, с. 200]. Лексической ошибки было немного меньше, чем синтаксической. Коммуникационная ошибка встречается реже всего.

При анализе частоты типов ошибок важно обратить внимание на то, что наибольшее значение коэффициента K наблюдается при ошибках в использовании частей речи, он составил $0,012819$. Следовательно, данный тип ошибок является самым частотным и типичным при овладении русским языком китайскими студентами [114, с. 200]. Ранжированные по частоте ошибки в использовании частей речи представлены на Рисунке 9.



Рисунок 9. Значения коэффициента K при различных типах ошибок в использовании частей речи

Из диаграммы, представленной на Рисунке 9, можно установить, что самыми сложными для китайских студентов являются глаголы, за ними следуют имена существительные, предлоги и местоимения [114, с. 200]. По сравнению с другими категориями, китайские студенты допустили меньше ошибок в использовании прилагательных, наречий и числительных [114, с. 200]. С одной стороны, это доказывает, что китайские студенты лучше понимают грамматические правила при использовании прилагательных и наречий, чем предлогов, которые не встречаются в китайском языке; с другой стороны, это также говорит о том, что у китайских студентов нет достаточного разнообразия в русских выражениях, они не могут активно использовать прилагательные и наречия для их обогащения [155, с. 72].

Для китайских студентов, которые недавно начали изучать русский язык, важнее целостность структуры предложений, а не богатство их лексического состава [114, с. 201]. Соответственно, китайские студенты уделяют больше внимания использованию существительных, глаголов, местоимений и предлогов в своих текстах для того, чтобы обеспечить целостность структуры предложений [114, с. 201].

Проанализируем ранжированный по частоте список ошибок, допускаемых китайскими студентами в использовании вышеперечисленных частей речи (см. Табл. 7).

Таблица 7

Значения коэффициента K при ошибках в использовании частей речи

Подкатегория ошибки	Название ошибки	Коэффициент K	Пример
ошибки в использовании глаголов	неверное использование возвратной формы	0,000195	После поездки в Москву я чувствовалась разнообразие мира.
	неверное использование несовершенного или совершенного вида	0,001401	Мы знали историю, культуру и погоду Москвы во время посещения. Мне нравится играть в баскетбол, слышать музыка , и кататься на велосипедах.
	неверное управление глагола	0,002267	Я побывала во многих красивых городов Китая. Книги могут способствовать образования наших детей.
	неверное использование деепричастия	0,000037	Гуля по городу, вы можете видеть архитектуры в мусульманском стиле.
	неверный выбор вида причастия	0,000110	Я больше всего люблю читать рассказы О. Генри, эти очень помогшие для меня.
ошибки в использовании имен существительных	неверное использование форм числа	0,000634	Жизни в общежитии никогда не кажется скучной: всё время знакомишься и общаешься.
	отсутствие согласования по роду	0,000158	В наше жизни интернет играет важное роль .
	неверное образование форм склонения	0,001645	Я обязательно вернулась домой с новыми знаниями после путешествия.
	нарушение падежных форм	0,000756	Из них я могу познакомиться с древней китайской архитектурой и чудесной историей Китай .
ошибки в использовании предлогов	неверное управление предлога	0,000890	На следующий день мы с родителей поехали в гости к бабушке и дедушке.
	использование неверного предлога	0,001133	У Пекина много известных достопримечательностей.

ошибки в использовании местоимений	неверный выбор местоимения	0.001694	Этот самый счастливый день для людей – день рождения.
------------------------------------	----------------------------	----------	---

В соответствии с данными, представленными в Таблице 7, наиболее частотные ошибки в использовании глаголов допускаются в их управлении, а ошибки в использовании имен существительных допускаются в образовании форм склонения [114, с. 201]. Затруднения в использовании предлогов возникли у китайских студентов из-за смешения значений предлогов [114, с. 201]. Использование местоимений чаще всего неверно.

Кроме того, частота неверного использования возвратной формы глагола у китайских студентов оценивается как $K_1 = 0,000195$, а частота неверного управления глагола и неверного использования глаголов несовершенного или совершенного вида оцениваются как $K_2 = 0,002267$, $K_3 = 0,001401$, соответственно. При сравнении K_1 с K_2 и K_3 , установлено, что значения K_2 и K_3 намного превышают значения K_1 . То есть во время изучения русского языка китайские студенты испытывают больше трудностей в изучении таких тем как управление глаголов и использование глаголов несовершенного или совершенного вида, но при этом они лучше понимают и усваивают использование возвратной формы глагола [114, с. 201].

Причина этого кроется во влиянии родного языка на процесс обучения иностранным языкам, т. е. порой учащиеся используют модели родного языка в процессе обучения иностранным языкам, что приводит к ошибкам. Это можно объяснить межъязыковой интерференцией, происходящей при конкуренции систем языков в сознании обучаемого, на что также обращают

внимание специалисты, занимающиеся преподаванием иностранного языка как второго [124, с. 86]. Явление межъязыковой интерференции определяется как взаимодействие языковых систем, в ходе которого система родного языка воздействует на изучаемый иностранный язык в процессе овладения им [2, с. 87].

Процесс изучения иностранного языка облегчается, когда структуры родного языка совпадают формально и семантически. Зная похожие явления в грамматике, лексике и фонетике родного языка, можно использовать их при обучении иностранным языкам [32, с. 44]. Например, возвратная форма глагола в русском языке имеет более близкое грамматическое соответствие в китайском языке, т. е. предложение со словом *被* (*служебное слово для образования пассивной формы глагола*) [173, с. 153], например, *Циндао также называются «городом пива» ... в китайском языке 青岛被称为啤酒之城...* т. е. предложения с возвратными глаголами в русском языке могут быть выражены в китайском языке предложениями со словом *被* [114, с. 202]. Поэтому китайским студентам при изучении русского языка легче понять правила использования возвратных форм, следовательно, они допускают меньше ошибок в данной теме.

Процесс усвоения иностранного языка тормозится там, где структуры родного и изучаемого языков различны. Грамматические правила из родного языка механически переносятся на обучаемый язык, что приводит к ошибкам [76, с. 1625]. Межъязыковая интерференция проявляет себя на всех уровнях языка – чаще всего на морфологическом [71, с. 3]. Поэтому китайские студенты

испытывают трудности в грамматических правилах спряжения и управления глаголов, склонении имен существительных и использовании предлогов, см., например,

Мне было 15 лет, когда я в первый раз <путешествовать в городе вр3> Чэнду.

Так что, в основном, <праздную вр2> его с товарищами и друзьями.

Когда спортсмены <завоевают вр2> золотые медали в Олимпиаде, я как китаянка, горжусь ими.

Каждый разговаривал с <другами пр3> <on-line fm2> и <на улице тихо sn3>.

В интернете мы можем создавать сайты, знакомиться с людьми из разных стран, зарабатывать <денеги пр3>, а потом, не выходя из дома, покупать <определённые fm1> вещи.

Мы с моим друзьями обычно не сразу <вернулись вр2> домой, а <погуляли вр2> <по парке пр1>.

Сейчас я учусь <во пр2> втором курсе на факультете русского языка.

Все это связано с отсутствием соответствующих грамматических структур в китайском языке.

Таким образом, изучение русского языка представляет особую сложность для китайских студентов в силу большой разницы между русским и китайским языками [100, с. 146]. Известно, что русский язык является флективным, то есть в нем преобладают синтетические формы выражения грамматических значений (вид, род, спряжение, управление и другие) [40, с. 131]. А китайский

язык – корневой с тенденцией к агглютинативности [62, с. 208], в нем грамматические отношения передаются через отдельные служебные слова, фиксированный порядок слов и контекст. В связи с этим китайский язык не имеет соответствующих русскому языку таких грамматических понятий как спряжение, управление, склонение и другие [123, с. 161].

Таким образом, китайским студентам, изучающим управление глаголов и использование несовершенного или совершенного вида глаголов, трудно понимать и усваивать эти грамматические правила из-за отсутствия соответствующих понятий в китайском языке [114, с. 202].

Более того, согласно данным Таблицы 7, китайские студенты сделали меньше ошибок в использовании деепричастия и причастия, при которых значения коэффициента K составили $0,000037$ и $0,000110$. Причина этого заключается не в том, что китайские студенты лучше усваивают эти грамматические правила, а в том, что они избегают использования таких сложных грамматических правил [114, с. 202].

Помимо морфологических ошибок, китайские студенты также допускают ошибки на синтаксическом и лексическом уровнях под влиянием родного языка. Проанализируем частотный список синтаксических и лексических ошибок, допускаемых китайскими студентами (см. Табл. 8).

Таблица 8

Значения коэффициента K при синтаксических и лексических ошибках

Подкатегория ошибки	Название ошибки	Коэффициент K	Пример
ошибка в использовании слова	употребление слова в несвойственном ему значении	0,001121	В детстве, когда у меня был день рождения, родители готовили мне много вкусных вещей .
	тавтология и повтор	0,000292	Следует различать различные подходы к проблеме.
	смещение синонимов и паронимов	0,001231	Мы тихо готовили ей подарок на день рождения, и я приготовила её любимую еду.
неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	0,001803	В конце экзамена все удивились, что я получила второе место .
	фразеологическая ошибка	0,000073	Верно говорят: в здоровом теле-здоровый .
нарушение структуры предложений	пропуск членов предложения	0,001633	Когда я учился в университете, обычно день рождения с друзьями .
	неверный порядок слов	0,000305	Туризм содействует не только развитию экономического сотрудничества, но и укрепляет социальные отношения.
нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	0,001523	Для современных людей это невозможно живут без интернета. Мне очень нравится его волосы .
	нарушение видовременной соотнесенности однородных членов или сказуемых в главном и придаточном предложениях	0,001048	Мы можем заниматься, разговаривать, обсуждать вопросы, на выходной день мы будем играть на гитаре .
пунктуационная ошибка	отсутствие знака препинания	0,000695	В изучении иностранного языка главное не теория а практика, надо больше говорить, читать.
	лишний знак препинания	0,000756	В прошлом году, я ездила с друзьями в горы.

	неверное использование знака препинания	0,003388	Пекин известен всему миру архитектурами, например: Запретный Город, Храм Неба, Великая китайская стена.
--	---	----------	---

Из представленного в Таблице 8 списка можно сделать вывод, что китайские студенты допускают больше ошибок при использовании знаков препинания, пропуске членов предложения, нарушении согласования подлежащего и сказуемого, смешении синонимов и паронимов и неверном использовании устойчивых выражений.

Синтаксис китайского языка характеризуется системой правил, тесно связанных с порядком слов, вариантами соединения слов в предложении [25, с. 57], а не согласованностью членов предложения. Это приводит к тому, что китайские студенты составляют предложения, не соответствующие требованиям русского языка из-за других правил порядка слов в китайском языке. Например, *Туризм <не только содействует sn2> развитию экономического сотрудничества, но и укреплению социальных отношений.* Союз *не только..., но и...* должен стоять перед однородными членами предложения. А в китайском языке этот союз часто стоит перед сказуемым в предложениях. И поэтому в китайском языке *不仅 (не только)* следует ставить перед глаголом *促进 (содействовать)*, *而且 (но и)* перед глаголом *加强 (укреплять)* независимо от наличия однокоренных слов в предложении, как *旅游不仅促进经济的发展, 而且加强社会联系.*

Одной из типичных ошибок при пропуске членов предложения является отсутствие союза в сложноподчиненных предложениях, поскольку в китайском языке отсутствует понятие подчиненного предложения [178],

поэтому китайские студенты легко игнорируют полную синтаксическую структуру подчиненных предложений, что приводит к ошибкам. Например, *этот праздник весны — одни из самых необычных дней, <произвели глубокое впечатление на меня sn1> ... вместо что произвели глубокое впечатление на меня...; <Я даже считаю, Пекин является своим вторым городом sn1.> вместо что Пекин является своим вторым городом.*

Влияние родного языка на изучаемый обусловлено не только отсутствием соответствующих грамматических понятий в родном языке, но и тем, что некоторые грамматические понятия присутствуют в родном языке, но отсутствуют в изучаемом, либо грамматические понятия присутствуют в обоих языках, но не совпадают [21, с. 81]. Так, в китайском языке после слова *例如 (например)* ставится двоеточие, а в русском – запятая или вообще ничего не ставится.

Кроме того, согласно правилам китайского языка запятая ставится после детерминанта [25], и поэтому появляются такие ошибки в текстах китайских студентов как ... *после ужина <, pt2> я хочу слушать музыку.* вместо ... *после ужина я хочу слушать музыку. ... в университете <, pt2> много студентов приехали из России...* вместо ... *в университете много студентов приехали из России.*

При анализе ошибок более очевидным оказалось, что китайские студенты неверно использовали устойчивые выражения из-за буквального перевода с китайского на русский язык, к примеру *<получила второе место ss1>* вместо *заняла второе место*, *<пользоваться знаниями ss1>* вместо *применять знания*,

<слышали такие вещи сс1> вместо *слышали об этом*, *<повышать рост сс1>* вместо *вырастить*, *<повышать кругозор сс1>* вместо *расширять кругозор*.

Итак, русский и китайский языки к разным языковым семьям и ветвям принадлежат, поэтому между ними существуют серьезные различия на всех структурных уровнях: фонетическом, лексическом, морфологическом и синтаксическом [6, с. 56], что приводит к трудностям в усвоении русского языка китайскими студентами, а также к большому количеству ошибок в таких областях как спряжение глаголов, склонение имен существительных, отсутствие членов предложения, неправильное использование пунктуации и т. д.

На эти ошибки необходимо обратить внимание преподавателей, чтобы разработать специальные программы обучения в процессе преподавания русского языка. Кроме того, следует делать акцент на сходствах и различиях между русским и китайским языками и с помощью положительного переноса навыков родного языка повышать восприимчивость студентов к русскому языку, тем самым снимая трудности в его изучении.

3.2. Ошибки, совершаемые китайскими студентами под влиянием английского языка

Процесс овладения иностранным языком в ходе обучения характеризуется сложным взаимодействием различных факторов, которые по-разному влияют на его конечную цель – приобретение иноязычной коммуникативной компетенции различного уровня [79, с. 161]. В дополнение ко влиянию родного языка, иностранный язык, который обучающиеся уже усвоили, а также их опыт и стратегии обучения, приобретенные при овладении иностранным языком, также влияют на процесс овладения новым иностранным языком [130, с. 430]. Так, например, кроме родного языка для китайских студентов английский язык также оказывает влияние на процесс усвоение русского языка, особенно на начальном этапе.

В системе образования Китая всегда уделялось большое внимание изучению первого иностранного языка, особенно английского [160, с. 1]. Китайские школьники в основном начинают изучать английский язык в третьем классе начальной школы (примерно в возрасте 8–9 лет), а в университетах до сих пор проводятся занятия по английскому языку, даже если студенты не специализируются в нем. Кроме того, английский язык является одним из трех основных предметов, обязательных для сдачи на вступительных экзаменах в китайские вузы, хотя в настоящее время учащиеся могут также выбрать русский, немецкий, французский, японский и испанский языки в качестве первого иностранного языка на вступительных экзаменах в вузы, большинство китайских учащихся по-прежнему выбирают английский [162].

Авторы текстов, зафиксированных в учебном корпусе, выбрали английский язык в качестве своего первого иностранного языка и в основном имеют 8–9-летний опыт изучения английского языка, что обеспечивает им определенную степень понимания английского языка и выработку собственных привычек в обучении иностранным языкам. Для анализа влияния английского языка как первого иностранного на усвоение русского языка студенты были разделены на четыре уровня в соответствии с их баллами по английскому языку на ЕГЭ и их тексты получили соответствующие обозначения, как показано в п. 2.3. Ранжированные списки по частоте ошибок, допускаемых китайскими студентами на разных уровнях английского языка представлены на Рисунке 10.



Рисунок 10. Значения коэффициента K при ошибках китайских студентов с разными уровнями владения английским языком

На основе ранжированного списка, представленного на Рисунке 10,

можно сделать вывод, что в целом, чем выше уровень английского языка студента, тем меньше ошибок он допускает при овладении русским языком [114, с. 203]. Английский и русский языки принадлежат к одной индоевропейской языковой семье. Кроме того, между ними существуют сходства в некоторых грамматических структурах и грамматических понятиях [84]. Следовательно, новички, которые знают английский язык, могут быстрее понять эти грамматические правила при овладении русским языком, в отличие от тех, кто не так хорошо владеет английским языком. Ср., например, частоту ошибок в использовании предлогов и нарушении видовременной соотнесенности в предложениях, допускаемых китайскими студентами с разными уровнями владения английским языком (см. Табл. 9) [114, с. 203].

Таблица 9

Значения коэффициента K при ошибке в использовании предлогов и нарушении видовременной соотнесенности в предложениях

уровень английского языка коэффициент K	A	B	C	D
ошибка в использовании предлогов	0,001566	0,001880	0,002164	0,002566
нарушение видовременной соотнесенности в предложениях	0,000696	0,000873	0,001369	0,001562

Из представленного в Таблице 9 списка можно сделать вывод, что по мере роста уровня владения английским языком значение коэффициента K при обеих этих ошибках снижается. Другими словами, китайские студенты с высоким уровнем английского языка лучше понимают использование предлогов и видовременную соотнесенность в предложениях [114, с. 203]. Из-за отсутствия соответствующих понятий в китайском языке использование

предлогов и видовременная соотнесенность в предложениях всегда представляли собой трудность для китайских студентов [114, с. 203]. Однако конструкция для выражения обстоятельства места в английском языке соответствует предложной конструкции в русском языке [14, с. 585] и проще для использования, в то же время в английском языке необходимо следовать соответствию видовременной соотнесенности между главными и придаточными предложениями [170, с. 179]. Именно в силу сходства русского и английского языков в этих конструкциях китайские студенты с высоким уровнем английского языка реже допускают эти два типа ошибок при овладении русским языком [114, с. 203].

Кроме того, у студентов, хорошо знающих английский язык, обычно более эффективный подход к изучению иностранного языка и хорошие учебные привычки [154, с. 31], что облегчает им адаптацию к изучению русского языка и решение проблем в этом процессе. См., например, частоту орфографических ошибок и неверного использования устойчивых выражений, допускаемых китайскими студентами с разным уровнем владения английским языком (см. Табл. 10) [114, с. 204].

Таблица 10

Значения коэффициента K при орфографических ошибках и неверном использовании устойчивых выражений

уровень английского языка коэффициент K	A	B	C	D
орфографические ошибки	0,008877	0,010596	0,013116	0,014503
неверное использование устойчивых выражений	0,001508	0,001626	0,002120	0,002231

Главная причина возникновения орфографических ошибок состоит в том, что у учащихся недостаточный уровень сформированности орфографического навыка и слабый запас слов [83, с. 134]. В учебном корпусе русского языка неверное использование устойчивых выражений в основном фиксируется как буквальный перевод с китайского языка на русский, например, *获得第二名(получить второе место)* вместо *занимать второе место* [114, с. 204]. Оба типа ошибок отражают отсутствие у учащихся эффективных и правильных методов и навыков обучения при овладении русским языком. В соответствии с данными, представленными в Таблице 10, китайские студенты, у которых выше уровень английского языка, реже допускают эти два типа ошибок по сравнению с теми, у кого уровень английского языка ниже. В процессе изучения английского языка они выработали хорошие навыки обучения, поэтому при овладении русским языком они более эффективно запоминают слова, выбирают правильные и подходящие методы и избегают таких ошибок как буквальный перевод [114, с. 204].

Однако английский язык также оказывает негативное влияние на усвоение русского языка, особенно на начальном этапе, например, очень часто возникают орфографические ошибки из-за путаницы латиницы и кириллицы, например, *<зНакомства fm2>*, *<есть fm2>*. Кроме того, возникают ошибки при использовании имен существительных и управлении глаголов у китайских студентов (см. Табл. 11).

Таблица 11

Значения коэффициента K при ошибке в использовании имен
существительных и неверном управлении глаголов

уровень английского языка коэффициент K	А	В	С	Д
ошибка в использовании имен существительных	0,003887	0,003311	0,002738	0,002566
неверное управление глагола	0,002611	0,002378	0,001987	0,001896

Согласно результатам, представленным на Таблице 11, студенты, хорошо владеющие английским языком, чаще допускают ошибки в русском языке при использовании имен существительных [114, с. 204]. Основная причина заключается в том, что грамматические понятия рода, числа и падежа имен существительных, которые также встречаются в английском языке, значительно отличаются от русского языка, что приводит к отрицательной интерференции [114, с. 204]. Род английских существительных в основном определяется в соответствии с полом одушевленных существительных, а род русских существительных определяется в основном по окончаниям [180], и является морфологической характеристикой слова.

Кроме того, по сравнению с английским языком формы числа и склонения имен существительных в русском языке намного сложнее. Управление глаголов также находится у китайских студентов под влиянием английского языка, например, в английском языке нет предлога в таких словосочетаниях как *play piano, play football* и т. д. И в этом случае китайские студенты допускают такие ошибки, как *играть пианино* вместо *играть на пианино*, *играть футбол* вместо *играть в футбол*. В результате в процессе изучения

русского языка английские грамматические правила мешают учащимся использовать соответствующие русские грамматические правила, особенно учащимся с высоким уровнем английского языка [114, с. 204].

Частота ошибок, допускаемых китайскими студентами с разным уровнем усвоения английского языка представлена в Приложении 5. Следует отметить, что студенты с более высоким уровнем владения английским языком допускают больше ошибок в использовании прилагательных, это объясняется тем, что студенты с высоким уровнем владения английским языком используют больше прилагательных для обогащения своих выражений и повышения читабельности своих текстов, а студенты с низким уровнем владения английским языком не имеют достаточного запаса прилагательных, и они постоянно используют одни и те же прилагательные, об этом свидетельствует, например, избыточное употребление слова *хороший*.

Основываясь на метаразметке китайского учебного корпуса русского языка, было проанализировано влияние английского языка на изучение русского языка [113, с. 93]. Основной вывод из этого анализа заключается в том, что в целом студенты, хорошо успевающие по английскому языку, лучше овладевают русским языком [114, с. 204]. В итоге, с учетом интерференции английского языка в процессе обучения русскому языку, преподавателям рекомендуется делать акцент на сопоставительном изучении русского и английского языков, направлять студентов на восприятие различий и сходств между разными языками, чтобы облегчить им понимание соответствующих грамматических правил и выбор подходящих методов [114, с. 204].

3.3. Ошибки, совершаемые китайскими студентами, в сравнении с данными носителей русского языка

Анализ ошибок, допускаемых учащимися, охватывает и систематизирует не только те, которые обнаруживаются по формальным признакам, но и не имеющие таковых, что делает ошибки второго рода «невидимыми» [50, с. 303]. Под невидимыми ошибками в работе понимаются неточности китайских студентов, возникающие при овладении русским языком, по сравнению с речью носителей русского языка, например расширение использования слов. Из-за того, что данный тип ошибок при анализе вручную достаточно сложно распознать, их необходимо обнаруживать при накоплении достаточного объема языковых материалов. Учебный корпус помогает извлечь именно такие ошибки. Охватывая большое количество аутентичных текстов, учебный корпус позволяет анализировать высокочастотные слова и коллокации учащихся с помощью корпусного менеджера. Эта полезная информация помогает проводить сопоставительный анализ языка тех, кто и изучает язык, и его носителей, а также обнаруживать невидимые ошибки учащихся, возникающие при овладении иностранным языком.

Для выявления невидимых ошибок, допускаемых китайскими студентами в процессе усвоения русского языка, материалом для сопоставительного анализа стал НКРЯ, который представляет собой значительное собрание различных источников [56, с. 167]. Проиллюстрируем выявление таких ошибок на примерах использования предлогов китайскими студентами и носителями русского языка. Ранжированные по частоте предлоги в речи

китайских студентов и носителей русского языка представлены в Таблице 12.

Таблица 12

Сравнение высокочастотных предлогов в текстах носителей русского языка и китайских студентов

Ранг	Тексты китайских студентов	Тексты НКРЯ
1	в	в
2	с	на
3	на	с
4	у	по
5	по	к
6	из	из
7	для	у
8	о	за
9	к	о
10	за	для

На основе ранжированного списка, представленного в Таблице 12, можно установить, что китайские студенты используют предлоги *у* и *для* гораздо чаще, чем носители русского языка. Ошибки при использовании предлога *у* вызваны расширением его использования на несоответствующий контекст, например: *<У Пекина pp2> много достопримечательностей*. Расширение использования предлога *для* связано с тем, что китайские студенты не имеют достаточного разнообразия слов при использовании русских выражений, поэтому не могут активно использовать синонимы для обогащения своего словарного запаса. В учебном корпусе русского языка предлог *для* появился 393 раза, а синонимы предлога *для*, такие как *в целях, с целью, для того, чтобы...* в учебном корпусе появились только с частотой 2, 5 и 9 соответственно.

Таким образом, при обучении русскому языку необходимо сосредоточиться на объяснении границ использования конструкции *у кого-чего* для напоминания китайским студентам о таких типах ошибок. Для

решения проблемы использования предлога *для* рекомендуется добавлять в учебные материалы синонимы и антонимы с целью расширения словарного запаса китайских студентов.

Кроме того, китайские студенты боятся употреблять некоторые слова из-за непонимания их точного значения. Частотные слова в текстах русского языка, написанных китайскими студентами, представлены в Приложении 6. Обращение к данным НКРЯ дает возможность сравнить десять наиболее частотных слов в текстах, написанных китайскими студентами и носителями русского языка (см. Табл. 13).

Таблица 13

Сравнение рангов высокочастотных слов в текстах носителей русского языка и в учебном корпусе текстов китайских студентов

Ранг	Тексты китайских студентов	Тексты НКРЯ
1	и	и
2	в	в
3	я	не
4	с	на
5	что	я
6	мы	быть
7	на	он
8	не	с
9	это	что
10	жизнь	а
11	человек	по
12	но	это
13	у	она
14	очень	к
15	город	но
16	много	они
17	по	мы
18	как	как
19	год	из
20	время	у

Согласно данным, представленным в Таблице 13, по сравнению с

носителями русского языка китайские студенты меньше употребляют глагол *быть*, который занимает 21 место по частоте употребления. Причина этого кроется в том, что китайские студенты не полностью владеют значением и употреблением этого слова, а в учебниках значение слов объясняется не так подробно (Рис. 11).

日常

现在时只用第三人称 *есть*(复也用*суть*〈旧〉; 否定用*нет*); 过*был, была, было*(*не был, не была, не было*); 将*буду, будешь*; 命*будь*; 副动*будучи*[未]

1. 有; 存在; 发生

(1). *Есть или нет?* 有没有?

(2). *У нас есть и будут друзья во всём мире. Мы во всём мире сейчас есть, в будущем обязательно будет много друзей.*

(3). *Тогда ему было сорок лет. В то время ему было сорок лет.*

(4). *Будет дождь. 将要下雨.*

(5). *Завтра не будет дождя. 明天没有雨.*

2. (现在时形式一般从略不用); 处在; 到来

(1). *Он дома. 他在家.*

(2). *Он вчера был дома. 昨天他曾在家.*

(3). *Он был у нас три раза. 他到我们这儿来过三次.*

(4). *Завтра я буду у вас. 明天我到您那里去.*

(5). *Прошу обязательно быть. 请务必到.*

(6). *Он был в Москве. 他到过莫斯科.*

(7). *Будьте всегда с массами. 要随时和群众在一起.*

(8). *Будьте как дома. 请和在家一样(不要客气).*

3. [用作合成谓语的系词](现在时形式一般不用); 系; 充当, 做

(1). *Он студент. 他是大学生.*

(2). *Он был рабочим. 他从前是工人.*

(3). *Он будет министром. 他将要当部长.*

(4). *Было холодно. (当时) 很冷.*

(5). *Неприятельский самолёт был сбит. 敌人的飞机当时被击落了.*

4. (*быть*的变位形式与未完成体动词不定式连用, 构成该动词的将来时)将

(1). *Буду писать. 我将写.*

(2). *Будем слушать радио. 我们将收听广播. ◇*

(3). *Была не была! 〈口语〉(表示决心)豁出来啦!不必顾虑啦!管它呢!*

(4). *Ну, была не была, поеду! 咳, 管它呢, 我去!*

Рисунок 11. Объяснение употребления слова *быть* в учебнике для китайских студентов (цит. по [171, с. 87])

На Рисунке 11 представлены объяснение употребления слова *быть* и примеры, но значение слова объясняется не так подробно и в учебнике отсутствуют примеры, которые тесно связаны с реальной жизнью студентов. Кроме того, как видно из Таблицы 12, китайские студенты употребляют предлог *к* значительно меньше, чем носители русского языка. Это также связано с тем, что китайские студенты не в полной мере владеют навыками употребления этого предлога, а учебные материалы в большей степени

направлены на объяснение основного значения предлога, не давая реальных примеров предложений, написанных носителями русского языка.

Ввиду этого при преподавании русского языка, а также при подготовке учебных материалов для иностранцев необходимо предоставлять как можно больше примеров использования русского языка его носителями, а также знакомить учащихся с реальными материалами носителей языка, чтобы учащиеся могли полностью и правильно овладеть навыками использования слов и грамматических правил.

3.4. Ошибки, совершаемые китайскими студентами, обусловленные разными темами текстов

С помощью программы AntConc на основе учебного корпуса русского языка было обнаружено 3176 ошибок в письменных текстах на русском языке, написанных китайскими учащимися. Количество и процентное соотношение орфографических ошибок, морфологических, лексических, синтаксических и коммуникационных ошибок представлены на Рисунке 12.

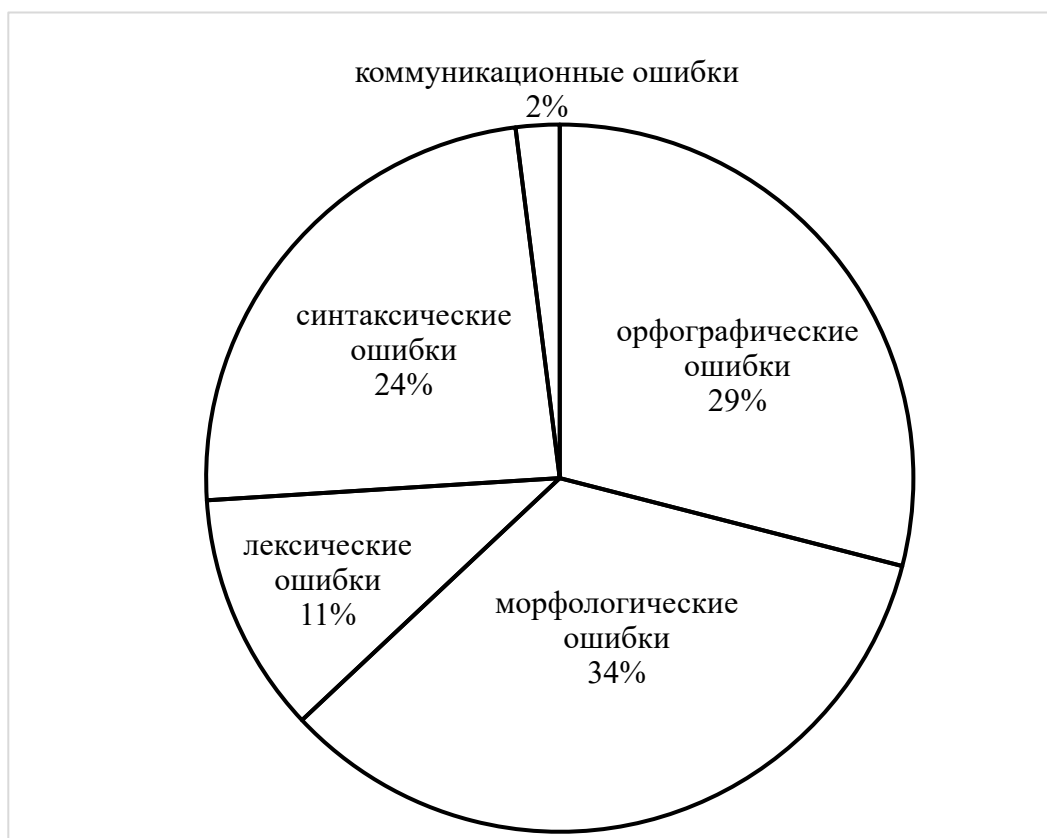


Рисунок 12. Процентное соотношение разных типов ошибок в учебном корпусе русского языка

Согласно результатам, представленным на Рисунке 12, наибольшая доля морфологических ошибок обусловлена негативным переносом с родного языка, далее следуют орфографические ошибки и синтаксические ошибки,

доля лексических ошибок составляет менее 20%, а доля коммуникационных ошибок самая низкая.

С целью установления влияния темы текста на число типичных ошибок в текстах с разными темами были проанализированы ошибки китайских студентов в текстах, написанных на разные темы. На основе результатов количественного анализа было проведено сравнение процента ошибок в текстах на разные темы и средний процент соответствующих ошибок. Частотный список ошибок в текстах, написанных на разные темы, представлен в Приложении 7.

Сравнение процентного соотношения морфологических ошибок в текстах, написанных на разные темы, со средним показателем представлено на Рисунке 13, список тем см. в разделе 2.2.

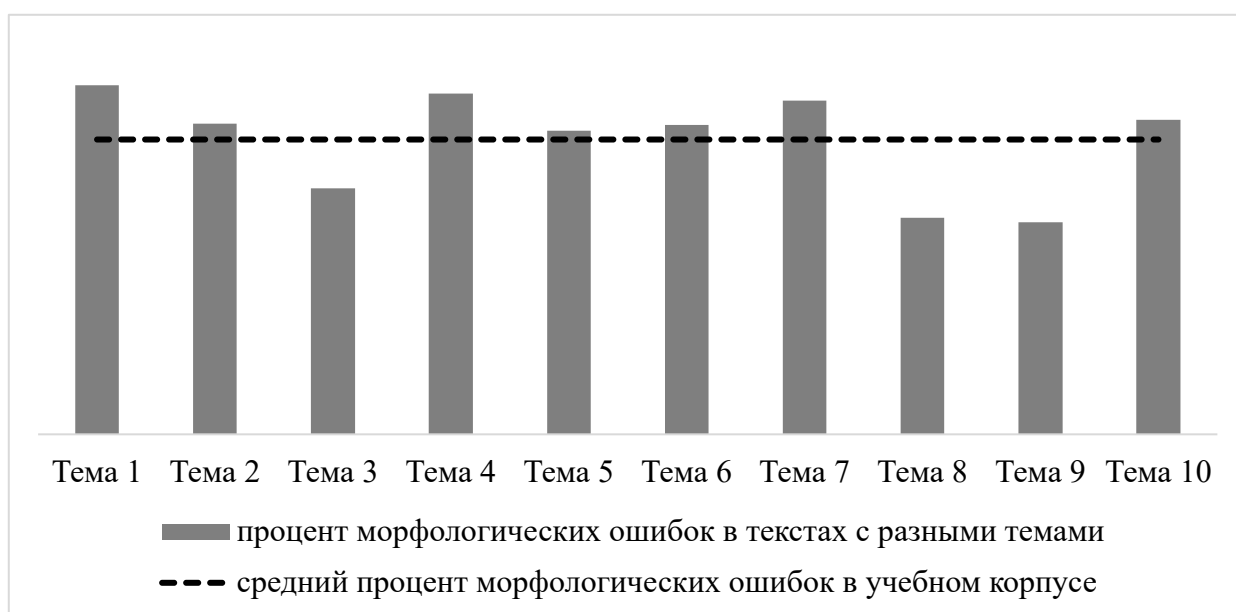


Рисунок 13. Сравнение процентного соотношения морфологических ошибок в учебном корпусе русского языка

На Рисунке 13 представлены процент морфологических ошибок на разные темы и средний процент морфологических ошибок в учебном корпусе русского языка. В соответствии с представленными на Рисунке 11 данными, морфологические ошибки в темах 1, 4, 7 и 10 значительно превышают средний процент. Ранжированный по частоте список ошибок в текстах с вышеперечисленными темами представлен в Таблице 14.

Таблица 14

Список высокочастотных ошибок в темах 1, 4, 7 и 10

Ранг	Тема 1: Мое интересное путешествие	Тема 4: Мои увлечения	Тема 7: Как я провел(а) свой день рождения	Тема 10: Самый любимый человек
1	орфографическая ошибка за счет незнания орфограмм (fm1)	орфографическая ошибка за счет незнания орфограмм (fm1)	неверное использование прописной буквы (fm3)	орфографическая ошибка за счет незнания орфограмм (fm1)
2	неверное использование прописной буквы (fm3)	неверное использование знака препинания (pt3)	орфографическая ошибка за счет незнания орфограмм (fm1)	неверное использование знака препинания (pt3)
3	<i>неверное управление глагола (vr3)</i>	<i>неверное управление глагола (vr3)</i>	пропуск членов предложения (sn1)	<i>неверное образование форм склонения (pr3)</i>
4	неверное использование знака препинания (pt3)	неверное использование прописной буквы (fm3)	<i>неверное использование форм несовершенного или совершенного вида (vr2)</i>	неверное использование прописной буквы (fm3)
5	<i>неверное образование форм склонения (pr3)</i>	<i>неверный выбор местоимения (pr1)</i>	неверное использование знака препинания (pt3)	<i>неверный выбор местоимения (pr1)</i>
6	<i>неверное использование форм несовершенного или совершенного вида (vr2)</i>	<i>неверное использование форм несовершенного или совершенного вида (vr2)</i>	<i>неверный выбор местоимения (pr1)</i>	<i>неверное управление глагола (vr3)</i>

7	<i>неверный выбор местоимения (pr1)</i>	нарушение согласования подлежащего и сказуемого (sn3)	неверное использование устойчивых выражений (cc1)	<i>неверное использование полной или краткой формы (aj1)</i>
8	<i>неверное использование полной или краткой формы (aj1)</i>	<i>неверное образование форм склонения (pr3)</i>	употребление слова в несвойственном ему значении (wd1)	употребление слова в несвойственном ему значении (wd1)
9	неверное использование устойчивых выражений (cc1)	лишний знак препинания (pt2)	<i>использование неверного предлога (pp2)</i>	<i>неверное использование форм числа имен существительных (nr1)</i>
10	смешение синонимов и паронимов (wd3)	<i>использование неверного предлога (pp2)</i>	нарушение согласования подлежащего и сказуемого (sn3)	неверное использование устойчивых выражений (cc1)

В Таблице 14 выделены курсивом и жирным шрифтом самые частотные типы морфологических ошибок в темах 1, 4, 7 и 10. Из представленных данных в Таблице 14 установлено, что в темах 1 и 4 китайские студенты больше допускают морфологические ошибки при управлении глаголов (*vr3*), в теме 7 частотным является неверное использование форм несовершенного или совершенного вида (*vr2*), в теме 10 — неверное образование форм склонения имен существительных (*nr3*).

В темах 1 и 4 речь шла о путешествиях и хобби, что заставило студентов использовать больше разных глаголов, поэтому типичной ошибкой в этих двух темах стало неверное управление глаголов, например *Бадминтон, этот самый <быстрый aj1> спорт в мире, <мне очень интересуют vr3 sn3>*. вместо... *меня очень интересует...* В теме 7 «Как я провел(а) свой день рождения» в основном используется прошедшее время, поэтому заметно, что китайские студенты склонны к ошибкам в выборе вида глаголов, например, *<За ночь pt1>*

до моего дня рождения я <покуплю *вр2*> торт, и ещё <свежий *аj1*> фрукты.
 вместо ... я купила торт... В теме 10 рассказывается о любимом человеке,
 больше всего ошибок появляется в описании личности и внешности
 персонажа, см., например, Я очень люблю <моей *pr1* мамы *пр3*> вместо ...
 люблю свою маму. Я больше люблю его <глазы *пр3*>: <эти *pr1*> острые
 чёрные <глазы *пр3*>. вместо глаза.

В связи с этим необходимо сознательно усилить практику по изучению
 управления глаголов, образованию форм склонения имен существительных
 при обучении подобным темам, а также приводить аутентичные примеры для
 студентов, чтобы помочь им правильно использовать грамматические правила,
 особенно те, которые не имеются в китайском языке.

Проанализируем сравнение процентного соотношения лексических
 ошибок в текстах, написанных на разные темы, со средним показателем (Рис.
 14).

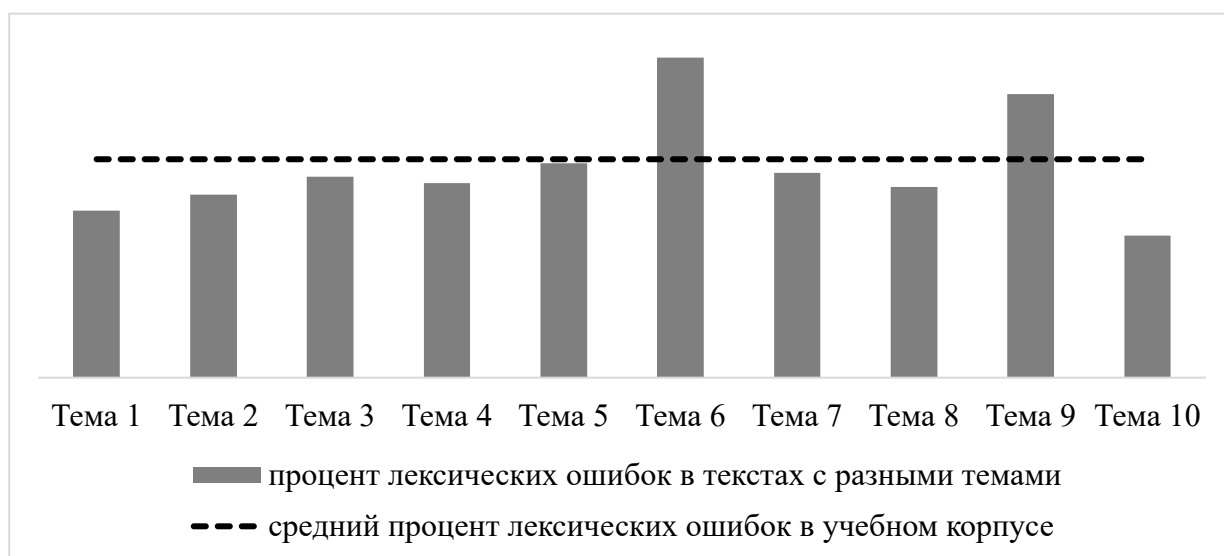


Рисунок 14. Сравнение процентного соотношения лексических ошибок в учебном корпусе русского языка

На Рисунке 14 представлены процент лексических ошибок в текстах на разные темы и средний процент лексических ошибок в текстах учебного корпуса русского языка. По данным Рисунка 12 в темах 6 и 9 процент лексических ошибок значительно выше среднего. Ранжированный по частоте список ошибок в темах 6 и 9 представлен в Таблице 15.

Таблица 15

Список высокочастотных ошибок в темах 6 и 9

Ранг	Тема 6: Кинга в нашей жизни	Тема 9: Как мы провели праздник Весны
1	орфографическая ошибка за счет незнания орфограмм (fm1)	неверное использование прописной буквы (fm3)
2	неверное использование знака препинания (pt3)	орфографическая ошибка за счет незнания орфограмм (fm1)
3	<i>неверное использование устойчивых выражений (cc1)</i>	<i>неверное использование устойчивых выражений (cc1)</i>
4	нарушение согласования подлежащего и сказуемого (sn3)	неверное использование знака препинания (pt3)
5	неверное управление глагола (vp3)	неверное управление глагола (vp3)
6	<i>употребление слова в несвойственном ему значении (wd1)</i>	<i>смешение синонимов и паронимов (wd3)</i>
7	неверное управление предлога (pp1)	нарушение согласования подлежащего и сказуемого (sn3)
8	неверный выбор местоимения (pr1)	употребление слова в несвойственном ему значении (wd1)
9	неверное использование прописной буквы (fm3)	нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях (sn4)
10	использование неверного предлога (pp2)	пропуск членов предложения (sn1)

В Таблице 15 выделены курсивом и жирным шрифтом частотные типы лексических ошибок в темах 6 и 9. Согласно данным, представленным в Таблице 15, типичной ошибкой в этих темах является неверное использование устойчивых выражений (*cc1*), т. е. буквальный перевод с китайского на русский язык. Например, в теме 9 «Как мы провели праздник Весны», где речь

идет о таких словосочетаниях, которые характеризуют особенности китайских праздников, как *放鞭炮* (запускать петарды), *贴春联* (наклеить парные надписи с новогодними пожеланиями), студенты часто используют буквальный перевод, что приводит к тому, что такие словосочетания не соответствуют требованиям русского языка, см., например, <срезать красивые цветы сс1> вместо *срезать бумажные украшения на окна*, <стрелять из бамбука сс1> вместо *запускать петарды*.

Таким образом, при изучении данных тем китайским студентам необходимо активизировать накопление выражений с китайской спецификой, чтобы позволить им точно понимать значение слов и как можно больше соприкоснуться с языковыми материалами носителей русского языка.

Перейдем к Рисунку 15 для анализа процентного соотношения синтаксических ошибок в текстах, написанных на разные темы.

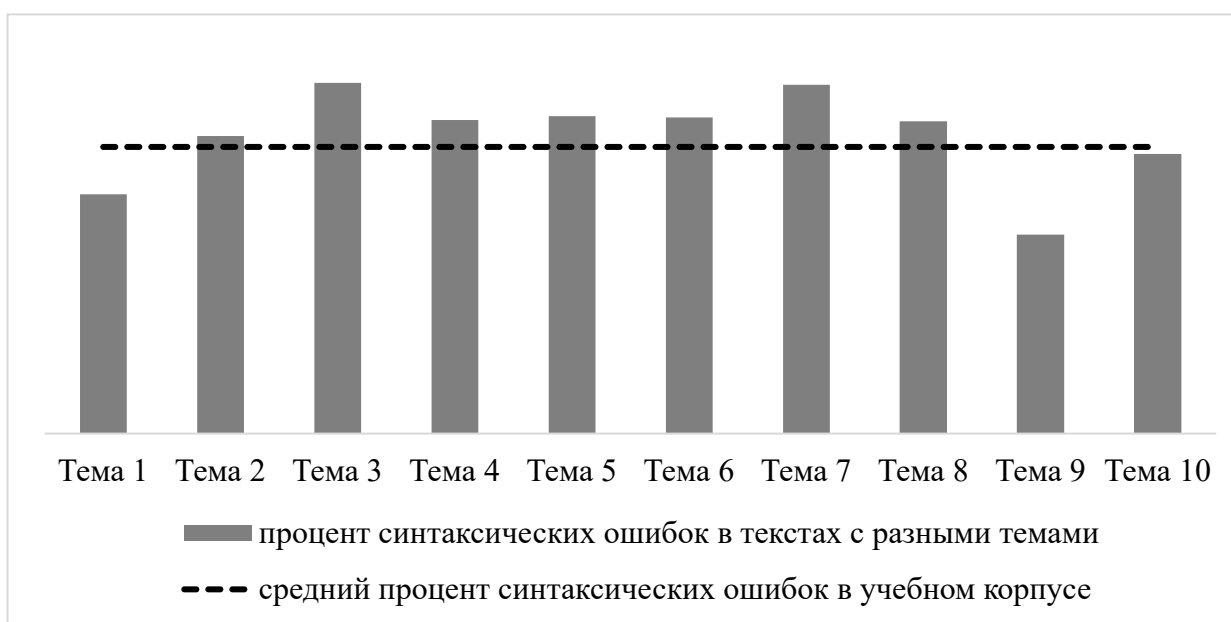


Рисунок 15. Сравнение процентного соотношения синтаксических ошибок в учебном корпусе русского языка

На Рисунке 15 представлены процент синтаксических ошибок в текстах на разные темы и средний процент синтаксических ошибок в учебном корпусе русского языка. Анализ процентного соотношения синтаксических ошибок в текстах на разные темы показал, что китайские студенты допускают больше синтаксических ошибок в текстах темы 3 и 7. Ранжированный по частоте список с наибольшим количеством ошибок в темах 3 и 7 представлен в Таблице 16.

Таблица 16

Список высокочастотных ошибок в темах 3 и 7

Ранг	Тема 3: Мой любимый город	Тема 7: Как я провел(а) свой день рождения
1	орфографическая ошибка за счет незнания орфограмм (fm1)	неверное использование прописной буквы (fm3)
2	неверное использование знака препинания (pt3)	орфографическая ошибка за счет незнания орфограмм (fm1)
3	<i>нарушение согласования подлежащего и сказуемого (sn3)</i>	<i>пропуск членов предложения (sn1)</i>
4	неверное использование прописной буквы (fm3)	неверное использование форм вида глаголов (vp2)
5	неверный выбор местоимения (pr1)	неверное использование знака препинания (pt3)
6	<i>пропуск членов предложения (sn1)</i>	неверный выбор местоимения (pr1)
7	неверное образование форм склонения (np3)	неверное использование устойчивых выражений (cc1)
8	неверное управление глагола (vp3)	употребление слова в несвойственном ему значении (wd1)
9	смещение синонимов и паронимов (wd3)	использование неверного предлога (pp2)
10	неверное использование форм несовершенного или совершенного вида (vp2)	<i>нарушение согласования подлежащего и сказуемого (sn3)</i>

В Таблице 16 выделены курсивом и жирным шрифтом частотные типы синтаксических ошибок в темах 3 и 7. На основе ранжированного списка,

представленного в Таблице 16, можно сделать вывод, что наиболее типичной синтаксической ошибкой в теме 3 оказалось нарушение согласования подлежащего и сказуемого в предложениях (*sn3*). Тема 3 посвящена влиянию Интернета на жизнь, по сравнению с другими темами она требует выражения собственного мнения и является более формальной, и поэтому в ней используется больше длинных предложений, что приводит к ошибкам студентов в несоответствии подлежащего и сказуемого.

Например: *Некоторые люди обманывают других <в Интернет pr1>, и <все больше людей становятся sn3> зависимыми от Интернета.* вместо ... *все больше людей становится...*

Интернет <способствует нам общению vr4> с разными людьми, даже если он за границей, <беседа по Интернету стоил sn3 sn4> бы гораздо дешевле. вместо ... *беседа по Интернету стоит ...*

Синтаксические ошибки в теме 7 в основном связаны с пропусками членов предложения (*sn1*). Причина этого кроется во влиянии родного языка, дело в том, что в китайском языке существуют такие предложения как *我九月生日*, в которых нет явного сказуемого [176, с. 132]. В этом случае китайские студенты допускают пропуск членов предложения из-за его буквального перевода с китайского языка на русский, такие как *<Мой день рождения в сентябре sn1.>* вместо *У меня день рождения в сентябре.* *<Этот день каждого года самым счастливым для меня sn1.>* вместо *Этот день каждого года является счастливым для меня.*

При этом нужно учесть, что типичные ошибки варьируются от темы к

теме, т. е., разные темы оказывают разное влияние на изучение русского языка китайскими студентами. Таким образом, при преподавании русского языка необходимо акцентировать особое внимание на предположении о возможных ошибках учащихся в текстах на разные темы, и вовремя корректировать направленность обучения в соответствии с разными темами, чтобы помочь учащимся эффективно избежать этих ошибок.

3.5. Методический потенциал учебного корпуса текстов.

Предварительные оценки

Создание пилотного проекта учебного корпуса текстов, написанных китайскими студентами, дает возможность его использования при решении различных методических задачи отборе примеров для разработки упражнений.

Результаты анализа ошибок, допускаемых китайскими студентами в процессе усвоения русского языка, показывают, что изучение русского языка для китайских студентов представляет особую сложность в силу большой разницы между русским и китайским языками [100, с. 146]. Именно поэтому необходимо обратить внимание преподавателей на сходства и различия между русским и китайским языками и с помощью положительного переноса навыков родного языка повысить восприимчивость студентов к русскому языку, тем самым снимая трудности в изучении русского языка.

Кроме того, с учетом интерференции английского языка при изучении русского языка, преподавателям рекомендуется обращать внимание на грамматические правила русского языка, которые значительно отличаются от английского языка, чтобы выбрать правильные и подходящие методы обучения, а также помочь студентам избегать таких ошибок как буквальный перевод [114, с. 205].

С помощью учебного корпуса русского языка было установлено, что разные темы оказывают разное влияние на изучение русского языка китайскими студентами, то есть, студенты допускают разные типы ошибок в текстах на разные темы. Таким образом, при преподавании русского языка

необходимо скорректировать содержание при работе с различными темами, разработать ориентированные задания по темам и также помочь студентам накопить примеры предложений с целью снижения трудностей в написания текстов на русском языке.

При сопоставительном исследовании текстов китайских студентов и речи носителей русского языка на основе учебного корпуса были обнаружены некоторые невидимые ошибки китайских учащихся, такие как расширение употребления предлогов и недостаточное использование слов из-за непонимания точного значения. В связи с выявлением этих ошибок и их фиксации в корпусе появляется возможность составлять учебные материалы, более соответствующие реальным потребностям студентов, например, в учебнике добавлять синонимы и антонимы слова, а в учебном словаре провести специальную разметку для того, чтобы напомнить учащимся о словах и грамматических правилах, в которых они часто допускают ошибки, к примеру:

Пример 1

*календарь: **阳性**, лунный календарь 农历, первый день **по** лунному календарю 农历第一天*

Пример 2

*играть: **во что** 玩 (球类, 棋牌 в мяч 打篮球, в карты 打牌, в компьютерные игры 玩电脑游戏), **на чем** 弹, 演奏 (乐器 на рояле 弹钢琴, на гитаре 弹吉他)。*

В примерах выделены жирным грамматические правила употребления

слов, в которых китайские студенты легко допускают ошибки.

Кроме того, учебный корпус текстов дает возможность предоставить в качестве учебных ресурсов понятные языковые материалы, написанные самими учащимися. Богатые языковые ресурсы на основе учебного корпуса текстов также позволяют преподавателям легко разработать такие упражнения, как задание для определения значения слова или словосочетания с помощью учебного корпуса русского языка (Рис. 16), задание для определения значения и употребления слова на основе учебного корпуса (Рис. 17) и т. д.

Задание: изучите данные конкорданса и определите значение словосочинения «так как»

евней столицей шести династий». Мне нравится путешествие, **так как** оно предоставляет мне возможность познать многогранные и
1> поехать за границу с родителями, особенно в Россию, **так как** там я могу получать языковую практику и
поделиться. Я не представляю свою жизнь без движения, **так как** мне нравится активный отдых, осмотр достопримечательности
печатления. Я не представляю свою жизнь без путешествий, **так как** посещение новых мест расширяет мне кругозор, обогащает
путешествовать, и не представляю свою жизнь без движения, **так как** мне нравится активный отдых. Мне очень хотелось
ем могут поделиться. Обожаю путешествовать на велосипеде, **так как** мне нравятся находиться в гармонии с природой,
путешествовать. Я не представляю свою жизнь без движения, **так как** мне нравится активный отдых, осмотр достопримечательности
себя. Именно тогда я подумала о сельском туризме, **так как** считала, что смена обстановки помогает восстановиться. Мо

Рисунок 16. Задание для определения значения словосочинения *так как* с помощью учебного корпуса русского языка

Задание: изучите данные конкорданса и сформулируйте правило употребления слова «Нравится»

Мне нравится **ездить** в новое место, знакомиться с разными
Мне нравится **её** внешность, но мне больше нравится **её**
Мне нравится **заворачивать** пельмени с мамой. Мы приготовили кле
Мне нравится **и** научная литература, особенно о космосе. Эти
Мне нравится **играть** на гитаре, я часто играю на
Мне нравится **мой** родной город, независимо от того, весна
Мне нравится **мой** родной город. Добро пожаловать в Хэби! <2.09.
Мне нравится **<мой rg1>** родной город. Мой родной город
Мне нравится **наблюдать** за жуками. Хотя некоторые люди боятся
Мне нравится **папа** с короткими прямыми волосами. Оказывается, н
Мне нравится **парк** Ихэюань больше всего. Там есть очень
Мне нравится **пекинская** кухня. Пекинская утка известна <мире пр
Мне нравится **проводить** выходные за чтением. Из всех писателей
Мне нравится **<проводить** выходные тренироваться vr4 с группой и
Мне нравится **проводить** выходные за чтением и за движением.
Мне нравится **путешествие**, так как оно предоставляет мне возмож
Мне нравится **русский** язык когда я только начала его

Рисунок 17. Задание для определения значения и употребления слова *нравится* на основе учебного корпуса русского языка

Такие языковые материалы и ресурсы в учебном корпусе текстов тесно связаны с обычной жизнью учащихся, что позволяет максимально снизить трудности усвоения русского языка китайскими студентами и помочь им запомнить, и усвоить значение и употребление слов.

Таким образом, при преподавании русского языка необходимо скорректировать содержание практик при работе с различными темами, разработать ориентированные задания по темам, а также помочь студентам накопить примеры предложений с целью снижения трудностей в написания текстов на русском языке.

В целом, как показали примеры использования, анализ ошибок китайских учащихся при овладении русским языком, выявленных на основе учебного корпуса русского языка, может повысить качество и улучшить методы преподавания русского языка, а также помочь учащимся легче овладеть грамматическими правилами русского языка.

Таким образом, как анализ ошибок, проведенный на основе учебного корпуса русского языка, так и аутентичные языковые материалы, представленные в учебном корпусе, подтверждают его методический потенциал для преподавания и изучения русского языка.

Выводы по Главе 3

Ошибки в процессе овладения иностранным языком представляют интерес для преподавателей иностранного языка, а также самих учащихся. В данной главе дан количественный анализ ошибок, допускаемых китайскими студентами в процессе создания текстов на русском языке под углом зрения интерференции китайского и английского языков, в сравнении с данными носителей русского языка и с учетом разных тем письменных текстов. Для оценки уровня знаний и частоты допускаемых ошибок китайскими студентами предлагается использовать относительное значение числа допускаемых ошибок, для вычисления которого вводится коэффициент K [114, с. 198]:

$$K = F_f / L$$

где F_f – частота ошибки типа f в исследуемом тексте,

L – объем текста в словах

После извлечения и проверки ошибок в учебном корпусе русского языка для обучения китайского языка было обнаружено, что общий объем письменных текстов составляет около 82064 словоупотребления, и в этих текстах всего было установлено 3176 ошибок [114, с. 199]. Из-за серьезных различий между китайским и русским языками китайские студенты сталкиваются с большими трудностями при овладении русским языком, особенно теми грамматическими правилами, которые не соответствуют китайскому языку, такие как склонение имен существительных, управление глаголов, использование глаголов несовершенного или совершенного вида, использование предлогов и другие [114, с. 200]. Кроме того, неверное

использование устойчивых словосочетаний, связанное с буквальным переводом с китайского на русский язык, также является типичной ошибкой китайских студентов. Все эти ошибки позволяют выявить языковые особенности китайских студентов в процессе усвоения русского языка и заслуживают особого внимания.

Авторы текстов в учебном корпусе русского языка имеют богатый опыт изучения английского языка, именно поэтому как первый иностранный язык английский язык также оказывает влияние на усвоение русского языка, особенно на начальных этапах. Основываясь на метаразметке учебного корпуса русского языка, можно сделать вывод, что в целом студенты, хорошо успевающие по английскому языку, лучше овладевают русским языком. Уже приобретенные студентами знания английского языка [114, с. 205], а также привычные методы обучения могут облегчить им изучение русского языка. Однако некоторые грамматические правила в английском языке, не соответствующие русскому языку, все же могут оказывать негативное влияние на изучение русского языка, например, из-за разного управления глаголов и рода имен существительных в русском и английском языках.

В связи с этим преподаватели должны в процессе усвоения русского языка направлять студентов на проведение сопоставительного изучения русского, китайского и английского языков с целью восприятия различий и сходств между языками, чтобы помочь им избежать ошибок и легко овладеть грамматическими правилами.

Кроме того, в процессе изучения русского языка встречаются невидимые

ошибки, которые обнаруживаются при достаточном накоплении языковых материалов. И в этом случае незаменимую роль играет учебный корпус. С помощью сопоставительного исследования и речи носителей русского языка были обнаружены некоторые невидимые ошибки китайских учащихся, такие как расширение употребления конструкции *у кого-чего* и недостаточное использование глагола *быть* из-за непонимания его точного значения. С учетом этих ошибок преподавателям рекомендуется поощрять студентов к чтению текстов, написанных носителями русского языка для пополнения запаса слов и составлять учебные материалы, более соответствующие реальным потребностям студентов, например, в учебнике и учебном словаре добавлять синонимы и антонимы слова.

Учебный корпус русского языка включает в себя письменные тексты по 10 различным темам. Распределение типичных ошибок варьируется от темы к теме. В таких темах как хобби и путешествия китайские учащиеся часто допускают ошибки в управлении глаголов, а при описании характера и внешности людей у них часто возникают проблемы со склонением существительных. В темах, связанных с использованием прошедшего времени, таких как проведение дня рождения и праздника Весны, учащиеся сталкиваются с трудностями при выборе вида глаголов. Также встречаются коллокации и предложения с китайской спецификой, в которых студенты склонны буквально переводить с китайского языка на русский, что приводит к ошибкам в употреблении слов, не соответствующих требованиям русского языка, а также в пропуске членов предложения.

В итоге учебный корпус текстов при обучении русскому языку обладает большим методическим потенциалом, то есть, анализ ошибок учащихся при овладении русским языком на основе учебного корпуса русского языка позволяет расширить методы преподавания русского языка, и также помочь учащимся точно и легко овладеть грамматическими правилами русского языка.

ЗАКЛЮЧЕНИЕ

Исследование проблемы анализа типичных ошибок, допускаемых китайскими студентами при обучении русскому языку на основе созданного специализированного исследовательского учебного корпуса русского языка, позволило получить следующие результаты:

1. Корпусная лингвистика является одним из наиболее перспективных направлений в современной теоретической и прикладной лингвистике [20, с. 118], активно внедряющимся в различные области науки, особенно в сфере образования. Характеризуясь репрезентативностью, сбалансированностью и ориентированностью, корпуса демонстрируют огромный потенциал в области лексикографии, лингводидактики, переводоведения и обучения языкам. Как особый тип корпуса, учебные корпуса все шире внедряются в обучение иностранным языкам с целью решения практических задач и повышения эффективности в этом процессе. Аутентичные языковые материалы учащихся, систематизированная разметка позволяют учебным корпусам играть незаменимую роль в преподавании и изучении иностранных языков.

2. По мере совершенствования информационных технологий, применяемых в педагогической деятельности, развивается создание и применение учебного корпуса в сфере преподавания русского языка как иностранного. Практика разработки и использования корпусов в этой сфере в Китае только начинает развиваться. При этом большинство исследований по овладению русским языком китайскими учеными используют материалы из НКРЯ, созданном в России. В этом случае отсутствие учебных корпусов,

составленных на основе текстов русского языка, написанных китайскими студентами по специальности «Русский язык», подтверждает востребованность разработки пилотного варианта такого корпуса для исследования реальных достижений китайских учащихся при обучении русскому языку.

3. Ориентируясь на цели данной работы, процедура создания учебного корпуса русского языка состояла из трех основных этапов:

1) Отбор письменных текстов на русском языке, написанных китайскими студентами для создания учебного корпуса.

В процессе создания пропедевтического учебного корпуса русского языка были установлены специальные критерии отбора письменных текстов: во-первых, в учебный корпус включались только обычные домашние задания с темами, близкими к жизни учащихся; во-вторых, в учебный корпус отбирались тексты, написанные китайскими студентами с разным, но определенным уровнем владения английским языком; в-третьих, учебный корпус, предназначенный для анализа ошибок китайских студентов при овладении русским языком, должен быть репрезентативен.

Согласно всем перечисленным выше критериям с целью создания учебного корпуса русского языка были использованы 457 письменных текстов по 10 темам, написанные 90 студентами, обучающимися по специальности «Русский язык» в китайских университетах.

2) Проведение разметки ошибок и метаразметки в учебном корпусе.

Разметка ошибок в пропедевтическом учебном корпусе русского языка

проводилась согласно стандарту *TEI*, в котором был установлен конкретный метод разметки метаинформации и информации о тексте. Опираясь на систему разметки *TEI*, виды разметки учебного корпуса русского языка были разделены на две части: первая — метаразметка (экстралингвистическая разметка), которая включает такую информацию как тема, возраст, пол и уровень владения английским языком авторами текстов [112, с. 128]. Вторая — разметка ошибок, которые помечаются строчными латинскими буквами и цифрами, заключенными в угловые скобки <...>. Например: *Я люблю путешествовать и <побывал много мест вр3> Китая* [112, с. 129].

3) Идентификация и классификация ошибок, допускаемых китайскими студентами при овладении русским языком.

В данном исследовании термин *ошибка* был определен как нарушение грамматичности, правильности и приемлемости речи с точки зрения носителя языка. То есть ошибка, которая допускается китайскими студентами в процессе усвоения русского языка, рассматривается как противоречие правилам и нормам русского языка и воспринимается как неправильность [110, с. 248]. На основе этого определения и с учетом языковых характеристик китайских учащихся и принятой в Китае типологии ошибок, ошибки в учебном корпусе русского языка разделены на 5 основных типов и 14 подкатегорий, всего выделено 34 варианта.

В целом, созданный исследовательский учебный корпус русского языка может использоваться разными пользователями с целью решения педагогических задач (для составления учебных материалов и словарей), и в

лингвистических целях (для исследования развития и изменения языка учащихся при изучении иностранного языка).

4. Использование корпусного менеджера AntConc позволило извлечь из учебного корпуса русского языка 3176 ошибок, в том числе 931 орфографическая ошибка, 1052 морфологические ошибки, 371 лексическая ошибка, 767 синтаксических ошибок, 55 коммуникационных ошибок. Относительное значение числа ошибок, рассчитываемое на основе предложенного коэффициента K , является ведущим методом оценки уровня знаний и частоты допускаемых ошибок китайскими студентами при создании текстов на русском языке.

5. Результат анализа ошибок китайских студентов при создании текстов на русском языке показал, что китайские студенты сталкиваются с большими трудностями при овладении русским языком, особенно с теми грамматическими правилами, которые не соответствуют китайскому языку, такие как склонение имен существительных, управление глаголов, использование глаголов несовершенного или совершенного вида, использование предлогов и другие. Это связано с большими различиями между китайскими русским языками.

Анализ ошибок китайских студентов на разных уровнях овладения английским языком показал, что английский язык в общем оказывает активное влияние на обучение русскому языку китайскими студентами. Таким образом, преподавателям рекомендуется делать акцент на сопоставительном изучении русского, китайского и английского языков для того, чтобы позволить

учащимся быстро овладеть соответствующим грамматическим правилам и повысить эффективность изучения русского языка.

6. Сравнение текстов китайских студентов в учебном корпусе русского языка и носителей русского языка показало, что китайские студенты допускают невидимые ошибки при употреблении глагола *быть*, предлогов *для*, *у* и *к*, что связано с недостаточным запасом слов у учащихся и непониманием точного значения слова. В этом случае учащиеся должны больше читать реальные тексты носителей русского языка для пополнения словарного запаса и четкого понимания значения слова. В то же время необходимо составлять учебные материалы, более соответствующие реальным потребностям китайских студентов.

7. Анализ типичных ошибок китайских студентов в текстах с разными темами, показал, что разные темы также оказывают различное влияние на усвоение русского языка, и китайские студенты сталкиваются с разными трудностями. В темах, связанных с хобби и путешествиями, китайские учащиеся часто допускают ошибки в управлении глаголов. В таких темах, как проведение дня рождения и праздника Весны, учащиеся сталкиваются с трудностями при выборе вида глаголов и использовании прошедшего времени. При употреблении слов и коллокаций с китайской спецификой в темах, связанных с праздником Весны, китайские студенты склонны буквально переводить с китайского языка на русский. Именно поэтому при обучении русскому языку необходимо оптимизировать курсы и содержание обучения по разным темам для преподавателей, и для студентов накопить хорошие

примеры текстов, особенно написанные носителями русского языка для обогащения своих выражений на русском языке.

Таким образом, проведенное исследование позволило идентифицировать и проанализировать ошибки китайских студентов в процессе усвоения русского языка на основе учебного корпуса русского языка. Созданный учебный корпус русского языка в качестве базы данных, и результаты, полученные в диссертационном исследовании, доступны для применения не только китайским и русским студентам, аспирантам, преподавателям, но и всем заинтересованным лицам.

В целом учебный корпус представляет собой мощный и перспективный инструмент для теоретического и практического обучения иностранным языкам. Технологии корпусов упрощают процесс анализа больших массивов данных и применяются в целях решения практических педагогических задач, а также подтверждают методический потенциал дальнейшего развития корпусной лингвистики в Китае.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Агафонова Л. И. Некоторые вопросы использования корпусных технологий как фактора повышения качества обучения иностранным языкам / Л. И. Агафонова // Известия Российского государственного педагогического университета им. А. И. Герцена. – СПб: Изд-во Российского государственного педагогического университета им. А. И. Герцена, 2009. – № 87. – С. 80–88.
2. Азимов Э. Г. Новый словарь методических терминов и понятий (теория и практика обучения языкам) / Э. Г. Азимов, А. Н. Щукин. – М.: ИКАР, 2009. – 448с.
3. Ачкасов А. В. Терминологический поиск в процессе перевода (на материале данных компьютерного мониторинга) / А. В. Ачкасов // Гуманитарные, социально-экономические и общественные науки. – Краснодар: Наука и образование, 2015. – № 10. – С. 160–165.
4. Базарова Б. Б. Корпусная лингвистика и преподавание иностранных языков / Б. Б. Базарова // Вестник Бурятского государственного университета. – Улан-Удэ: Изд-во БГУ имени Доржи Банзарова, 2015. – №. 15. – С. 88–92.
5. Баранов А. Н. Проблема репрезентативности корпуса данных (на примере политической метафорики) / А. Н. Баранов // Диалог-2001: Труды Международного семинара по компьютерной лингвистике и ее приложениям. Аксаково, 11–13 июня 2001 г. – М.: Изд-во НАУКА, 2001. – С. 68–74.
6. Баранова И. И. Учет межъязыковой интерференции при обучении

- студентов из Китая русскому языку / И. И. Баранова, К. М. Чуваева // Вестник Российского Университета Дружбы Народов: Русский и иностранные языки и методика их преподавания. – М.: Изд-во Российского Университета Дружбы Народов, 2012. – № 1. – С. 56–62.
7. Беляева Л. Н. Методы корпусной лингвистики при изучении глобального английского языка: этнолингвистический аспект / Л. Н. Беляева // Труды международной конференции «Корпусная лингвистика – 2006». Санкт-Петербург, 10–14 октября 2006 г. – СПб.: Изд-во Санкт-петербургского государственного университета, 2006. – С. 52–60.
 8. Беляева Л. Н. Корпусная лингвистика и перевод: потенциал и ограничения / Л. Н. Беляева // Труды международной конференции «Корпусная лингвистика – 2011». Санкт-Петербург, 27–29 июня 2011 г. – СПб.: Изд-во Санкт-петербургского государственного университета, 2011. – С. 87–91.
 9. Беляева Л. Н. Сетевые лингвистические технологии / Л. Н. Беляева, О. Н. Камшилова, К. Р. Пиотровская. – СПб: Изд-во Российского государственного педагогического университета им. А. И. Герцена, 2019. – 111 с.
 10. Беляева Л. Н. Машинный перевод в современной технологии процесса перевода / Л. Н. Беляева // Известия Российского государственного педагогического университета им. А. И. Герцена. – СПб: Изд-во Российского государственного педагогического университета им. А. И. Герцена (языкознание), 2022. – № 203. – С. 22–30.
 11. Богданова С. Ю. Возможности корпусной методологии в решении

- лингвистических задач / С. Ю. Богданова // Вестник Иркутского государственного лингвистического университета. – Иркутск: Изд-во Иркутского государственного лингвистического университета, 2012. – № 2(18). – С. 47–53.
12. Богданович Г. Ю. Русский язык в современном измерении / Г. Ю. Богданович // Ученые записки УО ВГУ им. П. М. Машерова. – 2018. – Т. 27. – С. 25–30.
13. Богоявленская Ю. В. Репрезентативность лингвистического корпуса: метод верификации достоверности полученных данных / Ю. В. Богоявленская // Политическая лингвистика. – Екатеринбург: ФГБОУ ВО «Уральский государственный педагогический университет», 2016. – № 4(58). – С. 163–166.
14. Болтакулова Г. Ф. (2015). Способы выражения обстоятельства времени в английском и русском языках и его место в предложении / Г. Ф. Болтакулова // Молодой ученый. – Казань: Молодой ученый, 2015. – № 14 (94). – С. 583–585.
15. Васильева Г. М. Лексико-синтаксический словарь сложноподчиненных предложений для иностранных студентов: теоретические и методические ориентиры / Г. М. Васильева, И. Н. Левина // Вопросы лексикографии. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2017. – № 12. – С. 5–23.
16. Веселовская Т. С. Иллюстративно-текстовый корпус учебников русского языка для детей младшего школьного возраста: концепция и методика

- создания / Т. С. Веселовская, А. Н. Лапошина, О. Ф. Купрещенко // Корпусная лингвистика – 2019: труды Международной конференции. Санкт-Петербург, 24–28 июня 2019 г. – СПб: Издательство СПбПУ, 2019. – С. 63–71.
17. Власова Е. А. Создание устного учебного русского как иностранного: первые результаты и перспективы / Е. А. Власова, Ю. В. Бец, Е. В. Каллистратидие // Актуальные проблемы филологии и педагогической лингвистики. – Северная Осетия: Северо-Осетинский государственный университет им. К. Л. Хетагурова, 2022. – № 2. – С. 118–130.
18. Вотинцев П. А. Использование формата TEI для обмена данными с полнотекстовой информационно-поисковой системой «Манускрипт» / П. А. Вотинцев // Материалы международной научной конференции. «Современные информационные технологии и письменное наследие: от древних рукописей к электронным текстам». Ижевск, 13–17 июля 2006 г. – Ижевск: Изд-во Ижевского государственного технического университета. – С. 30–31.
19. Вылегжанина Е. А. Использование информационно-коммуникационных технологий в образовательном процессе / Е. А. Вылегжанина, Н. Н. Мальцева // Актуальные задачи педагогики: материалы VI Международной научной конференции. Чита, январь 2015 г. – Чита: Молодой ученый, 2015. – С. 4–6.
20. Галинская Т. Н. Корпусный подход в реконструкции медиаобраза языковой личности политика / Т. Н. Галинская // Известия высших учебных

- заведений. Поволжский регион. Гуманитарные науки. – Пенза: Издательство Пензенского государственного университета, 2014. – № 2(30). – С. 114–124.
21. Гасконь Е. А. Некоторые ошибки китайских студентов, связанные с влиянием ранее освоенных языков на изучение нового (русского как иностранного) / Е. А. Гасконь // Филологические науки. Вопросы теории и практики. – Тамбов: Изд-во Грамота, 2016. – № 10(64). – С. 80–84.
22. Герасимова М. Н. Использование лингвистически аннотированного корпуса текстов на занятиях по дисциплинам «Русский язык» и «Русский язык и культура речи»: из опыта работы / М. Н. Герасимова, Е. В. Грудева // Вестник Череповецкого государственного университета. – Череповец: Изд-во Череповецкого государственного университета, 2014. – № 2(55). – С. 65–68.
23. Гончаровская О. Ю. Корпусная лингвистика. развитие корпусной лингвистике в истории лингвистики. исследовательские методы / О. Ю. Гончаровская // Современные научные исследования и разработки. – М: Научный центр Олимп, 2018. – Т. 2. – № 5 (22). – С. 147–150.
24. Горина О. Г. Инструменты корпусного анализа в обучении иностранному языку / О. Г. Горина // Вестник Томского государственного университета. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2018. – № 435. – С. 187–194.
25. Готлиб О. М. Практическая грамматика современного китайского языка / О. М. Готлиб. – М.: Восточная книга, 2012. – 288 с.

26. Гришина Е. А. Корпус звучащей русский речи в составе Национального корпуса русского языка / Е. А. Гришина, С. О. Савчук // Компьютерная лингвистика и интеллектуальные технологии по материалам ежегодной международной конференции Диалог – 2008. Бекасово, 4–8 июня 2008 г. – М: РГГУ, 2008. – С. 125–132.
27. Грудева Е. В. Корпусы ошибок: целевая аудитория, возможная архитектура корпуса / Е. В. Грудева, И. А. Бучилова, Н. А. Волкова // Вестник Череповецкого государственного университета. – Череповец: Изд-во Череповецкого государственного университета, 2018. – № 5 (86). – С. 63–72.
28. Добровольский Д. О. Использование корпусов текстов в двуязычной лексикографии / Д. О. Добровольский // Среди нехоженных путей: Сборник научных статей к юбилею А. А. Кретьова. – Воронеж: Изд-во Автономной некоммерческой организации по оказанию издательских и полиграфических услуг "НАУКА-ЮНИПРЕСС", 2012. – С.14-25.
29. Елохов Е. С. Создание хранилища научных публикаций с применением методов корпусной лингвистики / Е. С. Елохов [и др.]. // Открытые семантические технологии проектирования интеллектуальных систем. – Минск: БГУИР, 2014. – № 4. – С. 513–516.
30. Еремина В. В. Использование корпусных технологий при обучении иностранцев русскому языку / В. В. Еремина, О. С. Еремина // Актуальные вопросы профессионального образования. – Волгоград: Изд-во Волгоградского государственного технического университета, 2018. – № 1

- (10). – С. 9–12.
31. Жендаренко Е. В. Типичные ошибки, выявленные при тестировании по русскому языку как иностранному // В мире науки и искусства: вопросы филологии, искусствоведения и культурологии. – Новосибирск: Ассоциация научных сотрудников "Сибирская академическая книга", 2013. – № 29. – С. 60–84.
32. Журина А. С. Использование конструктивной и преодоление деструктивной интерференции при изучении второго иностранного языка / А. С. Журина // Ярославский педагогический вестник. – Ярославль: Изд-во Ярославского государственного педагогического университета им. К. Д. Ушинского, 2015. – № 4. – С. 42–45.
33. Замичева С. С. Тематическая разметка диалектного корпуса: опыт Томских диалектологов / С. С. Замичева, Е. В. Иванцова // Вестник Томского государственного университета. Филология. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2020. – № 67. – С. 45–61.
34. Захаров В. П. Корпусная лингвистика / В. П. Захаров, С. Ю. Богданова. – Иркутск: ИГЛУ, 2011. – 160 с.
35. Захаров В. П. Международные стандарты в области корпусной лингвистики / В. П. Захаров // Структурная и прикладная лингвистика. – Ленинград: Изд-во Ленинградского университета, 2012. – № 9. – С. 201–221.
36. Захаров В. П. Корпусная лингвистика: 2-е изд., перераб. и доп. / В. П.

- Захаров, С. Ю. Богданова. – СПб: Санкт-петербургский государственный университет. РИО. Филологический факультет, 2013. – 148 с.
37. Захаров В. П. Корпуса русского языка / В. П. Захаров // Труды Института русского языка им. В. В. Виноградова. – М.: Изд-во Института русского языка им. В. В. Виноградова, 2015. – №3(6). – С. 20–64.
38. Захаров В. П. Моделирование в корпусной лингвистике: специализированные корпуса русского языка / В. П. Захаров, И. В. Азарова [и др.]. – СПб: Изд-во Санкт-Петербургского государственного университета, 2019. – 208 с.
39. Захаров В. П. Корпусная лингвистика: учебник. 3-е изд., перераб. / В. П. Захаров, С. Ю. Богданова. – СПб: Изд-во Санкт-петербургского государственного университета, 2020. – 234 с.
40. Злобина Ю. И. Некоторые типичные ошибки в письменных работах китайских студентов: классификация и анализ / Ю. И. Злобина, М. О. Сорокина // Мир науки, культуры, образования. – Горно-Алтайск: Изд-во международного научного журнала «Мир науки, культуры, образования», 2018. – № 4(71). – С. 130–132.
41. Иवानян Е. П. Инновационные методы и проекты в освоении лингвистической русистики / Е. П. Иवानян // Историко-педагогические чтения. – Екатеринбург: Изд-во ФГБОУ ВО «Уральский государственный педагогический университет», 2019. – № 23. – С. 361–365.
42. Иргизова К. В. Корпусная лингвистика в отечественном и зарубежном языкознании на современном этапе / К. В. Иргизова // Огарев-online. –

- Саранск: Изд-во Национального исследовательского Мордовского государственного университета им. Н. П. Огарева, 2019. – № 6(127). – С. 1–8.
43. Иргизова К. В. Корпус английского языка как средство анализа ошибок в письменной речи учащихся / К. В. Иргизова, Сафонкина О. С. // Огарев-Online. – Саранск: Национальный исследовательский Мордовский государственный университет им. Н. П. Огарева, 2020. – № 12(149). – С. 1–9.
44. Калиниченко Э. Б. Роль языковых ошибок при обучении иностранному языку / Э. Б. Калиниченко [и др.] // Международный научно-исследовательский журнал. – Екатеринбург: Индивидуальный предприниматель С. М. Владимировна, 2016. – № 4(46). – С. 50–52.
45. Камшилова О. Н. Разработка корпуса текстов петербургских школьников: задачи и перспективы / О. Н. Камшилова, М. В. Колина, Е. А. Николаева // Материалы IV Международной научно-практической конференции «Прикладная лингвистика в науке и образовании», Санкт-Петербург, 27-28 марта 2008 г. – СПб.: Российский государственный педагогический университет им. А. И. Герцена, 2008. – С. 92–98.
46. Камшилова О. Н. Исследовательский потенциал корпуса английских текстов петербургских школьников: анализ интерязыка / О. Н. Камшилова // Известия Российского государственного педагогического университета им. А. И. Герцена. – СПб: Изд-во Российского государственного педагогического университета им. А. И. Герцена, 2009. – №104. – С. 114–

123.

47. Камшилова О. Н. Анализ ошибок в корпусе ученических текстов: исследовательский и технологический аспект / О. Н. Камшилова // Материалы Международной конференции «Корпусная лингвистика – 2011», 27–29 июня 2011, Санкт-Петербург. – СПб.: Изд-во Санкт-петербургского государственного университета, 2011. – С. 175–180.
48. Камшилова О. Н. Учебный корпус текстов: потенциал, состав, структура / О. Н. Камшилова. – СПб: Книжный Дом, 2012. – 56 с.
49. Камшилова О. Н. Корпусные технологии в обучении языкам / О. Н. Камшилова // Новые технологии в образовательном пространстве родного и иностранного языка. – Пермь: Изд-во Пермского национального исследовательского политехнического университета, 2013. – № 1. – С. 56–62.
50. Камшилова О. Н. Учебный корпус текстов: работа над ошибками / О. Н. Камшилова // Труды Международной конференции «Корпусная лингвистика – 2013», Санкт-Петербург, 25–27 июня 2013 г. – СПб: Изд-во Санкт-петербургского государственного университета, 2013. – С. 301–308.
51. Камшилова О. Н. LC-Технологии в исследовании освоения языка / О. Н. Камшилова // Известия Российского государственного педагогического университета им. А. И. Герцена. – СПб: Изд-во Российского государственного педагогического университета им. А. И. Герцена, 2014. – № 168. – С.103–109.
52. Камшилова О. Н. Планирование лонгитюдного исследования: большие

- проблемы маленького корпуса / О. Н. Камшилова, Е. Д. Смульская // Труды VIII Международной научной конференции Прикладная лингвистика в науке и образовании, Санкт-Петербург, 24–26 ноября 2016 г. – СПб: Изд-во Российского государственного педагогического университета им. А. И. Герцена, 2016. – С. 170–175.
53. Камшилова О. Н. Корпусное исследование прилагательных в иноязычной речи на разных этапах усвоения чужого языка / О. Н. Камшилова, В. П. Захаров // The II international conference “R. Piotrowski’s Readings 2017, Saint-Petersburg, 27 ноября 2017 г. – Saint-Petersburg: Creative Commons ССО, 2017. – С. 177–190.
54. Каштанова П. В. Методика работы с несвободными словосочетаниями как способ расширения профессиональной компетентности преподавателя русского языка / П. В. Каштанова, Г. П. Мосолова, Е. П. Прокаева // Гуманитарные науки и образование. – Саранск: Изд-во Мордовского государственного педагогического университета имени М. Е. Евсевьева, 2018. – Т. 9. – № 4 (36). – С. 65–72.
55. Ключихин В. В. Формирование коллокационной компетенции обучающихся на основе электронного лингвистического корпуса / В. В. Ключихин // Вестник Тамбовского государственного университета. – Тамбов: Изд-во Тамбовского государственного университета имени Г. Р. Державина, 2019. – Т. 24. – № 179. – С. 69–80.
56. Князева Ю. О. Корпусный анализ в исследовании синонимии / Ю. О. Князева // Вестник КГУ им. Н. А. Некрасова. – Кострома: Изд-во КГУ им.

- Н. А. Некрасова, 2008. – № 4. – С. 165–169.
57. Кондрашова Н. В. Прогнозирование и исправление студенческих ошибок при обучении иностранным языкам / Н. В. Кондрашова // Научный диалог. – Екатеринбург: Центр научных и образовательных проектов, 2015. – № 7(43). – С. 27–47.
58. Копотев М. В. Принципы создания Хельсинкского аннотированного корпуса русских текстов (ХАНКО) в сети Интернет / М. В. Копотев, А. Мустайоки // Научно-техническая информация. Серия 2. – М.: Всероссийский институт научной и технической информации, 2003. – № 6. – С. 33–36.
59. Корпус русских учебных текстов [Электронный ресурс]. – http://web-corpora.net/learner_corpus.
60. Кордер П. С. Значение ошибок учащегося / П. С. Кордер // Международный обзор прикладной лингвистики. 1976, – С. 161–170.
61. Котюрова И. А. Корпусные исследования с помощью сервиса Antconc в условиях работы в ВУЗе / И. А. Котюрова // Язык и культура. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2020. – № 52. – С. 36–50.
62. Крапивник Е. В. Языковая интерференция в аспекте национально-ориентированного обучения китайских студентов русскому языку / Е. В. Крапивник // Вестник Чувашского государственного педагогического университета им. И. Я. Яковлева. – Чебоксары: Изд-во Чувашского государственного педагогического университета им. И. Я. Яковлева, 2020.

– № 4(109). – С. 205–211.

63. Крылов С. А. Об усовершенствовании процедуры автоматического анализа и синтеза речевых отрезков (в связи с разработкой лингвистического процессора в интегрированной информационной среде STARLING) / С. А. Крылов // Аспекты компаративистики. – М.: Изд-во Российского государственного гуманитарного университета, 2005. – № 4. – С. 453–475.
64. Кузнецова Э. К. Возможности использования лингвистического корпуса в обучении иностранным языкам в неязыковом вузе / Э. К. Кузнецова // Казанский вестник молодых ученых. Педагогические науки. – Казань: Изд-во Казанского федерального университета, 2020. – № 4(4). – С. 35–46.
65. Куликова Е. В. Разработка тренировочных упражнений в курсе иностранного языка с использованием подходов корпусной лингвистики / Е. В. Куликова, М. С. Коган // Неделя науки СПбПУ: материалы научной конференции с международным участием, Санкт-Петербург, 13–19 ноября 2017 г. – СПб: Изд-во ФГАОУ ВО СПбПУ, 2017. – С. 84–87.
66. Куниловская М. А. Классификация переводческих ошибок для создания разметки в учебном параллельном корпусе / М. А. Куниловская // *Lingua Mobilis*. – Челябинск: Изд-во Челябинского государственного университета, 2013. – № 1 (40). – С. 141–158.
67. Куниловская М. А. Классификация переводческих ошибок и их электронная разметка в brat / М. А. Куниловская // Проблемы теории, практики и дидактики перевода. – Нижний Новгород: Изд-во Нижегородского государственного лингвистического университета, 2013. –

№ 1. – С. 59–71.

68. Куренко К. Н. Корпусная лингвистика в переводоведении / К. Н. Куренко // Иностранные языки: лингвистические и методические аспекты. – Тверь: Изд-во Тверского государственного университета, 2017. – № 37. – С. 318–323.
69. Лебедев В. Д. Грамматика английского языка как область знания с точки зрения ее исторического развития / В. Д. Лебедев // Молодой ученый. – Казань: Молодой ученый, 2019. – № 9(247). – С. 226–229.
70. Лопинова М. Н. Понятие «ошибка», причины возникновения и классификация ошибок в методике преподавания иностранного языка / М. Н. Лопинова [и др.] // Известия Великолукской государственной сельскохозяйственной академии. – Великие Луки: Великолукская ГСХА, 2016. – № 3. – С. 63–71.
71. Лукманова Р. Р. Межъязыковая Интерференция – Причина Нарушения Нормы Иностранного Языка (на Примере Китайско-Русского Билингвизма) / Р. Р. Лукманова, А. А. Утробина, Ван Ц // Russian Linguistic Bulletin. – Екатеринбург: Общество с ограниченной ответственностью «Цифра», 2023. – № 7(43). – С. 1–6.
72. Майорова А. Д. Корпусная лингвистика: исторический и лингводидактический аспекты [Электронный ресурс] / А. Д. Майорова // Международный научно-исследовательский журнал. – 2017. – № 5(59). – С. 42–46. – URL: <https://research-journal.org/archive/5-59-2017-may/korpusnaya-lingvistika-istoricheskij-i-lingvodidakticheskij-aspekty> (дата

обращения: 30.10.2023).

73. Мальцева М. С. Учебный корпус (Learner corpora) как база для лингвистического и лингводидактического анализа в рамках методики преподавания иностранных языков / М. С. Мальцева // Социально-экономические явления и процессы. – Тамбов: ТГУ им. Г. Р. Державина, 2011. – № 9 (31). – С. 209–212.
74. Мамонтова В. В. Корпусная лингвистика в современной языковедческой парадигме [Электронный ресурс] / В. В. Мамонтова // Актуальные вопросы современной науки. – Новосибирск: Общество с ограниченной ответственностью Центр развития научного сотрудничества, 2010. – № 12. – С. 230–238. – URL: <https://cyberleninka.ru/article/n/korpusnaya-lingvistika-v-sovremennoy-yazykovedcheskoy-paradigme> (дата обращения: 30.10.2023).
75. Мишланова С. Л. Корпусная лингвистика vs лингвистика стереотипного и творческого. Стереотипность творчества в тексте / С. Л. Мишланова, Т. М. Пермякова, Е. Ю. Аликина // Межвузовский сборник научных трудов Пермского государственного национального исследовательского университета, Пермь, 2011. – Пермь: Изд-во Пермского государственного национального исследовательского университета, 2011. – С. 289–301.
76. Муродова С. П. Межъязыковая интерференция / С. П. Муродова, Н. П. Панжиев // Science and education. – Source: Science and education, 2022. – № 6(3). – С. 1623–1628.
77. Нагель О. В. Корпусная лингвистика и ее использование в компьютерном языковом обучении / О. В. Нагель // Язык и культура. – Томск: Изд-во

Национального исследовательского Томского государственного университета, 2008. – № 4. – С. 53–59.

78. Национальный Корпус Русского Языка [Электронный ресурс]. – <https://ruscorpora.ru/>.
79. Новикова Н. Н. Отрицательная интерференция родного и первого иностранного языка при обучении второму иностранному языку / Н. Н. Новикова, Н. Н. Саклакова // Мир науки, культуры, образования. – Горно-Алтайск: Изд-во международного научного журнала «Мир науки, культуры, образования», 2019. – № 6(79). – С. 161–162.
80. Ольховская А. И. Корпус в преподавании русского языка и литературы / А. И. Ольховская, М. К. Парамонова // Перспективы науки и образования. – Тамбов: Фонд развития науки и культуры, 2018. – № 6 (36). – С. 187–195.
81. Ольховская А. И. Корпусное преподавание русского языка / А. И. Ольховская // Вестник Томского государственного педагогического университета. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2019. – № 2 (199). – С. 98–107.
82. Осипова Е. С. Формирование умений переводчика работать с ресурсами корпусной лингвистики при обучении переводу фразеологизмов / Е. С. Осипова // Индустрия перевода. – Пермь: Изд-во Пермского национального исследовательского политехнического университета, 2017. – Т. 1. – С. 200–206.
83. Пазова Л. М. Орфографическая безграмотность и пути ее преодоления / Л. М. Пазова, И. В. Щербашина // Вестник Майкопского государственного

- технологического университета. – Майкоп: Изд-во Майкопского государственного технологического университета, 2014. – № 3. – С. 130–134.
84. Пешеходько В. А. Этимологическое сходство русской и английской лексики / В. А. Пешеходько, О. В. Чаплыгина // Образование и наука в России и за рубежом. – М.: Общество с ограниченной ответственностью Московский двор, 2021. – № 5(81). – С. 145–148.
85. Пименова А. В. Специфика одноязычных учебных словарей (на материале Британских лексикографических издательств) / А. В. Пименова // Вестник Челябинского государственного университета. Филология. Искусствоведение. – Челябинск: Изд-во Челябинского государственного университета, 2009. – № 22(160). – С. 87–91.
86. Плотницкая С. В. Корпусная лингвистика: Национальный корпус языка / С. В. Плотницкая // Мова. – Одесса: Изд-во Одесского национального университета имени И. И. Мечникова, 2012. – № 17. – С. 23–26.
87. Плунгян В. А. Корпус как инструмент и как идеология: о некоторых уроках современной корпусной лингвистики / В. А. Плунгян // Русский язык в научном освещении. – М: Изд-во Института русского языка им. В. В. Виноградова, 2008. – № 2 (16). – С. 7–20.
88. Прилепская М. В. Использование Национального корпуса русского языка в преподавании иностранных языков / М. В. Прилепская // Альманах современной науки и образования. – Тамбов: Издательство Грамота, 2020. – № 11(42). – С. 110–113.

89. Раицкая Л. К. Дидактические возможности корпусных интернет-технологий в преподавании иностранного языка в высшей школе / Л. К. Раицкая // Вестник Московского государственного областного университета. Серия: Педагогика. – М.: Изд-во Московского государственного областного университета, 2009. – № 4. – С. 123–127.
90. Рахилина Е. В. Задачи и принципы семантической разметки лексики в НЛРЯ / Е. В. Рахилина, Г. И. Кустова, О. Н. Ляшевская и др. // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. – СПб: 2009. – С. 215–240.
91. Русская корпусная грамматика [Электронный ресурс]. – <http://rusgram.ru/>.
92. Рыков В. В. Корпус текстов как отражение состояния русского языка / В. В. Рыков // Русский язык: исторические судьбы и современность: труды Международного конгресса, Москва, 13-16 марта 2001 г. – М.: Изд-во Московского государственного университета им. М. В. Ломоносова, 2001. – С. 31–32.
93. Садовникова О. Э. Прямое и косвенное использование корпусов в зарубежной лингводидактике / О. Э. Садовникова // Научно-педагогический журнал Восточной Сибири. Magister Dixit. – Иркутск: Изд-во МГЛУ, 2013. – № 2. – С. 152–161.
94. Сербина А. С. Приграничное сотрудничество России и Китая / А. С. Сербина // Россия и Китай: стратегическое взаимодействие в XXI веке: Материалы международной школы-семинара, Новосибирск, 5–6 ноября 2015 г. – Новосибирск: Изд-во Новосибирского государственного

- технологического университета, 2015. – С. 82–85.
95. Серпутько А. С. Возможности и ограничения современной корпусной лингвистики в переводоведении / А. С. Серпутько, М. А. Ивлева // Язык и культура: сборник статей Международной научной конференции, Томск, 26–28 октября 2016 г. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2017. – С. 200–202.
96. Сибирцева В. Г., Хоменко А. Ю., Баранова Ю. Н. Национальный корпус русского языка как основа новаторских электронных учебников / В. Г. Сибирцева, А. Ю. Хоменко, Ю. Н. Баранова // Образовательные технологии и общество. – Казан: Изд-во Казанского национального исследовательского технологического университета, 2013. – № 16(3). – С. 508–520.
97. Скоринкин Д. А. Электронное представление текста с помощью стандарта разметки TEI / Д. А. Скоринкин // Вестник Московского университета. Серия 9: Филология. – М.: Изд-во Московского государственного университета им. М. В. Ломоносова, 2016. – № 5. – С. 90–108.
98. Словари, созданные на основе Национального Корпуса Русского Языка [Электронный ресурс]. – <http://dict.ruslang.ru/>.
99. Смульская Е. Д. Лонгитюдные исследования в лингвистике: опыт и перспективы / Е. Д. Смульская // Известия Российского государственного педагогического университета им. А. И. Герцена. – СПб: Изд-во Российского государственного педагогического университета им. А. И. Герцена, 2016. – № 182. – С. 53–58.
100. Соколова Г. Е. Наиболее распространенные виды ошибок китайских

- студентов при обучении русскому языку как иностранному / Г. Е. Соколова // Проблемы современного образования. – М.: Изд-во МПГУ, 2020. – № 2. – С. 146–154.
101. Солнышкина М. И. История развития корпусной лингвистики (на примере англоязычных корпусов) / М. И. Солнышкина, Г. М. Гатиятуллина // Вестник Томского государственного университета. Филология. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2020. – № 63. – С. 132–160.
102. Соснина Е. П. Корпусная лингвистика и корпусный подход в обучении иностранному языку [Электронный ресурс] / Е. П. Соснина // Прикладная лингвистика. Ульянов. гос. техн. ун-т. – Ульяновск, 2001. – URL: http://ling.ulstu.ru/linguistics/resources/literature/articles/corpus_linguistics_language_teaching/ (дата обращения: 30.10.2023).
103. Соснина Е. П. К анализу лингвистических ошибок учащихся на базе учебного корпуса текстов / Е. П. Соснина // Известия Самарского научного центра Российской академии наук. Социальные, гуманитарные, медико-биологические науки. – Самара: Изд-во Самарского научного центра Российской академии наук, 2017. – № 19(5). – С. 39–43.
104. Сысоев П. В. Лингвистический корпус в методике обучения иностранным языкам / П. В. Сысоев // Язык и культура. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2010. – № 1(9). – С. 99–111.
105. Тарнаева Л. П. Использование ресурсов корпусной лингвистики при

- подготовке переводчиков в сфере профессиональной коммуникации / Л. П. Тарнаева, Е. С. Осипова // Филологические науки. Вопросы теории и практики. – Тамбов: Издательство Грамота, 2016. – № 9–1(63). – С. 205–209.
106. Толстова Т. В. Корпусная лингвистика в обучении английскому языку делового общения / Т. В. Толстова // Сборник научных трудов «Вопросы прикладной лингвистики». – Москва: Изд-во Российского университета дружбы народов, 2009. – С. 66–70.
107. Фоменко Ю. В. Типы речевых ошибок / Ю. В. Фоменко. – Новосибирск: Изд-во НГПУ, 1994. – 60 с.
108. Фомичева Н. В. От умения взаимодействовать с информационными ресурсами к самостоятельной работе / Н. В. Фомичева // Вестник ПНИПУ. Проблемы языкознания и педагогики. – Пермь: Изд-во ПНИПУ, 2015. – № 4. – С. 127–134.
109. Фу Чжаочжэн О некоторых различиях между китайским и русским языками / Чжаочжэн Фу, Е. Ю. Казимирчик // Язык – Общество – Культура: сборник материалов международной студенческой научно-практической конференции в рамках 20-й международной научно-технической конференции БНТУ «Наука – образованию, производству и экономике», 28 апреля 2022 г. – Минск: БНТУ, 2022. – С. 106–107.
110. Хоу Линьсюе. Классификация ошибок для создания разметки в учебном корпусе русских текстов / Линьсюе Хоу // Известия Российского государственного педагогического университета им. А. И. Герцена. – СПб:

Изд-во Российского государственного педагогического университета им. А. И. Герцена (языкознание), 2022. – № 203. – С. 247–254.

111. Хоу Линьсюе. Принципы разметки ошибок при создании учебного корпуса русских текстов / Линьсюе Хоу // Актуальные проблемы лингвистики и литературоведения: сборник материалов IX (XXIII) Международной научно-практической конференции молодых ученых, Томск, 14–16 апреля 2022 г. – Томск: Издательство Национального исследовательского Томского государственного университета, 2022. – С. 213–218.
112. Хоу Линьсюе. Методы разметки учебного корпуса русских текстов / Линьсюе Хоу // Университетский научный журнал. Серия «Филологические и исторические науки, искусствоведение». – СПб: Книжный Дом, 2022. – № 69. – С. 126–132.
113. Хоу Линьсюе. Возможности применения учебного корпуса в обучении русскому языку в эпоху больших данных / Линьсюе Хоу // Российско-китайское сотрудничество: на пути к глобальному устойчивому развитию. Материалы VI международной научно-практической конференции, Новосибирск, 10–11 ноября 2022 г. – Новосибирск: Издательство Новосибирского государственного технического университета, 2022. – С. 91–94.
114. Хоу Линьсюе. Анализ ошибок на основе учебного корпуса русских текстов / Линьсюе Хоу // Вестник Московского городского педагогического университета. Серия «Филология. Теория языка. Языковое образование».

- М.: Изд-во Московского городского педагогического университета, 2023.
– № 3(51). – С. 196–207.
115. Цейтлин С. Н. Ошибки в письменной речи учащихся и способы их классификации / С. Н. Цейтлин // Русский язык в школе. – М.: Наш язык, 1984. – № 2. – С. 40–46.
116. Цейтлин С. Н. Речевые ошибки и их предупреждение: учеб. пособие Изд. стереотип. / С. Н. Цейтлин. – М.: Книжный дом ЛИБРОКОМ, 2017. – 192 с.
117. Чернявская В. Е. Интерпретация научного текста: учеб. пособие для студентов ст. курсов, магистрантов и аспирантов вузов, обучающихся по направлению 540300 (050300) Филол. образование / В. Е. Чернявская. – СПб.: Наука, 2004. – 127 с.
118. Чернякова Т. А. Использование лингвистического корпуса в обучении иностранному языку / Т. А. Чернякова // Язык и культура. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2011. – № 4(16). – С. 127–132.
119. Чернякова Т. А. Алгоритм формирования лексических навыков студентов на основе лингвистического корпуса / Т. А. Чернякова // Вестник Тамбовского государственного университета: Гуманитарные науки. Педагогика и психология. – Томск: Изд-во Национального исследовательского Томского государственного университета, 2012. – № 2(106). – С. 212–216.
120. Шевчук В. Н. Информационные технологии в переводе. Электронные

ресурсы переводчика / В. Н. Шевчук. – М.: ЗебраЕ, 2013. – 384 с.

121. Ширинкина К. М. Возможности применения корпусной лингвистики на примере сайта Cube Pronunciation Dictionary для изучения фонетики английского языка / К. М. Ширинкина // Международный студенческий научный вестник. – Пенза: Информационно-технический отдел Академии Естествознания, 2023. – № 1. – С. 100–105.
122. Щелокова, А. А. Типологизация речевых ошибок на основе понятий «конструктивность / деструктивность» / А. А. Щелокова // Мир науки, культуры, образования. – Горно-Алтайск: Изд-во международного научного журнала «Мир науки, культуры, образования», 2020. – № 4(83). – С. 389–391.
123. Щелокова А. А. Интерференция как лингвистическая основа формирования ошибок (на примере речи китайцев) / А. А. Щелокова // Эпомен: филологические науки. – Эпомен: Изд-во Эпомена, 2022. – № 1. – С. 159–167.
124. Щербакова М. В. Особенности обучения второму иностранному языку студентов языковых факультетов / М. В. Щербакова // Вестник ВГУ. Серия лингвистика и межкультурная коммуникация. – Воронеж: Изд-во Воронежского государственного университета, 2003. – № 2. – С. 84–87.
125. Anthony L. Tools and strategies for Data-Driven Learning in the EAP writing classroom / L. Anthony // Specialised English: New Directions in ESP and EAP Research and Practice. – Abingdon: UK. Routledge Press, 2019. – P. 179–194.
126. Baker M. Corpus-based Studies within the Larger Context of Translation

- Studies / M. Baker // *Génesis: Revista científica do ISAI*. – 2002. – № 2. – P. 7–16.
127. Baker P. *Glossary of Corpus Linguistics*. / P. Baker, A. Hardie, T. McEnery. – Edinburgh: Edinburgh University Press, 2006. – 192 p.
128. *Cambridge Advanced Learner's Dictionary*. – Beijing: Foreign Language Teaching and Research Press, 2002. – 767 p.
129. Cobb T. *Analyzing late interlanguage with learner corpora: Quebec replications of three European studies* / T. Cobb. // *The Canadian Modern Language Review*. – Toronto: University of Toronto Press, 2003. – № 59(3). – P. 393–423.
130. Forsyth H. *The Influence of L2 Transfer on L3 English Written Production in a Bilingual German / Italian Population: A study of Syntactic Errors* / Forsyth Helen // *Open Journal of Modern Linguistics*. – Los Angeles: Scientific Research Publishing, 2014. – № 4. – P. 429–456.
131. Foster M.E. *Corpus-based generation of head and eyebrow motion for an embodied conversational agent* / M.E. Foster, J. Oberlander // *Language Resources and Evaluation*. – Berlin: Springer, 2007. – № 41. – P. 305–323.
132. Granger, S. *A Bird's Eye View of Learner Corpus Research* / S. Granger, J. Hung // *Computer Learner Corpus, Second Language Acquisition and Foreign Language Teaching*. – Amsterdam, 2002. – P. 3–33.
133. Granger S. *Computer Learner Corpus Research: Current Status and Future Prospects* / S. Granger // *In Language and Computers Applied Corpus Linguistics: A Multidimensional Perspective*. – Amsterdam: Rodopi, 2004. – №

23. – P. 123–145.
134. Grigaliūnienė J. Corpora in the classroom / J. Grigaliūnienė. – Vilnius University, 2013. – 81 p.
135. James C. Errors in Language Learning and Use: Exploring Error Analysis / C. James. – Beijing: Foreign Language Teaching and Research Press, 2001. – 313p.
136. Johansson S. Some aspects of the development of corpus linguistics in the 1970-s and 1980-s / S. Johansson, A. Ludeling, M. Kyto., eds. // *Corpus Linguistics: An International Handbook*. – Berlin/New York: Walter de Gruyter, 2008. – P. 33–53.
137. Kemp J. Which words should I look up, in V. Vians (Ed.) / J. Kemp, L. Anthony // *New Ways in Teaching with Corpora*. – TESOL: TESOL Press, 2022. – P. 264–267.
138. Kemp J. How can I be more specific in my writing? in V. Vians (Ed.) / J. Kemp, L. Anthony // *New Ways in Teaching with Corpora*. – TESOL: TESOL Press, 2022. – P. 286–290.
139. Kennedy G. *An Introduction to Corpus linguistics* / G. Kennedy. – New York: Addison Wesley Longman limited, 1998. – 328 p.
140. Kunilovskaya M. A. Corpus-Based Quality Assessment of Error-Tagged Learner Translations / M. A. Kunilovskaya, T. A. Ilyushchenya, M. A. Kovyazina // *Tyumen State University Herald. Humanities Research. Humanities*. – Tyumen: Tyumen State University Herald, 2017. – № 3(3). – P. 94–112.

141. Laviosa-Braithwaite S. Corpus-based Translation Studies. Theory, Findings, Applications / S. Laviosa-Braithwaite. – Amsterdam/New York: Rodopi, 2002. – 138 p.
142. Leech G. Corpus annotation schemes / G. Leech // *Literary and Linguistic Computing*. – Oxford: Oxford University Press, 1993. – № 8(4). – P. 275–281.
143. Masciolini A. A query engine for L1-L2 parallel dependency treebanks / A. Masciolini // *Proceedings of the 24th Nordic Conference on Computational Linguistics, NoDaLiDa, May 22–24, 2023*. – NoDaLiDa: 2023. – P. 574–587.
144. McEnery T. *Corpus Linguistics: An Introduction 2nd edition* / T. McEnery, A. Wilson. – Edinburgh: Edinburgh University Press, 2001. – 235 p.
145. McEnery T. *Corpus-based Language Studies: An Advanced Resource Book* / T. McEnery, R. Xiao, Yukio Tono // Taylor & Francis. – London/New York: Routledge, 2006. – 386 p.
146. O'Donnell M. B. "From corpus to query and back again": The Survey of English Usage and ICECUP paradigm / Matthew Brook O'Donnell // *International Journal of Corpus Linguistics*. – Amsterdam: John Benjamins Publishing Company, 2008. – № 13(3). – P. 387–401.
147. Roger G. *Corpus Annotation: Linguistic Information from Computer Text Corpora*. / G. Roger, G. Leech, T. McEnery. – Longman: Addison Wesley Longman, 1997. – 281p.
148. Rundell M. Recent trends in pedagogical lexicography / M. Rundell // *International Journal of Lexicography*. – Oxford: Oxford University Press, 1998. – № 11(4). – P. 315–342.

149. Sanz C. Mind and content in adult second language acquisition: methods, theory, and practice / C. Sanz. – Georgetown: Georgetown University Press, 2005. – 322p.
150. Salakhyan E. The Tubingen Corpus of Eastern European English (TCEEE): From a small-scale corpus study to a newly emerging non-native English variety / Elena Salakhyan // *Token: A Journal of English Linguistics*. – Kielce: Jan Kochanowski University Press, 2012. – № 1. – P. 143–157.
151. Svartvik J. A corpus of English Conversation / J. Svartvik, R. Quirk. – Lund: Gleerup, 1980. – 893 p.
152. Volodina E. B. Towards privacy by design in learner corpora research: A case of on-the-fly pseudonymization of Swedish learner essays / E. B. Volodina, Y. A. Mohammed, S. Derbring, A. Matsson // *Proceedings of the 28th International Conference on Computational Linguistics, Barcelona*. – Spain: International Committee on Computational Linguistics, 2020. – P. 357–369.
153. Wang Renqiang. Perfect Match Between Lexicography and Corpus Linguistics: A Review of The Oxford Guide to Practical Lexicography / Renqiang Wang, Yu Zhou // *Chinese Foreign Languages*. – Beijing: Higher Education Press, 2010. – № 7(4). – P. 109–111.
154. 陈露.英语对零起点俄语专业学生的迁移影响探析[J].现代交际, 2021, 2: 30-33. [Чэнь Лу. Изучение влияния английского языка на китайских студентов по специальности русский язык / Лу Чэнь // *Современная коммуникация*. – Цзилинь: Академия общественных наук провинции Цзилинь, 2021. – № 2. – С. 30–33.]

155. 董爱华. 基于训练语料库的语言错误分析研究[J]. 北京语言大学学报. 2020, 28(04), 71–74. [Дун Айхуа. Исследование анализа языковых ошибок на основе учебного корпуса / Айхуа Дун // Вестник Пекинского полиграфического института. – Пекин: Изд-во Пекинского полиграфического института, 2020. – № 28(04). – С. 71–74.]
156. 丁树德. 浅谈西方翻译语料库研究[J]. 外国语, 2001, 5: 61–66. [Дин Шудэ. Введение в изучение западных переводческих корпусов / Шудэ Дин // Иностранный язык. – Шанхай: Изд-во Шанхайского университета иностранных языков, 2001. – № 5. – С. 61–66.]
157. 黄昌宁, 李涓子. 语料库语言学[M]. 北京: 商务印书馆, 2002: 280. [Хуан Чаннин. Корпусная лингвистика / Чаннин Хуан, Цзюаньцзы Ли. – Пекин: Коммерческий Пресс, 2002. – 280 с.]
158. 胡开宝. 语言智能视域下外语教育的发展: 问题与路径[J]. 中国外语, 2021, 18(6): 4–9. [Ху Кайбао. Развитие иноязычного образования в аспекте языковой интеллигенции: проблемы и пути / Кайбао Ху // Обучение русскому языку. – Пекин: Изд-во Пекинского университета иностранных языков, 2021. – № 18(6). – С. 4–9.]
159. 何婷婷. 语料库研究[D]. 武汉: 华中师范大学, 2003: 326. [Хэ Тинтин. Исследование на основе корпусной лингвистики / Тинтин Хэ // – Ухань: Центральный педагогический университет, 2003. – 326 с.]
160. 蒋洪新, 杨安. 一带一路倡议与中国外语教育改革[A]. 外语教学, 2020, 41(1): 1–2. [Цзян Хунсинь. Инициатива "Один пояс и один путь" и реформа образования в области иностранных языков в Китае / Хунсинь Цзян, Ань

- Ян // Обучение иностранным языкам. – Сиань: Изд-во Сианьского университета иностранных языков, 2020. – № 41(1). – С. 1–2.]
161. 柯平, 鲍川运. 世界各地高校的口笔译专业与翻译研究机构[J]. 中国翻译, 2002, 23(4):59-66. [Кэ Пин. Программы образования устного и письменного перевода и научно-исследовательские учреждения перевода в мире / Пин Кэ, Чуаньюнь Бао // Журнал Китайского перевода. – Пекин: Издательство иностранных языков Китая, 2002. – № 23(4). – С. 59–66.]
162. 刘道义. 外语教育的作用与高考改革[C]. 外国语, 2014, 37(6): 8–10. [Лю Даои. Роль обучения иностранным языкам и реформа системы экзамена высшего образования / Даои Лю // Иностранный язык. – Шанхай: Изд-во Шанхайского университета иностранных языков, 2014. – № 37(6). – С. 8–10.]
163. 刘国兵. 现代大型季度语料库的标注方法 — 以 CEC 政府文件子语料库标注为例[A]. 山东: 山东外语教学, 2008, 5: 53–59. [Лю Гуопинь. Методы разметки современного большого корпуса на основе подкорпуса CEC / Гуопинь Лю. – Шаньдун: Преподавание иностранных языков провинции Шаньдун, 2008. – № 5. – С. 53–59.]
164. 吕红周. 俄语名词语义范畴自动标注研究[J]. 外语研究, 2010, 4: 52–57. [Лу Хунчжоу. Исследование автоматического семантического аннотирования имен существительных в русском языке / Хунчжоу Лу // Исследование по иностранным языкам. – Нанкин: Изд-во Нанкинского института международных отношений, 2010. – № 4. – С. 52–57.]
165. 刘莉莉, 崔钰. 俄语教学理论与实践研究 [M]. 北京: 北京理工大学出版

- 社, 2012: 196. [Лю Лили. Теория и практика преподавания русского языка / Лю Лили, Юй Цуй. – Пекин: Изд-во Пекинского технического института, 2012. – 196 с.]
166. 刘丽芬, 吴娟. 专业俄语写作句法错误分析及其对策[J]. 外语研究. 2010(2): 59–63. [Лю Лифэнь. Анализ синтаксических ошибок в текстах китайскими студентами на русском языке и меры их устранения / Лифэнь Лю, Цзюань Ву // Исследование по иностранным языкам. – Нанкин: Изд-во Нанкинского института международных отношений, 2010. – № 2. – С. 59–63.]
167. 梁茂成. 语料库语言学研究的两种范式: 渊源、分歧及前景[J]. 外语教学与研究, 2012, 44(3): 323–335. [Лян Маочэн. Две парадигмы корпусных лингвистических исследований: источник, расхождение и перспектива / Маочэн Лян // Преподавание и исследование иностранных языков. – Пекин: Изд-во преподавания и исследования иностранных языков, 2012. – № 44(3). – С. 323–335.]
168. 李文新, 高凤兰. 基于语料库的俄语学习者词块分类特征分析及教学反思 [J]. 中国俄语教学, 2018, 37(4): 61–78. [Ли Вэньсинь. Анализ особенностей китайских учащихся по специальности русский язык на основе корпусных технологий / Вэньсинь Ли, Фэнлань Гао // Обучение русскому языку. – Пекин: Изд-во Пекинского университета иностранных языков, 2018. – № 37(4). – С. 61–78.]
169. 李邵哲. 基于语料库对 поезд 的历时研究 [J]. 俄罗斯语言文学与文化研究, 2012, 2: 21–26. [Ли Шаочжэ. Исследование по слову Поезд на основе

- корпусной лингвистики / Шаочжэ Ли // Русский язык и русская литература.
– Харбин: Изд-во Хэйлунцзянского университета, 2012. – № 2. – С. 21–26.]
170. 李小广. 英汉句法对比视角下的英语句子写作研究[J]. 长春师范大学学报, 2014, 33(6): 179–180. [Ли Сяогуан. Анализ предложений на английском языке с точки зрения сопоставления синтаксиса английского и китайского языков / Сяогуан Ли. // Вестник Чаньчуньского педагогического университета. – Чаньчунь: Изд-во Чаньчуньского педагогического университета, 2014. – № 33(6). – С. 179–180.]
171. 孙晓薇, 童丹. 高等学校俄语专业八级考试必备: 词汇篇 [M]. 外语教学与研究出版社, 285. [Сунь Сяовэй, Тун Дань. Лексика для экзамена по русскому языку для высших учебных заведений (уровень 8) / Сяовэй Сунь, Дань Тун. – Пекин: Изд-во преподавания и исследования иностранных языков, 2002. – 285 с.]
172. 吴娟. 专业学生俄语写作中的词汇错误分析及对词汇教学的启示 – 以石河子大学俄语专业为例[J]. 科教导刊. 2013(2): 94–95. [Ву Цзюань. Анализ лексических ошибок в текстах китайскими студентами на русском языке и меры их устранения – на примере магистрантов по специальности русский язык в университете Шихези / Цзюань Ву // Журнал науки и образования. – Хубэй: Ассоциация науки и технологий провинции Хубэй, 2013. – № 2. – С. 94–95.]
173. 王静. 俄汉被动结构对比[J]. 青年文学家, 2019(20):153-155. [Ван Цзин. Сравнение пассивных конструкций в русском и китайском языках / Цзин Ван // Молодые писатели. – Хэйлунцзян: Изд-во Молодых писателей, 2019.

– № 20. – С. 153–155.]

174. 王克非. 语料库翻译学探索[M]. 上海: 上海交通大学出版社, 2012: 249.
[Ван Кэфэй. Переводоведение на основе корпусной лингвистики / Кэфэй Ван. – Шанхай: Изд-во Шанхайского транспортного университета, 2012. – 249 с.]
175. 王立非, 孙晓坤. 国内外英语学习者语料库的发展:现状与方法[J]. 外语电化教学,2005,(5):19-24. [Ван Лифи. Развитие учебного корпуса английского языка в Китае и зарубежом: состояние и методы / Лифи Ван, Сяокунь Сунь // Электронное образование иностранных языков. – Шанхай: Изд-во Шанхайского университета иностранных языков, 2005. – № 5. – С. 19–24.]
176. 王萍, 王卫平. 中国大学生英语写作中中式英语的成因与对策[J]. 英语广场: 学术研究, 2011(Z2): 69–74. [Ван Пин. Причины ошибок в письменной речи на английском языке / Пин Ван, Вэйпин Ван // Платформа английского языка: научное исследование. – Хубэй: Изд-во Чанцзяна в провинции Хубэй, 2011. – 249с.]
177. 王臻. 俄语语料库语言学研究现状与瞻望[J]. 中国俄语教学, 2007, 2: 44–47. [Ван Чжэн. Современное состояние исследований и перспективы в области корпусной лингвистики русского языка / Чжэн Ван // Обучение русскому языку. – Пекин: Изд-во Пекинского университета иностранных языков, 2007. – № 2. – С. 44–47.]
178. 许家金. 语料库与中国学习者英语口语研究[M]. 北京: 外语教学与研究出版社 2020: 208. [Сюй Цзяцзинь. Корпусная лингвистика и исследование

- по устной речи английского языка китайскими студентами / Цзянцзинь Сюй. – Пекин: Изд-во преподавания и исследования иностранных языков, 2020. – 208 с.]
179. 杨惠中.基于 CLEC 语料库的中国学习者英语分析[M]. 上海外语教育出版社, 2005: 444. [Ян Хуэйчжун. Анализ английского языка китайских учащихся на основе корпуса CLEC / Хуэйчжун Ян. – Шанхай: Шанхайское издательство по образованию в области иностранных языков, 2005. – 444 с.]
180. 张春华.英语语言迁移对俄语的影响研究[J].人才培养, 2020(5):108-109. [Чжан Чуньхуа. Исследование влияния переноса английского языка на изучение русского языка как второй иностранный язык / Чуньхуа Чжан // Развитие талантов. – Чжэцзян: Чжэцзянский торговый университет, 2020. – № 5. – С. 108–109.]
181. 张禄彭,张超静.自建语料库在俄语教学研究者的应用[J].中国俄语教学, 2012, 31(3): 62–69. [Чжан Лупэн. Применение самостоятельно построенного корпуса в преподавание русского языка / Лупэн Чжан, Чаоцзин Чжан. Обучение русскому языку. – Пекин: Изд-во Пекинского университета иностранных языков, 2012. – № 31(3). – С. 62–69.]
182. 赵玮.中介语语料库词汇错误的标注问题及改进建议[J].云南师范大学学报. 2015, 13(2): 1–9. [Чжан Вэй. Разметка ошибок в межъязыковом корпусе и предложения по улучшению // Вестник Юньнаньского педагогического университета. – Юньнань: Изд-во Юньнаньского педагогического университета, 2015. – № 13(2). – С. 1–9.]

ПРИЛОЖЕНИЕ 1

ИНФОРМАЦИЯ ОБ АВТОРАХ ТЕКСТОВ В УЧЕБНОМ КОРПУСЕ

РУССКОГО ЯЗЫКА

Код студентов	Пол	Возраст	Тег уровня английского языка на ЕГЭ	Код студентов	Пол	Возраст	Тег уровня английского языка на ЕГЭ
1	F	19	C	46	F	19	B
2	F	18	B	47	M	19	C
3	F	18	C	48	M	20	A
4	F	19	B	49	F	20	C
5	F	18	C	50	F	19	A
6	F	20	D	51	F	19	B
7	F	19	D	52	M	19	A
8	F	19	B	53	M	19	B
9	F	19	A	54	M	20	A
10	F	19	C	55	F	20	B
11	F	19	A	56	F	19	B
12	F	18	B	57	F	19	B
13	F	19	B	58	F	18	A
14	F	19	A	59	M	21	D
15	M	19	B	60	F	19	B
16	F	18	C	61	F	19	B
17	F	20	C	62	F	19	A
18	F	18	C	63	F	19	B
19	F	18	C	64	F	19	A
20	M	19	B	65	F	19	B
21	F	20	C	66	F	19	C
22	M	19	B	67	F	19	C
23	F	22	D	68	F	19	B
24	M	19	B	69	M	20	A
25	M	19	A	70	F	19	C
26	F	19	D	71	F	19	C
27	M	19	B	72	F	20	A
28	F	20	B	73	M	19	B
29	F	18	B	74	M	19	B
30	F	19	B	75	F	19	A
31	F	18	C	76	F	19	A
32	F	19	D	77	F	20	C
33	F	19	B	78	F	19	A
34	F	20	C	79	F	19	B
35	F	19	B	80	M	19	A
36	F	20	B	81	F	19	B
37	F	18	B	82	F	20	B
38	F	18	C	83	F	20	A
39	F	19	C	84	F	19	C
40	F	21	A	85	F	19	B

41	F	19	B	86	M	19	B
42	F	19	A	87	F	19	B
43	F	18	B	88	F	22	D
44	M	19	B	89	F	19	C
45	M	18	B	90	F	19	C

ПРИЛОЖЕНИЕ 2

ПИСЬМЕННЫЕ ТЕКСТЫ КИТАЙСКИХ СТУДЕНТОВ НА РУССКОМ ЯЗЫКЕ (ФРАГМЕНТ)

Тема № 1: Мое интересное путешествие

Путешествия - это прежде всего новые впечатления. Это знакомства с новыми людьми, с которыми мы будем общаться не один год. Путешествия - это ещё и новые знания. Они открывают нам мир и делают нас богаче, делают нашу жизнь насыщеннее и ярче. Путешествия дают нам отойти от повседневных будней. Уехав подальше с того места, где живём и работаем, мы можем отдохнуть не только телом, но и душой.

Я особенно люблю путешествовать, и я часто использую свободное время для путешествия. Я побывала во многих городах, но всегда в Китае, я мечтаю, что когда-нибудь смогу поехать за границу. В прошлом году в праздник луны, я ездила с друзьями в горы, дорога в горах была очень трудна, так что мы почти весь день взобрались на вершину горы, в горах воздух был чист, на камне были нарисованы слова древнего поэта, Стоя на вершине горы, как будто можно дотронуться до облаков. Потом мы ночевали в горах и готовились к восходу солнца. На следующее утро в пять часов утра, солнце медленно поднимается, это действительно самая красивая картина в моей жизни. Туристы восхищаются красотой природы, фотографировать на память. Кроме того, я побывала в Ханчжоу, Янчжоу, видела огромное море, красивое Западное озеро. Это мой самый интересный и прекрасный туристический опыт.

Пока я живу на этой земле, я буду путешествовать. Я буду познавать мир, буду делать всё возможное, чтобы прочитать эту книгу.

Тема № 2: Мой любимый город

Мой любимый город - Лоян, и это также мой родной город. Лоян имеет историю, насчитывающую более 5000 лет, и был центром основания династии и деятельности в то время. Лоян является одним из мест рождения китайской цивилизации, восточной отправной точкой Великого шелкового пути и центром Большого канала династий Суй и Тан. В истории более десяти династий основали свою столицу в Лояне.

Лоян пересекает северные и южные берега среднего и нижнего течения реки Хуанхэ, а местность сложная. Лоян занимал удобное географическое положение, находился на пересечении важных транспортных путей – и водных и сухопутных. Известно, что город получил название от реки. В Лояне есть

много привлекательных мест, и многие древние исторические места. артефакты были сохранены нетронутыми, поэтому он также известен как древняя столица Девяти династий. Кроме того, пион Лояна также известен во всем мире: каждый год, когда цветет пион, он всегда может привлечь туристов со всего мира. Аромат пиона и его цвет любят туристы. Кроме того, Лоянский парк Ванчэн также очень известен. В нем много животных. Музей Лоян также собирает множество культурных реликвий и демонстрирует великолепную культуру. Кроме того, Лоян имеет много продуктов питания и имеет богатые местные особенности. Лоян - древний город с культурными особенностями.

Лоян - мой родной город и мой любимый город, Я люблю этот город, куда бы я ни шёл.

Тема № 3: Интернет в нашей жизни

Известно, что Интернет уже стал неотъемлемой частью нашей жизни и даёт нам возможность жить лучше и удобнее. Можно сказать, что без него нам не обойтись. Всё-таки какую роль Интернет играет в нашей жизни? Что приносит нам Интернет?

Прежде всего, Интернет приносит нам много пользы, и сильно влияет на нас во всех областях жизни. Благодаря ему мы можем найти любую нужную информацию, найти ответы на разные вопросы, немножко развлекаться, послушать музыку, поболтать с собеседниками. Ещё в Интернете можно купить всё, не выходя из дома. Кроме того, Интернет является отличным местом, где мы можем беседовать с друзьями, даже с незнакомыми людьми, находящимися в любом углу планеты.

Однако, злоупотребление Интернетом может повлечь много вопросов. С ростом развития всемирной паутины конфликт между виртуальным и реальным общением приобретает всё больше актуальность. Современные люди тратят много времени на виртуальное общение, целыми днями сидят у компьютера, никуда не ходят, ни с кем не общаются, а меньше уделяют внимание разговору глаза в глаза. Это сильно мешает умению живого общения. Вот специалисты уже ставят Интернет-зависимость в один ряд с табачной, алкогольной и даже наркотической. Между тем, Интернет приведёт к падению образованности студентов, так как они перестали мыслить и работать самостоятельно, так как с помощью Интернета ими легко найти правильные ответы на все вопросы, которые дал преподаватель.

В заключение хочу сказать, что Интернет – один из лучших изобретений человечества. Но его роль в жизни каждого человека сложно переоценить. Он является лучшим помощником для тех, кто правильно им распоряжается, а злом- кто уже был в его плену. Поэтому мы должны правильно относиться к

нему: не только использовать все его блага, но и обратить внимание на реальное общение, чтобы не заблудиться в Интернете.

Тема № 4: Мои увлечения

Недаром говорят, на вкус и цвет товарища нет, я считаю, мир увлечений безграничен, и каждому человеку важно иметь какое-то хобби. Я давно прочитала в статье, что интерес человека — это то, что он делает не потому, что надо, а потому, что ему это нравится и хочется делать. Поэтому я думаю, что любимые занятия приносят человеку удовольствие, радость и позволяет забыть о повседневных проблемах.

Круг моих интересов разнообразен, если речь идёт о моем любимом увлечении, я очень люблю танцевать. Современные танцы мне нравятся гораздо больше, чем классические, потому что они дают большую свободу самовыражения. Когда я поступила во Второй Пекинский университет иностранных языков, я познакомилась с группой людей, которые устраивают танцевальное собрание, и дважды в неделю, мы танцуем вместе. Можно сказать, что я много времени посвящаю танцам и когда я танцую, я чувствую счастье. Ещё я люблю заниматься спортом. Каждый вечер я бегаю на стадионе с моим подругой. На мой взгляд, заниматься спортом хорошо укрепляет сердце, тренирует лёгкие и силы воли.

По-моему, без увлечения жизни становится неинтересной. Хотя некоторые из моих ребят часто жалуются, что они некогда делают то, что нравится, но я никак не представляю своей жизни без увлечения.

Тема № 5: Моя студенческая жизнь

Меня зовут Катя. Сейчас я учусь во Втором Пекинском университете на втором курсе. Я живу в общежитии университета.

Наше общежитие находится в центре университета рядом с столовой и супермаркетом около стадиона. Нам можно дойти до учебного корпуса за десять минут.

Большое внимание уделяется безопасности студентов. Для этого в нашем общежитии функционируют системы видеонаблюдения и система противопожарной безопасности, круглосуточно работают сотрудники охраны. Мы входим в здание по пропускам и по сканирование лица.

У нас общежитие типа коридорного, и туалетные комнаты находятся на этаже. В каждой комнате бывают кровати, столы, стулья, книжные полки и шкафы. Студенты бакалавриата живут по пять людей.

Мы общаемся с соседями по комнате, стараюсь согласовать время работы и отдыха, поэтому мало мешаем друг другу. Это позволит мне понять, что мне

нужно учиться находить компромиссы. Обычно мы вместе делаем задания, слушаем музыку, читаем книги, смотрим сериалы и фильмы, и так далее. Эти незабываемые весёлые моменты будут всегда напоминать нам о молодости.

Одним словом, наш университет старается сделать всё, чтобы студенты жили комфортно и спокойно. И мне кажется, что жизнь в общежитии — это важная часть студенческой культуры.

Тема № 6: Книга в нашей жизни

Горький сказал: Книги - это лестница человеческой прогресса. Книги играют важную роль в нашей обычной жизни. Через чтение человек выражает себя, развивается. По-моему, невозможно воспитать по-настоящему культурных людей, если дома нет разных книг.

С раннего детства я начал знакомство с книгами: тогда я ещё очень молодой, и больше всего только читал детские рассказы, карикатуры и повести. Рос - я росла моя библиотека. После того, я стал любителем военного дела, который самый разный, чем другие люди. У меня в доме выше ста книг о том. Тогда мне нравится военное дело, просто потому, что внешность солдата воинственная, оружие красивые... А потом, после смотреть фильм «Рота Братьев», я начал думать о сути войны: Простая ли война, как мы часто думаем? Военный вид великолепный или кровавый? Эти проблемы придали мне способность думать самостоятельно, и я они по-прежнему использую.

Дальше, по требованиям экзамена я читал иностранные и внутренние книги к экзамену, на пример, «Сон в красном корпусе» и «Речные заводи» и так далее, но все они мне не нужно прочитать. Невесть отчего, я никак не могу найти интересы ко книгам. Сейчас по обусловлению библиотеки, я не могу ссылаюсь на классиков, когда говорю, и часто считаю, что у меня мало содержания, и может быть, оскорблю чьи чувства. У вас что-нибудь идея? Расскажи мне, если есть советы.

Тема № 7: Как я провел(а) свой день рождения

мой день рождения — 23 Мать. Мои родители и друзья празднуют меня каждый год.

В этот день мы обычно устраиваем дома небольшую вечеринку, приглашаем друзей к себе на ужин. Мы болтали, смотрели телевизор или играли вместе, все веселились. Когда мы играли, мои родители готовили нам вкусный ужин. Мы с друзьями очень любим эти блюда. Конечно, я тоже получил подарки от родителей и друзей. Я счастлива, потому что они всегда знают, чего я хочу.

В этом году я начала готовиться к своему дню рождения неделю назад, я

купила себе новую одежду и обувь, одета очень симпатично, моя мама готовила вкусное блюдо, в обеде мы просто съели лапшу, это долговечная лапша. Конечно, вареные яйца также необходимы, это означает удачу и благословения. В конце дня рождения, родители и друзья собираются со мной разделить торт. Торт, который мама со мной сделали, был моим любимым вкусом. После того я отослала друзей домой, в этот день мы веселились, поэтому мы договорились, что следующий день рождения будем вместе проводить.

это мой самый счастливый день рождения, я надеюсь, что каждый день рождения будет так счастлива.

Тема № 8: Моя мечта

Говорят, что на свете у каждого своя мечта: кто-то хочет стать инженером, кто-то желает быть экономистом, кто-то мечтает стать музыкантом и так далее. А для меня быть учительницей — это моя мечта.

С раннего детства я уважаю и люблю своих учителей, восхищаюсь их благородным и терпеливым трудом. Я не только получаю от них много знаний, но и всегда чувствую их заботу и внимание. Учителя, как свечи, освещают другим путь к счастью. Они учат своих учеников, как быть настоящими людьми, как стать бойкими и крепкими во всех случаях жизни, как стать полезными для Родины и народа. Всем известно, что учитель — инженер человеческих душ. И мне кажется, что педагогическая профессия — одна из самых священных и благородных профессий в мире. Вот почему я хочу стать учительницей.

Ведь многие опытные учёные говорят, что успеха можно достигнуть только при постоянных усилиях. Так, чтобы быть хорошей учительницей, самой надо хорошо учиться и много знать. Поэтому я сейчас стараюсь как можно лучше учиться, как можно больше обогащать свои знания, и как можно поднимать своё терпение. Одним словом, я всегда твёрдо уверена в том, что моя мечта обязательно сбудется в ближайшем будущем.

Тема № 9: Как мы провели праздник Весны

В Китае праздник весны - самый популярный и важный народный праздник. Праздник весны обычно начинает отмечать с конца января по начало февраля.

К празднику весны начинают готовиться заранее: начинают думать о подарках родным и друзьям, покупают продукты для новогоднего стола, готовят праздничную одежду, приглашают гостей... Перед праздником в каждом доме проводится тщательная уборка. Люди украшают дом

новогодними картинками с изображением символов «счастливой жизни» или древних и современных героев, вешают на ворота красные фонари.

Вечером накануне праздника вся семья собирается вместе за праздничным столом, чтобы проводить старый год и встретить новый. Каждый вспоминает всё хорошее, что было в старом году. Ровно в двенадцать часов ночи встречают новый год. На улицах, во дворах начинают рваться хлопушки, зажигается фейерверк. Все поздравляют друг друга с праздником, желают здоровья, счастья и успехов, дарят подарки родным, близким и друзьям. Повсюду разносится смех и радостные крики до самой поздней ночи. На следующий день с утра начинаются визиты к родственникам и друзьям.

Праздник весны - мой самый любимый праздник. Я очень рад провести этот праздник с семьей.

Тема № 10: Самый любимый человек

С раннего детства я увлекаюсь русской литературой, читаю много произведений русских писателей. Мой любимый русский писатель — Лев Николаевич Толстой.

Лев Николаевич Толстой — талантливый художник, на долю которого было создано много выдающихся произведений, как «Война и мир», «Анна Каренина», «Воскресение» и т. д. Ведь Лев Николаевич Толстой родился в усадьбе Ясная Поляна под Тулой в XIX веке, поэтому он сумел осмыслить и отразить жизнь простого русского народа в своём творчестве. Прежде чем Толстой познакомился со стихами Пушкина, он уже знал много народных песен, сказок, легенд.

В 60-е годы он написал роман «Война и мир», величайшее произведение русской литературы. В этом произведении проявляются сила и величие русского народа в победе над войной. После «Войны и мира» Толстым был написан роман о современной жизни «Анна Каренина». Устами героя, Анны Каренины, Толстой показал людям не только трагедию в поисках счастья и любви, но и в то же время современную семью, силу любви, и понимание долга. Главным произведением Льва Николаевича Толстого последних лет стал роман «Воскресение». Своим рассказом Толстой поведал о том, как жестоко и темно правили феодальные классы. Известно, Лев Николаевич Толстой относился к трудящимся благосклонно и добро, так как трудящиеся любят его многие произведения.

Одним словом, Лев Николаевич Толстой внёс огромный вклад в развитие русской и мировой литературы. Память о великом писателе будет вечной для всех людей.

ПРИЛОЖЕНИЕ 3

УЧЕБНЫЙ КОРПУС РУССКОГО ЯЗЫКА ДЛЯ ОБУЧЕНИЯ КИТАЙСКИХ СТУДЕНТОВ (ФРАГМЕНТ)

Тема № 1: Мое интересное путешествие

<1.01.19.F.C>

Путешествия <- pt3> это прежде всего новые впечатления. Это знакомства с новыми людьми, с <которыми fm1> мы будем общаться не один год. Путешествия <- pt3> это ещё и новые знания. Они открывают нам мир и делают нас богаче, делают нашу жизнь насыщеннее и ярче. Путешествия дают нам отойти от <повседневных будней wd1>. Уехав подальше с того места, где живём и работаем, мы можем отдохнуть не только телом, но и душой.

Я особенно люблю путешествовать, и я часто использую свободное время в университете для путешествия. Я побывала во многих городах, но всегда в <китае fm3>, я <мечтаю vr3>, что когда-нибудь смогу поехать за границу путешествовать.<В pp2> прошлом году в праздник луны <, pt2> я ездила с друзьями в горы, дорога в горах была очень трудна, так что мы почти весь день взобрались на вершину горы, в горах воздух был чист, на камне были нарисованы слова древнего поэта <, pt3> Стоя на вершине горы, как будто можно дотронуться до облаков. Потом мы ночевали в горах и готовились к восходу солнца. <на fm3> следующее утро в пять часов утра <, pt2> солнце медленно поднимается <, pt3> <это fm3> действительно самая красивая картина в моей жизни. <туристы fm3> восхищаются красотой природы, <фотографировать sn3> на память. Кроме того, я побывала в <ханчжоу fm3>, <янчжоу fm3>, <видела vr2> <огромное море ss1>, красивое Западное озеро. Это мой самый интересный и прекрасный туристический опыт.

Пока я живу на этой земле, я буду путешествовать. <Я буду познавать мир, буду делать всё возможное, чтобы прочитать эту книгу. cl2>

Тема № 2: Мой любимый город

<2.01.19.F.C>

Мой любимый город <- pt3> Лоян, и это также мой родной город. Лоян имеет историю, насчитывающую более 5000 лет, и был центром основания <династии pr4> и деятельности в то время. Лоян является одним из мест рождения китайской цивилизации, восточной отправной точкой Великого шелкового пути и центром Большого канала династий Суй и Тан. В истории более десяти династий основали свою столицу в Лояне.

Лоян пересекает северные и южные берега среднего и нижнего течения реки Хуанхэ, а местность сложная. Лоян занимал удобное географическое положение, находился на пересечении важных транспортных путей <- pt3> и водных и сухопутных. Известно, что город получил название от реки. В Лояне есть много привлекательных мест <, pt2> и многие древние исторические места. <артефакты fm3> были сохранены нетронутыми, поэтому он также известен как древняя столица Девяти династий. Кроме того, пион Лояна также известен во всем мире: каждый год, когда цветет пион, он всегда может привлечь туристов со всего мира. Аромат пиона и его цвет любят туристы. Кроме того, Лоянский парк Ванчэн также очень известен. В нем много животных. Музей Лоян также собирает множество культурных реликвий и демонстрирует великолепную культуру. Кроме того, Лоян имеет много продуктов питания и <имеет wd1> богатые местные особенности. Лоян <- pt3> древний город с культурными особенностями.

Лоян <- pt3> мой родной город и мой любимый город, Я люблю этот город, куда бы я ни шёл.

Тема № 3: Интернет в нашей жизни

<3.78.19.F.A>

Известно, что Интернет уже стал неотъемлемой частью нашей жизни и даёт нам возможность жить лучше и удобнее. Можно сказать, что без него нам не обойтись. Всё-таки какую роль Интернет играет в нашей жизни? Что приносит нам Интернет?

Прежде всего, Интернет приносит нам много <польз pr1>, и сильно влияет на нас во всех областях жизни. Благодаря ему мы можем найти любую нужную информацию, найти ответы на разные вопросы, немножко развлекаться, послушать музыку, поболтать с собеседниками. Ещё в Интернете можно купить всё, не выходя из дома. Кроме того, Интернет является отличным местом, где мы можем беседовать с друзьями, даже с незнакомыми людьми, находящимися в любом углу планеты.

Однако, злоупотребление Интернетом может повлечь много вопросов. С ростом развития всемирной паутины конфликт между виртуальным и реальным общением приобретает всё больше актуальность. Современные люди тратят много времени на виртуальное общение, целыми днями сидят у компьютера, никуда не ходят, ни с кем не общаются, а меньше уделяют внимание разговору глаза в глаза. Это сильно мешает умению живого общения. Вот специалисты уже ставят Интернет-зависимость в один ряд с табачной, алкогольной и даже наркотической. Между тем, Интернет приведёт к падению

образованности студентов, так как они перестали мыслить и работать самостоятельно, <так как wd2> с помощью Интернета ими легко найти правильные ответы на все вопросы, которые дал преподаватель.

В заключение хочу сказать, что Интернет – один из лучших изобретений человечества. Но его роль в жизни каждого человека сложно переоценить. Он является лучшим помощником для тех, кто правильно им распоряжается, а злом <- pt3> кто уже был в его плену. Поэтому мы должны правильно относиться к нему: не только использовать все его блага, но и обратить внимание на реальное общение, чтобы не заблудиться в Интернете.

Тема № 4: Мои увлечения

<4.57.19.F.B>

Недаром говорят, на вкус и цвет товарища нет, я считаю, мир увлечений безграничен, и каждому человеку важно иметь какое-то хобби. Я давно прочитала в статье, что интерес человека — это то, что он делает не потому, что надо, а потому, что ему это нравится и хочется делать. Поэтому я думаю, что любимые занятия приносят человеку удовольствие, радость и позволяет забыть о повседневных проблемах.

Круг моих интересов разнообразен, если речь идёт о моем любимом увлечении, я очень люблю танцевать. Современные танцы мне нравятся гораздо больше, чем классические, потому что они дают большую свободу самовыражения. Когда я поступила во Второй Пекинский университет иностранных языков, я познакомилась с группой людей, которые устраивают <танцевальное fm1> собрание, и дважды в неделю <, pt2> мы <танцуем vp2> вместе. Можно сказать, что я много времени посвящаю танцам и когда я танцую, я чувствую счастье. Ещё я люблю заниматься спортом. Каждый вечер я бегаю на стадионе с <моим pr1> подругой. На мой взгляд, заниматься спортом хорошо <укрепляет fm1> сердце, тренирует лёгкие и силы воли.

По-моему, без увлечения <жизни pr4> <становится fm1> неинтересной. Хотя некоторые из моих <ребят wd1> часто <жалуются vp3>, <что они некогда ad2> делают то, что нравится, но я никак не представляю своей жизни без <увлечения pr1>.

Тема № 5: Моя студенческая жизнь

<5.63.19.F.B>

Меня зовут Катя. Сейчас я учусь во Втором Пекинском университете на втором курсе. Я живу в общежитии университета.

Наше общежитие находится в центре университета рядом <с pr2> столовой и супермаркетом около стадиона. Нам можно дойти до учебного

корпуса за десять минут.

Большое внимание уделяется безопасности студентов. Для этого в нашем общежитии функционируют системы видеонаблюдения и система противопожарной безопасности, круглосуточно работают сотрудники охраны. Мы входим в здание по пропускам и <по сканирование pp2> лица.

У нас общежитие типа коридорного, и туалетные комнаты находятся на этаже. В каждой комнате <бывают wd1> кровати, столы, стулья, книжные полки и шкафы. Студенты бакалавриата живут по <пять nu1> людей.

Мы общаемся с соседями по комнате, стараюсь согласовать время работы и отдыха, поэтому мало мешаем друг другу. Это позволит мне понять, что мне нужно учиться находить <компромиссы np1>. Обычно мы вместе делаем задания, слушаем музыку, читаем книги, смотрим сериалы и фильмы, и так далее. Эти незабываемые весёлые моменты будут всегда напоминать нам о молодости.

Одним словом, наш университет старается сделать всё, чтобы студенты жили комфортно и спокойно. И мне кажется, что жизнь в общежитии – это важная часть студенческой культуры.

Тема № 6: Книга в нашей жизни

<6.46.19.F.B>

Горький сказал: Книги <- pt3> это лестница человеческой прогресса. Книги играют важную роль в нашей обычной жизни. Через чтение человек выражает себя, развивается. По-моему, невозможно воспитать по-настоящему культурных людей, если дома нет разных книг.

С раннего детства я начал знакомиться с книгами <: pt3> тогда я ещё очень <молодой wd1>, и больше всего только читал детские рассказы, карикатуры и повести. Рос я <- pt3> росла моя библиотека. После того <, pt2> я стал любителем <военного дела ss1>, который самый разный, чем другие люди. У меня в доме <выше ad1> ста книг о <том pr1>. Тогда мне нравится <военное дело ss1> <, pt2> просто потому, что внешность солдата воинственная, оружие красивые... А потом, <после смотреть pp1> фильм «Рота Братьев», я начал думать о сути войны: Простая ли война, как мы часто думаем? Военный вид великолепный или кровавый? Эти проблемы придали мне способность думать самостоятельно, и я <они по-прежнему использую vr3>.

Дальше, по требованиям экзамена я читал иностранные и <внутренние wd1> книги <к экзамену pp2>, например, «Сон в красном корпусе» и «Речные заводи» и так далее, <но все их мне не нужно прочитать sn4>. Невесть отчего, я никак не могу найти интересы ко книгам. Сейчас <по обусловлению pp2>

библиотеки, я не <могу ссылаюсь vr3> на классиков, когда говорю, и часто считаю, что у меня мало содержания, и может быть, оскорблю <чьи pr1> чувства. У вас <что-нибудь pr1> идея? Расскажи мне, если есть советы.

Тема № 7: Как я провел(а) свой день рождения

<7.06.20.F.D>

<мой fm3> день рождения — 23 <Мать fm1>. Мои родители и друзья празднуют меня каждый год.

В этот день мы обычно устраиваем дома небольшую вечеринку, приглашаем друзей к себе на ужин. Мы болтали, смотрели телевизор или играли вместе, все веселились. Когда мы играли, мои родители готовили нам вкусный ужин. Мы с друзьями очень любим эти блюда. Конечно, я тоже получил подарки от родителей и друзей. Я счастлива, потому что они всегда знают, чего я хочу.

В этом году я начала готовиться к своему дню рождения неделю назад, я купила себе новую одежду и обувь, одета очень симпатично <, pt3> <моя fm3> мама готовила вкусное блюдо, <в pr2> обеде мы просто съели лапшу, это долговечная лапша. Конечно, вареные яйца также необходимы, это означает удачу и благословения. <В конце дня рождения ss1>, родители и друзья <собираются со мной разделить торт sn4>. Торт, который мама со мной сделали, был моим любимым вкусом. После того я отослала друзей домой <, pt3> <в fm3> этот день мы веселились, поэтому мы договорились, что следующий день рождения будем вместе проводить.

<это fm3> мой самый счастливый день рождения, я надеюсь, что <каждый день рождения будет так счастлива sn3>.

Тема № 8: Моя мечта

<8.04.19.F.B>

Говорят, что на свете у каждого своя мечта: кто-то хочет стать инженером, кто-то желает быть экономистом, кто-то мечтает стать музыкантом и так далее. А для меня быть учительницей – это моя мечта.

С раннего детства я уважаю и люблю своих учителей, восхищаюсь их благородным и терпеливым трудом. Я не только получаю от них много знаний, но и всегда чувствую их заботу и внимание. Учителя, как свечи, освещают другим путь к счастью. Они учат своих учеников, как быть настоящими людьми, как стать бойкими и крепкими во всех случаях жизни, как стать полезными для Родины и народа. Всем известно, что учитель – инженер человеческих душ. И мне кажется, что педагогическая профессия – одна из самых священных и благородных профессий в мире. Вот почему я хочу стать

учительницей.

Ведь многие опытные учёные говорят, что успеха можно достигнуть только при постоянных усилиях. Так, чтобы быть хорошей учительницей, самой надо хорошо учиться и много знать. Поэтому я сейчас стараюсь как можно лучше учиться, как можно больше обогащать свои знания, и как можно поднимать своё терпение. Одним словом, я всегда твёрдо уверена в том, что моя мечта обязательно сбудется в ближайшем будущем.

Тема № 9: Как мы провели праздник Весны

<9.41.19.F.B>

В Китае праздник <весны fm3> <- pt3> самый популярный и важный народный праздник. Праздник <весны fm3> обычно начинает отмечать с конца января <по начало pp1> февраля.

К празднику <весны fm3> начинают готовиться заранее: начинают думать о подарках родным и друзьям, покупают продукты для новогоднего стола, готовят праздничную одежду, приглашают гостей... Перед праздником в каждом доме проводится тщательная уборка. Люди украшают дом новогодними картинками с изображением символов «счастливой жизни» или древних и современных героев, вешают на ворота красные фонари.

Вечером накануне праздника вся семья собирается вместе за праздничным столом, чтобы проводить старый год и встретить новый. Каждый вспоминает всё хорошее, что было в старом году. Ровно в двенадцать часов ночи <встречают новый год ss1>. На улицах, во дворах начинают <рваться wd1> хлопушки, зажигаться фейерверк. Все поздравляют друг друга с праздником, желают здоровья, счастья, успехов, дарят подарки родным, близким, друзьям. Повсюду разносится смех и радостные крики до самой поздней ночи. На следующий день с утра начинаются визиты к родственникам и друзьям.

Праздник <весны fm3> <- pt3> мой самый любимый праздник. Я очень рад провести <этот fm1> праздник с семьей.

Тема № 10: Самый любимый человек

<10.04.19.F.B>

С раннего детства я увлекаюсь русской литературой, читаю много произведений русских писателей. Мой любимый русский писатель – Лев Николаевич Толстой.

Лев Николаевич Толстой – талантливый художник, на долю которого было создано много выдающихся произведений, как «Война и мир», «Анна Каренина», «Воскресение» и т. д. Ведь Лев Николаевич Толстой родился в усадьбе Ясная Поляна под Тулой в XIX веке, поэтому он сумел осмыслить и

отразить жизнь простого русского народа в своём творчестве. Прежде чем Толстой познакомился со стихами Пушкина, он уже знал много народных песен, сказок, легенд.

В 60-е годы он написал роман «Война и мир», величайшее произведение русской литературы. В этом произведении проявляются сила и величие русского народа в победе над войной. После «Войны и мира» Толстым был написан роман о современной жизни «Анна Каренина». Устами героя, Анны Каренины, Толстой показал людям не только трагедию в поисках счастья и любви, но и в то же время современную семью, силу любви, и понимание долга. Главным произведением Льва Николаевича Толстого последних лет стал роман «Воскресение». Своим рассказом Толстой повествовал о том, как жестоко и темно правили феодальные классы. Известно, Лев Николаевич Толстой относился к трудящимся благосклонно и добро, так как трудящиеся любят его многие произведения.

Одним словом, Лев Николаевич Толстой внёс огромный вклад в развитие русской и мировой литературы. Память о великом писателе будет вечной для всех людей.

ПРИЛОЖЕНИЕ 4

СПИСОК КОЛИЧЕСТВА ОШИБОК В УЧЕБНОМ КОРПУСЕ

РУССКОГО ЯЗЫКА

Тип ошибки	Подкатегория ошибки	Название ошибки	Тег	Количество
орфографическая ошибка	нарушено правописание	орфографическая ошибка за счет незнания орфограмм	fm1	507
		неверное словообразование	fm2	29
		неверное использование прописной буквы	fm3	395
морфологическая ошибка	ошибка в использовании глаголов	неверное использование возвратной формы	vp1	16
		неверное использование форм несовершенного или совершенного вида	vp2	115
		неверное управление глагола	vp3	186
		неверное использование депричастия	vp4	3
		неверный выбор вида причастия	vp5	9
	ошибка в использовании имен существительных	неверное использование форм числа имен существительных	np1	52
		отсутствие согласования по роду	np2	13
		неверное образование форм склонения	np3	135
		нарушение падежных форм	np4	62
	ошибка в использовании предлогов	неверное управление предлога	pp1	73
		использование неверного предлога	pp2	93
	ошибка в использовании имен прилагательных	неверное использование полной или краткой формы	aj1	84
		неверное использование степеней сравнения	aj2	4

	ошибка в использовании местоимений	неверный выбор местоимения	pr1	139
	ошибка в использовании наречий	неверное использование степеней сравнения	ad1	21
		неверное использование форм долженствования	ad2	28
	ошибка в использовании числительных	неверное использование форм числительных	nu1	19
лексическая ошибка	ошибка в использовании слова	употребление слова в несвойственном ему значении	wd1	92
		тавтология и повтор	wd2	24
		смешение синонимов и паронимов	wd3	101
	неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	cc1	148
фразеологическая ошибка		cc2	6	
синтаксическая ошибка	нарушение структуры предложений	пропуск членов предложения	sn1	134
		неверный порядок слов	sn2	25
	нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	sn3	125
		нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях	sn4	86
	пунктуационная ошибка	отсутствие знака препинания	pt1	57
		лишний знак препинания	pt2	62
		неверное использование знака препинания	pt3	278
коммуникационная ошибка	логическое нарушение	нарушение логики при перечислении	lc1	29
		отсутствие формальных показателей связи между предложениями	lc2	26

ПРИЛОЖЕНИЕ 5

ЧАСТОТА ОШИБОК КИТАЙСКИХ СТУДЕНТОВ С РАЗНЫМ УРОВНЕМ ВЛАДЕНИЯ АНГЛИЙСКИМ ЯЗЫКОМ

Значения коэффициента К при ошибках китайских студентов с уровнем —
владения английским языком — А

Подкатегория ошибки	Название ошибки	Тег	Коэффициент К
нарушено правописание	орфографическая ошибка за счет незнания орфограмм	fm1	0.005106
	неверное словообразование	fm2	0.000406
	неверное использование прописной буквы	fm3	0.003365
ошибка в использовании глаголов	неверное использование возвратной формы	vp1	0.000232
	неверное использование форм несовершенного или совершенного вида	vp2	0.000870
	неверное управление глагола	vp3	0.002611
	неверное использование деепричастия	vp4	0.000116
	неверный выбор вида причастия	vp5	0.000116
ошибка в использовании имен существительных	неверное использование форм числа имен существительных	np1	0.000870
	отсутствие согласования по роду	np2	0.000406
	неверное образование форм склонения	np3	0.001973
	нарушение падежных форм	np4	0.000638
ошибка в использовании предлогов	неверное управление предлога	pp1	0.000870
	использование неверного предлога	pp2	0.000696
ошибка в использовании имен прилагательных	неверное использование полной или краткой формы	aj1	0.001218
	неверное использование степеней сравнения	aj2	0.000058
ошибка в использовании местоимений	неверный выбор местоимения	pr1	0.001857
ошибка в использовании наречий	неверное использование степеней сравнения	ad1	0.000058
	неверное использование форм долженствования	ad2	0.000522

ошибка в использовании числительных	неверное использование форм числительных	nu1	0.000232
ошибка в использовании слова	употребление слова в несвойственном ему значении	wd1	0.001276
	тавтология и повтор	wd2	0.000406
	смешение синонимов и паронимов	wd3	0.001276
неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	cc1	0.001508
	фразеологическая ошибка	cc2	0.000000
нарушение структуры предложений	пропуск членов предложения	sn1	0.001566
	неверный порядок слов	sn2	0.000348
нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	sn3	0.001683
	нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях	sn4	0.000696
пунктуационная ошибка	отсутствие знака препинания	pt1	0.000812
	лишний знак препинания	pt2	0.000986
	неверное использование знака препинания	pt3	0.003017
логическое нарушение	нарушение логики при перечислении	lc1	0.000232
	отсутствие формальных показателей связи между предложениями	lc2	0.000058

Значения коэффициента К при ошибках китайских студентов с уровнем
владения английским языком — В

Подкатегория ошибки	Название ошибки	Тег	Коэффициент К
нарушено правописание	орфографическая ошибка за счет незнания орфограмм	fm1	0.006532
	неверное словообразование	fm2	0.000181
	неверное использование прописной буквы	fm3	0.003883
ошибка в использовании глаголов	неверное использование возвратной формы	vp1	0.000181
	неверное использование форм несовершенного или совершенного вида	vp2	0.001475
	неверное управление глагола	vp3	0.002378
	неверное использование деепричастия	vp4	0.000030
	неверный выбор вида причастия	vp5	0.000090
ошибка в использовании имен существительных	неверное использование форм числа имен существительных	np1	0.000572
	отсутствие согласования по роду	np2	0.000120
	неверное образование форм склонения	np3	0.001746
	нарушение падежных форм	np4	0.000873
ошибка в использовании предлогов	неверное управление предлога	pp1	0.00111
	использование неверного предлога	pp2	0.000696
ошибка в использовании имен прилагательных	неверное использование полной или краткой формы	aj1	0.001023
	неверное использование степеней сравнения	aj2	0.000120
ошибка в использовании местоимений	неверный выбор местоимения	pr1	0.001806
ошибка в использовании наречий	неверное использование степеней сравнения	ad1	0.000120
	неверное использование форм долженствования	ad2	0.000572

ошибка в использовании числительных	неверное использование форм числительных	nu1	0.000151
ошибка в использовании слова	употребление слова в несвойственном ему значении	wd1	0.001355
	тавтология и повтор	wd2	0.000271
	смещение синонимов и паронимов	wd3	0.001264
неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	cc1	0.001626
	фразеологическая ошибка	cc2	0.000090
нарушение структуры предложений	пропуск членов предложения	sn1	0.001716
	неверный порядок слов	sn2	0.000181
нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	sn3	0.001626
	нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях	sn4	0.000873
пунктуационная ошибка	отсутствие знака препинания	pt1	0.000542
	лишний знак препинания	pt2	0.000542
	неверное использование знака препинания	pt3	0.002438
логическое нарушение	нарушение логики при перечислении	lc1	0.000391
	отсутствие формальных показателей связи между предложениями	lc2	0.000331

Значения коэффициента К при ошибках китайских студентов с уровнем
 владения английским языком — С

Подкатегория ошибки	Название ошибки	Тег	Коэффициент К
нарушено правописание	орфографическая ошибка за счет незнания орфограмм	fm1	0.006624
	неверное словообразование	fm2	0.000442
	неверное использование прописной буквы	fm3	0.006050
ошибка в использовании глаголов	неверное использование возвратной формы	vp1	0.000177
	неверное использование форм несовершенного или совершенного вида	vp2	0.001722
	неверное управление глагола	vp3	0.001987
	неверное использование деепричастия	vp4	0.000000
	неверный выбор вида причастия	vp5	0.000088
ошибка в использовании имен существительных	неверное использование форм числа имен существительных	np1	0.000574
	отсутствие согласования по роду	np2	0.000088
	неверное образование форм склонения	np3	0.001325
	нарушение падежных форм	np4	0.000751
ошибка в использовании предлогов	неверное управление предлога	pp1	0.000972
	использование неверного предлога	pp2	0.001192
ошибка в использовании имен прилагательных	неверное использование полной или краткой формы	aj1	0.000751
	неверное использование степеней сравнения	aj2	0.000000
ошибка в использовании местоимений	неверный выбор местоимения	pr1	0.001546
ошибка в использовании наречий	неверное использование степеней сравнения	ad1	0.000088
	неверное использование форм долженствования	ad2	0.000574

ошибка в использовании числительных	неверное использование форм числительных	nu1	0.000309
ошибка в использовании слова	употребление слова в несвойственном ему значении	wd1	0.000795
	тавтология и повтор	wd2	0.000353
	смещение синонимов и паронимов	wd3	0.001237
неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	cc1	0.002120
	фразеологическая ошибка	cc2	0.000132
нарушение структуры предложений	пропуск членов предложения	sn1	0.001502
	неверный порядок слов	sn2	0.000530
нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	sn3	0.001457
	нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях	sn4	0.001369
пунктуационная ошибка	отсутствие знака препинания	pt1	0.000662
	лишний знак препинания	pt2	0.000883
	неверное использование знака препинания	pt3	0.004151
логическое нарушение	нарушение логики при перечислении	lc1	0.000353
	отсутствие формальных показателей связи между предложениями	lc2	0.000442

Значения коэффициента К при ошибках китайских студентов с уровнем
 владения английским языком — D

Подкатегория ошибки	Название ошибки	Тег	Коэффициент К
нарушено правописание	орфографическая ошибка за счет незнания орфограмм	fm1	0.005913
	неверное словообразование	fm2	0.000669
	неверное использование прописной буквы	fm3	0.007921
ошибка в использовании глаголов	неверное использование возвратной формы	vp1	0.000223
	неверное использование форм несовершенного или совершенного вида	vp2	0.001339
	неверное управление глагола	vp3	0.001896
	неверное использование деепричастия	vp4	0.000000
	неверный выбор вида причастия	vp5	0.000223
ошибка в использовании имен существительных	неверное использование форм числа имен существительных	np1	0.000558
	отсутствие согласования по роду	np2	0.000000
	неверное образование форм склонения	np3	0.001450
	нарушение падежных форм	np4	0.000558
ошибка в использовании предлогов	неверное управление предлога	pp1	0.001004
	использование неверного предлога	pp2	0.001562
ошибка в использовании имен прилагательных	неверное использование полной или краткой формы	aj1	0.001339
	неверное использование степеней сравнения	aj2	0.000000
ошибка в использовании местоимений	неверный выбор местоимения	pr1	0.001339
ошибка в использовании наречий	неверное использование степеней сравнения	ad1	0.000000
	неверное использование форм долженствования	ad2	0.000112

ошибка в использовании числительных	неверное использование форм числительных	nu1	0.000335
ошибка в использовании слова	употребление слова в несвойственном ему значении	wd1	0.000781
	тавтология и повтор	wd2	0.000000
	смещение синонимов и паронимов	wd3	0.001004
неверное использование устойчивых словосочетаний	неверное использование устойчивых выражений	cc1	0.002231
	фразеологическая ошибка	cc2	0.000000
нарушение структуры предложений	пропуск членов предложения	sn1	0.001785
	неверный порядок слов	sn2	0.000112
нарушение норм формального синтаксиса	нарушение согласования подлежащего и сказуемого	sn3	0.001004
	нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях	sn4	0.001562
пунктуационная ошибка	отсутствие знака препинания	pt1	0.001116
	лишний знак препинания	pt2	0.000781
	неверное использование знака препинания	pt3	0.005689
логическое нарушение	нарушение логики при перечислении	lc1	0.000446
	отсутствие формальных показателей связи между предложениями	lc2	0.000446

ПРИЛОЖЕНИЕ 6

ЧАСТОТНЫЕ СЛОВА В ТЕКСТАХ КИТАЙСКИХ СТУДЕНТОВ НА РУССКОМ ЯЗЫКЕ (ФРАГМЕНТ)

Ранг	Слова	Ранг	Слова	Ранг	Слова
1	и	81	дома	161	иногда
2	в	82	родной	162	место
3	я	83	кажется	163	пельмени
4	с	84	стать	164	раньше
5	что	85	себя	165	считать
6	мы	86	путешествовать	166	Толстой
7	на	87	самый	167	здоровье
8	не	88	роль	168	несколько
9	это	89	чем	169	слово
10	жизнь	90	один	170	парк
11	человек	91	читать	171	вечер
12	но	92	хотеть	172	дать
13	у	93	университет	173	знание
14	очень	94	многий	174	мама
15	город	95	вы	175	некоторый
16	много	96	обычно	176	произведение
17	по	97	даже	177	узнать
18	как	98	например	178	народ
19	год	99	другой	179	особенно
20	время	100	кто	180	перед
21	быть	101	раз	181	почти
22	из	102	новый	182	проводить
23	для	103	спорт	183	сегодня
24	день	104	хорошо	184	составлять
25	мой	105	бы	185	нельзя
26	интернет	106	нужно	186	ни
27	он	107	ведь	187	столица
28	мочь	108	потом	188	большой
29	а	109	теперь	189	должен
30	который	110	думать	190	весело
31	есть	111	туризм	191	Гугун
32	это	112	надо	192	иметь
33	каждый	113	где	193	культура
34	когда	114	мечта	194	радость
35	то	115	найти	195	развитие
36	праздник	116	разный	196	час
37	о	117	более	197	важно
38	к	118	играть	198	известно
39	так	119	вид	199	наш
40	можно	120	глаза	200	прошлый
41	весь	121	же	201	среди
42	этот	122	сказать	202	торт

43	только	123	просто	203	баскетбол
44	весна	124	далее	204	благодаря
45	любить	125	конечно	205	два
46	являться	126	русский	206	занятие
47	путешествие	127	находиться	207	интересно
48	рождения	128	писатель	208	отдохнуть
49	за	129	хотя	209	польза
50	от	130	возможность	210	при
51	всегда	131	заниматься	211	сеть
52	больше	132	сторона	212	трудно
53	потому	133	лучше	213	однако
54	она	134	нет	214	провинция
55	они	135	приносить	215	утро
56	чтобы	136	страна	216	известен
57	поэтому	137	учиться	217	известный
58	Пекин	138	музей	218	общение
59	или	139	именно	219	тогда
60	вместе	140	истории	220	учитель
61	здесь	141	студенческий	221	хороший
62	любимый	142	через	222	дети
63	часто	143	информация	223	интересный
64	тоже	144	общаться	224	люди
65	уже	145	чтение	225	незабываемый
66	книга	146	делать	226	первый
67	после	147	жить	227	тысяча
68	кроме	148	обязательно	228	удовольствие
69	мир	149	родитель	229	часть
70	нравиться	150	хобби	230	вещи
71	если	151	говорить	231	дарить
72	свой	152	помочь	232	знакомиться
73	ещё	153	никогда	233	пора
74	друг	154	отдых	234	сердце
75	до	155	свободное	235	сначала
76	сейчас	156	язык	236	становиться
77	Китай	157	важный	237	быстро
78	также	158	Чжэнчжоу	238	вот
79	общепитие	159	впечатление	239	красивый
80	без	160	детство	240	Лоян

ПРИЛОЖЕНИЕ 7

СПИСОК ЧАСТОТНЫХ ОШИБОК В ТЕКСТАХ, НАПИСАННЫХ НА РАЗНЫЕ ТЕМЫ

Список частотных ошибок в текстах на темы 1 и 2

Ранг	Тема 1: Мое интересное путешествие	Тема 2: Мой любимый город
1	орфографическая ошибка за счет незнания орфограмм (fm1)	орфографическая ошибка за счет незнания орфограмм (fm1)
2	неверное использование прописной буквы (fm3)	неверное использование прописной буквы (fm3)
3	неверное управление глагола (vp3)	неверное использование знака препинания (pt3)
4	неверное использование знака препинания (pt3)	пропуск членов предложения (sn1)
5	неверное образование форм склонения имен существительных (np3)	неверное образование форм склонения имен существительных (np3)
6	неверное использование форм несовершенного или совершенного вида глагола (vp2)	неверный выбор местоимения (pr1)
7	неверный выбор местоимения (pr1)	неверное управление глагола (vp3)
8	неверное использование полной или краткой формы имен прилагательных (aj1)	неверное использование форм несовершенного или совершенного вида глагола (vp2)
9	неверное использование устойчивых выражений (cc1)	неверное использование устойчивых выражений (cc1)
10	смещение синонимов и паронимов (wd3)	использование неверного предлога (pp2)
11	нарушение падежных форм имен существительных (np4)	смещение синонимов и паронимов (wd3)
12	пропуск членов предложения (sn1)	нарушение согласования подлежащего и сказуемого (sn3)
13	неверное использование форм числа имен существительных (np1)	неверное управление предлога (pp1)
14	нарушение согласования подлежащего и сказуемого (sn3)	нарушение падежных форм имен существительных (np4)
15	нарушение видовременной соотнесенности в главном и придаточном предложениях (sn4)	лишний знак препинания (pt2)
16	отсутствие знака препинания (pt1)	нарушение логики при перечислении (lc1)
17	использование неверного предлога (pp2)	неверное использование форм числительных (nu1)

18	лишний знак препинания (pt2)	неверное использование полной или краткой формы имен прилагательных (aj1)
19	неверное управление предлога (pp1)	употребление слова в несвойственном ему значении (wd1)
20	употребление слова в несвойственном ему значении (wd1)	отсутствие знака препинания (pt1)
21	неверный порядок слов (sn2)	нарушение видовойременной соотнесенности в главном и придаточном предложениях (sn4)
22	неверное словообразование (fm2)	тавтология и повтор (wd2)
23	неверное использование форм долженствования (ad2)	неверный порядок слов (sn2)
24	нарушение логики при перечислении (lc1)	неверное использование форм числа имен существительных (np1)
25	неверное использование степеней сравнения наречий(ad1)	неверный выбор вида причастия глагола (vp5)
26	неверное использование возвратной формы глагола (vp1)	неверное словообразование (fm2)
27	тавтология и повтор (wd2)	неверное использование форм долженствования (ad2)
28	отсутствие формальных показателей связи между предложениями (lc2)	неверное использование степеней сравнения наречий (ad1)
29	неверное использование форм числительных (nu1)	отсутствие формальных показателей связи между предложениями (lc2)
30	отсутствие согласования по роду имен существительных (np2)	отсутствие согласования по роду имен существительных (np2)
31	неверное использование степеней сравнения прилагательных (aj2)	фразеологическая ошибка (cc2)
32	неверный выбор вида причастия глагола (vp5)	неверное использование возвратной формы глагола (vp1)
33	фразеологическая ошибка (cc2)	неверное использование степеней сравнения прилагательных (aj2)
34	неверное использование деепричастия глагола (vp4)	неверное использование деепричастия глагола (vp4)

Список частотных ошибок в текстах на темы 3 и 4

Ранг	Тема 3: Интернет в нашей жизни	Тема 4: Мои увлечения
1	орфографическая ошибка за счет незнания орфограмм (fm1)	орфографическая ошибка за счет незнания орфограмм (fm1)
2	неверное использование знака препинания (pt3)	неверное использование знака препинания (pt3)
3	нарушение согласования подлежащего и сказуемого (sn3)	неверное управление глагола (vp3)
4	неверное использование прописной буквы (fm3)	неверное использование прописной буквы (fm3)
5	неверный выбор местоимения (pr1)	неверный выбор местоимения (pr1)
6	пропуск членов предложения (sn1)	неверное использование форм несовершенного или совершенного вида глагола (vp2)
7	неверное образование форм склонения имен существительных (np3)	нарушение согласования подлежащего и сказуемого (sn3)
8	неверное управление глагола (vp3)	неверное образование форм склонения имен существительных (np3)
9	смещение синонимов и паронимов (wd3)	лишний знак препинания (pt2)
10	неверное использование форм несовершенного или совершенного вида глагола (vp2)	использование неверного предлога (pp2)
11	лишний знак препинания (pt2)	пропуск членов предложения (sn1)
12	неверное использование полной или краткой формы имен прилагательных (aj1)	смещение синонимов и паронимов (wd3)
13	неверное использование устойчивых выражений (cc1)	неверное использование полной или краткой формы имен прилагательных (aj1)
14	употребление слова в несвойственном ему значении (wd1)	неверное использование устойчивых выражений (cc1)
15	отсутствие формальных показателей связи между предложениями (lc2)	нарушение видовременной соотнесенности в главном и придаточном предложениях (sn4)
16	отсутствие знака препинания (pt1)	употребление слова в несвойственном ему значении (wd1)
17	тавтология и повтор (wd2)	отсутствие знака препинания (pt1)
18	нарушение падежных форм глагола (np4)	нарушение падежных форм глагола (np4)
19	неверное словообразование (fm2)	неверное управление предлога (pp1)
20	неверный порядок слов (sn2)	отсутствие формальных показателей связи между предложениями (lc2)
21	неверное использование форм числа имен существительных (np1)	неверное использование форм числа имен существительных (np1)

22	неверное использование степеней сравнения наречий (ad1)	тавтология и повтор (wd2)
23	неверное управление предлога (pp1)	неверный порядок слов (sn2)
24	неверное использование возвратной формы глагола (vp1)	фразеологическая ошибка (cc2)
25	нарушение логики при перечислении (lc1)	неверное словообразование (fm2)
26	отсутствие согласования по роду имен существительных (np2)	неверное использование возвратной формы глагола (vp1)
27	нарушение видовременной соотнесенности в главном и придаточном предложениях (sn4)	неверное использование степеней сравнения наречий (ad1)
28	неверный выбор вида причастия глагола (vp5)	нарушение логики при перечислении (lc1)
29	фразеологическая ошибка (cc2)	отсутствие согласования по роду имен существительных (np2)
30	неверное использование степеней сравнения прилагательных (aj2)	неверный выбор вида причастия глагола (vp5)
31	использование неверного предлога (pp2)	неверное использование степеней сравнения прилагательных (aj2)
32	неверное использование форм числительных (nu1)	неверное использование форм числительных (nu1)
33	неверное использование форм долженствования (ad2)	неверное использование форм долженствования (ad2)
34	неверное использование деепричастия глагола (vp4)	неверное использование деепричастия глагола (vp4)

Список частотных ошибок в текстах на темы 5 и 6

Ранг	Тема 5: Мое интересное путешествие	Тема 6: Кинга в нашей жизни
1	неверное использование прописной буквы (fm3)	орфографическая ошибка за счет незнания орфограмм (fm1)
2	орфографическая ошибка за счет незнания орфограмм (fm1)	неверное использование знака препинания (pt3)
3	неверное использование знака препинания (pt3)	неверное использование устойчивых выражений (cc1)
4	нарушение согласования подлежащего и сказуемого (sn3)	нарушение согласования подлежащего и сказуемого (sn3)
5	использование неверного предлога (pp2)	неверное управление глагола (vp3)
6	неверное использование устойчивых выражений (cc1)	употребление слова в несвойственном ему значении (wd1)
7	неверное управление глагола (vp3)	неверное управление предлога (pp1)
8	пропуск членов предложения (sn1)	неверный выбор местоимения (pr1)
9	неверное образование форм склонения имен существительных (np3)	неверное использование прописной буквы (fm3)
10	неверное управление предлога (pp1)	использование неверного предлога (pp2)
11	неверный выбор местоимения (pr1)	неверное образование форм склонения имен существительных (np3)
12	нарушение видовойременной соотнесенности в главном и придаточном предложениях (sn4)	отсутствие знака препинания (pt1)
13	употребление слова в несвойственном ему значении (wd1)	пропуск членов предложения (sn1)
14	неверное использование форм несовершенного или совершенного вида глагола (vp2)	неверное использование форм несовершенного или совершенного вида глагола (vp2)
15	смещение синонимов и паронимов (wd3)	нарушение видовойременной соотнесенности в главном и придаточном предложениях (sn4)
16	лишний знак препинания (pt2)	смещение синонимов и паронимов (wd3)
17	неверное использование полной или краткой формы имен прилагательных (aj1)	лишний знак препинания (pt2)
18	неверное использование форм числа имен существительных (np1)	неверное использование полной или краткой формы имен прилагательных (aj1)
19	отсутствие знака препинания (pt1)	неверное использование форм долженствования (ad2)
20	нарушение падежных форм имен существительных (np4)	неверное использование форм числа имен существительных (np1)

21	неверное словообразование (fm2)	нарушение падежных форм имен существительных (np4)
22	неверное использование степеней сравнения наречий (ad1)	неверное использование степеней сравнения наречий (ad1)
23	неверное использование форм числительных (nu1)	нарушение логики при перечислении (lc1)
24	неверное использование форм долженствования (ad2)	тавтология и повтор (wd2)
25	неверный порядок слов (sn2)	неверный выбор вида причастия глагола (vp5)
26	неверное использование возвратной формы глагола (vp1)	неверное использование деепричастия глагола (vp4)
27	нарушение логики при перечислении (lc1)	неверное словообразование (fm2)
28	отсутствие согласования по роду имен существительных (np2)	неверное использование форм числительных (nu1)
29	отсутствие формальных показателей связи между предложениями (lc2)	неверный порядок слов (sn2)
30	тавтология и повтор (wd2)	неверное использование возвратной формы глагола (vp1)
31	фразеологическая ошибка (cc2)	отсутствие согласования по роду имен существительных (np2)
32	неверный выбор вида причастия глагола (vp5)	отсутствие формальных показателей связи между предложениями (lc2)
33	неверное использование степеней сравнения прилагательных (aj2)	фразеологическая ошибка (cc2)
34	неверное использование деепричастия глагола (vp4)	неверное использование степеней сравнения прилагательных (aj2)

Список частотных ошибок в текстах на темы 7 и 8

Ранг	Тема 7: Как я провел(а) свой день рождения	Тема 8: Мои увлечения
1	неверное использование прописной буквы (fm3)	орфографическая ошибка за счет незнания орфограмм (fm1)
2	орфографическая ошибка за счет незнания орфограмм (fm1)	неверное использование знака препинания (pt3)
3	пропуск членов предложения (sn1)	неверное использование прописной буквы (fm3)
4	неверное использование форм несовершенного или совершенного вида глагола (vp2)	пропуск членов предложения (sn1)
5	неверное использование знака препинания (pt3)	смешение синонимов и паронимов (wd3)
6	неверный выбор местоимения (pr1)	неверный выбор местоимения (pr1)
7	неверное использование устойчивых выражений (cc1)	неверное управление глагола (vp3)
8	употребление слова в несвойственном ему значении (wd1)	нарушение видовременной соотнесенности в главном и придаточном предложениях (sn4)
9	использование неверного предлога (pp2)	неверное использование устойчивых выражений (cc1)
10	нарушение согласования подлежащего и сказуемого (sn3)	использование неверного предлога (pp2)
11	неверное управление глагола (vp3)	неверное использование полной или краткой формы имен прилагательных (aj1)
12	отсутствие знака препинания (pt1)	отсутствие согласования по роду имен существительных (np2)
13	неверное использование форм долженствования (ad2)	неверное использование форм несовершенного или совершенного вида глагола (vp2)
14	нарушение видовременной соотнесенности в главном и придаточном предложениях (sn4)	отсутствие формальных показателей связи между предложениями (lc2)
15	смешение синонимов и паронимов (wd3)	употребление слова в несвойственном ему значении (wd1)
16	неверное использование форм числа имен существительных (np1)	отсутствие знака препинания (pt1)
17	нарушение падежных форм имен существительных (np4)	неверное использование форм числа имен существительных (np1)
18	лишний знак препинания (pt2)	нарушение падежных форм имен существительных (np4)
19	неверное использование полной или краткой формы имен прилагательных (aj1)	нарушение логики при перечислении (lc1)

20	нарушение логики при перечислении (lc1)	нарушение согласования подлежащего и сказуемого (sn3)
21	неверное образование форм склонения имен существительных (np3)	неверное использование форм долженствования (ad2)
22	неверный порядок слов (sn2)	неверное управление предлога (pp1)
23	неверное управление предлога (pp1)	неверное использование степеней сравнения наречий (ad1)
24	тавтология и повтор (wd2)	неверное словообразование (fm2)
25	неверное использование форм числительных (nu1)	неверное образование форм склонения имен существительных (np3)
26	отсутствие формальных показателей связи между предложениями (lc2)	неверный порядок слов (sn2)
27	неверное использование степеней сравнения наречий (ad1)	тавтология и повтор (wd2)
28	неверный выбор вида причастия глагола (vp5)	лишний знак препинания (pt2)
29	неверное использование деепричастия глагола (vp4)	неверное использование форм числительных (nu1)
30	неверное словообразование (fm2)	неверный выбор вида причастия глагола (vp5)
31	неверное использование возвратной формы глагола (vp1)	неверное использование деепричастия глагола (vp4)
32	отсутствие согласования по роду имен существительных (np2)	неверное использование возвратной формы глагола (vp1)
33	фразеологическая ошибка (cc2)	фразеологическая ошибка (cc2)
34	неверное использование степеней сравнения прилагательных (aj2)	неверное использование степеней сравнения прилагательных (aj2)

Список частотных ошибок в текстах на темы 9 и 10

Ранг	Тема 9: Как мы провели праздник Весны	Тема 10: Самый любимый человек
1	неверное использование прописной буквы (fm3)	орфографическая ошибка за счет незнания орфограмм (fm1)
2	орфографическая ошибка за счет незнания орфограмм (fm1)	неверное использование знака препинания (pt3)
3	неверное использование устойчивых выражений (cc1)	неверное образование форм склонения имен существительных (np3)
4	неверное использование знака препинания (pt3)	неверное использование прописной буквы (fm3)
5	неверное управление глагола (vp3)	неверный выбор местоимения (pr1)
6	смещение синонимов и паронимов (wd3)	неверное управление глагола (vp3)
7	нарушение согласования подлежащего и сказуемого (sn3)	неверное использование полной или краткой формы имен прилагательных (aj1)
8	употребление слова в несвойственном ему значении (wd1)	употребление слова в несвойственном ему значении (wd1)
9	нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях (sn4)	неверное использование форм числа имен существительных (np1)
10	пропуск членов предложения (sn1)	неверное использование устойчивых выражений (cc1)
11	неверное управление предлога (pp1)	нарушение согласования подлежащего и сказуемого (sn3)
12	использование неверного предлога (pp2)	отсутствие знака препинания (pt1)
13	неверное образование форм склонения имен существительных (np3)	неверное словообразование (fm2)
14	неверное использование форм несовершенного или совершенного вида глагола (vp2)	нарушение видовременной соотнесенности однородных сказуемых или сказуемых в главном и придаточном предложениях (sn4)
15	нарушение падежных форм имен существительных (np4)	неверное использование форм долженствования (ad2)
16	неверный выбор местоимения (pr1)	пропуск членов предложения (sn1)
17	неверное словообразование (fm2)	лишний знак препинания (pt2)
18	неверное использование полной или краткой формы имен прилагательных (aj1)	неверное использование форм несовершенного или совершенного вида глагола (vp2)
19	неверное использование форм долженствования (ad2)	неверное управление предлога (pp1)

20	неверное использование возвратной формы глагола (vp1)	неверное использование возвратной формы глагола (vp1)
21	нарушение логики при перечислении (lc1)	нарушение падежных форм имен существительных (np4)
22	лишний знак препинания (pt2)	использование неверного предлога (pp2)
23	неверное использование форм числительных (nu1)	нарушение логики при перечислении (lc1)
24	неверное использование степеней сравнения наречий (ad1)	неверное использование степеней сравнения наречий (ad1)
25	тавтология и повтор (wd2)	смешение синонимов и паронимов (wd3)
26	неверный выбор вида причастия глагола (vp5)	неверный порядок слов (sn2)
27	отсутствие согласования по роду имен существительных (np2)	неверное использование форм числительных (nu1)
28	отсутствие формальных показателей связи между предложениями (lc2)	тавтология и повтор (wd2)
29	неверное использование форм числа имен существительных (np1)	отсутствие формальных показателей связи между предложениями (lc2)
30	неверный порядок слов (sn2)	неверное использование деепричастия глагола (vp4)
31	неверное использование деепричастия глагола (vp4)	неверный выбор вида причастия глагола (vp5)
32	отсутствие знака препинания (pt1)	отсутствие согласования по роду имен существительных (np2)
33	фразеологическая ошибка (cc2)	неверное использование степеней сравнения прилагательных (aj2)
34	неверное использование степеней сравнения прилагательных (aj2)	фразеологическая ошибка (cc2)