

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ТЮМЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

На правах рукописи

Башмакова Анастасия Юрьевна

СЕМАНТИЗАЦИЯ ИСТОРИКО-КУЛЬТУРНОГО КОМПОНЕНТА
В УЧЕБНОМ МУЛЬТИМОДАЛЬНОМ СЛОВАРЕ
(на материале русского / английского языков)

Специальность 10.02.20 — Сравнительно-историческое
типологическое и сопоставительное языкознание

Диссертация на соискание учёной степени
кандидата филологических наук

Научный руководитель:
доктор филологических наук,
доцент
Дрожащих Наталия Владимировна

Тюмень 2022

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	4
ГЛАВА 1. МУЛЬТИМОДАЛЬНОСТЬ КАК СОВРЕМЕННЫЙ ПРИНЦИП ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ В СЛОВАРЕ	19
1.1 Компьютерный словарь в современной лексикографии	19
1.2 Типология компьютерных лексикографических ресурсов	24
1.3 Эволюция компьютерных лексикографических ресурсов	28
1.4 Мультимодальный словарь как словарь для нового пользователя	35
Выводы по Главе 1	50
ГЛАВА 2. МУЛЬТИМОДАЛЬНЫЙ СЛОВАРЬ В ЛИНГВОДИДАКТИКЕ	51
2.1 Компьютерная лингводидактика	51
2.2 Учебный словарь для изучения иностранного языка	53
2.3 Создание и разработка учебного компьютерного словаря	58
Выводы по Главе 2	64
ГЛАВА 3. МУЛЬТИМОДАЛЬНЫЙ СЛОВАРЬ ИСТОРИКО- КУЛЬТУРОЛОГИЧЕСКОГО ТИПА КАК НОВЫЙ ТИП УЧЕБНОГО СЛОВАРЯ	67
3.1 Лингвокультурная лексикография в современном сопоставительном языкознании	67
3.2 Историко-культурный (диахронический) компонент лексики: история вопроса	71
3.3 Способы семантизации лексики в учебном словаре	81
3.4 Моделирование компьютерного учебного мультимодального словаря историко-культурологического типа	89
Выводы по Главе 3	97
ГЛАВА 4. РАЗРАБОТКА МОДЕЛИ УЧЕБНОГО МУЛЬТИМОДАЛЬНОГО СЛОВАРЯ ИСТОРИКО-КУЛЬТУРОЛОГИЧЕСКОГО ТИПА	100
4.1 Формирование концепции словаря	100
4.2 Тематическое поле «Образование / Education» в русском и английском языках: создание базы данных словаря	104
4.2.1 Особенности процедуры и этапы создания корпуса и анализа данных.	105

4.2.2 Составление русскоязычного корпуса текстов по теме «Образование»	113
4.2.3 Составление англоязычного корпуса текстов по теме «Education»	118
4.2.4 Этимологический и семантико-дефиниционный анализ тематического поля «Образование» в русском языке.....	123
4.2.5 Этимологический и семантико-дефиниционный анализ тематического поля «Education» в английском языке	129
4.2.6 Сопоставительный анализ тематических полей «Образование / Education»: сходства и различия	144
4.3 Проектирование компьютерного приложения модели учебного мультимодального словаря историко-культурологического типа.....	150
Выводы по Главе 4	159
ЗАКЛЮЧЕНИЕ	164
БИБЛИОГРАФИЧЕСКИЙ СПИСОК	169

ВВЕДЕНИЕ

Настоящее диссертационное исследование посвящено изучению особенностей и принципов семантизации историко-культурного компонента значения лексемы в учебном мультимодальном словаре, языковым материалом стал корпус новостных текстов про образование на русском и английском языках. **Проблематика** работы лежит на стыке компьютерной и учебной лексикографии, лингводидактики, лингвокультурологии, контрастивной лингвистики. Диссертация соответствует паспорту специальности *10.02.20 Сравнительно-историческое, типологическое и сопоставительное языкознание*, так как включает следующие области исследований:

- сравнение и сопоставление языков в диахронии и синхронии;
- хронологическая и пространственная локализация языковых явлений и состояний;
- сравнительная лексикография и терминография;
- сопоставительная семантика;
- дидактические и лингводидактические аспекты сравнительно-исторического, типологического, сопоставительного языкознания, переводоведения.

В современном мире происходит активное интегрирование информационных технологий и лингвистического знания (А.А. Бонч-Осмоловская, Д. Журавски, А.А. Кибрик, Н.В. Лукашевич, Р. Митков, В.П. Селегей). Данный процесс повлиял на появление совершенно новых стилей взаимодействия человека с компьютером, в которых естественный язык играет важную роль. Сформировалось новое поколение пользователей – digital natives [Prensky, 2001], людей, которые привыкли получать информацию через цифровые каналы. В результате это привело не только к появлению нового феномена – *электронного (цифрового) текста*, который представлен как в традиционной линейной форме, так и в формах, предназначенных для интерактивного взаимодействия и поиска (гипертекст, корпуса, компьютерные словари), но и к трансформации подходов в преподавании

иностранного языка и инструментов его изучения, появились различные компьютерные приложения, программы и словари, доступные как локально на компьютере пользователя, так и в сети Интернет. Кроме того, большинство гуманитарных исследований субъективны и тяжело поддаются верификации [Ильина, 2015; Шаповал, 2016], данный факт также стал одной из причин смены научной парадигмы и активного использования компьютерных методов для решения лингвистических и гуманитарных задач.

Лингводидактика, как лингвистическое исследовательское направление, ставит перед собой задачу поиска и создания новых более эффективных учебных ресурсов, которые будут способствовать успешному изучению иностранного языка и культуры. Лексический состав языка, как его самая обширная подсистема, в свою очередь, является репозиторием таких знаний о культуре и истории народа его носителя. В связи с этим внимание ученых стала привлекать **проблема семантизации слова как достояния определенной лингвокультуры**. Современные компьютерные лексикографические технологии предоставляют возможность совмещения различных мультимодальных средств в учебных целях. Опираясь на историю развития этих наук можно предположить, что учебный мультимодальный словарь, созданный на основе тематически отобранных данных из лингвистических корпусов, исторических словарей и тезаурусов, сможет стать интерактивным образовательным ресурсом, способным зафиксировать и визуализировать диахронические изменения в культуре.

Степень разработанности проблемы. Сегодня к поиску историко-культурного следа в текстах и лексической семантике, а также созданию и изучению корпусов, компьютерных словарей и приложений обращаются представители самых разных научных направлений – философии, социологии, информатики, корпусной лингвистики, культурологии, лингвокультурологии, культуромики, лингводидактики, цифровой гуманитаристики (М. Анандараджан, О.К. Ансимова, В.И. Беликов, М. Биесага, М.А. Большакова, А.А. Бонч-Осмоловская, Л. Борин, Е.Е. Голубкова, А.А. Зализняк, О.Н. Иванищева, Г. Кресс, А.Б. Кутузов, Н.В. Лукашевич, Н.А. Лукьянова, В.В. Миронов, Ж. Мишель,

Т. Нолан, С. Хилл, Э. Эрец). Тексты в цифровом формате могут служить источником автоматически извлекаемых сведений как о языке, так и о культуре. В связи с этим сформировалось новое научное направление – *культуромика*, которое позволяет с помощью статистического анализа употребления слов и словосочетаний в огромных текстовых массивах исследовать различные культурные и социальные процессы в языке.

В связи с тем, что термин *семантизация* используется не только по отношению к процессу знакомства с новой лексикой во время занятий по иностранному языку, но и для обозначения особенностей описания значения слова в словарной статье, изучением описанной выше проблемы занимаются как специалисты в области лингводидактики, так и лексикографы. Значимый вклад в развитие этого направления внесли работы отечественных и зарубежных исследователей, которые изучали *особенности семантизации лексических единиц*: в рамках *психолингвистического эксперимента* на примере тематически связанных слов [Трофимова, 2005]; с точки зрения *развития когнитивных характеристик межкультурной компетенции* [Соболева, Обдалова, 2015]; в аспекте *обучения чтению оригинальных иноязычных текстов* [Дакукина, 2015]; в вопросе *осмысления и изучения нового иноязычного материала* [Мальцев, 2018]. Ученые также предлагают различные подходы и способы семантизации иноязычной лексики: путем *анализа морфемных составляющих и словообразовательных элементов* [Атрашевская, 2012]; путем *анализа содержательной и формальной стороны понятия* (содержание, метаязык, типы толкований) [Левенталь, 2014]. С лексикографической точки зрения особенности семантизации рассматривались в контексте *развития речевой лексической компетенции в ситуациях иноязычного профессионального общения* на примере учебного терминологического словаря [Серова, Чайникова, 2015]. В терминологическом аспекте также представлены исследования семантизации: на материале *терминологии лингводидактики* [Ярославская, Астремская, 2014]; на материале *терминологии высшего образования в русском языке* [Айсакова, Филипская, 2018]; на примере *юридического дискурса* [Сысоев, Завьялов, 2018; Чикнаверова, 2019].

Вопросы *мультимодальности* первоначально были описаны в рамках теории социальной семиотики как взаимодействие различных вербальных и невербальных кодов и знаков [Maybury, 1999; Kress, 2000, 2004, 2010; Bezemer, Kress, 2015], в дальнейшем исследования данной проблемы стали развиваться в направлении лингвистики [Кибрик, 2010], теории дискурса и педагогики [Хутыз, 2016; Сорокина, 2017], а также лексикографии [Иванова, 2007; Lew, 2010; Sato, 2016; Гулевская, Мишланова, 2017]. Развитие теоретических основ и положений компьютерной и мультимодальной лексикографии оказало влияние на разработку словарей различной направленности: лингвокультурологических, идеографических, терминологических, переводных, иллюстративных, сопоставительных и других [Сивакова, 2004; Кантышева, 2011, 2012; Шафтельская, 2013; Волошина, 2015; Богачёва, Ольховская, Парамонова, 2017; Помаролли, 2018; Хомутова, Мамонтова, 2018; Сафиуллина, 2018, 2021].

Однако, несмотря на большое количество работ, ряд вопросов еще не получил должного освещения, в частности: *возможность использования мультимодальных средств представления информации для семантизации этимологических и лингвокультурологических особенностей лексики*, а также *необходимость создания компьютерного учебного лексикографического ресурса для сохранения культурного наследия, визуализирующего изменения в мышлении и культурном сознании носителей языка*. Таким образом, изучение мультимодальных инструментов семантизации на материале диахронических лингвистических корпусов, исторических словарей и тезаурусов, а также последующая разработка *модели компьютерного учебного мультимодального словаря историко-культурологического типа* для изучающих иностранный язык на примере тематического поля «Образование / Education» будет способствовать более глубокому осмыслению языковых и культурных изменений, а также выступит попыткой создания современного учебного словаря нового типа.

Актуальность темы диссертации определяется:

– стабильным вниманием научного сообщества к проблематике взаимосвязи языка, истории и культуры страны;

- сменой научных парадигм и ориентацией современной лингвистики на цифровизацию методов изучения языка и интерпретации текста;
- приложением методов компьютерной лингвистики к анализу семантических изменений в диахронии.

Кроме того, в условиях глобализации образования и взаимодействия образовательных систем разных стран увеличилось количество иностранных студентов как в российских университетах, так и русскоговорящих учащихся на англоязычных программах за рубежом. В связи с этим особую актуальность приобретает изучение различий в лексике сферы образования в русском и английском языках, обусловленных экстралингвистическими причинами. Такое исследование может помочь выявить историко-культурное своеобразие лексического состава языка.

В качестве **гипотезы** исследования выступает предположение о том, что принципы и технологии моделирования в современной мультимодальной лексикографии могут позволить создать учебный компьютерный словарь нового типа, способный визуализировать диахронические изменения в языке и семантизировать историко-культурные особенности лексики типологически разных языков.

Объект исследования – моделирование компьютерного учебного мультимодального словаря историко-культурологического типа для изучающих иностранные языки.

Предмет исследования – способы семантизации лексемы и ее историко-культурных изменений в тематическом поле «Образование / Education» в русском и английском языках.

Цель настоящего исследования заключается в выявлении принципов и особенностей семантизации историко-культурного компонента лексемы посредством разработки модели компьютерного учебного мультимодального словаря. Достижение поставленной цели предполагает решение следующих **задач**:

1. описать основные направления исследований семантизации в области учебной, компьютерной и мультимодальной лексикографии на примере существующих технологических решений;
2. составить список характеристик компьютерного учебного мультимодального словаря и способов семантизации историко-культурного компонента значения в словарной статье;
3. разработать модель компьютерного учебного мультимодального словаря и подобрать инструменты для ее реализации;
4. выявить состав тематического поля «Образование / Education» в русском и английском языках, описать его семантические характеристики с целью выявления историко-культурных изменений;
5. представить пилотную модель словаря для визуализации результатов исследования тематического поля «Образование / Education», описать этапы и технологию его разработки.

Методологической и теоретической основой исследования послужили работы российских и зарубежных исследователей:

- *по семасиологии, лексической семантике и когнитивной лингвистике:*
 К. Рейзиг [1839], А.А. Потебня [1864], М. Бреаль [1897], Р. Мерингер [1904],
 Г. Шухардт [1950], А.И Смирницкий [1952], М.М. Покровский [1959],
 О.С. Ахманова [1957], В.А. Звегинцев [1960], Г.П Пауль [1960],
 К.А. Левковская [1962], А.А. Уфимцева [1962, 1977], А. Марти [1965],
 И.В. Арнольд [1966], С. Ульман [1970], Р.А. Будагов [1974],
 Е.Г. Беляевская [1987], Х. Джексон [1988], Ч. Филлмор [1992],
 Дж. Пустейовский [1998], В.В. Виноградов [1999], Л. Кэмпбелл [2004],
 О.Б. Пономарева [2011], В.Д. Табанакова [2015], Д.Н. Ильин [2017],
 Н.Н. Белозерова, С.Н. Антонова [2017], Т.Н. Хомутова [2019],
 Н.Н. Бочегова [2019], С.А. Питина [2019], С.А. Песина [2021];
- *в области диахронической лингвистики:* А.А. Зализняк [2001],
 М. Григиэль [2005], Н.Н. Лыкова [2005, 2015], Н.В. Дрожащих [2012, 2018],

Г.В. Дуринова [2015], В.Л. Ибрагимова [2018], В.В. Фомин [2019], Ю.С. Родина [2019], А.Б. Кутузов [2018, 2020];

– *в области общей и учебной лексикографии*: Л.В. Щерба [1974], Р.А. Амслер [1980], В.В. Морковкин [1985], Д. Саммерс [1988], Й. Уилкс [1996], А.П. Кови [1999]; В.Д. Табанакова [2001];

– *в области компьютерной лексикографии, корпусной лингвистики и культурологии*: В.П. Селегей [2001], Р. Митков [2004], Н.Н. Белозерова [2004], Э. Эрец [2007], Е.Ю. Ваулина [2007], А.В. Русакова [2008], О.С. Рублева [2010], Р. Лью [2010], Ж. Мишель [2011], Н.В. Лукашевич [2011, 2016], М.В. Авдеева [2012], В.В. Миронов [2012], А.А. Зализняк [2013], Л.Ю. Щипицина [2013], О.В. Григоренко [2015], А.В. Палкова [2015], Е.В. Михалькова [2015], А.А. Бонч-Осмоловская [2015, 2018], В.И. Беликов [2016], Л.Н. Беяева [2016];

– *в области лингвокультурологии, лингвострановедения и лингвокультурной лексикографии*: Б.В. Беяев [1965], Н.М. Шанский [1969], Е.М. Верещагин, В.Г. Костомаров [1971], В.П. Берков [1975], В.Г. Гак [1988], О.Н. Иванищева [2005], О.Л. Петрушова [2009], О.К. Ансимова [2012, 2017], Н.Н. Бочегова [2018], Н.В. Халина, А.В. Жеребненко [2018], О.А. Бондарь, Н.А. Пластинина, Э.М. Рянская [2020], О.Л. Крамаренко, О.Ю. Богданова, Ю.В. Мошкина [2020, 2021];

– *в области общей и компьютерной лингводидактики*: К.Р. Пиотровская [1991], М. Леви [1997], М.А. Бовтенко [2000], В.М. Шаклеин [2008], А.Н. Шамов [2008], Р.Р. Зарипова [2014], А.М. Юсупова [2017], Е.О. Кузьмина [2019], Э.И. Клейман, Е.Ю. Мамонова [2019], С.К. Гураль, Э.П. Комарова, С.А. Бакленева, А.С. Фетисов [2020], Л.Ю. Щипицина [2020], Л.О. Бутакова [2021];

– *по теории мультимодальности, поликодовости и визуализации*: М. Мэйбери [1999], Г. Кресс [2010], А.А. Кибрик [2010], Р. Лью [2010], И.П. Хутыз [2016], Ю.В. Сорокина [2017], Т.Н. Хомутова [2018], Е.В. Рябцев,

Н.Н. Бочегова [2018], Н.В. Халина, Л.М. Комиссарова, Т.Н. Внучкова [2018], С.А. Песина [2020], С.А. Питина, А.Д. Щербов [2020];

– по сопоставительному исследованию: В.Г. Гак [1988], Н.В. Шафтельская [2013], И.И. Ярославская, Е.В. Астремская [2014], Д. Помаролли [2018], В.Л. Ибрагимова [2019], И.П. Конопелько [2019], С.О. Шереметьева, О.И. Бабина [2020];

– по лингвистическому и лексикографическому моделированию: В.В. Дубичинский [1998], В.Д. Табанакова [2001], М.А. Ковязина [2006], А.А. Прошина [2008], А.В. Русакова [2008], В.Д. Табанакова [2009], Н.А. Агапова, Н.Ф. Картофелева [2014], А.Ю. Левенкова [2016], О.Ю. Савина [2017], Е.И. Сьянова [2020], Е.Ю. Мамонова, А.Т. Фатуллаева [2020].

В работе были использованы следующие **методы**:

1. *теоретико-аналитический метод*, включающий обзор и анализ научной литературы;

2. *методы компьютерной и квантитативной лингвистики* (веб-скрейпинг, контент-анализ, частотный анализ, векторизация текста, машинное обучение, нейронные сети), направленные на автоматизированный сбор корпуса новостных статей, компьютеризированные предобработку текста и выявление ключевых слов;

3. *метод семантико-дефиниционного анализа*, применяемый для анализа лексем и выявления расхождений на семантическом уровне;

4. *метод этимологического анализа*, способствующий объяснению исходных словообразовательных связей;

5. *метод сопоставительного анализа*, позволивший выявить общие и дифференциальные признаки лексем русского и английского языка;

6. *описательно-аналитический метод*, позволивший обобщить, описать и представить полученные результаты;

7. *метод моделирования в компьютерной лингвистике*, заключающийся в разработке структуры компьютерного словаря и программировании интерактивного веб-приложения для его оболочки.

Материалом исследования является корпус новостных статей в тематике «Образование / Education», собранных с сайтов образовательного онлайн-портала «EDU-Inform» (русский язык) и журнала «Education Today Magazine» (английский язык). Корпус разделен на две части: *русскоязычная часть* насчитывает 1119 новостных статей (779793 словоупотребления); *англоязычная часть* состоит из 534 новостных статей (285320 словоупотреблений). При сборе обеих частей корпуса был использован весь опубликованный текстовый материал, поэтому для сохранения однородности данных, стиля и формата публикаций было принято решение не добавлять и не использовать дополнительно другие англоязычные источники с целью создания эквивалентного объема статистических показателей. Таким образом, различия в объеме русскоязычной и англоязычной частей корпуса рассматриваются как индивидуальные характеристики и особенности источников.

В качестве дополнительного систематизирующего материала использовались *толковые, учебные, этимологические, ассоциативные, идеографические и энциклопедические словари, интерактивные тезаурусы и корпуса* для русского и английского языков: Национальный корпус русского языка, Онлайн-тезаурус «Карта слов», Тезаурус RuWordNet, WordNet, Google Books Ngram Viewer, The Visual Thesaurus, Fine Dictionary, Oxford English Dictionary, Linguazza.com и др. На данный момент модель словаря включает 10 русских и 10 английских лексем из тематического поля «Образование / Education» в качестве наглядных иллюстрационных примеров.

Положения, выносимые на защиту:

1. *Компьютерный учебный мультимодальный словарь историко-культурологического типа* – это словарь для нового типа пользователя, который представляет собой современный многофункциональный образовательный ресурс. Его преимуществом является *новый формат семантизации* слов с использованием *мультимодальных ресурсов*, который позволяет визуально, более репрезентативно по сравнению с традиционными словарями, организовать культурологические и социально-исторические данные о лексеме, а также показать ее концептуальные, контекстуальные и экстралингвистические связи в языке.

2. *Историко-культурный (диахронический) компонент* значения лексемы в мультимодальном словаре является совокупностью информации об истории семантического развития слова, национально-культурных коннотациях, контексте и особенностях его употребления в определенную историческую эпоху, которая описывает слово как достояние определенной лингвокультуры.

3. Для *семантизации историко-культурного компонента* в словарной статье компьютерного учебного мультимодального словаря используются следующие элементы: *описательное вербальное толкование слова на иностранном языке; краткая этимологическая справка; экстралингвистический комментарий; мультимодальные ресурсы; наглядная визуализация языкового окружения лексемы; гипертекстовый доступ.*

4. Базовыми компонентами модели учебного мультимодального словаря историко-культурологического типа являются его макроструктура (*вводная часть, облака слов, словник*) и микроструктура, элементами которой являются *заголовочное слово; грамматическая помета; фонетическая транскрипция; аудиопрезентация слова; современные дефиниции слова с примерами употребления; этимологическая справка; историко-культурный комментарий; информация о частотности употреблений лексемы в диахронии; примеры употребления слова в корпусе; иллюстрации; видео-контент; разбор слова по составу (для русского языка); информация об уровне владения языком по системе CEFR для каждого значения лексемы (для английского языка); визуализация хронологии развития семантических изменений в значениях слов в формате «timeline» (для английского языка).*

5. Ключевая лексика *тематического поля «Образование / Education»* на материале текстов образовательных СМИ на русском и английском языках имеет отличительные характеристики, обусловленные *историческими* (ключевая лексика английского языка входит в употребление раньше, чем лексемы русского языка, что объясняет ее многозначность и наличие слов-когнатов), *социально-политическими* (прослеживается значительное влияние государства на стиль публикаций на русском языке, в обоих языках доминирует тематика пандемии

коронавируса) и *национально-культурными* (тематика благотворительности в английском языке, вопросы образовательного менеджмента в обоих языках) особенностями выбранных языков.

Научная новизна исследования определяется:

- комплексным описанием принципов и приемов беспереводной семантизации историко-культурного компонента лексемы в словарной статье с использованием мультимодальных инструментов представления информации;
- обоснованием необходимости создания учебного словаря нового типа, способного интегрировать в себе описание лексической единицы как динамичного образования, отражающего лингвокультурологические процессы в диахронии и синхронии;
- изучением и сопоставительным анализом диахронических семантических изменений ключевых слов образовательного дискурса русского и английского языков цифровыми методами лингвистического исследования;
- попыткой представить и визуализировать формализованную модель знаний о лексике тематического поля «Образование / Education» в русском и английском языках;
- построением модели компьютерного учебного мультимодального словаря историко-культурологического типа, основанной на технологиях компьютерной лексикографии (написание программного кода на языках программирования Python 3 и R).

Теоретическая значимость исследования обусловлена:

- уточнением и систематизацией принципов беспереводной семантизации историко-культурных изменений семантики лексемы в словарной статье в рамках теории мультимодальности;
- изложением методики компьютеризированного выявления ключевой лексики текстов образовательного дискурса и алгоритма ее визуализации;
- изучением и сопоставлением особенностей семантики, этимологии и частотных характеристик ключевых слов тематического поля

«Образование / Education» в русском и английском языках в рамках сравнительно-исторического подхода;

– конкретизацией методики построения словарной статьи за счет использования мультимодальных инструментов для семантизации лексики.

Полученные в ходе исследования выводы вносят вклад в развитие компьютерной, учебной и мультимодальной лексикографии в рамках лингводидактического подхода, а применяемая методика компьютеризированного выявления и изучения ключевой лексики может быть реализована при проведении исследований текстов других жанров и типов дискурса, что позволит расширить общетеоретическую значимость выдвинутых положений.

Практическая значимость работы заключается в том, что результаты проведенного исследования могут быть использованы для анализа особенностей и принципов семантизации, а также характера историко-культурных изменений лексики в сравнительно-историческом и лингвистическом аспектах в процессе преподавания таких дисциплин как основы лексикологии, лингвокультурология, лингводидактика, компьютерная лексикография и др. Разработанная модель компьютерного учебного мультимодального словаря историко-культурологического типа для студентов, изучающих русский и английский языки как иностранные, может быть применена для улучшения существующих компьютерных учебных словарей или создания аналогичных и более совершенных лексикологических ресурсов как для пары русский / английский язык, так и для других языков.

Степень достоверности результатов исследования определяется опорой на авторитетные теоретические положения отечественных и зарубежных ученых, применением методов анализа, соответствующих целям и задачам исследования, привлечением обширного фактического материала, согласованностью теоретических выводов с результатами практического исследования и возможностью их верификации на другом материале.

Апробация результатов исследования проводилась в форме отчетных выступлений на заседаниях кафедры английской филологии и перевода и кафедры английского языка Тюменского государственного университета. Основные положения исследования излагались в виде докладов и сообщений на научных конференциях:

– VII всероссийская молодежная научно-практическая конференция «Множественность интерпретаций: верификация гуманитарного знания», ТюмГУ, Тюмень, 2018 г.;

– Онлайн-конференция 68th Studentische Tagung Sprachwissenschaft (StuTS), Берлин, 2020 г.;

– Всероссийская междисциплинарная молодежная научная конференция с международным участием «IX Информационная школа молодого ученого», ЦНБ УрО РАН, Екатеринбург, 2021 г.;

– III Международная конференция «Синергия языков и культур: междисциплинарные исследования», СПбГУ, Санкт-Петербург, 2021 г.;

– Международная научно-практическая конференция «Язык. Культура. Перевод», МГИМО, Одинцово, 2022 г.;

– доклады на заседаниях лаборатории обработки естественного языка (NLP Lab) в Государственном университете Монтклера (Montclair State University), г. Монтклер, США, во время научно-исследовательской стажировки по программе У. Фулбрайта для выпускников вузов и аспирантов (Fulbright Foreign Student Program), 2021 г.

Основное содержание диссертации отражено в 6 публикациях автора, 4 из которых в ведущих рецензируемых изданиях.

Публикации в ведущих рецензируемых изданиях:

1. Пономарева, Е. Ю. Реализация концепта "international student" в дискурсе британских СМИ в условиях глобализации / Е. Ю. Пономарева, А. Ю. Башмакова // Вестник Тюменского государственного университета. Гуманитарные исследования. *Humanitates*. – 2018. – Т. 4. – № 3. – С. 40-51. – DOI 10.21684/2411-197X-2018-4-3-40-51.

2. Дрожащих, Н. В. Учебный мультимодальный словарь историко-культурологического типа в преподавании иностранного языка / Н. В. Дрожащих, А. Ю. Башмакова // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. 2021. Том 7. № 1 (25). С. 43-61. – DOI: 10.21684/2411-197X-2021-7-1-43-61

3. Башмакова, А. Ю. Выявление ключевых слов тематического поля "Образование/Education" / А. Ю. Башмакова // Филологические науки. Вопросы теории и практики. – 2021. – Т. 14. – № 12. – С. 3986-3990. – DOI 10.30853/phil20210598

4. Башмакова, А. Ю. Классические и современные подходы к изучению семантических изменений / А. Ю. Башмакова // Успехи гуманитарных наук. – 2022. – № 1. – С. 80-85.

Публикации в других изданиях:

5. Башмакова, А. Ю. Культуромика в корпусном исследовании / А. Ю. Башмакова // Множественность интерпретаций: верификация гуманитарного знания : Материалы VII молодежной научно-практической конференции, Тюмень, 19 декабря 2018 года / Научный редактор Е.В. Михалькова. – Тюмень: Тюменский государственный университет, 2019. – С. 17-22.

6. Башмакова, А. Ю. Выявление ключевых слов по теме «Education» (на материале корпуса текстов онлайн-издания «Education Today Magazine») / А. Ю. Башмакова // IX Информационная школа молодого ученого: Сборник научных трудов, Екатеринбург, 20–23 сентября 2021 года / Центральная научная библиотека УрО РАН. – Екатеринбург: ООО "Издательство УМЦ УПИ", 2021. – С. 400-407. – DOI 10.32460/ishmu-2021-9-0044.

Объем и структура диссертации. Диссертация состоит из введения, четырех глав («Мультимодальность как современный принцип представления информации в словаре», «Мультимодальный словарь в лингводидактике», «Мультимодальный словарь историко-культурологического типа как новый тип учебного словаря», «Разработка модели учебного мультимодального словаря историко-культурологического типа») с выводами по каждой главе, заключения

и библиографического списка, включающего в себя 273 источника, из них 92 источника на английском языке. Общий объем работы составляет 197 страниц печатного текста. Работа иллюстрирована таблицами, схемами и рисунками.

Материалы исследовательского корпуса, лингвистическая база данных модели словаря, программный код по автоматизированному сбору новостных статей и компьютеризированному выявлению ключевых слов на языке программирования Python 3 и код интерактивного веб-приложения модели словаря на языке программирования R расположены в репозитории GitHub в сети Интернет по ссылке¹.

¹ Репозиторий GitHub, URL:<https://github.com/nastuk/dissertation>

ГЛАВА 1. МУЛЬТИМОДАЛЬНОСТЬ КАК СОВРЕМЕННЫЙ ПРИНЦИП ПРЕДСТАВЛЕНИЯ ИНФОРМАЦИИ В СЛОВАРЕ

1.1 Компьютерный словарь в современной лексикографии

Нельзя недооценивать всеобъемлющую роль компьютерных технологий в жизни современного человека. Исследования в области искусственного интеллекта дали толчок стремительному развитию компьютерной лингвистики в целом и компьютерной лексикографии в частности. Ранее словари и компьютерные технологии не часто имели тесные междисциплинарные отношения в науке, однако за последние тридцать лет ситуация кардинально изменилась. Одной из первых и главных задач при разработке компьютерных систем стала автоматизация рутинных и энергозатратных процессов, которые ранее выполнялись человеком вручную. Составление словарей, как основное направление лексикографии, не стало исключением. В связи с этим в данном контексте важно понимать, что из методологии классической лексикографии может быть полезно для развития компьютерных технологий и наоборот, что новая цифровизация может принести для понимания естественных языков, таких как английский или русский. Настоящий параграф содержит теоретическое описание направлений исследований в компьютерной лексикографии и особенностей компьютерного словаря.

В начале для понимания задач данного научного направления необходимо обозначить термин *лексикография*, который этимологически объединяет в себе два слова греческого происхождения: «*lexikos*» – относящийся к слову, словарный и «*grapho*» – пишу. Таким образом, **лексикография** – «это научная дисциплина, занимающаяся теорией и практикой составления словарей» [Филиппович, Чернышева, 1999, с. 1], или в более узком значении «теория и практика составления словарей, главным образом языковых, лингвистических, в отличие от неязыковых, энциклопедических» [Авдеева, 2012, с. 26]. В результате развития информационных технологий ученые стали говорить о *компьютерной*

лексикографии, хотя в работах разных исследователей встречается и другая терминология: «машинная», «вычислительная», «электронная», «автоматическая» или «автоматизированная», и даже «кибернетическая» лексикография [Рублёва, 2010, с. 20]. Первое упоминание термина *компьютерная лексикография* (computational lexicography) появляется ещё в 1960-х годах, однако активно развиваться данное направление начало только в 1990-х годах [Wilks, Slator, Guthrie, 1996]. Для настоящего исследования мы выбрали использование термина «*компьютерная лексикография*», так как он выступил наиболее частотным и отражает непосредственно «средство создания электронных словарей и инструмент, позволяющий читателю использовать их возможности» [Палкова, 2015, с. 88].

Интересно, что толкование данного термина учеными различается в зависимости от того, на что делается акцент. С одной стороны, Лариса Юрьевна Щипицина обобщает, что «**компьютерная лексикография** представляет собой раздел прикладной лингвистики, нацеленный на создание компьютерных словарей, лингвистических баз данных и разработку программ поддержки лексикографических работ» [Щипицина, 2013, с. 65]. С другой, более практико-ориентированной, стороны, компьютерная лексикография понимается как «совокупность методов и программных средств обработки текстовой информации для создания словарей» [Авдеева, 2012, с. 28]. Но также встречается и определение компьютерной лексикографии как совокупности «словарей, доступных в электронной форме, независимо от способа их создания» [Ваулина, 2007, с. 189]. Зарубежные исследователи различают два направления компьютерной лексикографии: использование словарей, созданных человеком, для вычислительных целей и, наоборот, использование вычислительной техники для составления новых словарей [Mitkov, 2004, с. 49]. При всём многообразии значений термина, общим и важным звеном во всех определениях является сам словарь в его машиночитаемой форме.

Словарь является основным и наиболее используемым продуктом компьютерной лексикографии. Несмотря на широкую популярность

компьютерных словарей в современном мире, существует ряд синонимичных терминов для его названия: «электронный», «компьютерный», «автоматический», «машинный». В настоящем исследовании мы придерживаемся терминологии «компьютерный» как по отношению к лексикографии, так и к словарям. Одно из ранних и общих определений словарей такого рода дает Лев Львович Нелюбин: «электронный словарь – это любой упорядоченный, относительно конечный массив лингвистической информации, представленный в виде списка, таблицы или перечня, удобного для размещения в памяти ЭВМ и снабженного программами автоматической обработки и пополнения» [Нелюбин, 1983]. Позднее исследователи компании АВВУУ под руководством Владимира Павловича Селегея охарактеризовали **компьютерный словарь** как «особый лексикографический объект, в котором могут быть реализованы и введены в обращение многие продуктивные идеи, не востребованные по разным причинам в бумажных словарях» [Селегей, 2001, с. 1]. Марина Васильевна Авдеева определяет *компьютерный словарь* как «словарь в специальном машинном формате, предназначенный для использования на ЭВМ пользователем или компьютерной программой обработки текста» [Авдеева, 2012, с. 28]. Для целей настоящего исследования мы будем определять **компьютерный словарь** как *лексикографический ресурс в машиночитаемой форме, обладающий лексической базой данных для программ обработки естественного языка и оснащенный интерфейсом для пользователя-человека.*

Вопрос терминологии в этом направлении остается важным, так как учеными выделяются отдельно два разных типа компьютерных словарей в зависимости от целевой аудитории: *автоматические словари конечного пользователя (человека)* и *автоматические словари для программ обработки текста* [Walker, 1995; Авдеева, 2012]. *Словари для людей* могут быть представлены в виде самостоятельных приложений, в том числе онлайн, или встроенных систем в другие приложения (embedded dictionary systems), например дополнение к текстовому процессору для проверки орфографии. Так или иначе такие словари представляют информацию на дисплее компьютера или другого устройства для

непосредственного использования пользователем. Словарная статья может состоять только из текста или быть мультимодальной, включая в себя графические изображения, аудио- и видео файлы и другие интерактивные ресурсы. *Словари для программ* по своей сути являются лексическими базами данных для компьютерных приложений, которые поддерживают пользовательские интерфейсы на естественном языке [Walker, 1995, с. 208-214]. Пользователи не имеют прямого доступа к таким словарям, так как они используются непосредственно программами для решения задач, поставленных через интерфейс самого приложения, например поиск информации, распознавание речи, автоматический перевод, синтаксический анализатор, синтез речи и т. п. Чтобы успешно справляться с поставленными задачами такие словари должны обладать большой базой данных, основанной на многомиллионных лингвистических корпусах.

С развитием компьютерной лексикографии стал актуальным анализ сходств и различий, классификация, сравнение достоинств и недостатков компьютерных и традиционных словарей. Нередко, анализируя различия между традиционными (печатными) и компьютерными словарями, ученые выделяют ряд важных преимуществ последних. Во-первых, *актуальность информации*, так как «составление словарей на базе анализа письменных текстов отнимает много времени и к моменту публикации словаря на печатной основе его данные могут устаревать» [Рублёва, 2010, с. 20]. Во-вторых, *интерфейс пользователя* – «возможности показа содержания словарной статьи по разным критериям (различные «проекции» словаря), разнообразные графические средства, которые не используются в обычных словарях» [Селегей, 2001, с. 2]. Дополнительной особенностью компьютерных словарей является возможность использования *мультимедийных средств* «для представления различного рода информации (аудио-, видео-, радиоинформации, анимации, игр, форумов, заметок, подходящих под каждую категорию пользователей с различными запросами и интересами)» [Березовская, Сухова, 2014, с. 181]. Кроме этого, отмечают *реверсивность* словаря (возможность произвольного, по желанию пользователя, изменения входного и выходного языков), его *динамичность* (возможность постоянного пополнения,

коррекции, изъятия устаревших данных) и *объемность* (возможность включать в электронный словарь любое количество словарных единиц) [Шляхова, 2008, с. 182]. Лексикографы также воспользовались преимуществами, которые дают компьютеры при создании словарей, например с технической стороны составители больше не зависели от алфавитного порядка при наборе словарных статей для сдачи их в печать, что значительно улучшило и ускорило процесс обработки, последовательность экспертизы и редактуры словаря [Mitkov, 2004, с. 57]. Однако эти преимущества в большей степени проявляются по отношению человеку как пользователю или создателю словаря.

С другой стороны, компьютерные словари, предназначенные для человека как конечного пользователя, имеют ряд заметных недостатков при применении их в качестве лексических баз данных для компьютерных программ обработки естественного языка. Серьезной проблемой для компьютерных приложений является то, что такие словари сосредоточены на представлении списков значений для каждой лексемы с описанием тонких семантических отличий и не обладают информацией о том, как одно значение может отличаться от другого в тексте. Другим минусом становится отсутствие статистической информации о сравнительной частоте употребления разных слов и значений. Человек, использующий словарь, обладает фоновыми знаниями о языке и окружающем мире, которые помогают ему декодировать сообщение и дифференцировать одно значение от другого. Компьютерное приложение, как искусственно созданный человеком код, не обладает такими знаниями и умениями. Решение этой проблемы является одной из наиболее перспективных задач в современных исследованиях обработки естественного языка и создании искусственного интеллекта [Mitkov, 2004, с. 60]. Здесь принято говорить об исследованиях с применением нейронных сетей, векторной семантике, корпусной лингвистике, больших данных.

Таким образом, компьютерный словарь играет значительную роль в современной лексикографии и обладает множеством уникальных характеристик и особенностей, историю и типологию которых мы подробно рассмотрим в следующих параграфах.

1.2 Типология компьютерных лексикографических ресурсов

Вопросы типологии и классификации всегда остаются актуальными в любом научном направлении. Компьютерная лексикография не является исключением. Компьютерные словари так же, как и традиционные подвергаются классификации по различным основаниям. Существует множество подходов к типологии словарей, что вызывает сложность создания универсальной классификации из-за разных критериев их оценки. В настоящем параграфе мы опишем типологию компьютерных лексикографических ресурсов в обзоре работ отечественных и зарубежных ученых.

В начале важно отметить, что к инструментам компьютерной лексикографии кроме словарей ученые также относят компьютерные картотеки, лексические базы данных, лингвистические корпуса, встроенные в текстовые процессоры орфографические словари, а также любые другие программы для обработки текста на естественном языке, позволяющие в автоматическом режиме формировать словарные статьи, хранить словарную информацию и обрабатывать ее [Филиппович, Чернышева, 1999, с. 2; Авдеева, 2012, с. 28]. Словари же, в общем виде, варьируются в структурном отношении, в зависимости от мега-, макро- и микроструктуры словаря. Под *мегаструктурой* понимается введение, словник и словарная статья, под *макроструктурой* – структура словника, и под *микроструктурой* – непосредственно структура словарной статьи. Более детально можно классифицировать словари по количеству словарных статей, охвату лексики, количеству языков, назначению, оформлению информации, функциональной и культурологической направленности, порядку подачи лексического материала, источнику информации о слове, отношению к синхронии и диахронии и другим параметрам [Рублёва, 2010, с. 11]. Ряд ученых отмечают, что все лексикографические ресурсы можно поделить на две большие группы: *словари общего назначения* (general-purpose dictionary) и *специализированные словари* (special-purpose dictionary), которые в свою очередь классифицируются более узко [Anderman, Rogers, 1996; Jackson, 1998].

Так, **словари общего назначения** стремятся дать всесторонний охват словарного запаса языка и содержат всю необходимую лексикографическую информацию, которая может потребоваться среднестатистическому пользователю, такую как: значение слова, произношение, ударение, часть речи, этимология, стилистические и диалектные ограничения, область использования [Jackson, 1998, с. 159]. **Специализированные словари**, в свою очередь, либо обладают дополнительной более углубленной языковой информацией, либо ориентированы на особую группу пользователей. Примером специализированных словарей могут стать орфографические, этимологические, орфоэпические словари, словари топонимов, словари личных имен, словари сленга или словари по определенной специальности (медицина, информатика, ботаника) [Jackson, 1998, с. 165]. Также специализированными словарями являются словари для детей, словари для любителей кроссвордов и учебные словари для изучающих иностранный язык (*learner's dictionary*), которые представляют наибольший интерес для настоящего исследования.

Британский учёный, специалист по проблемам искусственного интеллекта Йорик Уилкс считает, что можно выделить по крайней мере пять разновидностей словарей: *традиционный словарь* (*standard dictionary*), *классический тезаурус*, *двуязычные словари*, *стилистические словари* (*style dictionaries*) и *конкордансы* [Wilks, Slator, Guthrie, 1996, с. 63]. Лингвисты, специалисты по переводоведению и лексикографии, Гунилла Андерман и Маргарет Роджерс полагают, что для репрезентации семантических особенностей и лексических связей, играющих важную роль при изучении иностранного языка, наиболее релевантными являются *тезаурусы* и *словари лексической сочетаемости* (*collocational dictionaries*) [Anderman, Rogers, 1996, с. 81]. Далее мы более подробно рассмотрим типы словарей, упомянутых выше.

В **традиционном словаре** значения слов одного языка объясняются средствами этого же языка, например списком синонимов, иллюстративными цитатами из литературы, реальными предложениями из корпуса (как в словаре *Collins COBUILD*) или выдуманными примерами (как в словаре *Longman Dictionary*

of Contemporary English – LDOCE), иногда смысл поясняется рисунками или схемами. Кроме того, в таких словарях могут присутствовать дополнительные лингвистические данные: грамматическая и орфографическая информация, произношение, этимология, стилистические пометы, сферы употребления и другие. В этой группе принято разделять современные описательные словари для носителей языка и учебные словари для изучающих иностранный язык. Последняя группа предоставляет большой объем синтаксической, семантической и прагматической информации, которая не считается необходимой для носителей языка [Béjoint, 1981; Wilks, Slaton, Guthrie, 1996, с. 64].

Тезаурусы в основном бывают двух видов: *лексикографический тезаурус*, в котором общеязыковая лексика организована с семантической точки зрения; и *тезаурус документации*, инструмент для индексации и поиска информации в определенных областях знаний. Здесь нас интересует только первый тип лингвистически ориентированного тезауруса. Тезаурус или сокровищница слов, по сути, представляет собой организацию списков частично синонимичных слов. Самый известный тезаурус – *Тезаурус Роже*, который является собранием слов и фраз, абстрактно классифицированных в соответствии с основными понятиями и значениями [Anderman, Rogers, 1996, с. 81]. Примерами тезаурусов для изучающих английский язык как иностранный являются *Longman Lexicon of Contemporary English*, *Longman Lexicon*, *Longman Language Activator*. Такие словари предлагают практическую и актуальную помощь учащимся, но не претендуют на раскрытие концептуальной системы языка и его универсалий. Кроме того, тезаурусы сыграли особую роль как инструмент вычислительного анализа языка.

Словари лексической сочетаемости (*collocation dictionaries*) играют важную роль для изучающих иностранные языки, в связи с тем, что содержат в себе ценную информацию о совместном использовании разных лексических единиц в контексте речи. Та информация о языке, что для его носителя является базовой, для переводчика или обучающегося может вызывать трудности. В особенности эта

проблема актуальна для тех лексических словосочетаний, которые не имеют переводческой эквивалентности [Anderman, Rogers, 1996, с. 83].

Двужычные словари. Это стандартные переводные словари-списки соответствий, которые стремятся связать два языка друг с другом на уровне эквивалентности значений отдельных слов [Wilks, Slator, Guthrie, 1996, с. 65].

Стилистические словари выступают своего рода расширенными грамматическими справочниками, например словарь *Fowler's Modern English Usage* 1926 года. Они играют особую роль, потому что дают не только примеры использования слов и правила, а также предоставляют большое количество примеров как отрицательных, так и положительных коннотаций лексемы [Wilks, Slator, Guthrie, 1996, с. 66].

Конкордансы являются особым лексикографическим форматом представления слов, который не дает объяснения значения, а только индексирует слова по их появлению в корпусе, обычно взятом из одного произведения или работ одного автора. Конкордансы являются специальным научным инструментом и, по сути, подвидом традиционного словаря, в котором не учитывается вся информация, кроме текстовых цитат [Wilks, Slator, Guthrie, 1996, с. 67].

В заключении, важно отметить, что приведенная выше классификация не является финальной или устоявшейся и представляет собой только попытку обобщить наиболее значимые результаты исследований отечественных и зарубежных ученых в области лексикографии. Многообразие типов и видов словарей представлено как в традиционной, так и в компьютерной лексикографии. Более того при разработке компьютерного словаря могут использоваться элементы и характеристики словарей разных типов и функциональной ориентированности, что обусловлено особыми запросами и требованиями новых пользователей. В следующем параграфе мы подробно рассмотрим историю развития компьютерных лексикографических ресурсов.

1.3 Эволюция компьютерных лексикографических ресурсов

Первоначально в основе создания компьютерных лексикографических ресурсов лежала идея научить компьютер «понимать» текст на естественном языке. Первые попытки автоматизировать словарную работу были предприняты в середине 60-х годов прошлого века. В большей степени они включали в себя подсчет частотности слов, но уже это было большим достижением, так как позволило лингвистам перейти от эмпирических наблюдений за языком к его статистическому анализу и способствовало активному развитию направления позже. Настоящий параграф включает в себя обзор литературы в области истории развития компьютерной лексикографии с особым вниманием на форматы репрезентации семантических отношений в словарях.

Активное развитие компьютерной лексикографии началось в 80-х годах прошлого века. Его результатом в первую очередь стал компьютерный словарь, который упростил использование лексикографического ресурса пользователем и значительно ускорил обработку и поиск информации. Здесь также важно упомянуть *Webster's Seventh New Collegiate Dictionary*, разработанный ранее. Он не был компьютерным словарем в современном его понимании, но стал одним из первых машиночитаемых словарей, перенесенных на перфокарты (ленты), и дал толчок к дальнейшему развитию направления [Wilks, Slator, Guthrie, 1996, с. 81]. Важной для развития компьютерной лексикографии в тот период считается новаторская работа Роберта Амслера 1980 года, который занимался вычислительной лексикологией и работал над лексической таксономией, которая послужила мотивацией для создания лексической базы **WordNet**, учитывающей особенности ментального лексикона как он определяется в рамках психолингвистики [Amsler, 1980]. Сам же ресурс WordNet был разработан в 1985 году Принстонским университетом и сейчас размещен в сети интернет. Он является сетевой базой данных, представляющей собой компьютерный тезаурус, в основе описания значений которого лежат представления лексической семантики [Fellbaum, 2010]. Весь лексикон делится на пять категорий (существительные,

глаголы, прилагательные, наречия и функциональные слова), которые в свою очередь сгруппированы в наборы когнитивных синонимов (синсетов), объединенных семантической сетью, которая выступает как модель человеческой памяти [Беляева, 2016, с. 64]. Разработчики WordNet создавали структуру синсетов, которая базируется на понятии синонимии так, что «два выражения являются синонимичными, если замена одного из них на другое в предложении не меняет значения истинности этого высказывания» [Лукашевич, 2011, с. 52]. Поэтому ресурс такого рода может быть особенно ценным для практики преподавания и изучения иностранных языков.

Для примера рассмотрим поисковый запрос лексемы «*study*» в тезаурусе WordNet (рис. 1), которая входит в тематическое поле английского языка «Education».

WordNet Search - 3.1
[- WordNet home page](#) - [Glossary](#) - [Help](#)

Word to search for:

Display Options:

Key: "S." = Show Synset (semantic) relations, "W." = Show Word (lexical) relations
 Display options for sense: (gloss) "an example sentence"

Noun

- [S: \(n\) survey](#), **study** (a detailed critical inspection)
- [S: \(n\) study](#), [work](#) (applying the mind to learning and understanding a subject (especially by reading)) "*mastering a second language requires a lot of work*"; "*no schools offer graduate study in interior design*"
- [S: \(n\) report](#), [study](#), [written report](#) (a written document describing the findings of some individual or group) "*this accords with the recent study by Hill and Dale*"
- [S: \(n\) study](#) (a state of deep mental absorption) "*she is in a deep study*"
- [S: \(n\) study](#) (a room used for reading and writing and studying) "*he knocked lightly on the closed door of the study*"
- [S: \(n\) discipline](#), [subject](#), [subject area](#), [subject field](#), [field](#), [field of study](#), [study](#), [bailiwick](#) (a branch of knowledge) "*in what discipline is his doctorate?*"; "*teachers should be well trained in their subject*"; "*anthropology is the study of human beings*"
- [S: \(n\) sketch](#), [study](#) (preliminary drawing for later elaboration) "*he made several studies before starting to paint*"
- [S: \(n\) cogitation](#), [study](#) (attentive consideration and meditation) "*after much cogitation he rejected the offer*"
- [S: \(n\) study](#) (someone who memorizes quickly and easily (as the lines for a part in a play)) "*he is a quick study*"
- [S: \(n\) study](#) (a composition intended to develop one aspect of the performer's technique) "*a study in spiccato bowing*"

Verb

- [S: \(v\) analyze](#), [analyse](#), [study](#), [examine](#), [canvass](#) (consider in detail and subject to an analysis in order to discover essential features or meaning) "*analyze a sonnet by Shakespeare*"; "*analyze the evidence in a criminal trial*"; "*analyze your real motives*"
- [S: \(v\) study](#) (be a student; follow a course of study; be enrolled at an institute of learning)
- [S: \(v\) study](#), [consider](#) (give careful consideration to) "*consider the possibility of moving*"
- [S: \(v\) learn](#), [study](#), [read](#), [take](#) (be a student of a certain subject) "*She is reading for the bar exam*"
- [S: \(v\) study](#), [hit the books](#) (learn by reading books) "*He is studying geology in his room*"; "*I have an exam next week; I must hit the books now*"
- [S: \(v\) study](#), [meditate](#), [contemplate](#) (think intently and at length, as for spiritual purposes) "*He is meditating in his study*"

Рисунок 1. WordNet². Поисковый запрос «study»

² WordNet, URL: <http://wordnetweb.princeton.edu/perl/webwn>

Данная лексема может использоваться в виде двух частей речи: как существительное и как глагол. По умолчанию пользователю показывается список значений слова (glosses), которые включают синонимы, и текст примера (example sentence). При необходимости можно раскрыть данные о синсете (direct hypernym / inherited hypernym / sister term / derivationally related form), а также из выпадающего списка добавить отображение частотности употребления лексемы (frequency counts), положение лексемы в базе данных (database location), порядковый номер значения лексемы (sense number, sense key), лексические данные (lexical filename, lexical file number) [WordNet]. Большая часть слов-синонимов представляют собой гиперссылки, по которым можно перейти на страницу с синсетами для другой лексемы и углубить поиск в семантической сети. Изначально WordNet был создан только для английского языка, но впоследствии по такому принципу были разработаны тезаурусы и для других языков, включая русский.

В продолжении изучения семантических ролей была предложена идея создания ресурса **FrameNet**, развитием которого занимались американские ученые в Международном институте информатики в Беркли, Калифорния (International Computer Science Institute – ICSI) под руководством Чарльза Дж. Филлмора [Fillmore, Atkins, 1992]. FrameNet основан на теории фреймовой семантики (frame semantics), которая предполагает, что для понимания значения слов в языке мы должны первоначально знать концептуальные структуры, то есть семантические фреймы, которые обеспечивают основу для существования лексики в языке и ее использования в дискурсе [Mitkov, 2004, с. 63]. В базе данных ресурса заложен корпус, состоящий из предложений, в которых вручную специальным образом аннотированы различные элементы семантического фрейма: события, объекты, участники и другие [Беляева, 2016, с. 76]. Для исследований в области обработки естественного языка такие семантические фреймы представляют собой уникальный набор обучающих данных для маркировки семантических ролей, используемый в таких задачах, как извлечение информации, машинный перевод, распознавание событий, анализ тональности и т.д.

С другой стороны, FrameNet (рис. 2) можно использовать для образовательных целей в качестве комбинаторного словаря. FrameNet, также как и WordNet, изначально создавался только для английского языка, но впоследствии были разработаны официальные версии для других языков: французского, китайского, немецкого, бразильского варианта португальского языка, испанского, японского, шведского и корейского. Ресурс открытый и размещен в сети интернет [FrameNet]. Для русского языка разрабатывается аналогичная система Russian FrameBank (ФреймБанк), которая является «дочерним» ресурсом Национального корпуса русского языка [Кашкин, Ляшевская, 2015].

[Lexical Unit Index](#)

Education_teaching

Definition:

This frame contains words referring to teaching and the participants in teaching. A **Student** comes to learn either about a **Subject**, a **Skill**, a **Precept**, or a **Fact** as a result of instruction by a **Teacher**. Some of the nouns (schoolmaster, -mistress) in this frame refer to administrative positions and do not take relevant frame elements; these will be moved.

Dad **TAUGHT** me that if you work hard, you will be OK - no matter what happens.

What Our Cat **TAUGHT** Me About Marketing!

For two years she **TAUGHT** me French

Mom and Dad **TAUGHT** me not to bullshit, because it always came back to bite me in the ass, without fail.

A young swiss girl up in a swiss chalet **TAUGHT** me how to yodel.

Young Murdock later met a blind martial arts master known as Stick, who taught him how to use his augmented senses and **TRAINED** him as a fighter.

Maria, a woman **STUDYING** to be a nun is sent from her convent to be the governess of the seven children of a widowed naval captain

Frame-Element relations: Several Frame element relations hold in this frame. There is a core-set {**Material**, **Teacher**} (possibly also including **Institution**); these FE's can fill the subject slot of the teach-type verbs. There is another core-set {**Precept**, **Subject**, **Skill**, **Fact**, **Role**} to do with what is learnt. In addition, the FE **Qualification** may occur with **Subject** and **Role**

My sister is **STUDYING** for a maths degree to be a teacher

She's **STUDYING** maths to be a teacher.

Phrases expressing **Qualification** also very often include information about the **Subject** (see below), which is to be annotated on the second layer.

Bill is **STUDYING** for a BS in physics/a maths degree.

Bill is **STUDYING** for a BS in physics/a maths degree.

Рисунок 2. FrameNet³. Поисковый запрос «Education»

Проблема множественности значений слов, то есть полисемия, стала отправной точкой для развития теории генеративного лексикона (generative lexicon), описанной в работах Джеймса Пустейовского [Pustejovsky, 1998]. Данная теория разработана на стыке лингвистической и компьютерной семантики и описывает законы дистрибутивности и логического построения естественных языков, а также изучает процесс формирования значения одного слова в различных контекстах употребления и возможность формального представления этих

³ FrameNet, URL: <https://framenet.icsi.berkeley.edu/fndrupal/>

отношений [Боярская, 2012]. Согласно Дж. Пустейовскому в генеративном лексиконе выделяется три основных элемента лексической структуры, влияющих на отражение семантической информации в синтаксисе: *аргументная структура* (argument structure), *структура событий* (event structure) и *смысловая структура* (qualia structure). Qualia-структура отображает: особенности объекта, отличающие его от других объектов; взаимосвязь объекта и его частей; описание назначения объекта и его функций; информацию об источниках и обстоятельствах возникновения объекта. По такому принципу в генеративном лексиконе семантические типы ограничивают значение других слов, например, английский глагол «*to eat*» подразумевает характеристику «*food*» к прямому дополнению, стоящему после него, независимо от того, что это дополнение обозначает [Mitkov, 2004, с. 63].

Следующим важным этапом развития компьютерной лексикографии стала разработка и широкое распространение лингвистических корпусов в научных исследованиях, которые поспособствовали формированию *корпусной лексикографии* – области «изучения языка на основе текстовых или акустических корпусов, при постоянном использовании компьютера в определенных фазах хранения, извлечения или анализа данных» [Палкова, 2015, с. 89]. В качестве примера можно привести тот факт, что в 1990-х годах британские издатели словарей, в особенности словарей для изучающих иностранные языки, вложили значительные средства для обновления своих продуктов, где сделали акцент на то, чтобы словари лучше соответствовали актуальным данным корпусов, как по составу словника, так и по значению и примерам употребления слов [Mitkov, 2004, с. 59]. Такое обновление словарного состава на основе корпусных данных включало в себя полное переписывание и реструктуризацию определений в поисках новых уровней обобщения, соответствующих имеющимся данным, что, в свою очередь, повлияло на представление семантических сетей и иерархий в компьютерных словарях. В первую очередь, основной причиной обновления имеющихся на то время словарей стало то, что корпуса являются хорошим источником примеров употребления слов в живом языке, который меняется

значительно быстрее, чем происходит лексикографическое оформление и обновление словарей. На данный момент существует несколько подходов в использовании корпусных данных в словаре: примеры употребления слов могут быть взяты из корпусов с минимальной модификацией и включены в словарь на стадии компиляции (*corpus-derived*), или представлять собой адаптированный материал из корпуса (*corpus-based*), либо стать комбинацией этих двух способов [Lew, 2010, с. 295]. Кроме этого, благодаря развитию компьютерной лексикографии, корпуса теперь сами могут составлять часть компьютерного словаря и, таким образом, предлагать полезный онлайн-источник дополнительных примеров словоупотребления.

Значительным достижением в этой области стала публикация корпусного словаря **COBUILD** (аббр. *Collins Birmingham University International Language Database*) в 1987 году, который сломал традиции классической лексикографии, объединив интуицию лексикографа с компьютерными инструментами. Особенность этого словаря состоит в том, что в его основе лежал самый крупный на тот момент корпус устных и письменных текстов. Корпус состоял из 20 миллионов слов, взятых из книг, журналов, газет, брошюр, бесед, радио- и телепередач, и представлял собой большую репрезентативную выборку языковых данных для английского языка [Wilks, Slator, Guthrie, 1996, с. 101]. Современная версия словаря доступна онлайн. Определения значений лексем в COBUILD представлены в виде предложений с использованием искомого слова (рис. 3). Такие предложения также служат для обозначения синтаксических или семантических характеристик, например определения исчисляемых существительных начинаются с неопределенного артикля. Кроме этого, в словарной статье приведены примеры употребления слова в корпусе, аудио с произношением, коллокации и другая информация, полезная для пользователя, изучающего иностранные языки.

Также на материале корпусных данных была разработана система *GDEX*⁴ (сокращение от «Good Dictionary EXamples»). Создателями проекта стала

⁴ GDEX, URL: <https://www.sketchengine.eu/guide/gdex/>

компания-разработчик программного обеспечения для управления корпусами и анализа текста *Sketch Engine* [Kilgarriff, 2008]. GDEX является системой оценки предложений с точки зрения их пригодности для использования в качестве хороших примеров употребления лексемы для словаря или в учебных целях. Такие системы и корпусные данные в целом могут стать успешным образовательным инструментом.

The screenshot shows the COBUILD online dictionary entry for the word 'study'. At the top, there are navigation tabs: 'study', 'Video pronunciation', 'English: study', 'American: study', 'Example sentences', 'COBUILD Collocations', and 'Trend:'. Below the tabs, the title 'Definition of 'study'' is followed by the word 'study' in a large font, with 'Collins COBUILD' underneath. To the right, there is a 'Word Frequency' indicator with five red dots and a share icon. The phonetic transcription '(stʌdi)' is shown with a speaker icon. The 'Word forms' section lists: plural, 3rd person singular present tense 'studies', present participle 'studying', past tense, and past participle 'studied'. The first section, '1. VERB', includes a definition: 'If you study, you spend time learning about a particular subject or subjects.' followed by three example sentences with their grammatical functions: '...a relaxed and happy atmosphere that will allow you to study to your full potential. [VERB]', 'He went to Hull University, where he studied History and Economics. [VERB noun]', and 'The rehearsals make it difficult for her to study for law school exams. [VERB + for]'. A 'Synonyms' section lists 'learn, cram [informal], swot (up) [British, informal], read up' and a link to 'More Synonyms of study'. The second section, '2. UNCOUNTABLE NOUN', includes a definition: 'Study is the activity of studying.' followed by an example sentence: '...the use of maps and visual evidence in the study of local history.' and a 'Synonyms' section listing 'examination, investigation, analysis, consideration' and a link to 'More Synonyms of study'.

Рисунок 3. COBUILD. ⁵Поисковый запрос «study»

Подводя итог, можно сказать, что в сферу исследований компьютерной лексикографии входят программы, базы данных и ресурсы различного рода. Так, лингвистический корпус обеспечивает более статистически достоверную базу данных для компьютерной обработки и изучения естественного языка, а такие ресурсы как WordNet, FrameNet и генеративный лексикон имеют безграничные возможности практического применения для разработки искусственного интеллекта. В свою очередь, компьютерные словари с удобным интерфейсом,

⁵ COBUILD, URL: <https://www.collinsdictionary.com/>

созданным для человека как финального пользователя, являются всего лишь частью компьютерной лексикографии. Однако для современного пользователя важен именно словарь, при этом чтобы он был не просто цифровой копией бумажной книги, а выступал как интерактивный образовательный ресурс. Компьютерные технологии позволяют реализовать такое желание пользователя. В следующих параграфах мы более подробно рассмотрим особенности мультимодального представления информации в компьютерных словарях.

1.4 Мультимодальный словарь как словарь для нового пользователя

В последние годы наблюдается стремительный рост разработки мультимедийных приложений, которые находят свое применение практически во всех сферах нашей повседневной жизни, таких как образование, развлечения, бизнес, транспорт и др. Развитие технологий и инструментов позволяет создавать большие мультимедийные архивы и разрабатывать совершенно новые стили взаимодействия, в которых естественный язык играет важную роль и становится неотъемлемой частью интерфейсов, так как, являясь одним из наиболее привычных способов коммуникации человека, он упрощает процесс взаимодействия с компьютером. Современный человек, как ежедневный пользователь компьютера, ставит довольно непростую задачу для создателей обучающих приложений и словарей. Как отмечает Юлия Владимировна Сорокина, «в современном мире передача и восприятие информации происходят посредством различных систем и каналов, способных порождать значения и взаимодействовать одновременно» [Сорокина, 2017, с. 168] таким образом, что коммуникация не сводится к какой-то одной знаковой системе, а представляет собой специфическую форму симбиотического взаимодействия различных модальностей. Этот фактор также оказывает прямое влияние на актуальные разработки в рамках компьютерной лексикографии. В настоящем параграфе мы рассмотрим особенности мультимодального представления информации в компьютерных словарях.

Для начала необходимо определить, что понимается под *мультимодальностью*. Этот термин ассоциируется и активно используется в теории коммуникации и социальной семиотике. **Мультимодальность** определяется как «описание общих законов и правил взаимодействия в коммуникативном акте вербальных и невербальных знаков, соединение различных кодов предъявления информации» [Сорокина, 2017, с. 168]. Большой вклад в изучение мультимодальности внёс австралийский семиотик Гюнтер Кресс. Ученый писал, что использование трех модальностей (модусов) в одном знаке – письмо, изображение и цвет – дает значительные преимущества при передаче сообщения. Каждый модус выполняет определенную функцию: изображение показывает то, что было бы слишком долго читать, текст записывает имена и названия, которые было бы трудно показать, цвет, в свою очередь, используется для выделения определенных аспектов сообщения в целом. Каждый из этих элементов вносит свой вклад для максимального эффекта на адресата – это и есть пример мультимодальности [Kress, 2010, с. 1].

В данном контексте необходимо отметить, что разделяются два основных термина, которые не являются синонимами: **медиа** как средство информации (*medium*) и **модус** как способ передачи (*mode*). Наиболее подходящей для настоящего исследования стала терминология Марка Мэйбери, который использует термины *модус* (*mode*) и *модальность* (*modality*) для обозначения различных видов воспринимаемых объектов (зрительных, слуховых, тактильных и обонятельных), а термин *медиа* (*medium*) относит к носителю информации (бумага, диск, экран, колонка, принтер) [Maybury, 1999; Mitkov, 2004, с. 651]. При этом, согласно Г. Крессу **модус** выступает как социально сформированный и культурно заданный семиотический ресурс для создания смысла, примером которого может выступать изображение, письмо, расположение объектов, музыка, жесты, речь, анимации и трехмерные объекты. В своей совокупности они отражают особенности языка, истории и культуры общества, которые оказали влияние на их формирование [Kress, 2010, с. 79]. В российской науке важной для изучения мультимодальности стала работа Андрея Александровича Кибрика

«Мультимодальная лингвистика», в которой он обозначил особенности коммуникации в целостном подходе к речевому действию и выделил три ее основных компонента – просодия, вербальность и жесты / мимика [Кибрик, 2010].

Ирина Павловна Хутыз отмечает, что мультимодальность как «симбиоз привычных модальностей с новыми цифровыми средствами» все чаще используется в образовательной среде и обучающих средствах, в том числе и в словарях, например, «когда студент одновременно читает текст, смотрит на графический объект (статичный или подвижный), получает информацию по слуховому каналу эффективно в процессе обучения» [Хутыз, 2016, с. 91]. В связи с этим в науке часто говорят о формировании «нового» типа пользователя – digital native «уроженца Интернета», то есть человека, привыкшего получать информацию через цифровые каналы [Голубкова, 2014, с. 75]. Такой пользователь требует создания кардинально нового словаря, способного включить в себя множество ресурсов разного рода и происхождения, содержащих культурно-значимую лексику, которая отражает ключевые слова эпохи. При этом отличительная черта электронных (цифровых) медиа – их мультимодальность, является той особенностью, которая может выступить значительным преимуществом при создании современных компьютерных словарей, содержащих кроме текста на естественном языке также и объекты различных модальностей, которые не могут быть перенесены на бумагу, например аудио, анимацию или видео.

С точки зрения лексикографии мультимодальность выражается в особенностях представления значения лексемы в словарной статье. Здесь ученые выделяют две основных модальности: *вербальную* и *невербальную* [Lew, 2010]. **Вербальная модальность** включает в себя непосредственно словарную дефиницию лексемы, а также ее аудиопрезентацию. Рассматривая словари для изучающих иностранный язык (learner's dictionaries) важно отметить, что среди лексикографов существует своего рода конвенция об использовании ограниченного словарного запаса для определения слов. Такой набор слов должен быть простым и понятным для иностранных обучающихся, например для английского языка этот список состоит примерно из 250 слов [Cowie, 1999]. Кроме

этого, текст в словарной статье имеет графические ресурсы, такие как шрифт, размер, начертание (полужирный, курсив, подчеркнутый), интервал и цвет [Kress, 2010, с. 79]. Аудиопрезентация, в свою очередь, также важна для пользователей словаря, не являющихся носителями языка. Она дает возможность не только услышать аутентичное произношение слова, но и несет информацию о просодических характеристиках речи, таких как ударение, ритм и интонация [Lew, 2010, с. 296].

Второй важной составляющей в словарной статье является **невербальная модальность**, которая может состоять из одного или комбинации нескольких следующих элементов:

- *аудиозаписи неязыковых звуков* (звукоподражательных слов), так называемые звуковые эффекты;
- *графические иллюстрации* (статичные). В основном используются как вспомогательный элемент в словарных статьях о животных, растениях, предметах, которые трудно объяснить словами (геометрические фигуры, формы и др.);
- *фотографии*;
- *графики*;
- *анимация*;
- *видеоклипы*.

Видео обладает высокой контекстуальностью и может передавать информацию о прагматическом и ситуативном контексте языкового общения, что может быть полезно в учебных словарях. Такое техническое решение было реализовано в словаре *Longman Dictionary of Contemporary English* (LDOCE4) [Lew, 2010, с. 297 - 303]. Дополнительно могут также использоваться и другие визуальные средства, например знаки препинания, пространство между словами или вокруг абзацев, блоки или рамки. Ожидается, что такая мультимодальная структура должна больше вовлекать пользователя словаря и способствовать лучшему пониманию и приобретению знаний.

Исследования в направлении мультимодальной лексикографии становятся всё более популярными, ими занимаются как отечественные, так и зарубежные

ученые. Такие словари меняют свой формат и принципы составления, но остаются важным направлением в компьютерной лексикографии и лингводидактике. Среди реализованных проектов и текущих исследований можно выделить *электронный словарь английских и русских фитонимов* [Сивакова, 2004], *терминологический электронный систематизирующий глоссарий «Экологический аудит»* [Кантышева, 2011], *мотивационно-сопоставительный словарь предметно-бытовой лексики русского и английского языков* [Шафтельская, 2013], *электронный переводной словарь-справочник естественнонаучных омонимичных терминов* [Волошина, 2015], *двуязычный англо-японский мультимодальный словарь для изучающих иностранный язык* [Sato, 2016], *электронный идеографический словарь: «Русский тематический словарь»* [Богачёва, Ольховская, Парамонова, 2017], *мультимодальный словарь фитонимов нового типа* [Гулевская, Мишланова, 2017], *русско-татарско-англо-испанский иллюстративный словарь для детей* [Сафиуллина, 2018], *сопоставительный лингвокультурологический словарь пищевой метафоры на материале русского и итальянского языков* [Помаролли, 2018], *татарско-английский онлайн-словарь проекта Glosbe* [Сафиуллина, 2021], также можно выделить исследование репрезентации культурных особенностей значения в учебном словаре для носителей китайского языка, изучающих английский как иностранный [Хуе, 2017], исследование влияния иллюстративных элементов на привлекательность лексикографического ресурса для пользователя [Biesaga, 2017] и др. Таким образом, вне зависимости от своего формата словари являются ценными хранилищами лексических знаний.

Далее мы проанализируем ряд русско- и англоязычных мультимодальных компьютерных словарей. Среди русскоязычных компьютерных лексикографических ресурсов можно выделить онлайн-тезаурус «*Карта слов*» и тезаурус *RuWordNet*, русскоязычный аналог WordNet. Англоязычные компьютерные мультимодальные лексикографические ресурсы в большей степени также основаны на технологии WordNet, поэтому мы рассмотрим два словаря

отличающихся друг от друга интерфейсом: интерактивный тезаурус «*The Visual Thesaurus*» и словарь «*Fine Dictionary*».

Проект «**Карта слов**» представляет собой онлайн-тезаурус в формате ассоциативной семантической сети. Пользователь может осуществлять поиск синонимов, сочетаемости слов и словесных ассоциаций, просматривать примеры употребления в контексте и цитаты из литературных произведений, толкования слов и устойчивых выражений, а также разбор слова по составу и сходные по морфемному строению слова. При создании ресурса разработчики пользовались последними достижениями в области компьютерной лингвистики, машинного обучения и искусственного интеллекта. Онлайн-тезаурус Карта слов содержит несколько полей (рис. 4):

- *толкование* (развёрнутое толкование значения слова или словосочетания с примерами употребления);
- *тезаурус* (синонимы и сходные по смыслу выражения, прямые и обратные ассоциации, информация о сочетаемости);
- *примеры* (примеры употребления в контексте из современных источников и из русской классической литературы);
- *справка* (правописание, таблицы склонения и спряжения, разбор по составу с графической схемой и списком сходных по морфемному строению слов);
- *цитаты* (высказывания известных людей, избранные цитаты из произведений культуры) [Карта слов].

Такой набор информации может быть полезным как для студента, изучающего русский язык как иностранный, так и для школьников-носителей русского языка. Кроме того, для зарегистрированных пользователей на сайте доступна бета-версия контекстного русско-англо-русского словаря (рис. 5), его структура отличается от русскоязычной версии. В контекстном словаре в результате поискового запроса лексемы отображаются иллюстративные примеры её употребления для каждого отдельного значения в формате параллельного корпуса с вариантами переводов.

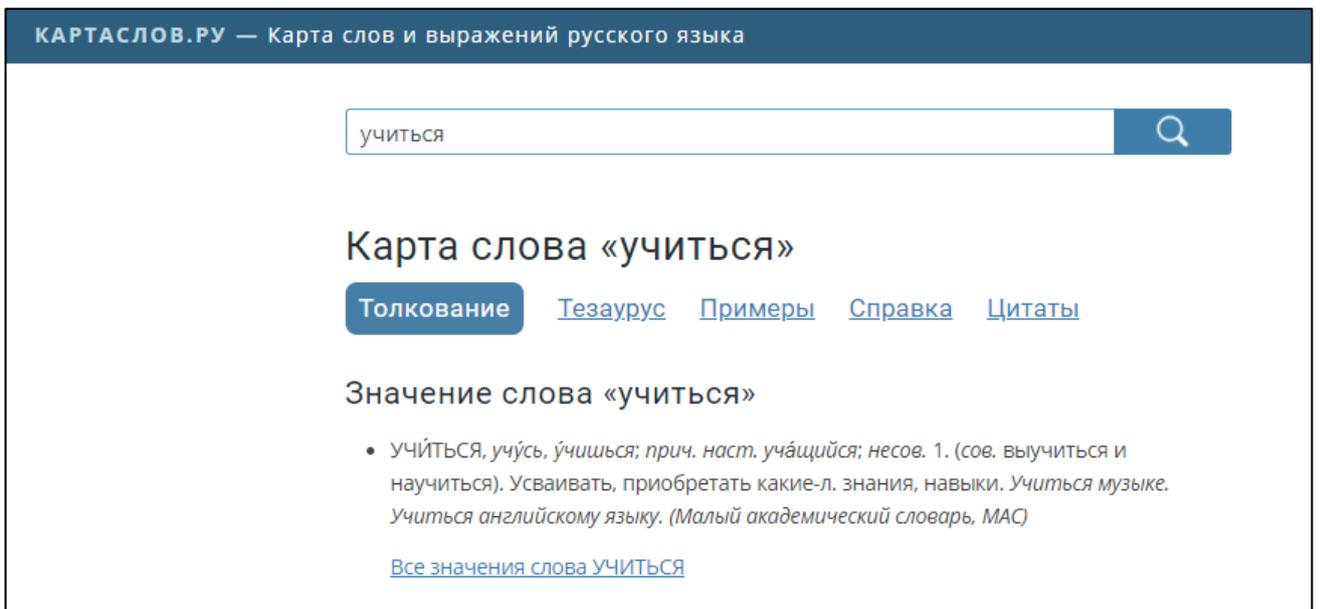


Рисунок 4. Карта слов⁶. Поисковый запрос «учиться»

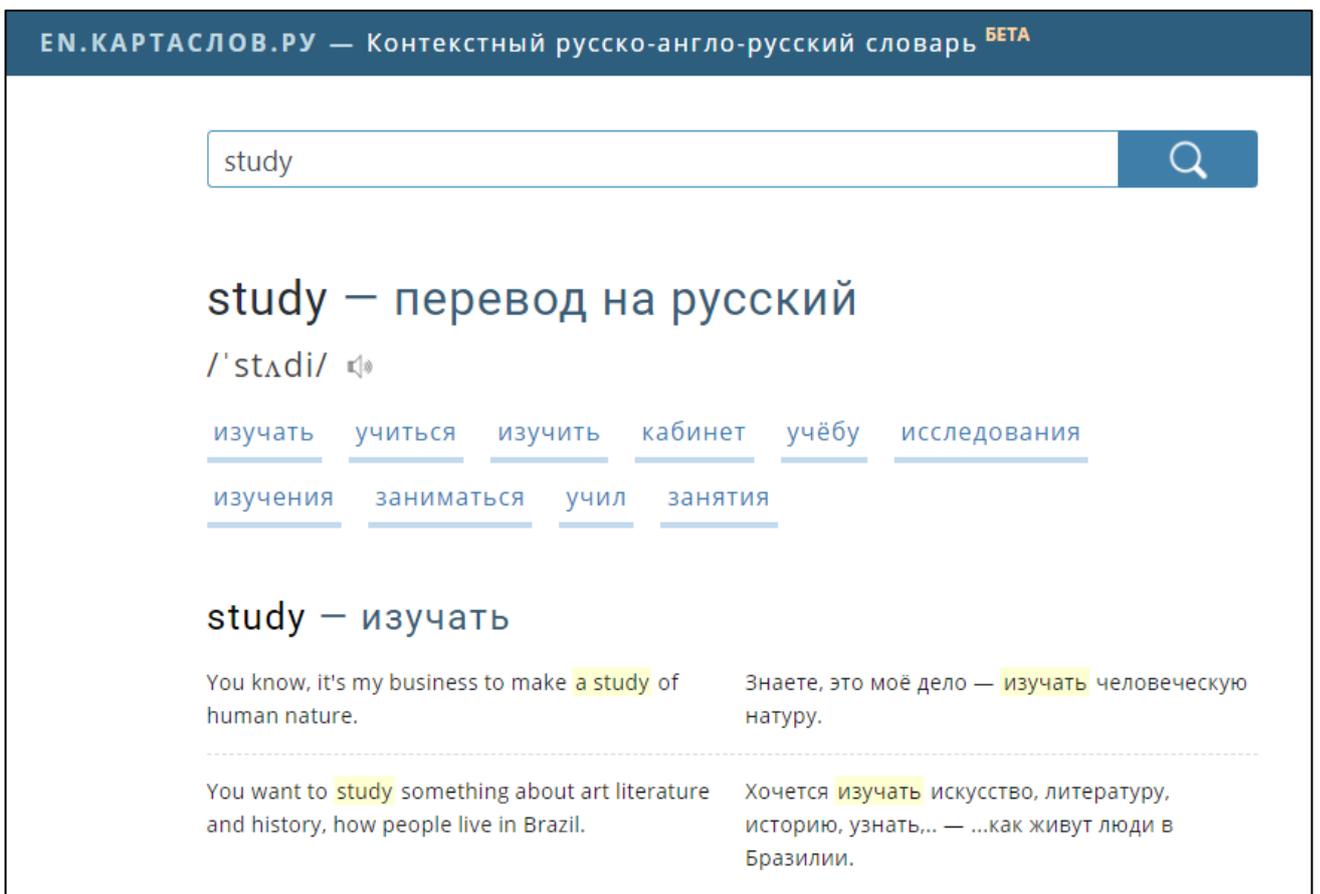


Рисунок 5. EN.Карта слов. Поисковый запрос «study»

Также на сайте есть два интерактивных сервиса, способствующих улучшению словаря и основанных на системе машинного обучения: «Научи бота!» и «Игра в ассоциации». В первом случае *лампочка* показывает пользователю слово

⁶ Карта слов, URL: <https://kartaslov.ru/>

и предлагает выбрать между тремя эмоциональными характеристиками лексемы: нейтральное, положительное или отрицательное понятие. Игра в ассоциации предлагает пользователю придумать максимальное количество ассоциаций к слову за 60 секунд, получая за каждый ответ баллы. Оригинальные ассоциации приносят больше баллов. В результате можно соревноваться с другими участниками и создать собственную карту слов, вести персональный словарь и участвовать в развитии проекта. Данные интерактивные ресурсы не только повышают мотивацию пользователя к изучению языка, но и способствуют расширению лексической базы данных словаря и обучению программы, лежащей в его основе.

В параграфе, посвященном истории развития компьютерной лексикографии, мы рассматривали тезаурус в формате WordNet, который описывает лексику английского языка в виде сети понятий (синсетов). Существует аналогичный ресурс для русского языка – RuWordNet. Разработкой **RuWordNet** занималась команда ученых под руководством Натальи Валентиновны Лукашевич. Русскоязычный ресурс был создан на основе автоматизированной трансформации тезауруса РуТез (рис. 6), разработанного этой же командой ранее [Лукашевич, 2016].

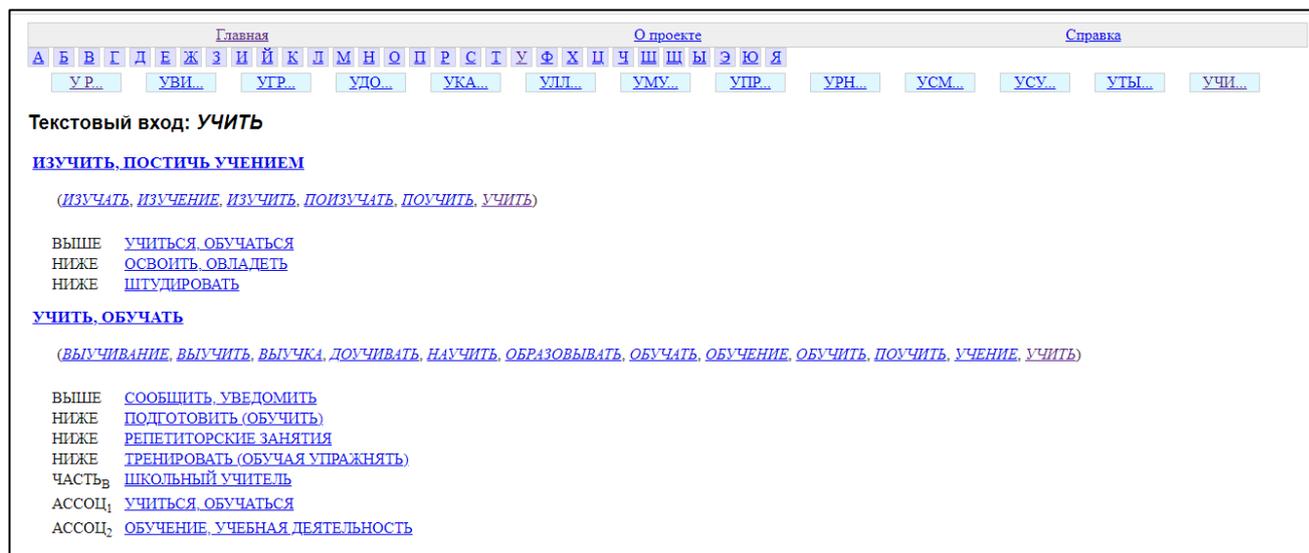


Рисунок 6. Тезаурус РуТез. ⁷Поисковый запрос «учить»

Тезаурус РуТез – лингвистический ресурс концептуального типа, представляющий собой иерархическую сеть понятий, к которым приписаны

⁷ Тезаурус РуТез, URL: <http://www.labinform.ru/pub/ruthes/index.htm>

текстовые выражения. По своей структуре он схож с тезаурусом WordNet, однако главное отличие состоит в том, что РуТез создавался как ресурс для автоматической обработки текстов, а не как словарь с пользовательским интерфейсом [Лукашевич, 2011]. До недавнего времени РуТез использовался только как корпоративный ресурс, сейчас он находится в открытом доступе в сети Интернет.

Сравнивая результаты поискового запроса лексемы «учить» в РуТез и лексемы «учиться» в тезаурусе RuWordNet (рис. 7), входящей в тематическое поле русского языка «Образование», можно увидеть четко выраженную структуру отношений между синсетам. В РуТез лексемы отображаются на основе родовидовых отношений: выше, ниже, часть, ассоц1, ассоц2. В тезаурусе RuWordNet, ориентированном на человека как на конечного пользователя, отношения указаны более эксплицитно: отношения гипоним-гипероним (род-вид), экземпляр-класс, отношение антонимии, часть-целое, причина, логическое следование, предметная область (домен), а также присутствует ссылка на синоним лексемы в англоязычном ресурсе (interlingual index). Важно отметить, что большая часть слов является гиперссылками, которые позволяют перемещаться в базе данных от одной лексемы к другой, переходя в синсете на уровень выше или ниже. Данные ресурсы, к сожалению, не обладают привлекательным пользовательским интерфейсом и наличием дополнительных мультимодальных ресурсов. Однако, несмотря на этот недостаток тезаурус RuWordNet является уникальным ресурсом, предоставляющим ценную информацию о семантических связях и отношениях лексем в русском языке.

Интерактивный интернет-ресурс **The Visual Thesaurus**, был разработан Нью-Йоркской компанией Thinkmap, Inc, специализирующейся на создании пользовательских интерфейсов и механизмах визуализации, которые позволяют конечным пользователям более эффективно просматривать и воспринимать сложную информацию. The Visual Thesaurus или Визуальный тезаурус представляет собой интерактивный словарь идеографического типа, который создает карты слов [The Visual Thesaurus]. Ресурс насчитывает более 145,000 слов

Тезаурус русского языка RuWordNet GB

УЧИТЬСЯ глагол

УЧИТЬСЯ учиться, обучаться

Синсет

1 **ВЫУЧИТЬСЯ, ДОУЧИВАТЬСЯ, ДОУЧИТЬСЯ, НАУЧАТЬСЯ, НАУЧИТЬСЯ, ОБУЧАТЬСЯ, ОБУЧИТЬСЯ, ПОУЧИТЬСЯ, ПРОУЧИТЬСЯ, УЧИТЬСЯ**
[Понятие RuTез: учиться, обучаться]

однокоренные слова

1 **ИЗУЧИТЬ, УЧИТЬ, ПОУЧИТЬ, ВЫУЧИТЬСЯ, ДОУЧИВАТЬСЯ, ДОУЧИТЬСЯ, НАУЧИТЬСЯ, ОБУЧИТЬСЯ, ПРОУЧИТЬСЯ, ПОУЧИТЬСЯ, ВЫУЧИВАНИЕ, ОБУЧИТЬ, ВЫУЧИТЬ, НАУЧИТЬ, УЧИТЬ, ДОУЧИВАТЬ, ПОУЧИТЬ, ПОДУЧИТЬСЯ, ПОДУЧИВАТЬСЯ, РАЗУЧИТЬСЯ, РАЗУЧИВАТЬСЯ, ОТУЧИТЬСЯ, ОТУЧИВАТЬСЯ**

Interlingual Index

1 i24750 learn.v.01 (learn, larn, acquire)
[Определение Wordnet: gain knowledge or skills]

домен

1 **НАРОДНОЕ ОБРАЗОВАНИЕ, ОБЛАСТЬ ОБРАЗОВАНИЯ, ОБЛАСТЬ ПРОСВЕЩЕНИЯ, ОБРАЗОВАНИЕ, ОБРАЗОВАТЕЛЬНАЯ ОБЛАСТЬ, ОБРАЗОВАТЕЛЬНАЯ СИСТЕМА, ОБРАЗОВАТЕЛЬНАЯ СФЕРА, ПРОСВЕЩЕНИЕ, СИСТЕМА ОБРАЗОВАНИЯ, СФЕРА ОБРАЗОВАНИЯ, СФЕРА ПРОСВЕЩЕНИЯ**
[Понятие RuTез: сфера образования]

гипероним

1 **ПОЗНАВАТЬ, ПОЗНАТЬ, ПОЛУЧИТЬ ЗНАНИЯ, ПОЛУЧИТЬ НОВОЕ ЗНАНИЕ, ПРИОБРЕСТИ ЗНАНИЕ, ПРИОБРЕТАТЬ ЗНАНИЕ**
[Понятие RuTез: познать, приобрести знание]

2 **ОСВАИВАТЬ, ОСВОИТЬ, ОСВОИТЬСЯ, УСВАИВАТЬ, УСВАИВАТЬ ИНФОРМАЦИЮ, УСВАИВАТЬСЯ, УСВОИТЬ, УСВОИТЬ ИНФОРМАЦИЮ, УСВОИТЬСЯ**
[Понятие RuTез: усвоить информацию]

гипоним

1 **НАБИТЬ РУКУ, НАВОСТРИТЬСЯ, НАЛОВЧИТЬСЯ, НАСОБАЧИВАТЬСЯ, НАСОБАЧИТЬСЯ, НАТОРЕТЬ, НАТОРЕТЬ В СВОЕМ ДЕЛЕ, ПОДНАТОРЕТЬ, ПРИОБРЕСТИ НАВЫК**
[Понятие RuTез: наловчиться, приобрести навык]

2 **НАТРЕНИРОВАТЬСЯ, ПОТРЕНИРОВАТЬСЯ, ТРЕНИРОВАТЬСЯ**
[Понятие RuTез: тренироваться]

3 **УСПЕВАТЬ, УСПЕВАТЬ В УЧЕБЕ**
[Понятие RuTез: успевать в учебе]

4 **ИЗУЧАТЬ, ИЗУЧИТЬ, ПОИЗУЧАТЬ, ПОУЧИТЬ, УЧИТЬ**
[Понятие RuTез: изучить, постичь учением]

5 **САМООБРАЗОВАТЬСЯ, САМООБРАЗОВЫВАТЬСЯ, САМООБУЧАТЬСЯ**
[Понятие RuTез: самообучение]

частеречная синонимия

1 **УЧЕБА, УЧЕНИЕ, УЧЕНЬЕ**
[Понятие RuTез: учиться, обучаться]

следовательно

1 **ГРЫЗТЬ ГРАНИТ НАУКИ, ПРЕПОДАВАТЬСЯ**
[Понятие RuTез: обучение, учебная деятельность]

Рисунок 7. RuWordNet⁸. Поисковый запрос «учиться»⁸ RuWordNet, URL: <https://ruwordnet.ru/ru>

английского языка и 115,000 значений и использует базу данных WordNet. Пользователь может задать в строку поиска интересующую его лексему, а тезаурус составит схему – «паутину» к связанным с ней словам. Пользовательский интерфейс (рис. 8) состоит из нескольких важных элементов: центрального окна, в котором отображаются слова и смежные значения; панели инструментов, позволяющей осуществлять навигацию, просматривать историю и выставлять настройки поиска; поля значений, представляющего морфологическую и грамматическую информация о слове, а также множество значений в зависимости от части речи.

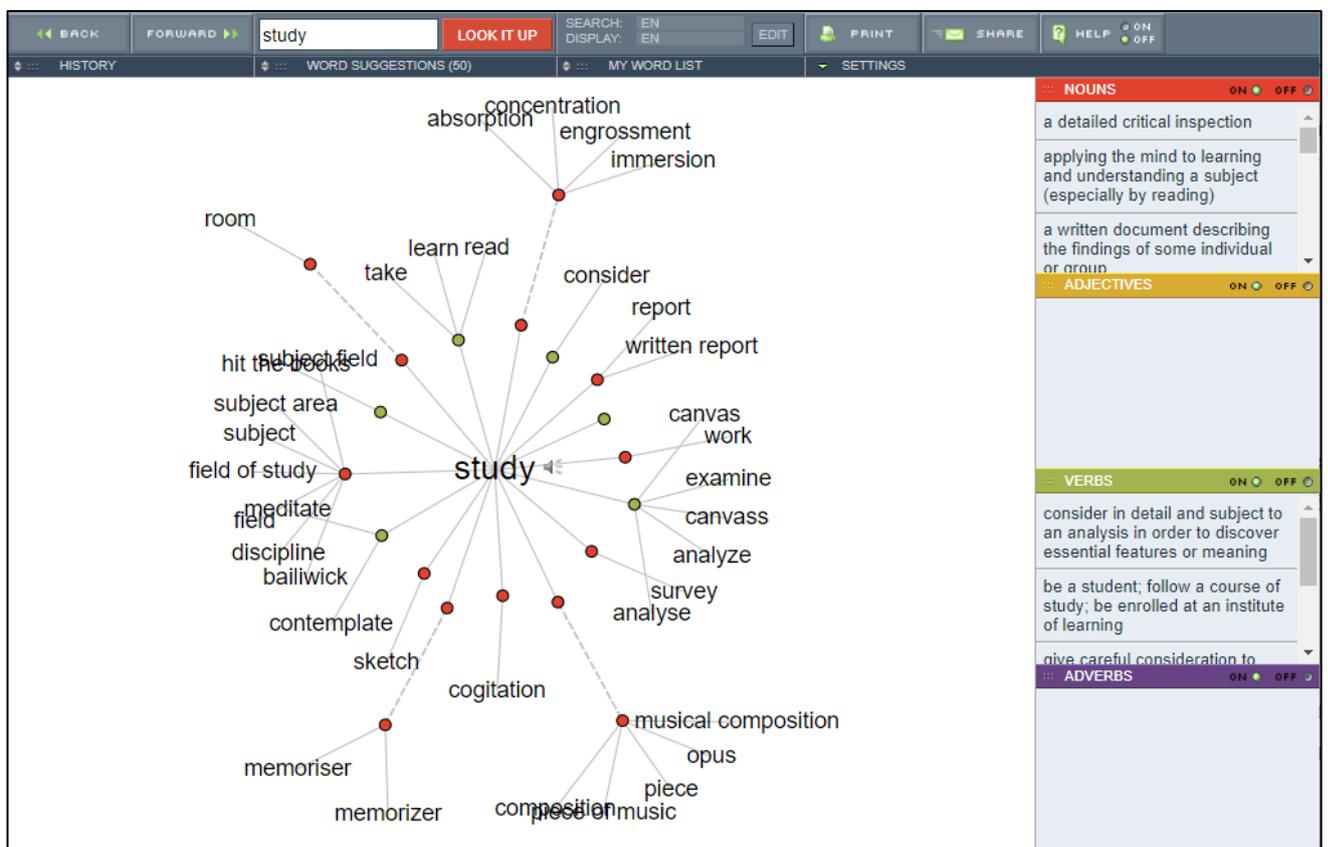


Рисунок 8. The Visual Thesaurus⁹. Поисковый запрос «study»

Инновационность данного ресурса состоит в том, что пользователь получает возможность интерактивно переключаться между лексемами, тем самым углубляясь в семантические связи и открывая новые карты слов, где центром может стать любое слово из периферии первоначального запроса. Кроме того, дополнительные инструменты всплывающих окон при изучении лексем позволяют

⁹ The Visual Thesaurus, URL: <https://www.visualthesaurus.com/>

прослушать произношение слова, совершить поисковый запрос изображения или какой-либо дополнительной информации о слове в сети Интернет. Создатели ресурса считают, что такой подход стимулирует мотивацию к изучению английского языка. Визуальный тезаурус в отличие от бумажного словаря позволяет исследовать семантические особенности лексики нелинейно, пользователь интуитивно перемещается от слова к слову и строит собственные ассоциативные связи, способствующие лучшему усвоению новой лексики.

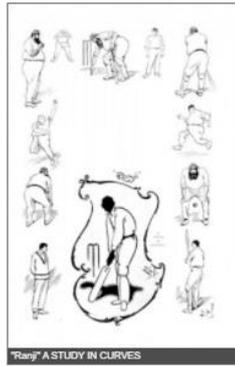
Другим мультимодальным онлайн-ресурсом английского языка является словарь **Fine Dictionary**. Словарь содержит множество различных элементов, которые не только дают толкования слов из четырех крупных словарей английского языка (WordNet 3.6, Webster's Revised Unabridged Dictionary, Century Dictionary and Cyclopedia, Chambers's Twentieth Century Dictionary), но и представляют другую культурологическую информацию. Результат поиска любого слова (рис. 9) представлен в виде длинной веб-страницы с разделами:

- *определение слова* (definitions), включая иллюстрации, фотографии и интересные факты;
- *цитаты* с использованием слова (quotations), каждая цитата сопровождается фотографией известной личности;
- *идиомы* (idioms) с объяснением их значения;
- *этимология* слова (etymology);
- *использование слова* (usage): примеры предложений из литературных произведений, поэзии, научных текстов и новостных статей;
- *дополнительная информация* о связанных словах с гиперссылками (related words): синонимы, гиперонимы, гипонимы, обратный словарь, опечатки [Fine Dictionary].

Такой ресурс позволяет не просто изучить толкование лексемы, но и расширить знания о культуре страны языка. Разнообразные примеры употребления слова в зависимости от контекста раскрывают разные сферы употребления лексемы, цитаты и крылатые выражения показывают, что слово употребляется в живой речи. Картинки в таком словаре используются как средство

Study

Definitions



WordNet 3.6

1. (v) **study** be a student of a certain subject "She is reading for the bar exam"
2. (v) **study** learn by reading books; I must hit the books now" "He is studying geology in his room", "I have an exam next week"
3. (v) **study** be a student; follow a course of study; be enrolled at an institute of learning
4. (v) **study** consider in detail and subject to an analysis in order to discover essential features or meaning "analyze a sonnet by Shakespeare", "analyze the evidence in a criminal trial", "analyze your real motives"
5. (v) **study** think intently and at length, as for spiritual purposes "He is meditating in his study"
6. (v) **study** give careful consideration to "consider the possibility of moving"
7. (n) **study** a detailed critical inspection
8. (n) **study** preliminary drawing for later elaboration "he made several studies before starting to paint"
9. (n) **study** a room used for reading and writing and studying "he

knocked lightly on the closed door of the study"

10. (n) **study** a state of deep mental absorption "she is in a deep study"

11. (n) **study** applying the mind to learning and understanding a subject (especially by reading) "mastering a second language requires a lot of work", "no schools offer graduate study in interior design"

12. (n) **study** attentive consideration and meditation "after much cogitation he rejected the offer"

13. (n) **study** a branch of knowledge "in what discipline is his doctorate?" "teachers should be well trained in their subject", "anthropology is the study of human beings"

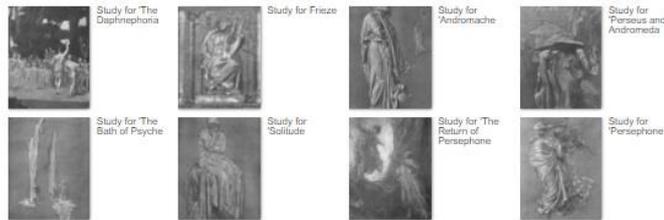
14. (n) **study** a composition intended to develop one aspect of the performer's technique "a study in spiccato bowing"

15. (n) **study** a written document describing the findings of some individual or group "this accords with the recent study by Hill and Dale"

16. (n) **study** someone who memorizes quickly and easily (as the lines for a part in a play) "he is a quick study"

* * *

Additional illustrations & photos:



Quotations



Harry S. Truman

"Study men, not historians."



Albert Goldman

"The Jews have always been students, and their greatest study is themselves."



Confucius

"Study the past if you would divine the future."



W. Clement Stone

"If you employed study, thinking, and planning time daily, you could develop and use the power that can change the course of your destiny."



Fyodor Dostoevsky

"Realists do not fear the results of their study."



Pierre Charron

"The true science and study of man, is man himself."

Etymology

Webster's Revised Unabridged Dictionary

OE. *studġe*, L. *studium*, akin to *studere*, to study; possibly akin to Gr. *haste*, zeal, to hasten; cf. OF. *estudie*, *estude*, F. *étude*, Cf. *Etude Student Studio Study* (v. i.)

Usage

In literature:

Coates, "said he, "I wish you'd come to my study and help me choose half a dozen trout-flies, there's a good fellow."

Dig and I will get a study after Christmas.

And above all, study, study, study!

The study of men's skulls then is a study which is strictly physical, a study of facts over which the will of man has no direct control.

Early in February a study circle was formed, under the leadership of Mrs. Simon Litman, for the study of post-Biblical Jewish history.

Рисунок 9. Fine Dictionary¹⁰. Поисковый запрос «study»

¹⁰ Fine Dictionary, URL: <http://www.finedictionary.com>

визуализации, а дополнительная информация расширяет семантическую сеть ассоциаций и позволяет пользователю составлять свою собственную ментальную карту слов. Данный словарь не обладает множеством интерактивных инструментов, но предоставляет обширную и разностороннюю информацию о слове, включая культурологическую справку, которая может быть особенно полезна для изучающих английский язык как иностранный.

Для сравнения мы постарались выбрать мультимодальные компьютерные лексикографические ресурсы русского и английского языков, различающиеся по принципу представления информации, объему лингвистических данных и пользовательскому интерфейсу. Ниже приведена сводная таблица мультимодальных ресурсов, используемых в представленных анализе словарях (таблица 1). Мультимодальность словарей оценивалась по наличию следующих модусов: *текст* (вербальная составляющая), *графика* (использование графических ресурсов: шрифт, размер, начертание, интервал), *цвет*, *аудио*, *изображение*, *визуализация языкового окружения лексемы* (семантические сети, сочетаемость, сферы употребления), *интерактивность* (гиперссылки). Наличие мультимодального ресурса в той или иной степени отображено в таблице знаком «+», а отсутствие знаком «-».

Таблица 1. Мультимодальные ресурсы русско- и англоязычных компьютерных словарей

Словарь / Модус	Карта слов	RuWordNet	Visual Thesaurus	Fine Dictionary
Текст	+	+	+	+
Графика	+	—	+	+
Цвет	—	—	+	+
Аудио	+	—	+	—
	(для версии EN.Карта слов)			
Изображение	—	—	—	+
Языковое окружение	+	+	+	+
Интерактивность	+	+	+	+

Как можно заметить, англоязычные компьютерные словари обладают сочетанием большего количества мультимодальных ресурсов. Так *The Visual Thesaurus* не хватает только наличия изображений-иллюстраций, но при необходимости пользователю доступен внешний запрос через встроенную поисковую функцию интерфейса словаря. В *Fine Dictionary*, в свою очередь, отсутствуют аудиоматериалы, что может стать значительным минусом для иностранного обучающегося. В целом можно отметить, что русскоязычные ресурсы обладают сочетанием меньшего количества мультимодальных ресурсов. *RuWordNet* имеет наименее привлекательный пользовательский интерфейс, из-за чего поиск слова не всегда интуитивен и может вызывать сложности как у иностранца, так и у носителя языка. *Карта слов* имеет значительные преимущества в этом плане по сравнению с *RuWordNet*. В общем выбранные словари имеют ряд плюсов и минусов, анализ которых ляжет в основу разработки структуры модели учебного мультимодального словаря в настоящем исследовании.

Результаты изучения теоретического материала и анализа существующих компьютерных лексикографических ресурсов показали, что мультимодальность является неотъемлемой составляющей повседневной жизни человека, включающей в себя особенности восприятия информации и коммуникацию. Таким образом, словарь, как образовательный ресурс, не может оставаться статичным справочником, состоящим только из текста. Требования современного общества и цифровизация ставят перед лексикографией новые задачи по созданию компьютерных мультимодальных ресурсов, включающих различные форматы представления информации, такие как аудио- и видеозаписи, иллюстрации, интерактивные элементы и другие. Компьютерная лексикография, как активно развивающееся научное направление, уже может представить большое разнообразие мультимодальных ресурсов, часть из которых была рассмотрена в настоящем параграфе. Завершая изучение исследований в области компьютерной лексикографии и особенностей мультимодального представления информации в словаре, в следующем параграфе мы подведем итоги первой главы.

Выводы по Главе 1

Изучение теоретических основ компьютерной лексикографии и теории мультимодальности, а также определение базовых понятий работы: *компьютерная лексикография, компьютерный словарь, мультимодальность, модус, модальность* позволило нам сделать следующие выводы:

1. *Компьютерная лексикография*, как научное направление, появилась ещё в 1960-х годах и начала активное развитие в 1990-х годах. В настоящее время разработка компьютерных словарей представляет собой перспективное направление современной лексикографии.

2. В рамках нашего исследования мы определяем *компьютерный словарь* как лексикографический ресурс в машиночитаемой форме, обладающий лексической базой данных для программ обработки естественного языка и оснащенный интерфейсом для пользователя-человека. Как и в традиционной лексикографии существует множество типов и видов компьютерных словарей, различающихся по своей структуре, размеру, способу представления информации, языковому материалу в своей основе и многим другим факторам.

3. *Мультимодальность* является важной составляющей компьютерного словаря. Термины *модус (mode)* и *модальность (modality)* используются для обозначения различных видов воспринимаемых объектов (зрительных, слуховых, тактильных и обонятельных). Возможности современных компьютерных технологий позволяют разработчикам создавать образовательные приложения и словари с использованием мультимодальных ресурсов.

4. Использование в словаре элементов различных модальностей, таких как текст, изображения, аудио- и видеоконтент, как полагают ученые, способны облегчить процесс изучения иностранного языка и повысить интерес и вовлеченность обучающихся.

Описанные выше особенности будут более подробно рассматриваться в следующей главе, в которой мы перейдем к изучению роли и функций мультимодального словаря в лингводидактических целях.

ГЛАВА 2. МУЛЬТИМОДАЛЬНЫЙ СЛОВАРЬ В ЛИНГВОДИДАКТИКЕ

2.1 Компьютерная лингводидактика

Мы уже не раз отмечали, что компьютеры, плотно вошедшие в повседневный мир современного человека, оказывают влияние на все сферы жизни, включая образование и обучение иностранным языкам. Поэтому современному преподавателю иностранного языка необходимо не только развивать иноязычную коммуникативную компетенцию студентов посредством аудиторных занятий, но и обучать их самостоятельному использованию информационных ресурсов, в том числе и компьютерных словарей. Методические основы, теорию и практику преподавания иностранного языка изучает лингводидактика. В рамках лингводидактики для нашего исследования более актуальным представляется изучение особенностей применения инструментов и продуктов компьютерной лексикографии в преподавании английского и русского языков как иностранных. В настоящем параграфе мы рассмотрим историю развития направления компьютерной лингводидактики и ее взаимодействия с компьютерной лексикографией.

Лингводидактика – относительно молодая научная дисциплина, которая возникла во второй половине XX века на стыке лингвистики, педагогики, методики, психологии и психолингвистики. Как наука, лингводидактика «представляет собой лингвистическую базу обучения иностранным языкам во взаимосвязи с перечисленными научными дисциплинами» [Шаклеин 2008, с. 24]. В большей степени данный термин употребляется в русскоязычной традиции, в том числе по отношению к преподаванию русского языка как иностранного (РКИ). Он был введен академиком Николаем Максимовичем Шанским в 1969 году [Шанский 1969; Габдулхаков, Гареева, Гарифуллина, 2012; Ярославская, Астремская, 2014]. Для английского языка более характерен термин *прикладная лингвистика* (applied linguistics).

Компьютерная лингводидактика, как можно предположить из определения, нацелена на эффективную интеграцию и использование информационных технологий и компьютерных средств в процессе обучения иностранному языку. В отечественной науке данный термин был предложен Ксенией Раймондовной Пиотровской в 1991 году [Пиотровская, 1991]. Рината Раисовна Зарипова определяет **компьютерную лингводидактику** как «область лингводидактики, изучающую теорию и практику использования компьютеров в обучении языку», отмечая, что использование информационных технологий в обучении языку принято разделять на обучение родному языку и обучение иностранному (второму) языку [Зарипова, 2014, с. 4]. Компьютерная лингводидактика включает в себя, но не ограничивается особенностями дистанционного обучения и организации самостоятельной работы студентов, мультимедийными технологиями и компьютерными средствами обучения, электронными изданиями и ресурсами, образовательными порталами и сайтами, электронными библиотеками, словарями, энциклопедиями и переводчиками. В теоретических и прикладных аспектах компьютерной лингводидактики выделяют три направления исследований: *разработка теоретических аспектов использования компьютеров в обучении языку* (методологические и психолого-педагогические проблемы, вопросы дизайна обучающих компьютерных систем и типологии учебных материалов, проблемы оценки качества и другие), *экспериментальная работа по созданию и использованию компьютерных учебных материалов и интеграция компьютерного обучения в общий процесс обучения языку* [Зарипова, 2014, с. 7].

В зарубежной традиции используется другая терминология, где широкомасштабное внедрение компьютеров в процесс обучения языку в 1980-е годы получило название *Computer Assisted Language Learning (CALL)* [Levy, 1997], а более усовершенствованная версия с использованием систем машинного обучения – *Intelligent Computer Assisted Language Learning (ICALL)* [Anderman, Rogers, 1996, с. 85]. CALL – это широкое научное направление, значительно крупнее, чем компьютерная лингвистика или компьютерная лексикография. Его развитию способствовал тот факт, что традиционные методы

изучения и преподавания иностранных языков требуют большого количества аудиторной работы с преподавателем, что не всегда может быть доступно обучающемуся по разного рода причинам. В то время как системы CALL становятся хорошей альтернативой, предоставляющей возможность самостоятельной языковой практики и дополнительные учебные материалы на иностранном языке, способствующие ускорению процесса обучения. Такие компьютерные образовательные приложения можно использовать как для самостоятельной работы дома, так и в дополнении к классическим занятиям с преподавателем [Mitkov 2004, с. 671-673].

Таким образом, возможности компьютерных программ позволили разработчикам образовательных приложений усовершенствовать процесс изучения иностранного языка не только в качественном, но и в количественном аспектах. Более того, в рамках компьютерной лингводидактики активно используются элементы различных модальностей, такие как текст, изображения, аудио- и видеоконтент, с целью облегчения процесса изучения новой лексики иностранного языка и повышения интереса и вовлеченности обучающихся. Так образовательный процесс перестал быть линейным и зависеть непосредственно от аудиторных занятий, а компьютерный словарь смог выступить самостоятельным обучающим ресурсом. В следующем параграфе мы изучим особенности учебной лексикографии и основные характеристики учебного словаря для изучения иностранного языка.

2.2 Учебный словарь для изучения иностранного языка

При обращении к словарям пользователи имеют разные потребности, которые зависят не только от их личных предпочтений, но также и от конкретной задачи поиска, их родного языка, национальности и уровня владения иностранным языком. Ряд ученых отмечает, что цели и частота использования словарей носителями языка и иностранными обучающимися различаются, и, как результат, содержание двух типов таких словарей тоже должно отличаться [Lew, 2010, с. 291].

В связи с этим многие годы ведется активная разработка учебных компьютерных словарей с акцентом на их использование при изучении иностранного языка, так называемых *learner's dictionaries*. В настоящем параграфе мы изучим историю развития и особенности учебных лексикографических ресурсов в призме их применения в преподавании иностранного языка.

Отдельным направлением лексикографии является разработка учебных словарей для разного рода пользователей, в том числе и для изучающих иностранный язык. Как отмечает Вера Дмитриевна Табанакова, учебная лексикография является «быстро развивающимся, едва ли не самым представительным разделом современной английской лексикографии» [Табанакова, 2001, с. 8]. Созданием учебных словарей специально для целей преподавания иностранных языков, начали активно заниматься в 70-80-е года прошлого столетия. Стремительное развитие произошло в связи с растущей популярностью исследований в направлении преподавания английского языка как иностранного: TESOL (Teaching English to Speakers of Other Languages) и EFL (English as a Foreign Language). В данном контексте с акцентом на компьютерную лексикографию ключевыми являются словари *Oxford Advanced Learners Dictionary of Current English* (OALDCE, 1974 г.), *Longman Dictionary of Contemporary English* (LDOCE, 1978 г.) и *Collins COBUILD* (1987 г.), новые форматы которых были разработаны именно для иностранцев, изучающих английский язык, и подчеркнули различия, существующие между двуязычными и одноязычными словарями, а также между словарями для носителей языка и для изучающих иностранный язык [Wilks, Slator, Guthrie, 1996, с. 5; Carter, McCarthy, 1988, с. 52].

В большей степени ввиду глобализации и других исторических причин это исследовательское направление представлено словарями для английского языка, но учебные словари для других языков также разрабатываются и играют важную роль в лингводидактике. В отечественной науке большой вклад в развитие учебной лексикографии для иностранцев внес Валерий Вениаминович Морковкин и его идея создания *Базового учебного словаря русского языка для учащихся определенной национальности* [Морковкин, Брандт, 1985]. Примером учебной

лексикографии для современного русского языка является работа Марины Николаевны Аникиной: *Русский словарь. Учебный словарь русского языка для иностранцев* [Аникина, 2014].

Лексикографу для создания эффективного учебного словаря важно понимать для чего иностранный учащийся может использовать такой ресурс. С этой целью в 1980 году департамент словарей Longman ELT провел исследование потребностей в использовании словаря учащимися средних школ и колледжей в Великобритании, Японии, Германии, США, Мексике и Нигерии на материале LDOCE. В рамках проекта был проведен анализ частоты и целей использования словаря, который показал, что иностранные учащиеся использовали словарь чаще, чем носители языка. Основными задачами были проверка орфографии, получение дефиниции слова и увеличение словарного запаса [Summers, 1988, с. 113]. Кроме этого, как отмечают другие ученые, при изучении иностранного языка словари можно использовать для проверки произношения, получения грамматической справки о частях речи, изучения этимологии, проверки существования того или иного слова или формы слова [Jackson, 1998, с. 192]. Делла Саммерс, лексикограф и издатель словарей Longman, считает, что учебный словарь может использоваться не только как инструмент восприятия новой информации на иностранном языке (receptive skills), то есть для чтения текстов и изучения новой лексики, но и, в дополнение к другим стратегиям обучения, для продуцирования речи (productive skills) [Summers, 1988, с. 111].

Отечественные ученые, в свою очередь, подчеркивают, что словарь играет важную роль при изучении новой лексики иностранного языка и с успехом может выступать в качестве учебного пособия. Так Калерия Викторовна Пиотух подчёркивает, что словарь обладает дидактической направленностью, так как «объектом дидактики является обучение, в ходе которого осуществляется образование, воспитание и развитие тех, на кого данный словарь ориентирован» [Пиотух, 2017, с. 185]. Гульнара Раисовна Чумарина отмечает, что словарь является ресурсом «универсальным, полифункциональным, реализующим

систематизирующую, нормативную, справочную, учебную функции» [Чумарина, 2010, с. 9].

С другой стороны, важным вопросом среди методистов остается целесообразность и частота использования словарей при изучении иностранного языка и их типы. Например, Джеймс Бакстер, отмечал, что двуязычные словари обычно используются на начальных этапах изучения языка, а по мере развития навыков студенты переходят на использование одноязычных словарей, так как длительная зависимость от двуязычных словарей, имеет тенденцию замедлять развитие языковых компетенций [Baxter 1980]. Ученые также отмечают, что изучающие иностранный язык могут испытывать трудности при интерпретации словарных статей, поскольку им требуется гораздо более подробная информация, чем носителям языка, в особенности о взаимосвязях между словами [Anderman, Rogers, 1996, с. 80]. Марина Анатольевна Бовтенко пишет о том, что в таком случае «учебный словарь должен предлагать описание лексики на основе потребностей конкретных групп учащихся с учетом их родного языка и ориентироваться на практическое владение языком – использование лексических единиц в речи» [Бовтенко, 2000, с. 28].

Кроме этого, важно отметить, что при создании учебного словаря и разработке его структуры необходимо учитывать основные особенности и этапы обучения новой лексике. Так овладение новыми словами и расширение словарного запаса состоит из двух последовательных частей: *восприятие* и *продуцирование*. **Восприятие** новой лексики включает в себя интерпретацию нового слова и фиксирование его в памяти. В то время как **продуцирование** требует извлечения этих знаний из памяти и использования слова в речи в соответствующей ситуации. Для успешного овладения новой лексикой студентам необходимо выполнять определенные упражнения по работе со словами, а преподавателю при этом необходимо учитывать эти особенности при составлении плана занятия [Nattinger, 1988, с. 64]. Например, как пишет Джуди Кегл, американский лингвист, в педагогической традиции США существует методика обучения словарным навыкам (*dictionary skills training*). В большей степени этот прием является

особенностью начального образования, но частично данные подходы были адаптированы и для преподавания английского как иностранного. Помимо обучения алфавиту с целью поиска слов, студентам предлагаются различные упражнения, предназначенные для ознакомления с использованием словаря в качестве инструмента для пополнения словарного запаса [Kegl, 1995, с. 257]. Среди таких упражнений можно выделить несколько основных, например, *задача разрешения неоднозначности*, в которой студенту дается предложение, содержащее многозначное слово: он должен найти это слово в словаре и определить, какое из множества значений является наиболее подходящим. Другим примером является *задача продуцирования*, в которой студенту необходимо использовать найденное слово или словосочетание в собственном предложении. Комбинацией такого рода задач может стать *упражнение на замещение*, в котором студента просят подобрать синонимы к выбранному слову так, чтобы это не изменило смысл исходного предложения [Kegl, 1995, с. 261-263]. Эти и другие упражнения способствуют расширению словарного запаса обучающегося и развитию навыков самостоятельной работы со словарем.

При этом несмотря на то, что словарь рассматривается как хранилище языка и культуры, мы должны помнить, что так или иначе он остается коммерческим продуктом, который должен отвечать запросам потребителей. Система образования стимулирует использование словарей различного рода отчего возрастает потребность в них на рынке. В результате разработчики современных компьютерных словарей вынуждены находиться в постоянной конкуренции, развивая и добавляя новые инструменты и функции, например различные мультимодальные ресурсы. Хорошо спроектированный компьютерный словарь имеет ряд преимуществ по сравнению с его предшественником – традиционным бумажным словарем [Cowie, 1983; Kipfer, 1987; Summers, 1988]. Например то, что раньше было трудоемким перелистыванием страниц по алфавитному указателю, теперь может стать простым щелчком мыши или даже наведением курсора на слово для получения информации посредством всплывающего окна. Кроме того, интеллектуальный интерфейс компьютерного словаря приведет слово к его

начальной форме (лемме), так что вероятность перепутать искомое слово с похожим или отвлечься на нерелевантную запись стремится к нулю. Однако при этом как преподавателю иностранного языка, так и обучающемуся необходимо критически подходить к анализу словаря, который они планируют использовать в учебных целях.

Подводя итог, можно сказать, что в настоящее время разработка словарей, в частности компьютерных, представляет собой перспективное и активно развивающееся направление современной учебной лексикографии. Компьютерные словари имеют большой потенциал использования в лингводидактических целях и в рамках подхода CALL. Так, например, они могут стать лингвистической базой данных для учебных приложений или же быть полностью интегрированы с другими автоматизированными интерактивными справочными материалами. При этом нельзя забывать и о традиционных форматах работы со словарем для изучения новой лексики иностранного языка и практики использования ее в реальной речи. В следующем параграфе мы разберем различные подходы к созданию и разработке учебных компьютерных словарей.

2.3 Создание и разработка учебного компьютерного словаря

Словарь, как лексикографическое произведение, может содержать не только лингвистическую информацию, но и отражать другие экстралингвистические данные, например особенности истории и состояние культуры стран носителей языка. Для совершенствования учебной функции, презентации расширенного объема образовательной информации и повышения эффективности работы со словарем необходимо разработать оптимальную структуру данных, позволяющую осуществлять поиск по заданным пользователем параметрам, а также создать интуитивно понятный интерфейс для удобной навигации. В связи с этим вопросы создания, разработки и использования словарей не теряют свою актуальность среди исследователей в течение длительного периода времени. В настоящей работе мы делаем акцент на создании оптимального лексикографического ресурса для

использования его на занятиях иностранного языка, поэтому настоящий параграф включает в себя обзор теоретических положений вопросов разработки словаря, изучение особенностей его структуры, а также описание принципов визуализации информации в нем.

Учеными как в России, так и за рубежом предпринимались многочисленные попытки создания словарей разной направленности. Особенности лингвистического и лексикографического моделирования, а также различные аспекты их изучения описываются в работах ряда исследователей: В.Д. Табанакова [2001], Н.А. Сивакова [2004], М.А. Ковязина [2006], Е.В. Ягаева [2006], А.А. Прошина [2008], А.В. Русакова [2008], Н.Г. Кантышева [2011, 2012], И.С. Волошина [2015], А.Ю. Левенкова [2016], А.А. Потапенко [2018] и др. Так при разработке лексикографических ресурсов изучались факторы, которые необходимо учитывать перед началом работы по созданию концепции словаря, ключевые элементы словаря, компоненты его структуры, особенности базы данных словаря, его функции, а также технические требования и инструментарий. При этом до настоящего времени не сложилось единой универсальной целостной системы, описывающей особенности и этапы моделирования словаря. Более того, вопрос совмещения мультимодальных инструментов, визуализации и историко-культурной информации в словаре остается неразработанным.

В русскоязычной терминографической традиции выделяют несколько основных функций словарей: *учебная, систематизирующая, справочная и нормативная*, при этом в классической лексикографии также говорят об *информативной, социальной* и ряде других функций, являющихся подвидами основных [Гринев, 1990; Мартынюк, 2004; Герд 2005]. Определение основных функций словаря является важным этапом в разработке его концепции и определении адресатов и сфер использования. Так японский ученый Макото Нагао, описывая особенности работы над **концепцией словаря**, подчеркивал, что важно с начального этапа разработки иметь четкое представление о типе создаваемого словаря. Для этого необходимо рассмотреть следующие **факторы**:

– *научная область будущего словаря*: лингвистический словарь; энциклопедический словарь широкой направленности; терминологический словарь по определенной предметной области и др.;

– *особенности организации словника и словарной статьи*: поиск осуществляется по отдельному слову, словосочетанию или фразе; каждое слово имеет собственную словарную статью, или узкие понятия объясняются в описании более широких понятийных терминов (в том числе насколько подробно дается толкование слова);

– *компоненты словарной статьи*: определение значения; примеры; деривативы; статистическая информация и др.;

– *уровень объяснения*: строгое научное определение; объяснение для непрофессионала или обычного пользователя; вводное объяснение для учащихся;

– *особенности доступа*: создание гиперссылок между словами; система и особенности индексации лексики;

– *основная цель использования*: глубокое научное понимание; быстрое поверхностное понимание; изучение особенностей употребления слова; проверка орфографии и произношения [Nagao, 1994, с. 399-400].

Изучение структуры словаря и содержания словарной статьи также вызывает множество вопросов среди ученых как в России, так и за рубежом. Например, американский лексикограф Сидни Ландау выделял следующие **ключевые элементы словарей** и других языковых справочников:

– *алфавитный указатель*: чаще всего словари упорядочивают по алфавиту, такое расположение упрощает поиск слова пользователем. Однако также существуют частотные словари или словари идеографического типа, принцип построения которых связан с частотой употребления лексемы или с ее семантическими и понятийными отношениями в языке;

– *количество заголовочных единиц словаря*: существует две традиции подсчета объема словаря: *британская*, в которой считается количество заголовочных слов, и *американская*, которая включает в статистику не только

заголовочные слова, но и другие слова или фразы, которые определены явно или неявно в тексте словарной статьи и обычно выделяются жирным шрифтом;

– *грамматическая справка*: особенно важна для пользователей, изучающих иностранный язык, может включать в себя информацию о части речи лексемы, переходности глаголов, падеже, числе и роде существительного и др.;

– *произношение*: наиболее широко используется фонетическая система международного фонетического алфавита (IPA), которая основана на манере артикуляции и может использоваться для воспроизведения звуков любого языка, даже того, с которым человек совершенно незнаком;

– *этимология*: чаще всего представлена в диахронических словарях, имеет цель продемонстрировать развитие формы и значения слова в языке в течение длительного периода времени;

– *синонимы*: демонстрация слов со схожей семантикой для расширения словарного запаса пользователя;

– *иллюстрации*: ряд объектов, таких, как архитектурные формы, животные, растения, геометрические фигуры, легче понять с помощью иллюстрации, чем словесного описания. Использование иллюстраций зависит от целевой аудитории словаря и может стать преимуществом для учебных словарей иностранных языков;

– *вводный материал и приложения*: для пользователя словаря, изучающего иностранный язык, краткое руководство в начале словаря является неотъемлемой частью для его понимания и использования. Приложения могут содержать различную дополнительную справочную, грамматическую и экстралингвистическую информацию (список неправильных глаголов, особенности орфографии, таблицы порядковых и количественных чисел, правила пунктуации, таблицы с мерами веса и длины) [Landau, 1984, с. 76-119].

Позднее Владимир Владимирович Дубичинский также описывал **универсальные составляющие словаря любого типа**, включая в них следующие элементы: *словник*; *заголовочная единица*; *толкование* (дефиниция или переводной эквивалент); *системность языкового материала*, описываемого в словаре;

нормативность – допустимая ненормативность языковых единиц; *определённый порядок подачи словарного материала* [Дубичинский, 1998, с. 32].

В структурном отношении, как мы отмечали ранее, в словаре выделяют *мега-*, *макро-* и *микроструктуры*. В работах В.В. Дубичинского описаны только две из них: микро- и макроструктура. **Макроструктуру** словаря образуют «общие принципы структуры лексикографического произведения: *вводная статья, предисловие, отношения и связи словарных единиц*» [Дубичинский, 1998, с. 37]. Рассматривая **микроструктуру**, мы говорим непосредственно о *словарной статье*, которая состоит из *заголовочной единицы* и её *описания*. Описание может включать *грамматическую справку, фонетическую характеристику, семантику дефиниций, этимологию, информацию о сочетаемости и употреблении лексемы, лексикографические пометы, дополнительную справочную информацию и другие примечания* [Дубичинский, 1998, с. 32].

С точки зрения компьютерной лексикографии словари более структурированы, чем любой другой текст. Команда итальянских ученых под руководством Николетты Кальцолари одними из первых начали разработку универсальной **структуры базы данных** словаря, отмечая, что в современном компьютерном словаре должны быть доступны следующие действия: *доступ к текстовым корпусам* (функция просмотра примеров, статистика), *лексическая база данных с функцией редактирования* (стандартизированная структура словарной статьи, просмотр имеющихся данных, добавление новой информации), *интегрированный доступ к другим ресурсам* [Calzolari, Picchi, 1994, с. 439]. В настоящее время нормой является соблюдение определенных особенностей структурирования базы данных и разметки текста, например TEI (инициативы по кодированию текста), совместимой с языками программирования SGML и HTML [Mitkov, 2004, с. 57].

Вопросы визуализации информации становятся актуальными и для компьютерной лексикографии в связи с растущей популярностью мультимодального формата репрезентации данных. Цифровые платформы активно используются в качестве новых ресурсов для проектирования учебной среды.

Ведущее преимущество таких платформ состоит в том, что они делают доступными модули, которые остаются недоступными для проектирования традиционных печатных словарей и учебников, например аудио, видео и анимацию [Bezemer, Kress, 2015, с. 68]. С точки зрения лингводидактики педагоги часто используют дополнительные материалы и различные средства визуализации с целью задействовать одновременно несколько чувственных рецепторов при изучении нового материала для успешного усвоения информации обучающимися. **Визуализация** здесь понимается как представление информации (текстовой или числовой) в виде изображений, диаграмм, графиков, таблиц, карт или схем. Визуальная информация дополняет текст или устную лекцию и улучшает восприятие нового материала, так как различные исследования ученых утверждают, что около 90% информации человек воспринимает через зрение [Линдгрэн, 1962; Тобажнянский, Романовский, Бондаренко, Пономарев, Черванева, 2005]. Исследователи выделяют множество различных техник визуализации в образовательном процессе, которые также могут быть имплементированы в компьютерный словарь, например, *таймлайн* («временная шкала, прямой отрезок, на который в хронологической последовательности наносятся события, *timeline*»), *интеллект-карта* («ментальная карта, диаграмма связей, карта мыслей, ассоциативная карта, *mind map*»), *скрайбинг* («визуализация информации при помощи графических символов, просто и понятно отображающих ее содержание и внутренние связи»), *инфографика* и многие другие [Сорока, Васильева, 2012].

Следующим важным этапом при создании компьютерного словаря является разработка его оболочки – *пользовательского интерфейса*, который играет ключевую роль в привлечении внимания к продукту и повышения вовлеченности пользователя. С этой целью интерфейс должен быть интуитивно понятным и простым в использовании. Под термином **интерфейс** в компьютерной лексикографии принято понимать именно пользовательскую оболочку словаря, которая включает в себя «средства взаимодействия пользователя с программой через графический дизайн, панели выбора меню, подсказки и др.»

[Лаврищева, 2007, с. 126]. На сегодняшний день самой распространенной формой взаимодействия человека с компьютерным словарем является *графический пользовательский интерфейс*, который отличается интуитивностью, наглядностью, простотой восприятия и ввода информации. Для взаимодействия с таким интерфейсом часто достаточно обычной компьютерной мыши.

В заключении можно сделать вывод, что вопросы создания и разработки словарей до сих пор вызывают множество дискуссий среди исследователей и ученых, работающих в области лексикографического моделирования. Это обусловлено как разнообразием научных лексикографических школ, так и широким спектром задач, для решения которых ведется разработка словарей. В следующей главе настоящего исследования мы предпримем попытку описать содержательные и структурные компоненты моделирования компьютерного учебного мультимодального словаря историко-культурологического типа для изучающих иностранный язык, которые могут внести вклад в развитие данного направления. Завершая изучение особенностей компьютерной лингводидактики и возможностей применения компьютерного словаря в целях изучения иностранного языка, в следующем параграфе мы подведем итоги второй главы.

Выводы по Главе 2

Изучение основных положений компьютерной лингводидактики и учебной лексикографии позволило нам сделать следующие выводы:

1. Компьютерные технологии активно используются в образовательных целях в общем, и для изучения иностранных языков в частности, данное направление получило названия *компьютерная лингводидактика* или *computer-assisted language learning*. Компьютерные словари могут выступать образовательными ресурсами для знакомства с новой лексикой иностранного языка, с помощью которых студент принимает активное участие в изучении и формировании концептуальных языковых структур, а также создает собственные семантические связи.

2. Учебный словарь, нацеленный на использование при изучении иностранного языка, должен обладать особыми характеристиками, направленными на развитие мотивации к самостоятельной учебной деятельности студента для овладения иноязычной лексикой. Такой учебный словарь становится *многофункциональным*, объединяя *справочную, систематизирующую, учебную и нормативную* функции.

3. Преимущества компьютерных словарей по сравнению с традиционными в преподавании иностранного языка были доказаны учеными ещё в 80-х годах прошлого века, среди них выделяют большой объем данных доступный для презентации более подробной информации о каждой лексеме, интеграцию мультимедийных ресурсов, связь с корпусами, семантические сети по типу WordNet и другие элементы.

4. Основные различия между одноязычными словарями для носителей языка и учебными словарями для иностранцев заключаются в особенностях и содержании информации в словарной статье. Для пользователя-иностранца основная цель – предоставить информацию о контексте функционирования лексемы и ее значениях, которая позволит в дальнейшем продуктивно использовать слово в реальной речи. Этим объясняется важность наличия в словарной статье достаточного количества примеров употребления слова в различных контекстах, которые необходимы как для расширения понимания языка учащимися, так и для предоставления формальной синтаксической модели использования лексемы в речи. Оптимальным ресурсом для этого являются материалы лингвистических корпусов и других компьютерных лексикографических ресурсов, представляющих тезаурусные иерархические структуры лексических отношений.

5. Перед началом разработки концепции словаря необходимо учитывать следующие особенности: *научная область* будущего словаря, *адресат* словаря, *основная цель использования*, *тематическая ориентация* словаря, *организация словника и словарной статьи* (макро- и микроструктурные характеристики), *принципы отбора единиц словника*, *компоненты словарной статьи*, *уровень*

объяснения, особенности доступа и другие элементы, влияющие на структуру и содержание будущего словаря.

6. Основными компонентами *макроструктуры* словаря являются *вводная статья, предисловие, структура организации словника*. *Микроструктура* словаря, или, другими словами, структура словарной статьи, представлена *заголовочной единицей* и её *описанием* (грамматическая информация, фонетические особенности, дефиниция, стилистические пометы, этимология, примеры употребления лексемы и ее сочетаемости с другими словами, другие примечания и дополнительная справочная информация).

7. Для улучшения учебной функции словаря могут быть использованы различные элементы *визуализации информации* в качестве дополнительных обучающих материалов. Выделяются следующие техники визуализации в образовательном процессе: *изображения, диаграммы, графики, таблицы, карты, схемы, таймлайн, интеллект-карты, скрайбинг, инфографика* и др. Такие дополнительные элементы словаря задействуют одновременно несколько чувственных рецепторов учащихся при изучении нового материала и способствуют его успешному усвоению. Инструменты визуализации реализуются в графическом пользовательском интерфейсе компьютерного словаря. Эти и другие мультимедийные ресурсы, такие как аудио-, видеоматериалы, анимации и гиперссылки, составляют *мультимодальность* репрезентации данных в компьютерном словаре.

Таким образом, компьютерный словарь, включающий в себя различные мультимодальные форматы представления данных (текстовый, визуальный, аудио, видео и т.д.) с опорой на лингвистические корпуса, может стать образовательным инструментом, способствующим интеграции и использованию информационных технологий в процессе обучения иностранному языку. В следующей главе мы перейдем к изучению особенностей семантизации историко-культурного компонента лексемы в компьютерном учебном мультимодальном словаре для иностранных учащихся, а также описанию содержательных и структурных компонентов моделирования такого словаря.

ГЛАВА 3. МУЛЬТИМОДАЛЬНЫЙ СЛОВАРЬ ИСТОРИКО-КУЛЬТУРОЛОГИЧЕСКОГО ТИПА КАК НОВЫЙ ТИП УЧЕБНОГО СЛОВАРЯ

3.1 Лингвокультурная лексикография в современном сопоставительном языкознании

Применение современных методик и инструментов в преподавании иностранных языков становится основным фактором успешного овладения языковыми компетенциями. Это особенно актуально в связи с тем, что языки различаются по своей структуре и другим лингвистическим характеристикам. При изучении иностранного языка важно понимать, что такие различия между языками могут быть связаны с особенностями в истории и культуре страны его носителей, которые изучаются в рамках сопоставительного языкознания. Разработчикам учебных словарей для изучающих иностранный язык важно учитывать такую культурную и историческую информацию при формировании словарной статьи для того, чтобы сделать словарь не просто лингвистическим справочником, а полноценным учебным ресурсом. В настоящем параграфе мы рассмотрим роль лингвокультурологической информации о слове при разработке учебного словаря для иностранных учащихся, а также возможность ее использования для сопоставительного исследования.

Всё, что окружает нас, общество, культура образуют тесно интегрированное целое, оказывающее прямое влияние на язык и его развитие. Слова в языке служат для обозначения «предметов, процессов, явлений объективной действительности, отраженных в общественном сознании», такой процесс называется *номинацией* [Беляевская, 1987, с. 24]. В данном процессе исследователи-лингвисты выделяют *первичную* и *вторичную* номинации. **Первичная номинация** – «соотнесение отражаемого в сознании фрагмента внеязыковой действительности и звуковой оболочки, впервые получающей функцию названия» [Серебренников, Уфимцева, 1977, с. 73]. Тогда как для нашего исследования более интересной

представляется **вторичная номинация** – «использование уже имеющихся в языке номинативных средств в новой для них функции наречения» [Будагов, 1974, с. 129]. В таком случае принято говорить о различиях в *денотате* и *коннотате* лексемы. **Денотат** «указывает на то, какие конкретно признаки предмета избираются в процессе номинации и закрепляются в языковой системе для его обозначения» [Беляевская, 1987, с. 29], тогда как **коннотат** «относится к ассоциациям, которые слово имеет сверх своего основного значения (денотата)» [Jackson, 1998, с. 58]. Таким образом, на развитие коннотата и процесс вторичной номинации слова могут оказывать влияние различные факторы: социальные, религиозные, этнические, географические, психологические, индивидуальные личностные, развитие науки и техники, медиа и другие.

Выражение личного оценочного отношения к объекту или предмету обсуждения также является частью коммуникации, в связи с этим часто говорят об эмоциональном значении слова (эмотивности), оценочности и интенсивности его коннотаций [Беляевская, 1987, с. 49-51]. Например, в любой культуре в определенную эпоху существуют слова, которые часто используют в контексте политики или сфере массовой информации для обозначения и разграничения положительных или отрицательных ценностей. В результате такие «эмоциональные» подтексты становятся более важными в слове, чем его прямое значение (например: *империализм, революция, свобода, республика, равенство, прогресс*). При этом важно отметить, что несмотря на разнообразие и субъективность коннотаций, они в первую очередь формируются и оказывают влияние на само слово (лексему), а не на предмет или процесс, которое оно обозначает. Более того, коннотации, разделяемые отдельной группой носителей языка или целым языковым сообществом, являются частью их культуры и могут быть «унаследованы» потомками вместе с самим языком [Jackson, 1998, с. 59].

Важность изучения истории и культуры страны изучаемого языка подчеркивают многие лингвисты, методисты и исследователи как в России, так и за рубежом на протяжении длительного периода времени [Верецагин, Костомаров, 1971; Щерба, 1974; Берков, 1975; Beheydt, 1987; Гак, 1988;

Иванищева, 2005; Петрушова, 2009; Ансимова, 2019; Полонникова, 2020]. В отечественной науке для описания данного направления используют термин *лингвострановедение*, предложенный в работе Евгения Михайловича Верещагина и Виталия Григорьевича Костомарова [Верещагин, Костомаров, 1971], а за границей в более обобщенном значении применяют термины *лингвокультурология* или *культурная лингвистика* (cultural linguistics) [Sharifian, Palmer, 2007]. Так или иначе основная цель заключается в использовании страноведческой информации в процессе изучения иностранного языка и ознакомлении учащихся с новой для них языковой культурой. Такие знания способствуют развитию иноязычной коммуникативной компетенции, навыков межъязыкового и межкультурного общения и адекватного понимания культуры речи собеседника или аутентичных текстов. В этом процессе словари играют важную роль, что подтверждает Ольга Леонидовна Петрушова, говоря, что «задача культурологического направления в учебной лексикографии состоит в лексикографическом представлении культурологической полифонии, имеющем целью преодоление возможных социокультурных барьеров» [Петрушова, 2009, с. 17].

Впоследствии в результате более глубокого изучения возможностей применения продуктов лексикографии в лингвострановедческих и лингвокультурологических исследованиях было предложено использовать термины *лингвокультурная лексикография* или *лингвокультурология*, под которым предлагается объединять лингвострановедческие, лингвокультурологические и этнолингвистические словари [Лукиянова, 2005; Ансимова, 2012; Агапова, Картофелева, 2014]. Ольга Константиновна Ансимова пишет, что «объектом **лингвокультурологии** являются языковые и неязыковые (например, мимика, жесты) единицы, содержащие культурную информацию, знание которой необходимо для успешной коммуникации в рамках определенного лингвокультурного сообщества; предметом – параметры описания данных единиц в соответствующих словарях, т.е. их лексикографическая интерпретация» [Ансимова, 2012, с. 12]. При этом исследователь отмечает, что главная цель таких

словарей – посредством представления лексики с преобладанием национально-культурной семантики помочь иностранным учащимся познакомиться с особенностями общественной жизни страны изучаемого языка, характерными явлениями повседневной действительности и культуры. В состав словника такого словаря могут входить лексемы, «характеризующие историю культуры страны, обладающие национально-культурной спецификой, имена исторических деятелей, географические названия, историко-культурные эргонимы, даты» [Ансимова, 2017, с. 72]. Однако лексика с историко-культурным компонентом может вызывать у иностранных студентов трудности в понимании в связи с тем, что, как считает Татьяна Анатольевна Дакукина, «с национально-маркированным словом вводится обширная культурно-страноведческая, лингвокультурная и социокультурная информация», отражающая особенности народа и страны изучаемого языка [Дакукина, 2015, с. 60]. Поэтому словари такого типа, имеющие цель визуализировать историю и культуру, требуют создания особой структуры и выбора определенного подхода в представлении лексической информации в словарной статье.

С другой стороны, лингвокультурные словари выступают неотъемлемой частью не только лингводидактики, но и сопоставительного языкознания и могут стать уникальными справочниками лингвистических и экстралингвистических данных о слове важных для контрастивного исследования. *Сопоставительное языкознание*, которое также называют *контрастивной лингвистикой*, направлено на исследование и описание языка через его системное сравнение с другим языком, а также выявление различий между двумя сравниваемыми языками [ЛЭС, 2002]. Поэтому лингвокультурная лексикография, как научное направление, также может внести значительный вклад в сопоставительные исследования, так как в свои задачи ставит фиксирование и описание не только лингвистических, но и культурных особенностей лексики, что позволяет провести анализ сходных и различных черт между несколькими языками.

Таким образом, учебный словарь, включающий в себя информацию как о культуре страны, так и об истории языка, может стать уникальным

образовательным ресурсом для иностранных учащихся. При этом инструменты компьютерной лексикографии дают технические возможности посредством мультимодальных ресурсов визуализировать на материале лексики культурные изменения в языке с течением времени. В следующих параграфах мы более детально изучим важность изучения историко-культурной информации о слове, а также особенности семантизации лексемы в словаре и на занятиях иностранного языка.

3.2 Историко-культурный (диахронический) компонент лексики: история вопроса

С течением времени и развитием языка слова получают способность накапливать знания об истории и культуре его народа-носителя, становясь своего рода лингвистическим культурным репозиторием. Вся накопленная информация отражается путем изменений в семантике лексемы. Чаще всего такие особенности и семантические характеристики слова не отражены в традиционных толковых словарях из-за своей избыточности. По этой причине для изучения семантических изменений в историческом контексте необходимо обращаться к специализированным этимологическим словарям или историческим тезаурусам. Учебные словари для изучающих иностранные языки также редко обладают этимологической информацией, так как имеют основную цель репрезентации современного живого языка в синхронии. В свою очередь, мы считаем, что для полноценного освоения лексического материала обучающемуся может быть также полезна информация об исторических семантических изменениях, которая будет способствовать более глубокому изучению культуры страны изучаемого языка и формировать ассоциативные связи, способствующие успешному запоминанию слов. В настоящем параграфе мы рассмотрим историографию вопроса диахронического изучения лексики с целью исследования историко-культурных особенностей ее значений.

Язык как живая система подвержен различным изменениям как в количественном, так и в качественном аспектах. Наиболее заметны исторические диахронические изменения в семантике лексических единиц языка, так как лексика является самой обширной подсистемой языка. Как писал профессор Александр Иванович Смирницкий: «слово выступает не только как единица словарного состава, но и как центральная, узловая единица языка вообще» [Смирницкий, 1952, с. 183]. Каждая историческая эпоха имеет собственные понятия и слова, которые воспринимаются через призму определенного периода, при этом с течением времени такие слова могут накапливать различные значения как слои или же, наоборот, изменять их под влиянием различных факторов [Дуринова, 2015]. Именно поэтому история значений слов интересует ученых-лингвистов на протяжении длительного периода времени.

О важности описания исторических семантических изменений лексики также писал профессор Джеффри Хьюз в своей работе по истории английского языка «*A history of English words*», отмечая, что каждое отдельное слово и весь лексикон в целом становится «историческим хранилищем» и «лексической картой» языка. Эта идея также была выдвинута ранее сэром Джеймсом Мюрреем, создателем Оксфордского словаря английского языка – *Oxford English Dictionary* (OED), основанного на исторических принципах описания лексики, который писал, что «изучая эволюцию английской лексики, мы обнаружим, что слова – это и окаменелости, в которых хранится культура прошлого, и жизненный организм, реагирующий на давление настоящего» [Hughes, 2000, с. 1]. Мюррей использовал образ словаря, подобного галактике и предложил структурировать всю лексику английского языка в виде схемы (рис. 10). Схема организации лексики в языке Дж. Мюррея (Murray's lexical configuration) или, другими словами, карта слов, по сути, изображает важное социолингвистическое понятие регистра (стиля), который отображает особенности выбора слова в зависимости от социальной ситуации или контекста [Hughes, 2000, с. 3]. Такая информация важна не только для лексикографов или лингвистов, но и для студентов, изучающих иностранный язык, в особенности обладающих высоким уровнем языковой компетенции.

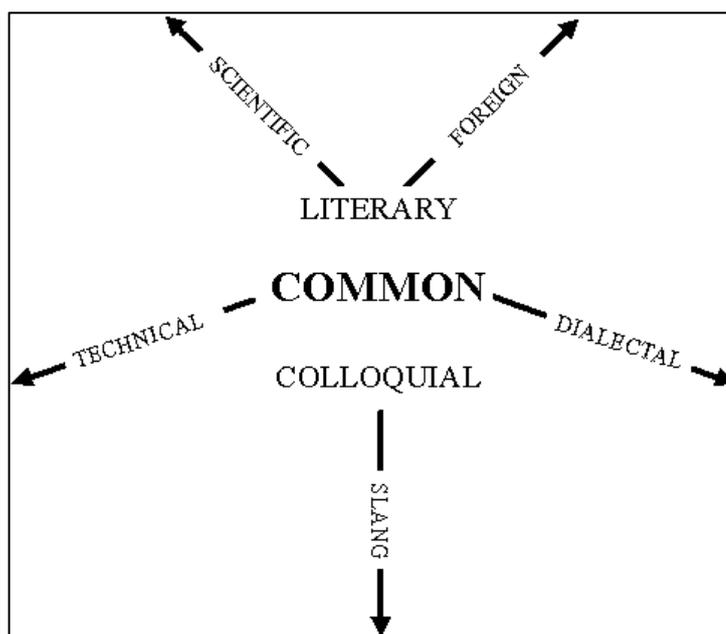


Рисунок 10. Структура организации лексики в языке согласно Дж. Мюррею

В целом можно сказать, что проблема диахронического изучения семантики – одна из старейших в языкознании, внимание лингвистов она начала привлекать в начале XIX века. Термин «*семантика*» для обозначения исследований, связанных с содержательной стороной понятия, впервые появился в трудах французского лингвиста Мишеля Бреалья, в частности в его работе 1897 года «*Essai de Sémantique*» [Breal, 1897; Беляевская, 1987, с. 10]. Учеными было выдвинуто два подхода к изучению значения слова: *семасиологический* (от слова к понятию) и *ономасиологический* (от понятия к слову). В нашем исследовании мы рассматриваем семасиологический подход. *Семасиология* как специальный раздел исторической лексикологии изучает вопросы, связанные с причинами и формой изменений смысловой структуры слов [Grygiel, 2005]. Исследования по этому направлению представлены в работах таких ученых как К. Рейзиг, Г. Пауль, А. Марти, М. Бреаль, С. Ульман, В.А. Звегинцев, А.А. Потебня, М.М. Покровский и других.

Считается, что первым начал развивать данное направление немецкий ученый Карл Рейзиг в своей работе «*Лекции по латинскому языкознанию*», опубликованной посмертно в 1839 году, в которой он выявил классификацию типов изменения значений слов (метафора, метонимия, синекдоха, изменение переходности глаголов, переход пространственных отношений во временные,

изменения значения слов с предлогами). Основной целью такого изучения было совершенствование процедуры разработки словарей, при которой бы использовались семантические принципы при составлении каждой словарной статьи [Reisig, 1839; Grygiel, 2005; Колокольникова, 2012]. Таким образом, можно отнести начало активного изучения изменений в семантике слова к середине – концу XIX века.

Начиная со второй половины XIX века широко распространилось психологическое направление в лингвистике, представители которого рассматривали слово как символ психических явлений в жизни человека и противопоставляли языковую норму (узуальное) и индивидуальные особенности (окказиональное). Одним из представителей данного направления был немецкий лингвист Герман Пауль, который считал, что «изменение значения слов осуществляется путем отклонений в индивидуальном употреблении узуального и путем превращений такого индивидуального употребления в узуальное» [Пауль, 1960, с. 94].

Мишель Бреаль, в свою очередь, занимался поиском закономерностей в семантических изменениях. Он вывел несколько законов, в частности закон распределения / размежевания синонимов, согласно которому слова с тождественными значениями развивают различия и спустя время перестают быть взаимозаменяемыми [Breal, 1897; Колокольникова, 2012]. Позднее исследователи *Школы «слов и вещей»* Рудольф Мерингер и Гуго Шухардт развивали идеи о том, что для понимания семантики и истории слова необходимо также ставить задачу изучения экстралингвистических факторов, таких как история самой вещи, которую оно обозначает. Выделялись три группы причин семантических изменений: исторические, социальные и языковые [Meringer, 1904; Шухардт, 1950].

В отечественной школе основополагающей является работа Александра Афанасьевича Потебни «О связи некоторых представлений в языке», где ученый высказал идею об исторических изменениях целых рядов слов, группирующихся вокруг одних и тех же представлений при анализе смыслового родства рядов слов

[Потебня, 1864]. Он также разработал теорию «внутренней формы» слова, где под значением слова понимается его этимология так, что многозначность слова не предполагается возможной [Беляевская, 1987, с. 13]. В дальнейшем Михаил Михайлович Покровский подчеркивал, что семасиологические процессы распространяются на целые категории слов и выдвинул идею их изучения как целостной системы, так как семасиологические изменения объединяют социальные и психологические факторы [Покровский, 1959]. Он предложил сопоставительный семантический анализ единиц языка, который позволял сравнивать друг с другом слова близкие по значению или имеющие общность обозначающего понятия [Беляевская, 1987, с. 16]

Позже английский ученый Стивен Ульман обращал внимание, что для анализа семантических изменений в семасиологии важны оба плана: синхронический и диахронический, он пытался выявить «универсалии исторической семантики». Для классификаций семантических изменений он использовал сопоставительную методику и выделял исторические (внелингвистические) факторы, лингвистические причины с психологической интерпретацией и другие явления, такие как табу, эвфемизмы и ухудшения значений [Ульман, 1970].

Во второй половине XX века становится очевидно, что информация об изменениях и эволюции языка играет важную роль в понимании его нынешнего значения и функции, так как в семантической структуре слова содержатся сведения, хранящие «память» о ее прошлых состояниях. Об этом писал академик Виктор Владимирович Виноградов, подчеркивая важность применения методов синхронии и диахронии совместно, исходя из их взаимосвязи друг с другом [Виноградов, 1999]. Он также ввел понятие «смысловой структуры слова как множества и единства его лексико-семантических вариантов (ЛСВ)» [Беляевская 1987, с. 22]. В дальнейшем изучением лексико-семантического варьирования занимались такие советские ученые как О.С. Ахманова, И.В. Арнольд, А.А. Уфимцева, К.А. Левковская и другие [Ахманова, 1957; Уфимцева, 1962; Левковская, 1962; Арнольд, 1966]. В последствии появилась

концепция создания каталога семантических переходов, с 1990-х годов ее разработкой занималась группа ученых: А.А. Зализняк, М.С Булах, Д.С. Ганенков, И.А. Грунтов и другие [Зализняк, 2013]. По мнению Анны Андреевны Зализняк диахронические исследования семантических изменений лексем способствуют более глубокому изучению процессов семантической деривации (порождения производного значения), семантического сдвига (перехода) и семантической параллели (смещения значения в нескольких словах) [Зализняк, 2001].

В настоящее время на изучение историко-культурных характеристик лексики обратили внимание специалисты и исследователи компьютерной лингвистики, где одним из первых этапов в изучении изменений в семантике стал дискурсивный поиск употребления лексем в необычном контексте или количестве. В задачах обработки текстов на естественном языке (Natural Language Processing, NLP) можно говорить о таких направлениях, как *дистрибутивная (векторная) семантика* (word embeddings), *культуромика* и *цифровая гуманитаристика*, которые позволяют изучить языковые данные и оценить результаты исследования статистическими методами [Антоненко, 2011; Миронов, 2012; Беликов, 2016; Башмакова, 2019].

Дистрибутивная семантика – это область лингвистики, которая «изучает способы определения семантической близости слов на основе их распределения в большом корпусе текстов» [Потапенко, 2018, с. 13]. Для определения такой близости исследователю необходим большой массив лингвистических данных (корпус), в котором каждому слову присваивается свой индивидуальный *контекстный вектор* – числовой набор координат в многомерном пространстве. Семантическое расстояние между словами вычисляется как косинусное расстояние между такими векторами. Другими словами, понятия, схожие по значению или связанные одним тематическим полем, будут находиться рядом друг с другом в многомерном словесном пространстве всего корпуса, так как будут иметь схожий набор векторных координат, и обратная ситуация складывается для семантически не связанных слов. В основе этой теории лежат математические вычисления линейной алгебры и матричный подход к тексту [Моченов, Бледнов,

Луговских, 2006]. Впоследствии ряд ученых, Д. Журафски, У. Гамильтон, Ю. Лесковец, А.Б. Кутузов и другие, начали исследование возможностей применения контекстных векторов для изучения семантических изменений в диахронии [Hamilton, Leskovec, Jurafsky, 2016; Kutuzov, 2018]. В такой модели использовались показатели косинусной близости для пар лексем на материале корпуса одного исторического периода для сравнения с корпусом другого исторического периода, так как контекстные векторы одного и того же слова в разных корпусах не совпадают, а косинусное расстояние между векторами одной и той же пары лексем может отображать даже небольшие изменения в их контекстном употреблении.

Научное направление **культуромика** начало развиваться в течение последних десятилетий и стремительно становится популярным в сфере гуманитарных исследований. Его рождение стало возможно благодаря появлению в машиночитаемой форме огромного числа культурно значимых текстов. Впервые этот термин появился в статье «*Quantitative Analysis of Culture Using Millions of Digitized Books*» американских исследователей Жана-Батиста-Мишеля и Эреза Либермана Эйдена, которые впоследствии создали проект Google Ngram Viewer, помогающий проанализировать цифровую библиотеку для изучения культурных моделей в использовании языка в течение заданного периода времени [Lieberman, 2007; Michel, 2011]. Результаты исследования, опубликованные в журнале Science [Michel, 2011], показывают тенденции изменения лексики английского языка в период с 1900 до 2000 гг. Так в начале века насчитывалось 540 тысяч слов, к середине – 600 тысяч, а к концу тысячелетия более миллиона слов литературной речи. Для сравнения приводится крупнейший словарь Вебстера, который насчитывает сейчас лишь 350 тысяч английских слов [Michel, 2011]. Такие данные показывают, что большая часть слов живого литературного языка ускользает от лингвистов. Среди других результатов исследования: прослеживание эволюции грамматики, выявление подъемов и спадов внимания к различным объектам и понятиям, а также возможность обнаружения деятельности цензуры. Компьютерный анализ текстов позволяет объективными

математическими методами исследовать вопросы, которые ранее были доступны только для гуманитарных подходов.

Для демонстрации того, как работает Google Ngram Viewer рассмотрим поисковый запрос четырех лексем английского языка из тематического поля «Education»: *study*, *research*, *learn*, *teach* (рис. 11). На графике представлен период с 1800 по 2019 год. Каждая из цветных линий показывает частотность употребления лексемы в указанный период времени в процентном отношении ко всему корпусу. Можно обратить внимание, что частотность употребления лексем «*learn*» и «*teach*» относительно стабильна на протяжении всего графика и не показывает резких изменений, тогда как «*study*» и «*research*» имеют значительный рост и ряд резких скачков. Для более детальной интерпретации таких статистических данных исследователи обращаются к изучению исторического и социального контекста той эпохи, где наблюдаются нетипичные изменения в частотности употребления. После качественного анализа таких данных появляется возможность сделать предположения или выводы о причинах изменений или культурных трендах, повлиявших на семантику слов.

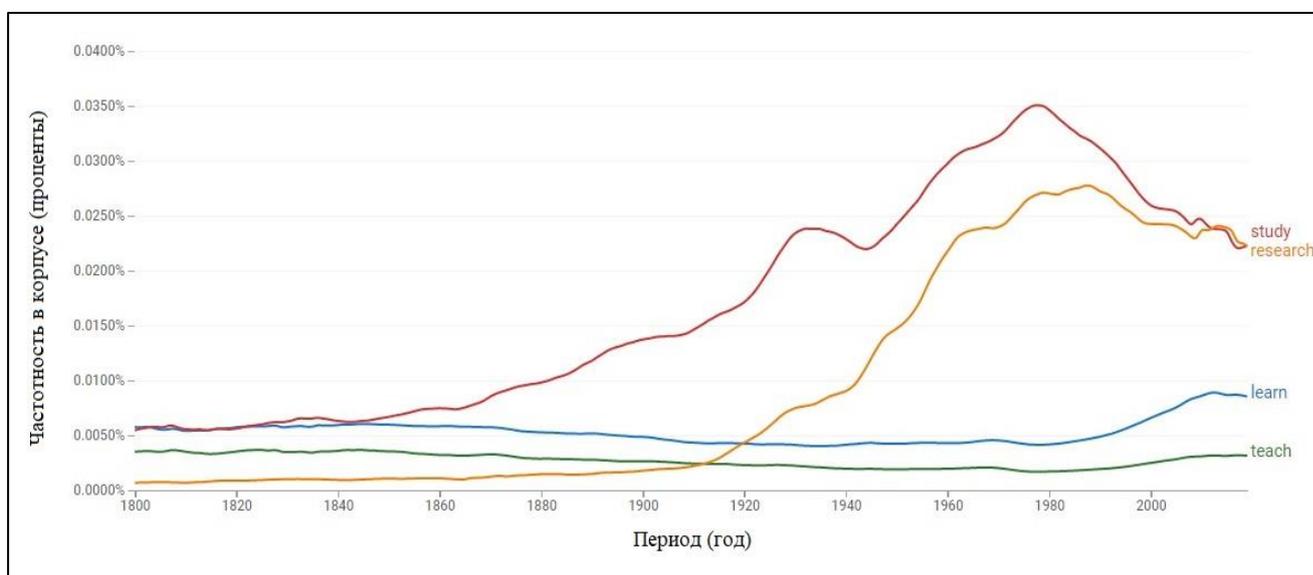


Рисунок 11. Google Ngram Viewer.¹¹ Лексемы «*study*, *research*, *learn*, *teach*»

Среди исследований семантики методами культуромики можно также выделить работы исследовательской группы Språkbanken Гётеборгский университета (Швеция). В противопоставление пилотному проекту команды

¹¹ Google Ngram Viewer, URL: <https://books.google.com/ngrams>

Google, который критиковали за недостаточную связь и опору на современные работы в области языкознания, исследователи предлагают так называемый *knowledge-based* подход, основанный на методологии лингвистических исследований. Цель работы заключается в изучении и создании инструментов для обработки большого объема оцифрованных текстов на шведском языке. При этом основное внимание уделяется семантическому анализу: распознаванию именованных сущностей (NER), определению семантических отношений, кореференции и др. [Borin, 2013; Tahmasebi, 2015]. Такой подход также использовали и другие исследователи для изучения трендов в дискурсе СМИ на материале газетных текстов [Phani, Lahiri, Biswas, 2012; Suchanek 2014], для визуализации геопространственных изменений в тональности текстов посредством создания *тепловых карт* (heatmaps) методом GIScience [Shook, 2012], для выявления культурно-обусловленных диахронических сдвигов в значении слов методами векторной семантики [Kutuzov, 2018; Rodina, 2019; Fomin 2019; Kutuzov, 2020].

В России исследованиями в области **цифровой гуманитаристики** занимается Центр цифровых гуманитарных исследований Национального исследовательского университета «Высшая школа экономики» под руководством Анастасии Александровны Бонч-Осмоловской. Сотрудники центра уделяют особое внимание компьютерному анализу художественного текста, созданию жанровых корпусов и электронных инструментов для гуманитарных исследований. Среди реализованных проектов можно выделить «Имена Времени» в 2014 году, посвященный анализу сочетаемости эпитетов русского языка на материале выборки из Национального корпуса русского языка. Исследователи рассматривали конструкции, включающие название десятилетия и прилагательное, определяющее его, предполагая, что частотные списки сочетаемости слов помогут восстановить мнемонический образ каждого из десятилетий советской и постсоветской истории. Различная семантика прилагательных формирует ассоциативные ряды и показывает разницу между восприятием десятилетий, в результате получились следующие словосочетания: *атеистические двадцатые, страшные тридцатые,*

стиляжные пятидесятые, романтические шестидесятые, лихие девяностые и т.д. [Бонч - Осмоловская, 2018]. Для презентации проекта была подготовлена инфографика, наглядно иллюстрирующая изменения в употреблении культурно значимых для описания российской действительности эпитетов и определений. Главный вывод исследователей заключается в том, что политические режимы и исторические события накладывают свой отпечаток на особенности вербализации культурных ценностей [Бонч-Осмоловская, 2015].

Таким образом, семантика слова являлась важным предметом изучения лингвистов на протяжении длительного периода времени. Несмотря на существование различных методов и подходов к изучению значения слова, его характеристик и содержания, семантические особенности лексемы в диахронии являются важной информацией для процесса семантизации новой лексики иностранного языка, в том числе при составлении словарей, в особенности для изучающих иностранный язык. Рассмотрев историю изучения семантического развития лексики и влияние социально-культурных факторов на процесс изменения языка, для целей настоящего исследования мы можем определить **историко-культурный (диахронический) компонент** значения лексемы как *совокупность информации об истории семантического развития слова, национально-культурных коннотациях, контексте и особенностях его употребления в определенную историческую эпоху, которая описывает слово как достояние определенной лингвокультуры.*

В завершение отметим, что, на наш взгляд, знание семантических особенностей в контексте употребления слов является важным аспектом при изучении иностранных языков. В учебных целях инструментом визуализации такой информации может стать компьютерный словарь, включающий в себя различные элементы (текст, изображения, аудио, графики и др.). В следующем параграфе мы исследуем способы семантизации лексики в учебном словаре с целью визуализации историко-культурного компонента значения лексемы в диахронии.

3.3 Способы семантизации лексики в учебном словаре

Освоение новой лексики иностранного языка является одной из самых важных составляющих развития иноязычной коммуникативной компетенции. Как отмечал Борис Васильевич Беляев ещё в середине прошлого столетия: «из всех основных аспектов иностранного языка, которые должны практически усваиваться учащимися в процессе обучения, наиболее важным и существенным с психологической точки зрения следует считать лексику, потому что без запаса слов, хотя бы и незначительного, владеть языком невозможно» [Беляев, 1965, с. 118]. Однако данный аспект и обучение ему может повлечь за собой определенные трудности, вызванные различными факторами: *влиянием родного языка, когнитивными способностями учащегося, психологическими особенностями* восприятия информации, а также *уровнем языковой подготовки* [Кузьмина, 2019]. Именно поэтому при разработке учебного словаря для иностранных учащихся необходимо правильно подбирать способы описания значения лексемы в словаре и инструменты визуализации всех ее лингвистических и экстралингвистических особенностей. В настоящем параграфе мы изучим способы семантизации лексики в учебном словаре для иностранных учащихся с целью описания историко-культурных особенностей ее значений.

Процесс описания значения лексемы в словаре, а также объяснение новой лексики иностранного языка во время занятия с преподавателем называется *семантизацией*. Посредством *семантизации* новая для обучающегося лексема приобретает свое значение и становится осмысленной языковой единицей, которую в дальнейшем можно использовать в речи, при этом студент «не только узнает лексическое значение слова, но и знакомится с особенностями его употребления (сочетаемость, использование в контексте)» [Кузьмина, 2019, с. 172]. Вилфрид Деку, изучая вопрос семантизации с точки зрения дидактики, отмечал, что проблема передачи смысла при изучении иностранного языка стояла перед учеными и переводчиками на протяжении длительного периода времени. Начиная с античных времен менялись подходы и методы, но необходимость изучения

особенностей лексической семантики в иноязычной коммуникации и переводе только усиливалась в условиях растущей глобализации [Decoo, 1985]. Так, например, в последние годы с развитием сетевых технологий, семантизацию стали рассматривать также и как процесс представления семантической информации доступной в сетевой среде, который потенциально позволяет более точно осуществлять поиск и машинную обработку данных на естественном языке в сравнении с традиционными методами организации данных [Sosińska - Kalata, 2014].

В настоящем исследовании мы используем термин **семантизация** в контексте лингводидактики, поэтому для современного определения данного понятия обратимся к *Новому словарю методических терминов и понятий (теория и практика обучения языкам)*, в котором он определяется как «выявление смысла, значения языковой единицы; процесс и результат сообщения необходимых сведений о содержательной стороне языковой единицы» [Азимов, Щукин, 2009]. Александр Николаевич Шамов, в свою очередь, определяет *семантизацию* как «раскрытие значения слова и фразеологического сочетания, показ особенностей его употребления» [Шамов, 2008]. В зарубежных исследованиях семантизации важной считается работа бельгийского лингвиста Лудо Бехейдта 1987 года «*The semantization of vocabulary in foreign language learning*». В контексте психолингвистического подхода он определял *семантизацию* как процесс изучения форм слова с их многозначностью в различных контекстах употребления. При этом ученый подчеркивал важность морфологической и синтаксической информации о лексеме, также как и данных о контексте ее употребления и коллокациях, как важных компонентов процесса семантизации, без которых учащийся не может в полной мере семантизировать слово [Beheydt, 1987, с. 55].

Сам процесс семантизации, как отмечают исследователи, включает в себя *две* последовательные *фазы*: *первая* предполагает идентификацию денотативного значения слова, *вторая* объясняет национально-культурный компонент семантики данной лексической единицы [Neuner, 1994, с. 58]. Также семантизация может быть подвержена влиянию *двух факторов*: языковых особенностей лексики

и психологического процесса обучения [Stowell-Ruzicka, 1991]. В целом ученые различают два способа семантизации лексических единиц: *переводный* (однословный перевод лексемы на родной язык или перевод с дополнительным объяснением значения слова) и *беспереvodный* (использование вербальных средств иностранного языка или других невербальных инструментов).

Переводный способ признается традиционным, в особенности на начальных этапах изучения иностранного языка, однако он имеет ряд недостатков, например «часто встречающаяся асимметрия лексических единиц двух языков, связанная со спецификой культурных компонентов слов, которая не дает полного представления о содержательной стороне лексической единицы» [Дакукина, 2015, с. 60]. В нашем исследовании мы делаем акцент на одноязычных словарях с наличием мультимодальных ресурсов, поэтому предпочтительным становится именно **беспереvodный способ** семантизации, который требует больше времени для объяснения и понимания, а также может вызывать трудности на начальных этапах обучения иностранному языку, но при этом имеет ряд значительных преимуществ, так как «развивает языковую догадку, увеличивает практику в языке, создает опоры для запоминания, усиливает ассоциативные связи» [Дакукина, 2015, с. 60]. Обобщая различные классификации ученых, можно выделить следующие приемы *беспереvodной семантизации*:

– *описательное вербальное толкование значения слова на иностранном языке*. Основной классический метод, применяемый в учебных словарях, но он может использоваться только при условии минимального владения лексикой и грамматикой изучаемого языка;

– *наглядная визуализация*. Самый обширный подход, включающий в себя использование изображений, фотографий, карточек-картинок, коллажей, контекстных ситуационных иллюстраций, презентаций, видео, анимации, демонстрацию реальных объектов, жестов и движений в качестве примеров и др. Такой способ подходит для всех уровней владения иностранным языком и особо эффективен на начальном этапе, когда учащиеся ещё не владеют большим словарным запасом, а также для демонстрации объектов или реалий, которые

тяжело было бы описать словами (животные, еда, здания, цвета, геометрические формы);

– *словообразовательный анализ формы слова.* Обучение словообразовательному анализу лексем помогает учащимся без перевода на родной язык определять грамматическую форму слова, выделять морфемы и предполагать их значение, что является основой для дальнейшего самостоятельного понимания новых слов;

– *этимологический анализ лексемы.* Изучение особенностей происхождения слов может стать дополнительным инструментом при словообразовательном анализе, так как знания об исторических изменениях в структуре лексем и их семантике может помочь учащимся выявить схожие и различные черты новых слов иностранного языка, предполагая их значение без перевода;

– *анализ контекстного употребления.* Студент учится строить предположения о значении лексемы по контексту ее употребления, в том числе на материале корпусных данных и частотности употребления. Контекст, в котором употребляется слово, может рассказать много о его значении по анализу семантики соседних слов. Такой подход развивает языковую догадку, которая необходима для успешного овладения иностранным языком;

– *группировка лексики на основе семантических отношений* (лексико-семантические и тематические поля, синонимические и антонимические ряды, родовидовые понятия, однокоренные слова). Такой метод систематизирует знания учащихся на основе сходства или различия семантики слов, помогая формировать свой собственный тематический тезаурус изученной лексики;

– *демонстрация устойчивых ассоциаций,* связанных со словом. Данный подход является производным, или своего рода сочетанием, изучения контекстного употребления лексемы и ее семантических отношений, что способствует установлению абстрактных логических понятийных отношений и парадигматических связей слов. Для изучения новых слов в рамках этого подхода

предлагается использовать ассоциограммы, визуализацию семантических сетей, коллажи с представлением примеров коллокаций и типичных словосочетаний;

– *использование лингвострановедческой информации.*

Экстралингвистические данные о культуре и традициях страны изучаемого языка способствуют более успешному усвоению национально-маркированной лексики, например реалий, которые не имеют адекватного эквивалентного перевода на родной язык. Здесь могут использоваться лингвокультурные справочники, лингвострановедческие или этнографические комментарии, документальные фильмы и др. [Beheydt, 1987; Процик, 2007; Дакукина, 2015; Юсупова, 2017; Кузьмина, 2019].

Кроме этого, как предлагают ученые, частью процесса семантизации при изучении новой лексики иностранного языка, помимо использования учебного словаря, может стать индивидуальная работа студента, например ведение собственного глоссария, создание и использование карточек со словами, заметки на полях страниц учебника или любого другого текста [Lawson, Hogben, 1996]. Так, например, в контексте компьютерной лексикографии существует исследование по изучению эффективности различных подходов семантизации лексики с использованием данных конкордансов для создания учебных глоссариев [Poole, 2011]. При этом важно отметить, что выбор приема беспереводной семантизации зависит от ряда разных факторов, как лингвистических, например индивидуальных характеристик слова, его формы и значения (абстрактное / конкретное / реалия), так и психолого-педагогических, таких, как возрастные особенности, уровень владения и степень способностей к изучению иностранного языка, а также внутренняя мотивация обучающегося [Айдарова, 2013].

Лексика языка отражает многогранность своих значений посредством употребления в устной речи и письменных текстах таким образом, что проследить какие-либо историко-культурные изменения представляется возможным только при анализе контекста. В связи с этим различные форматы объединения слов на основе их лексического окружения могут выступить преимуществом для процесса

семантизации лексики иностранного языка. Исходя из этого важной частью выбора способа семантизации лексики является ее системный анализ, что подчеркивает Анна Анфилофьевна Уфимцева, которая пишет, что «единицей, подвергшейся историко-семасиологическому рассмотрению, является слово во всем многообразии его лексико-фразеологических форм и семантико-синтаксических связей в системе языка» [Уфимцева, 1962, с. 4]. В данном контексте учеными выделяются различные способы объединения лексики на основе системных семантических отношений, существующих между словами: *семантические классы, синонимические и антонимические ряды, лексико-семантические группы и поля, тематические поля и группы* и другие [Глобина, 1995].

Первоначально теория *семантического поля* разрабатывалась немецкими филологами: в *ономазиологическом* подходе Йостом Триром, который делал акцент на парадигматические отношения лексем, и в *семасиологическом* подходе Вальтером Порцигом, который считал важным выявлять синтагматические отношения лексики [цит. по Лысякова, 2005, с. 5]. Позднее при описании лексических единиц их стали также объединять в *тематические группы*, которые были описаны в работах Ольги Сергеевны Ахмановой, Ирины Владимировны Арнольд и других ученых [Ахманова, 1957; Арнольд, 1966]. Так Федот Петрович Филин противопоставлял *тематические группы* (ТГ) *лексико-семантическим* (ЛСГ). Основная причина состояла в том, что лексемы, входящие в ЛСГ, объединены по принципу семантической однородности значений, а в ТГ такие отношения основываются «не на лексико-семантических связях, а на классификации самих предметов и явлений» [Филин, 1957, с. 526]. В связи с этим в тематические группы допустимо объединять слова с экстралингвистической семантикой, которые обладают только общностью предметной сферы и не обязательно схожим лексическим значением. Отличие тематического поля от группы в большинстве случаев проявляется в большем размере первого. Для целей настоящего исследования мы будем понимать **тематическое поле (группу)** как *совокупность слов разных частей речи, семантически объединенных одной общей темой или связанных с определенным понятием в широком смысле.*

В лингводидактических целях, для задач семантизации новой лексики, наиболее рациональным представляется объединение слов именно в тематические поля и группы, которые обладают наибольшей системообразующей силой и соответствуют практическим целям изучения иностранного языка, так как лексемы, входящие в них, могут находиться не только в отношениях синонимии и антонимии, но и в связях ассоциативного характера. Как подчеркивает Ольга Николаевна Иванищева, «тематический принцип подачи материала создает условия для более полного представления культурного компонента значения слова» [Иванищева, 2005, с. 15]. Кроме этого, Наталья Александровна Аксарина отмечает, что «именно изучение системных отношений в лексической семантике играет важнейшую роль в грамотном и осознанном формировании индивидуального лексикона личности». При этом «лексико-семантические и тематические группы слов, организованные в лексико-семантические поля, выстраивающие всю лексическую систему языка в его понятийной сложности, обладают высочайшим лингвометодическим потенциалом» [Аксарина 2015, 20], а работа с лексико-семантическими и тематическими группами позволяет студентам развить навыки самостоятельной работы по изучению языка, семантизации и определению родовидовых отношений и связей слов друг с другом. В настоящей работе мы остановились на использовании именно *тематического* формата группировки лексических единиц ввиду ряда описанных выше преимуществ.

Обобщая теоретические положения, исследователи отмечают, что в качестве *учебных словарей* для целей изучения иностранного языка более предпочтительны *тезаурусы*, которые стремятся показать концептуальные и тематические связи между лексемами в формате лексических / семантических полей и групп [Anderman, Rogers, 1996; Jackson, 1998]. Это вызвано тем, что традиционные словари имеют определенные недостатки, которые связаны с организацией содержащегося в них лексического материала. Например, алфавитный порядок расположения словарных статей из-за своей произвольности не показывает лексические и семантические отношения между словами, которые важны для

иностранных обучающихся. Кроме этого, студенты нуждаются в информации, которая включает особенности использования слова и его значений в зависимости от контекста или типа текста: синтагматические и парадигматические отношения, а также другие прагматические аспекты.

Также важно учитывать, что для восприятия новой информации, например изучения лексики иностранного языка, у студентов должны быть задействованы разные рецепторы: визуальные, аудиальные, кинестетические или различные их комбинации. Как отмечает Елена Олеговна Кузьмина «исходя из знания специфики каждого вида восприятия, преподаватель должен использовать разные способы объяснения новой информации» таким образом, чтобы все студенты получили возможность лучше понять информацию и запомнить ее [Кузьмина, 2019, с. 168]. При изучении лексики основной базой данных и инструментом являются словари во всём их многообразии видов и форм. Именно поэтому одной из задач современной компьютерной лексикографии является создание оптимального компьютерного учебного словаря для изучающих иностранный язык, способного облегчить процесс семантизации новой лексики. Такой словарь должен включать в себя мультимодальные ресурсы и задействовать максимальное количество рецепторов обучающегося.

Подводя итог, можно сделать вывод, что инструменты компьютерной лексикографии имеют значительные преимущества в разработке учебного словаря для изучающих иностранный язык. Возможность использования мультимодальных ресурсов в современном компьютерном словаре является важным инструментом для беспереводной семантизации историко-культурных характеристик лексемы в словарной статье, а также репрезентации семантических изменений значения слова и связей языкового окружения лексики. В следующем параграфе мы опишем содержательные и структурные компоненты моделирования компьютерного учебного мультимодального словаря историко-культурологического типа для изучения иностранного языка.

3.4 Моделирование компьютерного учебного мультимодального словаря историко-культурологического типа

Знание особенностей семантики слов важно для изучения иностранных языков, поскольку оно играет ключевую роль в развитии навыков и умений составлять высказывания, адекватно отражающие основную идею говорящего. Более того, понимание характеристик развития значения слова с исторической точки зрения помогает обучающимся построить ассоциативные связи и лучше понять культуру носителей изучаемого языка, что является одной из важнейших задач, которые разрабатывает современная лингводидактика. Важную роль при этом играет информативная репрезентация языковых особенностей слова и их визуализация. Для этой цели оптимальным форматом может выступить правильно смоделированный интерактивный компьютерный учебный словарь. В настоящем параграфе мы обозначим основные содержательные и структурные компоненты моделирования компьютерного учебного мультимодального словаря, направленного на изучение и визуализацию историко-культурного (диахронического) компонента значения лексемы в рамках лингводидактического подхода.

Одним из фундаментальных принципов изучения особенностей, структуры и свойств какого-либо объекта в науке является *моделирование*. **Моделирование** – это метод «воспроизведения и исследования определённого фрагмента действительности (предмета, явления, процесса, ситуации) или управления им, основанный на представлении объекта с помощью его копии или подобия – *модели*» [Стёпин, Гутнер, Голдберг, 2010]. В контексте лингвистики моделирование «позволяет изучать факты языка в условиях, управляемых и контролируемых исследователем» [Дорофеева, 2015, с. 41], а в рамках компьютерной лексикографии в большей степени заключается в *выявлении и разработке структуры компьютерного словаря*.

Основной идеей при разработке современных компьютерных словарей становится *антропоцентризм* – ориентация на пользователя будущего словаря.

Компьютерный словарь, как один из самых популярных и доступных к использованию сервисов, обладает дидактической, культурологической и познавательной ценностью, так как кроме знаний о языке содержит в себе полезную экстралингвистическую информацию, в том числе и о культуре страны изучаемого иностранного языка. Именно поэтому современные учебные словари должны обладать дополнительными функциями и инструментами, и кроме базовой лингвистической информации о слове давать примеры его употребления, синтаксические шаблоны, дополнительные семантические пометы и другую полезную культурологическую информацию о языке [Дрожащих, 2021].

Опираясь на положения, изученные в теоретических главах настоящей работы, можно сделать вывод, что хороший учебный словарь может помочь студенту систематизировать знания о языке, например, при помощи семантических сетей, коллокаций, семантической связи слов через этимологию, морфологию или фонологию. Кроме этого, словарь должен быть способен обеспечить дальнейшее раскрытие слова в других контекстах, с другими сочетаниями и конструкциями, заставив учащегося формировать собственные лексические связи и семантические сети. Следовательно, представляется возможным выделить основные **характеристики**, которыми должен обладать современный компьютерный учебный словарь для изучающих иностранный язык:

- обладать структурой, которая стремится показать концептуальные связи между лексемами в формате лексических / семантических полей и групп;
- предоставлять грамматическую и морфологическую информацию о склонении, спряжении и пр.;
- предоставлять информацию об употреблении лексемы с предлогами, другими словами, словосочетаниями или конструкциями;
- помогать студенту декодировать новые встречающиеся слова с помощью контекстных подсказок;
- предоставлять дополнительную информацию, помогающую студенту погрузиться в контекст употребления лексемы с ее историческими и культурологическими особенностями;

– мотивировать учащегося к дальнейшему изучению языка и формированию собственных лексических связей и семантических сетей.

На основе этого мы выделяем основные **принципы и особенности семантизации историко-культурного компонента** в учебном мультимодальном словаре:

– *описательное вербальное толкование значения слова на иностранном языке* (беспереvodный способ семантизации помогает пользователю словаря с самого начала его использования погрузиться в языковое окружение лексемы иностранного языка);

– *краткая этимологическая справка* (необходима для демонстрации диахронического развития значения слова);

– *экстралингвистический комментарий* (показывает социально-исторический и национально-культурный компонент в употреблении лексемы);

– *мультимодальные элементы* (иллюстрации, видео, аудиофайлы, которые содержат культурно-значимую информацию о слове или понятии);

– *наглядная визуализация языкового окружения лексемы* (схемы, графики, облака слов, частотные и иерархические списки);

– *гипертекстовый доступ* (соединяет словарь с корпусами и базами данных, лежащими в его основе, и другими справочными компьютерными лексикографическими ресурсами, тем самым предоставляя возможность и мотивируя учащегося продолжить самостоятельное изучение новой лексики иностранного языка);

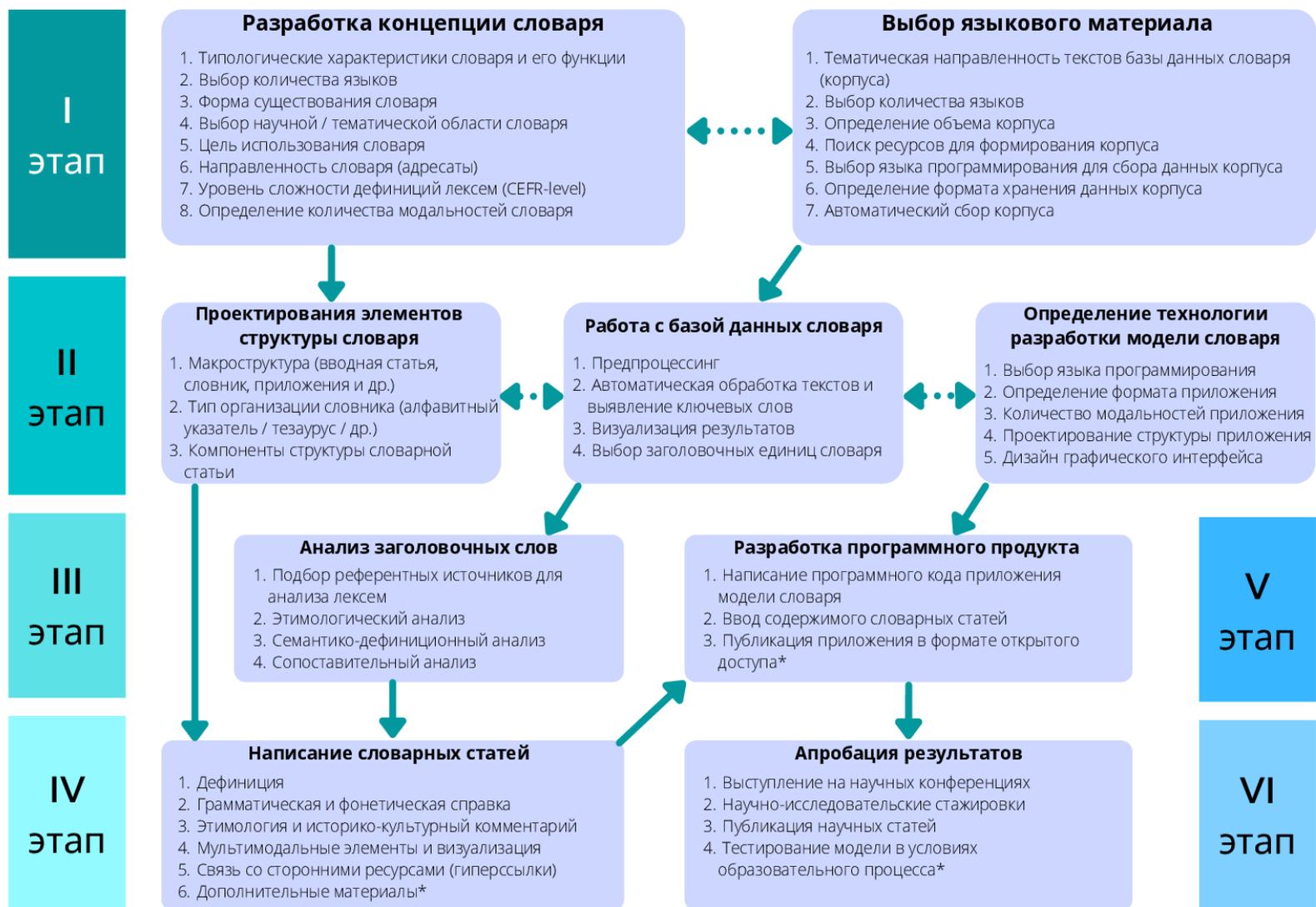
В дальнейшем, проанализировав теоретические исследования и различные методики моделирования словарей, представленные авторами изученных работ [Прошина, 2008; Русакова, 2008; Большакова, 2013; Агапова, Картофелева, 2014; Левенкова, 2016; Сьянова, 2020], мы определили три основных концептуально различимых направления работы над моделью компьютерного словаря: (1) *моделирование концепции и структуры словаря*, (2) *работа с языковым материалом-основой* и (3) *разработка программного продукта*.

Внутри описанных выше направлений можно выделить **содержательные и структурные компоненты процесса моделирования** лексикографического продукта, разрабатываемого в настоящем исследовании (рис. 12): *разработка концепции словаря; выбор языкового материала – основы словаря; определение технологии разработки модели словаря; проектирование элементов структуры словаря; работа с базой данных словаря; анализ заголовочных слов; написание словарных статей; разработка программного продукта; апробация результатов.* Данные компоненты были распределены на шесть последовательных этапов. При этом, важно отметить, что некоторые этапы включают в себя несколько равных по значимости компонентов, последовательность разработки которых может определяться исследователем в каждом конкретном случае в зависимости от различных факторов, как внешних, так и внутренних.

I этап состоит из *разработки концепции словаря и выбора языкового материала.* На стадии **разработки концепции словаря** необходимо рассмотреть следующие вопросы:

- типологические характеристики словаря (традиционный / тезаурус / энциклопедический / частотный и др.) и его функции;
- определение количества языков, представленных в словаре (одноязычный / двуязычный / переводной / многоязычный словарь);
- форма существования словаря (компьютерный словарь / традиционный словарь или совмещение форматов);
- определение научной или тематической области словаря (для дальнейшего поиска и подбора языкового материала – основы);
- цель использования словаря;
- направленность словаря и определение возможных адресатов будущего лексикографического продукта;
- уровень сложности дефиниций лексем (по классификации *Системы общеевропейских компетенций владения иностранным языком – CEFR Levels*);
- определение количества модальностей словаря (аудиальная, визуальная, кинестетическая).

Содержательные и структурные компоненты процесса моделирования словаря



* опционально

Рисунок 12. Содержательные и структурные компоненты процесса моделирования словаря

При **выборе языкового материала** необходимо определить ряд особенностей и характеристик базы данных будущего словаря, а также выполнить следующие процедуры:

- тематическая направленность текстов, которые выступят базой данных будущего словаря в формате лингвистического корпуса;
- определение количества языков, которые будут представлены в текстах корпуса (должно соответствовать количеству языков, представленных в словаре);
- определение объема корпуса (в количестве текстов / словоупотреблений / предложений или в другой единице измерения);
- поиск ресурсов для формирования корпуса (оцифрованные книги / литературные произведения / новостные тексты газет / веб-сайты и др.);
- выбор языка программирования для автоматического сбора данных корпуса (при необходимости использования метода веб-скрейпинга);
- определение формата и формы хранения базы данных корпуса (файлы в расширении .csv, .txt, .json или др.);
- автоматический сбор корпуса (опционально, если необходимо).

II этап состоит из трех компонентов: *проектирование элементов структуры словаря, работа с базой данных словаря и определение технологии разработки модели словаря*. **Проектирования элементов структуры словаря** требует работы над следующими вопросами:

- макроструктура словаря (вводная статья, словник, приложения и др.);
- тип организации словника (алфавитный указатель / тезаурус / семантические сети и др.);
- определение компонентов структуры словарной статьи (дефиниция, грамматическая справка, аудиопрезентация слова, транскрипция, примеры употребления слова, этимология, визуализация языкового окружения, иллюстрации, видео, экстралингвистическая информация в виде историко-культурного комментария, гиперссылки на сторонние справочные ресурсы или их интеграция в словарь, другие дополнительные материалы и визуализации).

Работа с базой данных словаря требует решения следующих задач:

- предпроцессинг текста (токенизация, стандартизация, удаление стоп-слов, лемматизация);
- автоматическая обработка текстов и выявление ключевых слов, которые станут заголовочными единицами словаря (частотный анализ / модели машинного обучения / использование нейросетей);
- визуализация результатов автоматической обработки (графики / списки / таблицы / облака слов и др.);
- отбор заголовочных единиц словаря для дальнейшего анализа (по материалам собранного корпуса).

Определение технологии разработки модели словаря состоит из:

- выбора языка программирования (зависит от выбранной на первом этапе формы существования словаря);
- определения формата приложения (компьютерное приложение, требующее установки / онлайн веб-ресурс / встроенная утилита и др.);
- количество модальностей приложения (связано с выбранным на первом этапе количеством модальностей словаря, особенностями формата приложения и возможностями языка программирования);
- проектирование структуры приложения (меню, основное окно, вкладки, страницы и др.);
- дизайн графического интерфейса (цветовая схема, шрифты, размер текста, раскладка страницы, иконки и др.).

III этап представляет собой **анализ заголовочных слов** будущей модели словаря, в который входят:

- поиск и анализ референтных источников для анализа лексем (толковые, энциклопедические, идеографические, этимологические, специализированные словари и лингвокультурные справочники, лингвистические корпуса, компьютерные лексикографические ресурсы и др.);
- этимологический анализ выбранных лексем;
- семантико-дефиниционный анализ выбранных лексем;

– сопоставительный анализ выбранных лексем двух и более языков (в соответствии с выбранным на первом этапе количеством языков словаря).

IV этап состоит из непосредственного **написания словарных статей**, согласно структуре, определенной на втором этапе работы и включающей следующие возможные элементы:

- дефиниция слова (по данным толковых словарей);
- грамматическая и фонетическая справка (по материалам учебных пособий и словарей для изучающих иностранный язык);
- этимология слова (краткая справка о происхождении слова);
- историко-культурный комментарий (энциклопедическая справка лингвострановедческого характера об особенностях истории развития слова и социально-культурных контекстах его употребления);
- выбор и анализ мультимодальных элементов словаря, имеющих историко-культурную и лингвострановедческую ценность для пользователя, изучающего иностранный язык (иллюстрации, видео- и аудиоматериалы);
- визуализация языкового окружения и семантического диахронического развития слова (таймлайн / облако слов / графики и др.);
- выбор и анализ сторонних лексикографических, справочных, образовательных и других ресурсов, которые могут быть использованы в учебных целях в моделируемом словаре (в формате гиперссылок или встроенных виджетов);
- поиск и добавление любых других дополнительных материалов учебной направленности (опционально).

V этап состоит из **разработки программного продукта**, включающей в себя следующие элементы:

- написание программного кода приложения модели словаря;
- ввод содержимого словарных статей, подготовленных на предыдущем этапе работы над моделью;
- публикацию приложения в формате открытого доступа (опционально, если выбран формат онлайн веб-приложения).

VI этап заключается в **апробации результатов** исследования, которая может проходить в различных форматах:

- выступления на научных конференциях;
- научно-исследовательские стажировки;
- публикация научных статей и тезисов;
- тестирование модели в условиях образовательного процесса (опционально, если исследование нацелено на экспериментальную работу и изучение дидактических особенностей и учебных функций словаря).

В заключение, на наш взгляд, представленные в настоящем параграфе содержательные и структурные компоненты процесса моделирования компьютерного словаря позволяют разработать лексикографический ресурс, соответствующий требованиям мультимодального представления информации, который способен семантизировать историко-культурный компонент лексемы. Такой словарь может быть использован в качестве учебного ресурса нового типа для обучения русскому или английскому языку как иностранному, так как соответствует определенным характеристикам – требованиям, чтобы успешно справляться с поставленной задачей. Завершая изучение основных теоретических положений, послуживших обоснованием практической части настоящего исследования, в следующем параграфе мы подведем итоги третьей главы.

Выводы по Главе 3

Изучение основных теоретических положений лингвокультурной лексикографии, особенностей семантизации и лексикографического моделирования позволило нам сделать следующие выводы:

1. *Лингвокультурная лексикография и лингвострановедение* подчеркивают важность изучения истории и культуры страны изучаемого языка, которые участвуют в развитии иноязычной коммуникативной компетенции, навыков межъязыкового и межкультурного общения и адекватного понимания культуры речи собеседника или аутентичных текстов. Поэтому при разработке

учебного словаря также важно учитывать *историко-культурный (диахронический) компонент* значения лексемы, который является совокупностью информации об истории семантического развития слова, национально-культурных коннотациях, контексте и особенностях его употребления в определенную историческую эпоху, которая описывает слово как достояние определенной лингвокультуры.

2. Исторические изменения, происходящие в любом языке, наиболее заметны в семантике лексических единиц языка. Современные компьютерные методы культуромики и цифровой гуманитаристики оптимальны для проведения исследований, связанных с изучением историко-культурных изменений лексики в диахронии. Данные статистического анализа употребления слов и словосочетаний в огромных текстовых массивах способны визуализировать различные культурные и социальные процессы, протекающие в языке.

3. *Семантизация* представляет собой процесс описания значения лексемы в словаре или объяснение новой лексики иностранного языка во время занятия с преподавателем. Для *семантизации историко-культурного компонента* значения лексемы мы выделили наиболее значимые принципы и особенности: *описательное вербальное толкование значения слова на иностранном языке, краткая этимологическая справка, экстралингвистический комментарий, мультимодальные элементы, наглядная визуализация языкового окружения лексемы и гипертекстовый доступ.*

4. Учебные словари для изучения иностранного языка должны соответствовать определенным требованиям, чтобы успешно справляться с поставленной задачей. Первой особенностью можно выделить то, что содержание словаря должно отвечать целям развития коммуникативной иноязычной компетенции и навыков межкультурной коммуникации. Кроме этого, состав и количество слов должны соответствовать уровню владения языком, а лексемы, представленные в словаре, обладать коммуникативной и лингвокультурной ценностью. Также важно, чтобы структура словаря могла давать иностранцам возможность самостоятельно использовать словарь без каких-либо трудностей и мотивировала студентов к изучению иностранного языка.

5. Содержательные и структурные компоненты процесса моделирования учебного мультимодального словаря историко-культурологического типа состоят из *разработки концепции словаря, выбора языкового материала – основы, определения технологии разработки модели, проектирования элементов структуры словаря, работы с базой данных словаря, анализа заголовочных слов, написания словарных статей, разработки программного продукта и апробации результатов.*

Подводя итог, мы считаем, что создание компьютерного учебного словаря, включающего различные мультимодальные форматы представления данных (текстовый, визуальный, аудио, видео и т.д.) с опорой на лингвистические корпуса и формальные модели языковой семантики позволяет говорить о словаре будущего, востребованном в современной теории образования. Такой словарь может стать инновационным инструментом преподавания иностранного языка в рамках компьютерной лингводидактики, так как он направлен на пользователя и содержит богатый экстралингвистический материал, имеющий большую культурологическую ценность.

В следующей главе мы перейдем к практической части настоящего исследования, в которой подробно опишем макро- и микроструктуру, элементы пользовательского интерфейса, особенности программного кода и языкового материала модели словаря. Лингвистической базой данных словаря выступил корпус новостных публикаций по теме «Образование / Education» на русском и английском языках, тексты которого были автоматически обработаны методами компьютерной лингвистики, с помощью использования нейросети BERT был составлен список ключевых слов, которые подверглись в дальнейшем этимологическому, семантико-дефиниционному и сопоставительному анализу. Ключевые слова с подробной семантической, историко-культурной и экстралингвистической информацией стали базой данных для модели учебного мультимодального словаря историко-культурологического типа.

ГЛАВА 4. РАЗРАБОТКА МОДЕЛИ УЧЕБНОГО МУЛЬТИМОДАЛЬНОГО СЛОВАРЯ ИСТОРИКО-КУЛЬТУРОЛОГИЧЕСКОГО ТИПА

4.1 Формирование концепции словаря

В качестве практического доказательства изученных теоретических положений и с целью визуализации результатов настоящего исследования была разработана оригинальная модель компьютерного учебного мультимодального словаря историко-культурологического типа, представленная в практической главе настоящего исследования. В настоящем параграфе мы уделим внимание содержательному описанию концепции моделируемого лексикографического продукта.

Моделируемый в настоящем исследовании *компьютерный учебный мультимодальный словарь историко-культурологического типа* представляет собой формализованную модель знаний о лексике тематического поля «Образование / Education» в русском и английском языках. Настоящий словарь комбинирует типологические и функциональные признаки различных словарей: *идеографического, толкового, учебного, частотного, двуязычного, компьютерного, мультимодального*. Словарь реализован в формате автономного веб-ресурса и не требует инсталляции, доступ осуществляется через сеть Интернет по ссылке¹².

Спроектированный словарь реализует базовые функции словарей: *учебную, справочную, систематизирующую и нормативную*. Доминирующей является **учебная функция**, которая выражена в представлении информации о современном толковании лексемы, ее грамматических характеристиках и семантических связях, а также во включении вербальных, аудиальных и графических иллюстраций. Кроме этого, особое место в моделируемом словаре занимает **справочная функция**, выраженная в качестве включенных дополнительных сведений об особенностях семантического развития и этимологии слова, лингвострановедческого

¹² Модель учебного мультимодального словаря, URL:https://na-stuk.shinyapps.io/multimodal_dict/

комментария и образовательных видео, способствующих глубокому познанию истории и культуры страны носителей изучаемого языка. **Нормативная функция**, как одна из основных для любого словаря, реализована в качестве фиксации значения и нормы употребления слов. **Систематизирующая функция** словаря проявляется в попытке построения карты слов для визуализации частотности состава ключевой лексики тематического поля «Образование / Education» в русском и английском языках.

В моделируемом словаре представлены два языка: *русский* и *английский*. При этом используются принципы *беспереводной семантизации* лексем. В словарной статье преобладает использование только **одного языка** (русского или английского соответственно). Дефиниция лексемы, справочные комментарии, дополнительные аудио- и видеоматериалы были оформлены исходя из этого условия. При выборе источника дефиниции мы руководствовались принципами *полноты, достаточности, минимальности, четкости*, а также *простоты* и *доступности изложения*. **Уровень объяснения** лексемы: вводное объяснение для учащихся, студентов, с дополнительной информацией энциклопедического характера. Дефиниции терминов взяты из одноязычных толковых, этимологических, исторических, энциклопедических и специализированных словарей. Для **русского языка**: *Большой толковый словарь русского языка С.А. Кузнецова, Морфемно-орфографический словарь А.Н. Тихонова, Этимологический словарь современного русского языка М. Фасмера, Этимологический словарь современного русского языка А.К. Шапошникова, Этимологический словарь русского языка: Русский язык от А до Я, Этимологический словарь русского языка Г.А. Крылова, Портал «Грамота.ру», Портал «Академик. Словари и энциклопедии», Онлайн-тезаурус «Карта слов», Новый словарь методических терминов и понятий (теория и практика обучения языкам), Энциклопедия русской цивилизации.* Для **английского языка**: *Oxford English Dictionary, Cambridge Advanced Learner's Dictionary (online), Merriam-Webster: Dictionary and Thesaurus (online), The Oxford Dictionary of Word Histories, Arcade Dictionary of Word Origins by John Ayto, The Oxford Dictionary of Word Origins (2 ed.) edited by Julia Cresswell, The Merriam-*

Webster New Book of Word Histories, A Dictionary of Education (2 ed.) Susan Wallace, Online Etymology Dictionary (online), Dr. Goodword AlphaDictionary (online).

Так как моделируемый словарь является полифункциональным, он может быть предназначен для широкого круга пользователей. **Адресатами** словаря могут стать пользователи, изучающие русский или английский языки как иностранный на продвинутом уровне, лингвисты, занимающиеся исследованием этимологии слов, а также студенты-специалисты в области педагогики и студенты-лингвисты. **Основной целью использования словаря** может стать изучение особенностей употребления слов тематического поля «Образование / Education», их семантики и этимологии.

На уровне **макроструктуры** словаря в качестве базового принципа организации материала было решено использовать **алфавитный указатель** для отобранных лексем из частотного списка ключевых слов тематического поля «Образование / Education» в русском и английском языках. **Особенности организации словника:** поиск осуществляется по отдельному слову, каждое слово имеет собственную словарную статью. Таким образом, макроструктура моделируемого компьютерного словаря во многом повторяет макроструктуру традиционного бумажного словаря, включая:

- *вводную часть:* предисловие, правила пользования словарем, список источников словаря, список используемых сокращений, ссылка на репозиторий материалов исследования;

- *облака слов* ключевой лексики тематического поля «Образование / Education» в русском и английском языках, составленные по материалам частотного анализа корпуса текстов СМИ.

- *словник:* словарные статьи для слов, входящих в зеркальный список ключевой лексики тематического поля «Образование / Education» в русском и английском языках.

На уровне **микроструктуры** словарные статьи для русского и английского языков имеют несколько различий, вызванных особенностями используемого материала. При этом общие элементы микроструктуры словарной статьи

компьютерного учебного мультимодального словаря историко-культурологического типа для обоих языков включают:

- *заголовочное слово;*
- *грамматическую помету;*
- *фонетическую транскрипцию;*
- *аудиопрезентацию слова;*
- *современные дефиниции слова с примерами употребления;*
- *этимологическую справку;*
- *историко-культурный комментарий;*
- *информацию о частотности употреблений лексики в диахронии;*
- *примеры употребления слова в корпусе;*
- *иллюстрации;*
- *видео-контент.*

Дополнительным в словарной статье на русском языке является *раздел «Морфемы»*, а для английского языка информация на каком уровне владения языком, согласно *Системе общеевропейских компетенций владения иностранным языком (CEFR Levels)*, студент должен овладеть тем или иным значением слова. Кроме этого, в англоязычной словарной статье присутствует *блок «Timeline»*, который состоит из двух частей: первая *текстовая часть* описывает хронологию развития семантических изменений в значениях слов, вторая представляет собой *визуализацию* этих данных в виде интерактивной панели в формате таймлайн.

Таким образом, описанные нами макро- и микроструктуры соответствуют общепринятым стандартам построения словаря. *Мультимодальный аспект* подачи информации выражается в виде нелинейного представления тематического поля «Образование / Education» в русском и английском языках в формате облака слов, аудиопрезентации лексем, иллюстраций, видео, визуализации хронологии развития семантических изменений в формате таймлайн, использовании цветowych акцентов выделения информации, гиперссылок на другие лексикографические и справочные ресурсы. Выбранная подача материала представляется оптимальной для реализации *учебной функции словаря* и повышения его *интерактивности*.

В результате моделируемый словарь может использоваться на занятиях иностранного языка, так как выполняет функцию семантизации историко-культурного компонента лексемы и соответствует определенным ранее в обзоре теоретических предпосылок исследования характеристикам, которыми должен обладать учебный мультимодальный словарь: *интерактивность, ориентация на пользователя, функция системного многофункционального справочного пособия, наличие экстралингвистической информации, этимологии и контекста употребления лексемы, мультимодальность, гипертекстуальность*. В следующем параграфе мы приступим к работе над языковым материалом – основой базы данных моделируемого словаря.

4.2 Тематическое поле «Образование / Education» в русском и английском языках: создание базы данных словаря

Для объективной репрезентации семантической информации в учебном словаре требуется большое количество примеров реального употребления слова в языке в виде текстов, аудио- и видеозаписей, скриптов и т. д. В настоящее время такой материал могут предоставить лингвистические корпуса во всём их многообразии, которые сами по себе уже представляют собой уникальный образовательный ресурс. Это в первую очередь связано с тем, что лингвистические корпуса, в особенности диахронические, содержат в себе ценную информацию об историческом развитии и культурной эволюции языка. Каждый текст отражает особенности эпохи, в которую он был написан, данные о культуре людей, живших в тот период, исторические факты и семантические особенности лексики языка.

Целью настоящего этапа работы является исследование современного состояния тематического поля «Образование / Education» в русском и английском языках и определение ключевых лексем, которые способны наиболее репрезентативно представить основные понятия и темы, характеризующие его, а также стать основой базы данных моделируемого словаря. Для достижения поставленной цели, необходимо решить ряд задач:

– создать корпус релевантных текстовых данных дискурса образовательных СМИ на русском и английском языках, которые охватывают разнородные и актуальные темы, связанные со сферой образования.

– выбрать инструменты автоматической обработки и выделения ключевых слов тематического поля «Образование / Education»;

– установить список самых частотных ключевых слов и выделить из них лексемы для дальнейшего анализа, которые войдут в модель словаря как заголовочные единицы;

– изучить значения отобранных лексем в существующих словарях русского и английского языков и провести этимологический, семантико-дефиниционный и сопоставительный анализ;

– написать статьи для разрабатываемой модели словаря;

Далее в параграфе подробно описывается процедура автоматизированного сбора корпуса новостных статей с сайтов исследуемых онлайн-изданий посредством языка программирования Python 3, изложена методика выявления ключевых слов с использованием машинного обучения и моделей нейронной сети BERT, представлены результаты этимологического, семантико-дефиниционного и сопоставительного анализа выбранных лексем, а также визуализация ключевой лексики тематического поля «Образование / Education» в формате облака слов, автоматически сгенерированного на основе частотности.

4.2.1 Особенности процедуры и этапы создания корпуса и анализа данных

Развитие компьютерных технологий оказало значительное влияние на стремительное распространение и растущий интерес к исследованиям в рамках компьютерной лингвистики. В последние годы набирают популярность методы и подходы машинного обучения в обработке естественного языка, разрабатываются библиотеки и специальные инструменты для анализа больших текстовых данных. В ряд задач таких проектов входит классификация, кластеризация и синтез текстов, прогнозное моделирование (predictive modelling),

интеллектуальный анализ данных (data mining) и другие [Grus, 2019]. При этом важной составляющей таких исследований является хорошо подобранный корпус, способный дать репрезентативные результаты. В настоящем параграфе мы опишем процедуру и этапы работы над созданием исследовательского корпуса, а также рассмотрим особенности автоматической обработки текстов для дальнейшего выявления и анализа ключевых слов тематического поля «Образование / Education» в русском и английском языках, которые выступают в качестве заголовочных слов модели компьютерного учебного мультимодального словаря историко-культурологического типа.

Большую популярность в работах по компьютерной лингвистике приобрел язык программирования Python 3, на котором разработано множество специальных инструментов для решения задач обработки естественного языка (NLP). Для целей настоящего исследования также был выбран Python 3, мы использовали следующие библиотеки и модули:

- для извлечения данных с веб-сайтов: *requests*, *beautifulsoup4*;
- для выявления ключевых слов: *sklearn*, *sentence-transformers*;
- для формирования базы данных (датафрейма): *pandas*;
- для предпроцессинга и лемматизации текстов: *string*, *re*, *pymorphy2*, *nlTK*, включая *punkt*, *stopwords*, *corpus*, *tokenize*, *wordnet*;
- для визуализации результатов: *matplotlib.pyplot*, *collections*, *PIL*, *wordcloud*.

Далее мы подробно опишем особенности, процедуру и инструменты, использованные для каждого из этапов работы.

Создание корпуса. Использование текстовых данных, доступных в сети Интернет, в исследованиях естественного языка становится все более популярным. Это происходит ввиду многочисленных преимуществ, среди которых можно выделить открытость, бесплатный доступ, удобный формат хранения информации. Например, если данные содержатся на веб-странице, то они могут быть автоматически собраны с помощью *метода веб-скрейпинга (web-scraping)* или, другими словами, *парсинга* сайтов [Anandarajan, Hill, Nolan, 2019, с. 35].

Веб-страницы написаны на языке HTML, в котором текст размечен на элементы и их атрибуты, обозначенные тегами. В каждом таком элементе содержится информация о новостной статье (заголовок, текст, дата публикации, категории и др.), те метаданные, которые необходимы для создания исследовательского корпуса в формате датафрейма. *Датафрейм* (dataframe) – это двумерная маркированная структура данных. Визуально она представляет собой таблицу, состоящую из фиксированного числа строк и столбцов [Grus, 2019, с. 134]. Такой формат оптимально подходит для быстрой и удобной работы с большими данными.

Автоматическая обработка текстовых данных. Предварительная обработка текста (preprocessing) необходима для приведения языковых данных к единому формату и уменьшения размерности. Процесс предобработки включает в себя: *токенизацию* (разделение текста на единицы анализа: слова, предложения, N - граммы), *стандартизацию* (преобразование текстовых данных из разных документов в единый сопоставимый формат: удаление пунктуации, перевод в нижний регистр), *удаление стоп-слов* (служебные части речи и другие высокочастотные слова, которые не несут смысловой нагрузки), *стемминг* (приведение всех однокоренных слов к общей основе посредством удаления суффиксов слова для уменьшения количества токенов) или *лемматизацию* (приведение всех словоформ к начальной форме – лемме). Лемматизация является более предпочтительным форматом при изучении семантических особенностей текстовых данных, так как основана на частеречной разметке слов с учетом их значения [Anandarajan, Hill, Nolan, 2019, с. 45-57]. В процессе предобработки программа получает на входе исходный текст; далее на каждом из этапов удаляется ненужная для анализа информация; в результате возвращается список очищенных токенов. Правильная предварительная обработка текста является залогом успеха автоматизированного анализа данных естественного языка. Для настоящего исследования предобработка необходима в первую очередь для того, чтобы диверсифицировать финальный список ключевых слов и избавиться от множественных словоформ одной и той же лексемы, например: *teacher-teachers*. В работе мы использовали следующие инструменты для предобработки текстов

новостных статей: приведение всех слов к нижнему регистру, чистка текста регулярными выражениям от ссылок на сайты, email-адресов, цифр, классических и нетипичных знаков пунктуации, токенизация, лемматизация и удаление стоп-слов.

Автоматическое выявление ключевых слов. Для автоматизированной обработки текстов новостных статей на русском и английском языках и дальнейшего выявления ключевых слов были использованы предобученные модели **BERT** (*Bidirectional Encoder Representations from Transformers*), которые представляют собой нейросеть, разработанную командой Google, и предварительно обученную на огромном объеме данных. Модель является мультязычной (104 языка) [Devlin, 2018], ее используют для специфических задач обработки естественного языка, таких, как классификация текстов, определение эмоциональной окраски (тональности) предложений, создания вопросно-ответных систем (чат-ботов), извлечения данных, построения выводов по тексту, информационного поиска, анализа высказываний и т.д. [Rogers, Kovaleva, Rumshisky, 2020]. В связи с широким спектром возможностей практического применения данной модели голландский исследователь Маартен Гроотендорст (Maarten Grootendorst) реализовал проект по выявлению ключевых слов в текстах, на основе которого впоследствии была создана библиотека для языка программирования Python 3, названная KeyBERT [Sharma, 2019; Grootendorst, 2020]. Именно этот алгоритм был использован в настоящем исследовании.

Процедура выявления ключевых слов состоит из нескольких этапов. В начале необходимо создать список ключевых слов-кандидатов или ключевых фраз из документа на основе N-грамм. Для этого использовался Scikit-Learn Count Vectorizer, который позволяет определять длину ключевых слов и превращать их в ключевые фразы. Для нашего исследования мы использовали униграммы (одно слово), но при необходимости можно также определить ключевые словосочетания, например биграммы (два слова) или триграммы (три слова). Следующим шагом является векторизация текста – перевод слов в числовой

машиночитаемый вид, где каждому слову назначается уникальный вектор в многомерном пространстве текста и корпуса. Здесь была использована модель BERT, в частности *distilbert — base-nli-stsb-mean-tokens*. На последнем этапе вычисляется сходство между словами-кандидатами и текстом посредством измерения косинусной близости между векторами, так как такой метод хорошо работает для высокой размерности.

Частотный анализ и визуализация данных. Визуализация результатов в компьютерных исследованиях естественного языка занимает важное место, поэтому одним из наглядных способов может стать создание *облака слов* на основе их частотности. Специальная компьютерная программа анализирует все словоупотребления в тексте и чем чаще слово встречается, тем больший размер оно принимает в облаке. При этом важно отметить, что, рассматривая частотность слов в тексте, принято говорить об *абсолютной частоте* и *относительной частоте* слов. **Абсолютная частота** «показывает количество словоупотреблений, или сколько раз данное слово (или словоформа) повторяется в выборке, она является целым числом», в то время как **относительная частота** «получается из абсолютной, если ее поделить на объем выборки, и является дробным числом из промежутка от 0 до 1, которое показывает, какую долю данное значение составляет от всего объема выборки» [Studiorum]. При частотном анализе текстов чаще всего используют относительную частоту, которая позволяет сравнить употребление одной и той же лексемы в нескольких текстах разного объема. Для создания облаков слов компьютерные системы используют такой показатель как *оценка веса ключевого понятия* или соотношение TF (term frequency) к IDF (inverse document frequency). «Показатель *TF-IDF* является статистическим и используется для оценивания важности конкретного слова в контексте всего документа или текста, входящего в общую коллекцию или корпус» [Shkurina, Sabinin, 2019]. Здесь TF – это относительная частота слова, о которой мы говорили раньше, а IDF – его инверсия. Таким образом понижается вес широко используемых общеупотребительных слов, таких, как предлоги, союзы, местоимения и другие. Показатель TF-IDF будет выше, если определенное слово с большой частотой

используется в конкретном тексте, но редко – в других документах [Jones, 2004; Sammut, Webb, 2010].

Визуализация таких объектов, как облака слов, варьируется от целей и задач исследования. При этом существует множество различных бесплатных онлайн-ресурсов, которые позволяют визуализировать загруженный текст без дополнительных знаний в программировании (Облакослов.рф, Word's Cloud и др.). Встречаются облака, в которых важность слова подчеркивается цветом или же где можно индивидуально задать форму и цветовую гамму получаемого изображения. На интернет-сайтах более популярен другой формат облака слов – облако тегов, где каждое слово есть тег, который является гиперссылкой, при этом такое облако динамично и способно отображать изменения в частоте поисковых запросов или содержания страниц сайта, например, новостных тем. Однако, несмотря на различия исходных данных, принцип построения таких объектов остается одинаковым [Kaser, Lemire, 2007]. В качестве примера представлено облако слов на основе произведений А.С. Пушкина (рис. 13). На заключительном этапе настоящего исследования результаты анализа и подсчета частотности объединенных списков ключевых слов в русско- и англоязычном корпусах будут также оформлены в облака слов с использованием библиотек wordcloud и wordcloud2 на языках программирования Python 3 и R соответственно.

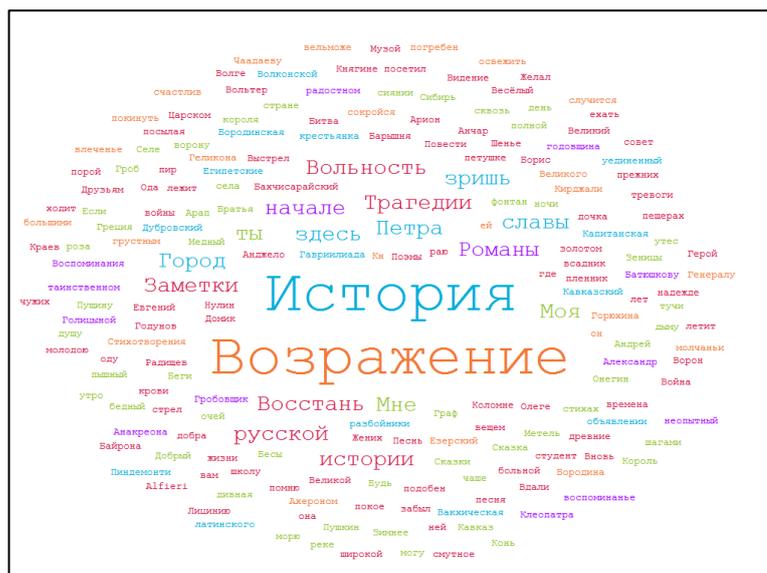


Рисунок 13. Облако слов на основе произведений А.С. Пушкина ¹³

¹³ Облако слов, URL: <https://облакослов.рф/>

Анализ ключевых слов для написания словарных статей. Самым объемным этапом стал «ручной» анализ ключевых слов, выявленных с помощью автоматизированных методов обработки текстовых данных исследовательского корпуса. Он включал в себя определение объема выборки ключевых слов на основе их частотности, изучение особенностей семантических диахронических изменений каждой лексемы в сравнительно-историческом аспекте и сопоставление полученных данных для русского и английского языков. С этой целью для анализа лексики тематического поля «Образование / Education» в настоящем исследовании использовались следующие методы:

– *метод этимологического анализа лексики*, который применяется для изучения исторического развития какого-либо явления или системы в языке, он способствует объяснению исходных словообразовательных связей и направлен на «выявление происхождения и первичного значения слов с целью изучения источников и процессов формирования словарного состава языка» [ЛЭС, 2002; Моргачёва, 2016, с. 107]. В настоящем исследовании используется *прием культурно-исторической интерпретации* и поэтапной *хронологизации семантических изменений*. Историческое исследование языка обязательно включает в себя освещение семантических изменений значения слов с течением времени. Семантический анализ лексики языка доказывает, что значения слов, в особенности многозначных, не являются автономными статическими данными. Слово выступает как универсальный символ, смысл которого формируется индивидуально для каждого конкретного случая употребления и зависит в первую очередь от окружения лексемы [Beheydt, 1987, с. 55]. Эту особенность семантики могут продемонстрировать результаты семантико-дефиниционного анализа лексемы;

– *метод семантико-дефиниционного анализа* применяется для анализа лексем и выявления расхождений между ними на семантическом уровне. В основе данного метода заложена идея разложения значения лексемы на несколько дифференцирующих компонентов, которые могут описывать индивидуальные характеристики значения слова, по аналогии с фонологической

классификацией фонем. Компоненты имеют различительную функцию и служат для отделения значения одной лексемы от других, семантически связанных с ней [Jackson, 1998, с. 80-86];

– *метод сопоставительного анализа* используется для исследования и описания языка через его системное сравнение с другим языком, часто он также называется *контрастивным методом*. Как отмечают ученые Воронежской школы сопоставительных исследований, «*сопоставительно-семантический метод* определяется задачей получения семантического описания сравниваемых единиц или явлений и выявления конкретных сходств и различий в их семантике», тогда как *контрастивный метод*, в свою очередь, обычно применяется для сопоставления родного языка исследователя с иностранным и «отражает сопоставление отдельных единиц одного языка со всеми возможными соответствиями этой языковой единицы в языке сравнения для выявления их национальной семантической специфики» [Конопелько, 2019, с. 40]. Использование данного метода в настоящем исследовании позволит выявить общие и дифференциальные признаки лексем русского и английского языков в тематическом поле «Образование / Education» с акцентом на их историко-культурные особенности.

Подводя итог, важно отметить, что методы и инструменты компьютерной лингвистики были выбраны для настоящего исследования, так как они способны статистически подтвердить, являются ли субъективные предположения исследователя о тексте верными или нет. Кроме этого, данные частотного анализа способны показать языковые особенности и характеристики, о которых возможно и не предполагалось на этапе постановки цели и задач. В следующих параграфах мы более детально опишем выбранные новостные источники, статьи которых легли в основу исследовательского корпуса, а также рассмотрим результаты, полученные при его обработке и анализе. Список автоматически выявленных ключевых слов пройдет этимологический, семантико-дефиниционный и сопоставительный анализ.

4.2.2 Составление русскоязычного корпуса текстов по теме «Образование»

В настоящем исследовании для изучения русского языка используются материалы новостных статей образовательного портала «*EDU-Inform*»¹⁴. Данный ресурс является информационно-справочным порталом, который содержит сведения об образовательных заведениях России всех уровней, их структуре, направлениях деятельности и перечне предоставляемых услуг, включая дошкольное, школьное, высшее, профессиональное и дополнительное образование. В том числе на сайте портала представлены дополнительные сервисы: доска объявлений, пресс-релизы и дни открытых дверей ВУЗов, актуальные новости образования в регионах РФ и в России в целом. Таким образом, данный ресурс наиболее оптимально удовлетворяет цели и задачам настоящего исследования.

На момент составления корпуса раздел сайта во вкладке «Новости образования» насчитывал 112 страниц и 1 119 новостных статей, охватывающих период с июня 2011 года по август 2021 года. Для сбора корпуса был использован метод веб-скрейпинга, с помощью написанного кода на языке программирования Python 3. Для оценки размера корпуса важна его статистическая информация (таблица 2). Общий список ключевых слов составляет 11 190 единиц, по 10 слов для каждой статьи в подкорпусе. На основе частотного анализа и группировки списка ключевых слов была выявлена 2 281 уникальная лексема, которые наиболее характерно отражают тематику публикаций в сфере образования на русском языке.

Таблица 2. Статистика корпуса новостных статей портала «*EDU-Inform*»

Параметр	Количество
Веб-страницы	112
Новостные статьи	1 119
Предложения	34 960
Словоупотребления (токены)	779 793
Уникальные слова	49 199
Уникальные ключевые слова	2 281

¹⁴ *EDU-Inform*, URL: <http://edu-inform.ru/>

Важно также отметить, что тексты новостных статей, используемых в настоящем исследовании, были насыщены канцеляризмами и именами собственными, которые из-за высокой частотности употребления попадали в список ключевых слов. Такое свойство может быть объяснено как индивидуальными характеристиками источника материала исследования, так и общими особенностями публицистического стиля русского языка. В связи с этим некоторые из таких лексем были дополнительно добавлены к стандартному списку стоп-слов библиотеки *rumorphy2* с целью диверсификации результатов, хотя частично общеупотребительная лексика всё же осталась и рассматривается в качестве погрешности. Таким образом были исключены следующие лексемы (имена собственные, эргонимы, названия городов, имена), которые обладали высокой частотностью, но не отражали релевантной для общей характеристики тематического поля «Образование» информации: «государственный», «представитель», «Санкт - Петербург», «санкт-петербургский», «Москва», «московский», «новосибирский», «дальневосточный», «южноуральский», «Лобачевский», «правительство», «рособнадзор», «Минпросвещение», «Минобрнауки», «министерство», «негосударственный», «общественный», «распоряжение», «постановление», «предусматривать», «административный», «ежемесячный», «соответствующий», «здравоохранение».

Русскоязычная часть корпуса в формате датафрейма *pandas* (таблица 3) состоит из пяти основных элементов, включающих также метаданные:

- Столбец «*Title*» отображает заголовок статьи в виде одного предложения.
- Столбец «*Date*» отображает дату публикации новостной статьи.
- Столбец «*Link*» содержит прямую ссылку на статью на сайте портала «EDU-Inform».
- Столбцы «*Text*» и «*Text_preprocessed*» содержат соответственно оригинальный и предобработанный для выявления ключевых слов тексты новостной статьи.

делового и публицистического стилей, а также специфике репрезентации информации в прессе государственных структур. Так, например, можно выделить преобладание сложных прилагательных по сравнению с другими частями речи, обилие канцеляризмов и нейтральной общеупотребительной лексики: «*пресс-служба*», «*необходимость*», «*количество*», «*использование*», «*продолжительность*», «*большинство*», «*обязательный*», «*действительно*», «*предварительный*», «*опубликовать*», «*предоставлять*», «*организовать*», «*соответствовать*», «*воспользоваться*», «*обсуждение*», «*завершиться*», «*прокомментировать*».

Далее мы рассмотрим подробнее 60 наиболее частотных лексем в списке ключевых слов для определения основных тематических направлений новостных статей об образовании на русском языке (таблица 4). Для начала рассмотрим лексику, которая описывает основную действительность процесса образования, его типы, особенности и участников: «*образование*», «*просвещение*», «*профобразование*», «*специальность*», «*преподаватель*», «*руководитель*», «*образовательный*», «*общественно-профессиональный*», «*студенческий*», «*общеобразовательный*», «*педагогический*», «*преподавательский*», «*учебно-методический*», «*университетский*», «*экзаменационный*», «*вступительный*», «*исследовательский*», «*научно-исследовательский*», «*обществознание*», «*контрольно-измерительный*», «*профессиональный*», «*технологический*», «*дополнительный*», «*художественный*». В контексте российской образовательной традиции государство играет важную роль в регулировании процесса обучения, что может объяснить появление следующих ключевых слов: «*премьер-министр*», «*муниципальный*», «*правоохранительный*», «*законодательство*», «*демографический*», «*подведомственный*», «*предоставление*», «*строительство*». Кроме этого, важно отметить, что ситуация, сложившаяся в современном мире в связи с глобальной эпидемиологической обстановкой, оказала своё влияние в том числе и на процесс образования, что также отразилось в изменении новостных тем, в которых частотными стали следующие слова, отражающие новый формат обучения и проведения различных образовательных мероприятий:

«безопасность», «распространение», «компьютерный», «искусственный», «психологический», «доступный», «международный», «взаимодействие», «конкурентоспособность», «самостоятельный», «самостоятельно».

Таблица 4. Частотный список ключевых слов по теме «Образование»

№	Ключевое слово	Часто та	№	Ключевое слово	Часто та
1	учебно-методический	1 104	31	специальность	26
2	общественно-профессиональный	1 104	32	взаимодействие	25
3	профессиональный	1 061	33	соответствовать	25
4	образовательный	1 040	34	премьер-министр	25
5	дополнительный	123	35	большинство	25
6	завершиться	110	36	доступный	24
7	образование	92	37	научно-исследовательский	23
8	педагогический	88	38	предварительный	23
9	организовать	81	39	демографический	23
10	общеобразовательный	71	40	строительство	22
11	искусственный	65	41	контрольно-измерительный	22
12	руководитель	65	42	муниципальный	21
13	законодательство	59	43	международный	20
14	обсуждение	55	44	экзаменационный	20
15	опубликовать	50	45	подведомственный	19
16	просвещение	47	46	конкурентоспособность	19
17	обязательный	46	47	художественный	19
18	преподаватель	44	48	безопасность	19
19	вступительный	43	49	университетский	19
20	обществознание	43	50	предоставлять	19
21	использование	37	51	продолжительность	18
22	пресс-служба	36	52	количество	18
23	исследовательский	36	53	предоставление	18
24	студенческий	36	54	прокомментировать	18
25	технологический	35	55	воспользоваться	18
26	психологический	32	56	правоохранительный	18
27	необходимость	32	57	распространение	17
28	действительно	28	58	самостоятельный	17
29	профобразование	26	59	преподавательский	17
30	самостоятельно	26	60	компьютерный	17

Для дальнейшего изучения были выбраны десять слов из списка ключевой лексики для русского языка: «образование», «университет», «школа», «студент», «преподаватель», «учебный», «наука», «экзамен», «профессиональный», «руководитель». Принцип отбора основывался на нескольких параметрах, которые включают в себя семантику слова, соответствующую тематическому полю «Образование», историко-культурные особенности значения и употребления слова, частотность в общем списке ключевых слов, а также наличие лексем с идентичным значением в списке ключевых слов для английского языка. В связи с этим у некоторых лексем были изменены части речи, например прилагательное «студенческий» было заменено на существительное «студент». Такие замены были мотивированы преимущественно большим количеством значений слова у существительных, чем у прилагательных, так как для моделируемого словаря в первую очередь было важно представить наибольшее разнообразие значений лексем. Таким образом, материалом для подготовки словарных статей модели учебного мультимодального словаря историко-культурологического типа станут два зеркальных списка лексем для русского и английского языков. В следующих параграфах полученные данные подвергнутся более глубокому и качественному этимологическому, семантико-дефиниционному и сопоставительному анализу.

4.2.3 Составление англоязычного корпуса текстов по теме «Education»

Следующим шагом стало создание подкорпуса текстовых данных для английского языка. Для определения ключевых лексем тематического поля «Education» в английском языке в настоящем исследовании используются материалы новостных статей с сайта онлайн-издания журнала «*Education Today Magazine*»¹⁵. Выбор пал на данный британский журнал в связи с тем, что он ориентирован на весь сектор образования от начального (primary) до образования для взрослых (adult education). Здесь публикуются статьи на различные темы: освещение текущих новостей, правительственные стратегии в сфере образования,

¹⁵ Education Today Magazine, URL: <https://www.education-today.co.uk/>

особенности и критерии учебных программ, обзоры выставок, конкурсов, различных продуктов и ресурсов. Кроме того, журнал является членом Британской ассоциации поставщиков образовательных услуг BESA (British Educational Suppliers Association). Таким образом данный ресурс оптимально удовлетворял целям и задачам настоящего исследования и обладал схожей направленностью с русскоязычным источником.

На момент составления корпуса новостной раздел сайта во вкладке «Latest News» насчитывал 54 страницы и 534 новостных статьи, которые охватывали период с июня 2014 года по август 2021 года. Статистическая информация англоязычной части корпуса (таблица 5): 8 071 предложение, 285 320 словоупотреблений (токенов), из которых 18 853 уникальных слова, в том числе содержатся метаданные к каждой статье: *заголовок, тематическая категория новости, дата публикации, ссылка на web-страницу, оригинальный текст статьи, предобработанный текст и список ключевых слов*. Общий список ключевых слов составляет 5 340 единиц, по 10 слов для каждой статьи в корпусе. На основе частотного анализа и группировки списка ключевых слов было выявлено 1 755 уникальных лексем, которые наиболее характерно отражают тематику публикаций в сфере образования на английском языке.

Таблица 5. Статистика корпуса новостных статей журнала «Education Today Magazine»

Параметр	Количество
Веб-страницы	54
Новостные статьи	534
Предложения	8 071
Словоупотребления (токены)	285 320
Уникальные слова	18 853
Уникальные ключевые слова	1 755

Необходимо отметить, что англоязычная часть корпуса в сравнении с русскоязычной обладает рядом отличий, такими как временной период публикаций, общее количество статей, и, как следствие, статистические показатели корпуса. В обоих случаях, при сборе русскоязычной и англоязычной частей корпуса, был использован весь имеющийся текстовый материал опубликованных

статей, поэтому для сохранения однородности данных, стиля и формата публикаций было принято решение не добавлять и не использовать дополнительно другие англоязычные источники с целью создания эквивалентного объема статистических показателей. В связи с тем, что данная особенность не оказывает негативного влияния на работу модели машинного обучения по выявлению ключевых слов, в настоящем исследовании описанные различия рассматриваются как индивидуальные характеристики источников текстового материала разных языков.

Англоязычная часть корпуса в формате датафрейма pandas (таблица 6) состоит из шести основных элементов, включающих также метаданные:

Таблица 6. Корпус новостных статей журнала «Education Today Magazine»

Title	Category	Date	Link	Text	Text_preprocessed	Keywords
How could AI affect your organisation? – Educa...	['News']	1 week ago	https://www.education-today.co.uk/how-could-ai...	Artificial intelligence is rapidly revolutioni...	artificial intelligence is rapidly revolutioni...	[innovative, student, intelligence, education,...
Keep pupils active this summer with Public Hea...	['Educational event', 'England', 'Health, Hygi...	1 week ago	https://www.education-today.co.uk/keep-pupils-...	Evidence suggests that regular physical activi...	evidence suggests that regular physical activi...	[swim, daily, fun, curriculumlinked, favourite...
Distinguished schools recognised for maths att...	['Maths', 'News', 'Partner content']	2 weeks ago	https://www.education-today.co.uk/distinguishe...	Fourteen UK schools and almost forty in countr...	fourteen uk school and almost forty in country...	[academic, tutor, mexico, teacher, math, award...
Warwickshire CCC and Sandwell College form par...	['BTEC', 'Community engagement', 'FE colleges'...	3 weeks ago	https://www.education-today.co.uk/warwickshire-...	Warwickshire CCC and Sandwell College have ann...	warwickshire ccc and sandwell college have ann...	[cricketing, charity, friday, training, cricke...

– *столбец «Title»* отображает заголовок статьи в виде одного предложения;

– *столбец «Category»* состоит из списка тегов-категорий, которые присваиваются новостной статье для ее классификации на сайте, количество тегов у каждой статьи индивидуально;

– *столбец «Date»* отображает дату публикации новостной статьи, основной трудностью здесь выступило отсутствие в html-коде веб-страницы точной даты публикации для статей не старше одного месяца, в связи с этим часть дат отображается как «3 days ago», «2 weeks ago» и т.д. Впоследствии после

отражающих культурные особенности образовательного процесса в Великобритании. Первая подгруппа обозначает участников и людей, вовлеченных в процесс образования и научные исследования: «*teacher*», «*student*», «*professor*», «*educator*», «*principal*», «*head-teacher*», «*parent*», «*scientist*». Следующей подгруппой выделяются лексемы, описывающие процесс обучения, дисциплины, места и названия образовательных учреждений: «*education*», «*science*», «*literacy*», «*teaching*», «*learning*», «*training*», «*teach*», «*class*», «*classroom*», «*school*», «*college*», «*university*», «*library*», «*academic*», «*professional*», «*educational*», «*innovation*», «*curriculum*», «*mathematics*».

Таблица 7. Частотный список ключевых слов по теме «Education»

№	Ключевое слово	Частота	№	Ключевое слово	Частота	№	Ключевое слово	Частота
1	teacher	197	21	June	32	41	scientist	18
2	education,	122	22	parent	32	42	Monday	18
3	student	112	23	head teacher	31	43	Wednesday	17
4	teaching	103	24	July	30	44	college	16
5	charity	100	25	April	30	45	literacy	16
6	classroom	92	26	principal	29	46	month	16
7	online	76	27	professor	28	47	innovation	16
8	competition	67	28	website	27	48	download	15
9	university	63	29	educator	27	49	February	15
10	CEO	57	30	school	25	50	champion	15
11	learning	54	31	pandemic	24	51	winning	15
12	new	54	32	finalist	24	52	celebration	14
13	mathematics	48	33	volunteer	24	53	compete	14
14	academic	44	34	wellbeing	21	54	professional	14
15	manager	44	35	award-winning	21	55	November	14
16	annual	41	36	coronavirus	21	56	Christmas	14
17	science	37	37	festival	19	57	world-class	13
18	Friday	35	38	fundraising	19	58	educational	13
19	class	35	39	library	19	59	celebrating	13
20	training	34	40	curriculum	19	60	teach	13

Также важно отметить, что благотворительность является неотъемлемым от повседневной жизни процессом в Великобритании, в котором участвуют как

взрослые, так и дети. Здесь можно говорить о различных ежегодных конкурсах, ярмарках, концертах и других мероприятиях, чем объясняется частотность следующих лексем: «*charity*», «*fundraising*», «*festival*», «*annual*», «*celebration*», «*celebrating*», «*volunteer*». Кроме этого, в последние годы важным становится направление образовательного менеджмента и развития управленческих и лидерских качеств как у сотрудников учебных учреждений, так и у обучающихся, это можно проследить по лексемам: «*CEO*», «*manager*», «*world-class*», «*competition*», «*compete*», «*finalist*», «*champion*», «*winning*», «*award-winning*». Отдельной группой в ключевой лексике английского языка выделяются слова, характерные для публикаций в контексте пандемии коронавируса, которые отражают новые цифровые реалии образовательного процесса: «*coronavirus*», «*pandemic*», «*wellbeing*», «*online*», «*website*», «*new*», «*download*».

Для дальнейшего более детального изучения из списка ключевой лексики английского языка были выбраны десять слов, которые не только имеют высокую частотность, но и являются эквивалентным переводом русскоязычных лексем, отобранных ранее: «*education*», «*university*», «*academic*», «*student*», «*school*», «*teacher*», «*science*», «*examination*», «*professional*», «*manager*». В следующих параграфах мы перейдем к этимологическому, семантико-дефиниционному и сопоставительному анализу зеркального списка ключевых слов с изучением особенностей историко-культурного компонента значения лексем.

4.2.4 Этимологический и семантико-дефиниционный анализ тематического поля «Образование» в русском языке

Для реализации этимологического и семантико-дефиниционного анализа лексики важным этапом является подбор оптимальных справочных источников для детального изучения семантических изменений в значениях слов. Для решения задач, поставленных в настоящем исследовании, мы остановились на использовании обобщенных данных из словарей разных типов: *современных толковых, этимологических, исторических, энциклопедических*

и *специализированных*. Такие справочные источники способны представить подробные сведения о семантических изменениях в толкованиях и употреблении слов. В настоящем параграфе представлен этимологический и семантико-дефиниционный анализ выбранных ключевых слов тематического поля «Образование» для русского языка: «*наука*», «*образование*», «*преподаватель*», «*профессиональный*», «*руководитель*», «*студент*», «*университет*», «*учебный*», «*школа*», «*экзамен*».

Для фиксации современных значений лексем и морфолого-орфографической информации мы использовали материалы *Большого толкового словаря русского языка С.А. Кузнецова (БТСРЯ)*, *Морфемно-орфографического словаря А.Н. Тихонова (МОС)* и *онлайн-тезауруса «Карта слов»*. При выборе материала для анализа, который впоследствии станет основой составления словарной статьи в моделируемом словаре, мы придерживались принципа, что справочные данные о лексеме должны быть краткими, информативными и релевантными для студента, изучающего русский язык как иностранный. В связи с этим часть слишком детальных и сложно дифференцируемых толкований были сокращены или объединены в одно тезисно. Таким образом, для описания современного значения лексемы за основу берутся наиболее смысловозначительные единицы в значении слова по материалам БТСРЯ.

С целью изучения этимологии лексем мы обратились к следующим словарям: *Этимологический словарь современного русского языка М. Фасмера (ЭССРЯ)*, *Этимологический словарь современного русского языка А.К. Шапошникова (ЭССРЯ(Ш))*, *Этимологический словарь русского языка: Русский язык от А до Я*, *Этимологический словарь русского языка Г.А. Крылова (ЭСРЯ(К))*. Данные, которые составят раздел «Этимология», представляют собой краткую этимологическую справку, обобщенную по материалам словарных статей референтных этимологических словарей. Кроме этого, в качестве справочных источников для исторического и культурологического комментариев в словарную статью моделируемого словаря были использованы онлайн-порталы

«Грамота.ру»¹⁶ и «Академик. Словари и энциклопедии»¹⁷, а также словари и энциклопедии *Новый словарь методических терминов и понятий (теория и практика обучения языкам)* (НСМТП), *Энциклопедия русской цивилизации* (ЭРЦ).

Далее мы перейдем к более детальному описанию и анализу ключевой лексики. Этимологически слово «наука», означающее «определенную область знаний», является исконно русским по происхождению: *от праславянского *na- + *učiti* и *от древнерусского укъ* в значении «учение» [ЭССРЯ; ЭСРЯ]. Как отмечено в Энциклопедии русской цивилизации центрами науки на Руси были монастыри, при которых нередко существовали школы. Как свидетельствуют памятники письменности и техники, особенно высокий уровень знаний был достигнут в зодчестве, различных ремеслах и металлообработке [ЭРЦ]. В современном контексте важно отметить, что самостоятельной наука считается в том случае, «если у нее имеется свой объект, предмет исследования и своя терминология» [НСМТП]. В современном русском языке лексема «наука» имеет три основных дифференцируемых значения, согласно БТСРЯ:

1. «система знаний о закономерностях развития природы и общества и способах воздействия на окружающий мир»;
2. «то, что поучает, даёт опыт, урок», используемое преимущественно в *разговорной речи*;
3. *устаревшее* «обучение какому-либо ремеслу» [БТСРЯ].

Лексема «образование» происходит *от русского церковнославянского слова образованиѣ*, восходящего к *праславянскому* глаголу **образовати* в значении «создавать образ, образы». В XIX веке слово «образование» стало использоваться в русском литературном языке как калька с немецкого языка «Bildung» [ЭССРЯ(Ш)]. В современном русском языке лексема имеет два основных значения:

1. «то, что образовалось, возникло в результате какого-либо процесса»;
2. «процесс усвоения знаний; обучение, просвещение», именно в этом значении слово рассматривается в настоящем исследовании [БТСРЯ].

¹⁶ Грамота.ру, URL; <http://gramota.ru/slovari/online/>

¹⁷ Академик. Словари и энциклопедии, URL: <https://dic.academic.ru/>

Слово «преподаватель» образовано от глагола преподавать (пре- + по- + дать) *от праславянского *dātī; *dājātī; *dāvātī* в значении «дать, давать» [ЭССРЯ]. В современном русском языке слово имеет одно значение: «тот, кто занимается преподаванием чего-либо» [БТСРЯ]. В основном слово «преподаватель» употребляется по отношению к работникам средне специальных или высших учебных заведений, в общеобразовательной школе используется слово «учитель» [НСМТП].

Прилагательное «профессиональный» является производным от слова «профессия», которое пришло в русский язык в XVIII веке *из французского языка*, где было заимствовано от *латинского professiō* в значении «публичное заявление о роде своей деятельности» [ЭСРЯ; ЭССРЯ]. В современном русском языке лексема имеет два значения:

1. «относящийся к профессии, связанный с какой-либо профессией»;
2. «относящийся к профессионализму, являющийся профессионалом» [БТСРЯ].

Слово «руководитель» происходит от существительного «рука» и глагола «водить». Слово «рука» существует в русском языке с XI века и образовано *от общеславянского roka*, семантически родственного *греческому* слову в значении «ладонь, кисть руки», «собираю, складываю» [ЭССРЯ]. В современном русском языке лексема имеет одно значение: «тот, кто руководит кем-, чем-либо, направляет деятельность кого-, чего-либо» [БТСРЯ]. Руководитель в сфере образования – это тот, кто несёт ответственность за всё, что происходит в любом образовательном учреждении. Название должностей при этом отличается: в школе – директор или завуч, в университете, колледже или техникуме – ректор, проректор, декан, заместитель декана, заведующий кафедрой, в детском саду – заведующий.

Слово «студент» происходит *от латинского studens* (род. п. studentis) «старющийся», причастия от глагола *studere* в значении «усердно работать», которое произошло из *праиндоевропейского *steu-* «толкать, бить». В русском языке слово «студент» впервые было использовано в Духовном регламенте

1721 года: «народный скудѣнт», вероятно под влиянием слова «ску́дный» и было заимствовано через польский **student** или немецкий языки **Student** [ЭССРЯ]. В современном русском языке лексема употребляется в одном значении: «учащийся высшего учебного заведения» [БТСРЯ].

Слово «университет» было заимствовано из немецкого языка – **Universität** и восходит к латинскому **universitas** в значении «совокупность». В русском языке широкое распространение слово получило с 1755 года в связи с открытием первого университета в Москве. Диалектная форма существительного – «наверститут» – образовалась на основе двух самостоятельных слов – «наверстать» и «институт» [НСМТП; ЭСРЯ(К); ЭСРЯ; ЭССРЯ]. В современном русском языке лексема употребляется в одном значении: «высшее учебно-научное заведение с различными гуманитарными и естественно-математическими отделениями (факультетами)» [БТСРЯ].

Прилагательное «учебный» имеет схожее происхождение с описанным выше существительным «наука» и образовано от праславянского ***učiti** [ЭССРЯ]. В современном русском языке лексема употребляется в значении «связанный, соотносящийся по значению с существительным учёба (период с сентября по июнь) или с организацией процесса обучения» [БТСРЯ].

Слово «школа» происходит от латинского **schola** и древнегреческого **σχολή** в значении «досуг; задержка, занятие в свободные часы, чтение, лекция, школа», восходящих к праиндоевропейской основе ***segh-** в значении «держат, обладать, иметь». Древние греки в свободное от работы время предпочитали посещать беседы философов и таким образом пополнять объем знаний. Исходным значением было «досуг», позже «непринуждённая беседа», а затем уже «место для беседы». Благодаря этому появилось слово «школа» в значении «место, где можно получить образование». Слово «школа» получило широкое распространение довольно поздно – в XVI–XVII вв., хотя известно было и раньше, но до этого времени его заменяли синонимом – словом «училище». Древнерусская школа «училище», начиная с 1388 г., было заимствовано через польский язык [ЭСРЯ; ЭСРЯ(К); ЭССРЯ]. В современном русском языке лексема имеет четыре значения:

1. «учебное заведение, которое осуществляет общее образование и воспитание молодого поколения, или где даются специальные знания, профессиональные навыки»;

2. «приобретение знаний, выучки, опыта и система приёмов, обязательных упражнений, видов работ при изучении чего-либо, при овладении каким-либо мастерством, искусством»;

3. «направление, течение в науке, искусстве, литературе, общественно-политической мысли и т.п., обладающее характерными свойствами, методами, приёмами»;

4. «питомник, где выращивают растения (сеянцы или саженцы) до посадки на постоянное место», используемое *в отрасли сельского хозяйства* [БТСРЯ].

Существует две версии происхождения слова «экзамен», согласно которым в русский язык оно пришло либо из *польского* – **egzamen**, либо было заимствовано из *немецкого* – **Examen** в значении «испытание». Оба этих слова восходят к *латинскому* слову **examen** в значении «вереница, куча, толпа; исследование, испытание», которое поначалу имело значение «стрелка у весов (приходящая в движение при взвешивании)», затем – «взвешивание», а уж потом – «испытание». В русском языке это слово известно с начала XVIII века, т.е. с эпохи Петра I и означало «проверка знаний учащихся по какому-либо предмету» [ЭСРЯ; ЭССРЯ]. В современном русском языке лексема имеет значение «проверочное испытание по какому-либо учебному предмету или испытание чьих-либо знаний и навыков» [БТСРЯ].

Таким образом, ключевые слова тематического поля «Образование» для русского языка можно разделить на две группы по этимологическим признакам. Первая часть слов имеют исконно-русское или славянское происхождение, такие, как «наука», «образование», «преподаватель», «руководитель», «учебный». Вторая часть является словами-когнатами для английского языка, обладающими общими этимологическими признаками и появившимися в результате заимствований из латинского, греческого, немецкого или французского языков, например: «профессиональный», «студент», «университет», «школа», «экзамен». Большая

часть изученных слов не являются многозначными и имеют только одно значение, используемое по отношению к конкретному объекту или феномену, связанному со сферой образования. Такая особенность семантики ключевой лексики облегчает изучение новых слов иностранными студентами, предоставляя узкий и тематически точный контекст употребления. В следующем параграфе мы изучим лексико-семантические особенности зеркального списка ключевых слов тематического поля «Education» в английском языке.

4.2.5 Этимологический и семантико-дефиниционный анализ тематического поля «Education» в английском языке

В настоящем параграфе представлен подробный этимологический и семантико-дефиниционный анализ выбранных ключевых слов тематического поля «Education» для английского языка: «*academic*», «*education*», «*examination*», «*manager*», «*professional*», «*school*», «*science*», «*student*», «*teacher*», «*university*». Материалом выступили различные *толковые, этимологические, исторические и специализированные* словари английского языка.

Для фиксации современных значений лексем мы использовали материалы онлайн-изданий двух толковых учебных словарей английского языка: *Cambridge Advanced Learner's Dictionary*¹⁸ (CALD) и *Merriam-Webster: Dictionary and Thesaurus*¹⁹ (MW), которые являются многофункциональными лингвистическими справочными ресурсами, имеющими широкий список дополнительных образовательных инструментов, таких, как этимологическая справка, список примеров употребления предложений из корпусов, обобщенное представление значений из базы данных различных словарей, стилистические пометы и др. При выборе материала для анализа, который впоследствии станет основой составления словарной статьи в моделируемом словаре, мы придерживались принципа, что справочные данные о лексеме должны быть краткими, информативными

¹⁸ Cambridge Advanced Learner's Dictionary, URL: <https://dictionary.cambridge.org/>

¹⁹ Merriam-Webster: Dictionary and Thesaurus, URL: <https://www.merriam-webster.com/dictionary>

и релевантными для студента, изучающего английский язык как иностранный. В связи с этим часть слишком детальных и сложно дифференцируемых толкований были сокращены или объединены в одно тезисно. Таким образом, в качестве «эталонных» для изучения и сравнения семантических изменений в английском языке берутся за основу наиболее смысловоразличительные единицы в современном значении слова по материалам онлайн-словарей CALD и MW.

С целью подробного изучения истории развития семантики лексики мы обратились к Оксфордскому словарю английского языка, *Oxford English Dictionary* (OED), в формате компьютерного приложения, версия 4.0. Данный словарь является крупнейшим лексикографическим ресурсом, демонстрирующим историческое развитие английского языка с акцентом на этимологию и датирование возникновения новых семантических характеристик слова, в связи с этим он выступил основным справочным ресурсом для анализа семантических изменений ключевых слов. В качестве единиц сравнения исторических семантических изменений в английском языке рассматривается отобранный набор значений слова, представленных в словаре OED. При этом для датирования семантических значений в формате *таймлайн* (*timeline*) дата вхождения того или иного значения фиксируется как дата первого примера употребления слова из словаря OED. В случае, если такое значение не совпадает с современным толкованием слова, то дата окончания его использования фиксируется как дата последнего примера употребления слова в указанном значении из словаря OED.

Кроме этого, в качестве справочных ресурсов для изучения истории развития семантики лексем и истоков их происхождения, а также для подготовки этимологического и историко-культурологического комментариев в словарную статью моделируемого словаря (разделы «Etymology» и «Cultural comment») были использованы следующие исторические и специализированные словари, как в традиционном печатном формате, так и в виде компьютерного онлайн-издания: *The Oxford Dictionary of Word Histories* (ODOWH), *Arcade Dictionary of Word Origins by John Ayto* (ADWO), *The Oxford Dictionary of Word Origins (2 ed.) edited by Julia Cresswell* (ODOWO), *The Merriam-Webster New Book of Word Histories* (MWNBWH),

*A Dictionary of Education (2 ed.) Susan Wallace (ODE), Online Etymology Dictionary*²⁰ (EDO), *Dr. Goodword AlphaDictionary*²¹ (AD). Данные, которые составят раздел «Etymology», представляют собой краткую этимологическую справку, обобщенную по материалам статей в референтных этимологических словарях.

Далее мы перейдем к более детальному описанию и анализу ключевой лексики английского языка. Слово «**academic**» в английском языке употребляется как прилагательное и как существительное, при этом набор значений различается для каждой части речи. Оно является лексемой для сопоставления со словом «*учебный*» в русском языке. Этимологически слово «*academic*» было заимствовано в английский язык из *среднефранцузского* **achadémique, académique** и *латинского языка* **Acadēmicus** в значении «of the school of the philosopher Plato» (относящийся к школе философа Платона), которые произошли от *греческого* **Akadēmeia** – названия места, где преподавал Платон. Первое упоминание слова появляется в 1580-х гг. [ADWO; MWNBWH; MW; ODOWO]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. как *прилагательное* слово «*academic*» в значении «belonging to an academy or institution for higher learning» [OED] используется для обозначения отношения чего- или кого-либо к образовательному процессу и учебным заведениям. Это значение появилось в 1588 году и сохранилось до настоящего времени в форме «relating to schools, colleges, and universities, or connected with studying and thinking, not with practical skills» (относящийся к школам, колледжам и университетам или связанный с обучением и мышлением, а не с практическими навыками) [CALD];

2. значение *прилагательного* «belonging to the Academy, the school or philosophy of Plato» является *устаревшим*, оно использовалось в более узком контексте напрямую по отношению к школе философа Платона приблизительно в период с 1610 по 1777 год [OED];

²⁰ Online Etymology Dictionary, URL: <https://www.etymonline.com/>

²¹ Dr. Goodword AlphaDictionary, URL: <https://alphadictionary.com/>

3. следующее значение *прилагательного* «not leading to a decision, unpractical, theoretical, formal, or conventional» [OED] появилось в 1886 году и употребляется до настоящего момента в интерпретации «based on ideas and theories and not related to practical effects in real life» [CALD] для описания теоретических, не практико-ориентированных вещей, разговоров и идей;

4. также в современном английском языке выделяют значение *прилагательного* «used to describe someone who is clever and enjoys studying» [CALD] для описания личностных качеств и черт характера человека, любящего учебу;

5. как *существительное* слово «*academic*» имеет два значения: *первое* – «a senior member of the academic staff of a college or university or a member of a society for promoting art or science» [OED], появившееся в 1587 году и сохранившее свое употребление до настоящего времени в форме «someone who teaches at a college, or who studies as part of their job» [CALD] для обозначения человека, чья рабочая деятельность связана с преподаванием, исследованиями или обучением в высшем учебном заведении; и *второе устаревшее* значение «an ancient philosopher of the Academy, an adherent of the philosophical school of the philosopher Plato (c. 428-347 BC), a Platonist» (античный философ, приверженец философской школы Платона), использовавшееся приблизительно в период с 1586 по 1846 гг. [EDO].

Слово «**education**» в английском языке используется как существительное и является лексемой для сопоставления со словом «*образование*» в русском языке. Этимологически оно было заимствовано в английский язык из *латинского языка* и имеет связь с двумя глаголами: **ēducāre** в значении «bring up, rear» (воспитывать, выращивать) и **ēducere** в значении «lead out» (вести за собой, выводить). Первоначально метафора, вероятно, заключалась в том, чтобы вывести кого-то из детства во взрослую жизнь. Первое упоминание слова появляется в 1531 году [AD; ADWO; MW; ODOWH]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. одними из первых значений слова «*education*» были «the process of nourishing or rearing a child or young person, an animal» и «the process of ‘bringing up’

young persons» [OED], которые использовались примерно в период с 1531 по 1661 гг. для обозначения процесса воспитания и кормления детей и животных и напрямую связаны с этимологией слова;

2. значение «the systematic instruction, schooling or training given to the young in preparation for the work of life» [OED] появилось в 1616 году и сохранилось до настоящего времени в форме «the process of teaching or learning, especially in a school or college, or the knowledge that you get from this» [CALD]. Такой переход и конкретизация значения показывают, как от «воспитания детей» и «дрессировки животных» слово эволюционировало в «обучение подрастающего поколения социальным нормам и правилам»;

3. также существует два особых случая в употреблении слова «*education*»: *специализированное* значение в сфере шелковой промышленности «the rearing of silkworms» (выращивание тутового шелкопряда) [OED], первое упоминание которого появилось в 1888 году, и современное «the study of methods and theories of teaching» [CALD], используемое преимущественно в американском варианте английского языка по отношению к образовательным программам по педагогике.

Слово «**examination**» в английском языке употребляется как существительное и является лексемой для сопоставления со словом «*экзамен*» в русском языке. Этимологически слово было заимствовано в среднеанглийский язык из *старофранцузского* **examiner**, происходящего от *латинского* **examinare** в значении «weigh, test» (взвешивать, испытывать), восходящего к **examen**, что дословно значит «tongue of a balance» (стрелка у весов). Первое упоминание слова появляется в XIV веке, а сокращенная форма слова – «*exam*» в 1848 году как студенческий сленг [ADWO; EDO; MW; ODOWH]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. одно из первых значений слова «*examination*» – «judicial inquiry into the guilt or innocence of an accused person by a standard or rule» [OED], появилось в 1387 году и использовалось для обозначения судебного расследования для доказательства виновности или невиновности обвиняемого приблизительно до 1861 года;

2. значение «the action of investigating the nature, qualities, or condition of any object by inspection or experiment» [OED] появилось в 1510 году и употребляется до настоящего момента в интерпретации «the act of looking at or considering something carefully in order to discover something» [CALD] для описания процесса детального изучения чего-либо посредством опытов и экспериментов;

3. дефиниция «the process of testing, by questions oral or written, the knowledge or ability of pupils, or of candidates for office, degrees, etc.», использующееся в сфере образования для формальной проверки знаний учащихся, появилось в 1612 году и сохранилось до настоящего времени в форме «a formal test that you must pass in order to earn a particular qualification or to be allowed to do a particular job» [CALD].

Слово «**manager**» в английском языке употребляется как существительное и является лексемой для сопоставления со словом «*руководитель*» в русском языке. Этимологически слово было заимствовано в английский язык из *итальянского языка* – **maneggiare**, восходящего к *латинскому manus* «hand» (рука). Первоначальное значение глагола «*to manage*» было «put (a horse) through the paces of the manège = to handle or train a horse» (провести (лошадь) через манеж, тренировать лошадь). Таким образом сфера употребления слова «*manager*» была расширена от верховой езды, до финансов и бизнеса и, впоследствии, до сферы образования. Первое упоминание слова появляется в 1588 году [MW; ODOWO; ODOWN]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. основное значение слова «*manager*» – «one whose office it is to manage a business establishment or a public institution» [OED], относящееся к человеку, вовлеченному в руководство каким-либо предприятием, появилось в 1588 году и сохранилось до настоящего времени в форме «the person who is responsible for managing an organization or whose job is to organize sports team / singer / actor», расширив контекст употребления и включив в себя также сферы шоу-бизнеса и спорта [CALD];

2. дефиниция «one of several members of either house of parliament appointed for the performance of some specified duty connected with the functions of the two houses» [OED] является *устаревшей*, она использовалась в более узком политическом значении приблизительно в период с 1666 по 1863 год для обозначения должности членов палат парламента, назначенных для выполнения определенных обязанностей;

3. также слово «*manager*» имеет два *специализированных* значения: в *компьютерной сфере* для обозначения определенного типа программ – «computer program or system that helps you to control or organize a particular type of activity» [CALD]; и в *сфере права и юриспруденции* *устаревшее* значение «a person appointed, ordinarily by a court of chancery, to control, carry on and account for any business which may have fallen into the hands of the court for the benefit of creditors or others» [OED], использовавшееся приблизительно в период с 1793 по 1880 год для описания должности человека, выполняющего ряд определенных работ в суде.

Слово «**professional**» в английском языке употребляется как прилагательное и существительное, при этом набор значений различается для каждой части речи. Оно является лексемой для сопоставления со словом «*профессиональный*» в русском языке. Этимологически слово было заимствовано в английский язык из *латинского языка* через причастие **prōfessus** от глагола **profitēri** в значении «to declare publicly» (заявлять публично). Таким образом этимологически «*professor*» (профессор) это человек, который «публично заявляет» о своих знаниях в определенной области, а «profession» (профессия) – сфера деятельности, в которой он «profess» (преподает) какие-либо умения или компетенции. Первое упоминание слова появляется в 1616 году [ADWO; MW; ODOWH]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. одно из первых значений слова «*professional*» как *прилагательного* «pertaining to or marking entrance into a religious order» [OED] сейчас является *устаревшим*, оно использовалось в религиозном контексте приблизительно

в период с 1420 по 1881 год, а затем было вытеснено светской сферой употребления лексемы;

2. как *прилагательное* слово «*professional*» в значении «engaged in one of the learned or skilled professions, or in a calling considered socially superior to a trade or handicraft» [OED], используемое для обозначения отношения человека к определенному виду деятельности, появилось в 1747 году и сохранилось до настоящего времени в форме «relating to work that needs special training or education» (относящийся к работе, требующей специальной подготовки или образования) [CALD]. Такая же ситуация сложилась и для формы *существительного* со схожим значением «one who belongs to one of the learned or skilled professions» [OED], которое начало употребляться в 1848 году и сохранилось как «a person who has the type of job that needs a high level of education and training» (человек, выполняющий работу, требующую высокого уровня образования и подготовки) [CALD];

3. в современном английском языке слово «*professional*» имеет значение *прилагательного* «used to describe someone who does a job that people usually do as a hobby» и значение *существительного* «a person who does a job that people usually do as a hobby» [CALD], которые используются для обозначения профессионального характера деятельности человека, которая обычно является любительским занятием на уровне хобби (искусство, музыка, танцы, спорт). Такая сфера употребления лексемы появилась в 1811 году в форме «one who makes a profession or business of any occupation, art, or sport, otherwise usually or often engaged in by amateurs as a pastime» [OED].

4. современное значение *прилагательного* «having the qualities that you connect with trained and skilled people, such as effectiveness, skill, organization, and seriousness of manner» [CALD] связано со значением *существительного* «someone who has worked hard in the same type of job for a long time and has become skilled at dealing with any problem that might happen» [CALD], которые используются в большей степени в неформальном контексте для описания человека и его профессионализма в той или иной сфере деятельности;

5. среди более частных и редких случаев употребления слова «*professional*» можно выделить значение *существительного* «a prostitute» (проститутка, женщина легкого поведения) [OED], примеры использования которого датируются периодом с 1861 по 1977 год, а также значение прилагательного «disparagingly applied to one who pursues relentlessly an activity or belief that is regarded with disfavour; inveterate, habitual, ruthless» [OED], появляющееся в текстах в 1879-1978 гг и использующееся в пренебрежительном контексте по отношению к человеку, который занимается деятельностью не одобряемой обществом.

Слово «**school**» в английском языке употребляется как существительное и как глагол, при этом набор значений различается для каждой части речи. Оно является лексемой для сопоставления со словом «школа» в русском языке. Этимологически слово «*school*» пришло в *древнеанглийский* язык как **scōl, scolu** в значении «place for instruction» (место для обучения) из *латинского* языка через *греческий* **skholē** со значением «leisure, philosophy, lecture-place» (досуг, философия, место для лекций). Первоначально слово обозначало «досуг», но позже постепенно изменило значение на «leisure used for intellectual argument or education» (досуг, используемый для интеллектуальных споров или обучения), а затем на «*school*» как «the building» (здание). Первое упоминание слова появляется ранее XII века [AD; ADWO; MWNBWH; MW; ODOWH; ODWO]. Изучая изменения в семантике лексемы по материалам OED и CALD можно отметить, что слово «*school*» является многозначным и обладает 11 четко дифференцируемыми значениями для существительного и 5 для глагола. Таким образом, можно выделить следующие семантические особенности лексемы:

1. основное значение слова «*school*» как *существительного*, «an institution in which instruction of any kind is given (whether to children or adults)» [OED] для обозначения учебного учреждения появилось в 1000 году и сохранилось до настоящего времени в форме «a place where children go to be educated or a where people can study a particular subject» [CALD];

2. следующее значение «the body of persons that are or have been taught by a particular master (in philosophy, science, art, etc.)» [OED] появилось в 1612 году и употребляется до настоящего момента в интерпретации «a group of painters, writers, poets, etc. whose work is similar, especially similar to that of a particular leader» [CALD] для описания группы людей (художников, писателей, поэтов, исследователей и т.д.), имеющих общий стиль, направление и идеи работ, или возглавляемых одним общим лидером;

3. *существительное* «*school*» также используется для обозначения части структурного подразделения высшего учебного заведения, отличающегося особой предметной направленностью – «a part of a college or university specializing in a particular subject or group of subjects» [CALD]. Данное значение появилось в 1400 году как «the faculties composing a college or university, universities in general» [OED] и сохранилось до настоящего момента;

4. слово «*school*» как *существительное* имеет ряд *устаревших* значений. Часть из них связана с философией: «the place in which an ancient Greek or Roman philosopher taught his hearers» (место, в котором древнегреческий или римский философ учил своих слушателей) [OED], которое использовалось в период с 1375 по 1781 год, и значение «the scholastic philosophers and theologians collectively» [OED] в качестве имени собственного для обозначения схоластического философского течения в текстах периода 1614 – 1684 гг. Кроме этого, одним из ранних примеров использования лексемы в период с 816 по 1450 год является значение «a hostelry at Rome for the reception of pilgrims» (гостиница для паломников в Риме) [OED], которое предположительно трансформировалось позднее посредством генерализации в «a public building, gallery, or the like» (общественное здание, галерея) [OED], использовавшееся в 1400-1601 гг.

5. слово «*school*» как *существительное* также имеет ряд частных случаев употребления: в *сленге* существует значение «a company of thieves or beggars working together» [OED] для обозначения группы воров или нищих, работающих вместе, зафиксированное в период с 1812 по 1976 год, а также значение «a group of persons drinking together in a bar or public house, and taking turns to buy the drinks»

(группа пьющих людей в баре, покупающих напитки по очереди) [OED], появляющееся в текстах 1890-1971 гг. Другим современным случаем *специализированного* употребления лексемы можно выделить значение *существительного* «the periodical examinations for the degree of B.A. at Oxford University», появившееся в 1800 году и использующееся до настоящего времени для обозначения регулярных экзаменов, являющихся особенностью образовательного процесса в Оксфордском университете.

6. слово «*school*» как *глагол* имеет одно основное современное значение – «to train a person or animal to do something» (обучать человека или животное делать что-либо) [CALD], которое с развитием объединило в себе два значения вместе: «to educate, train (a person, his mind, powers, tastes, etc.)» [OED], появившееся в 1573 году, и «to train or exercise (a horse) in movements» [OED] от 1869 года.

7. *глагол* «*school*» также имеет *устаревшее* значение «to give a lesson to (a person) by punishment, to chastise» (проучить наказанием, наказывать), зафиксированное в текстах с 1592 по 1628 год, и узкое *специализированное* значение в области биологии и сельского хозяйства «to rear (a plant) in a nursery» (выращивать (растение) в питомнике) [OED], упоминание которого датируется 1902 годом.

8. редким случаем употребления слова «*school*» является значение *глагола* «to ride straight across country» [OED] от 1885 года, которое перекликается со значением *существительного* «a cross-country ride» (поездка по пересеченной местности), появившемся немного позднее в 1892 году.

9. дополнительно необходимо отметить, что слово «*school*» имеет омоним, не связанный со сферой образования. *Существительное* «*school*» в значении «multitude (of fish)» (множество, косяк рыб), которое используется преимущественно в американском английском вместо британского вариант «*shoal of fish*» [AD].

Слово «*science*» в английском языке употребляется как *существительное* и является лексемой для сопоставления со словом «*наука*» в русском языке. Этимологически оно было заимствовано в английский язык через

старофранцузский scientia «knowledge, a knowing; expertness» (знание, мастерство) от *латинского* глагола *scire* «know» (знать). Первоначально слово обозначало «knowledge gained by study» (любые знания, полученные в процессе обучения) и только потом, в начале XVIII века, стало использоваться для обозначения исследований в области биологии, физики, математики и технических наук. Первое упоминание слова появляется в XIV веке [ADWO; EDO; MW; ODWH]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. слово «*science*» имеет одно современное значение «(knowledge from) the careful study of the structure and behaviour of the world, especially by doing experiments» [CALD] для обозначения знаний, полученных в результате тщательного изучения чего-либо путем проведения экспериментов, которое появилось в 1289 году в форме «knowledge acquired by study, acquaintance with or mastery of any department of learning» [OED];

2. слово «*science*» как *существительное* имеет ряд *устаревших* значений: «a craft, trade, or occupation requiring trained skill» [OED], которое использовалось для обозначения ремесла или занятия, требующего определенных навыков, приблизительно в период с 1480 по 1660 год, а также значение *сленга* «the noble science (of defence): the art of boxing or that of fencing» по отношению к военной науке и боевым искусствам, появляющееся в текстах 1588-1839 гг.

Слово «**student**» в английском языке употребляется как *существительное* и является лексемой для сопоставления со словом «*студент*» в русском языке. Этимологически слово было заимствовано в английский язык через *старофранцузский estudiant* «student, scholar, one who is studying» (студент, ученый, тот, кто учится) от *латинского* существительного **studium** «zeal, painstaking application» (усердие, кропотливость), которое относилось к процессу обучения. Вероятно это слово восходит к индоевропейскому корню ***steud-**, ***teud-** «hit» (удар), который также стал основой *латинского* слова **tundere** «hit» (ударить) и немецкого *stossen* «shove, hit» (толкнуть, ударить). Таким образом, понятие, лежащее в основе слова «*study*» (учеба), обозначает «application of extreme effort»

(приложение особого усилия) [ADWO; ODOWH]. В *древнеанглийском языке* было слово-аналог **leorningcild** в значении «student, disciple» (студент, ученик) [EDO]. Первое упоминание слова «student» появляется в XV веке [MW]. Изучая изменения в семантике лексемы по материалам OED и CALD, можно выделить следующие особенности:

1. слово «*student*» имеет одно современное значение «a person who is learning at a school, college or university» (человек, который учится в школе, колледже или университете) [CALD], которое появилось в 1398 году в форме «a person who is undergoing a course of study and instruction at a university or other place of higher education or technical training» [OED];

2. слово «*student*» как *существительное* имеет также ряд частных случаев употребления слова: в Оксфордском университете существует значение «a member of the foundation, corresponding to the ‘fellow’ or ‘scholar’ of other colleges» [OED], использовавшееся преимущественно в колледже Крайст-Чёрч (Christ Church) в период с 1651 по 1858 год для обозначения члена совета; в американском *сленге* наркоторговцев существует значение слова «an inexperienced user of illegal drugs, one who takes small or occasional doses» (неопытный человек, употребляющий запрещенные наркотики в небольших дозах), которое упоминается в текстах периода 1936–1951 гг.

Слово «**teacher**» в английском языке употребляется как *существительное* и является лексемой для сопоставления со словом «*преподаватель*» в русском языке. Этимологически слово восходит к *праиндоевропейской* основе ***deik-**, которая также прослеживается в *греческом* слове **deiknunai** со значением «show» (показывать) и *латинском* **dīcere** «say» (говорить). *Древнеанглийский* глагол **tācan** обозначал «to show, point out, declare, demonstrate» (показать, указать, объявить, продемонстрировать) от *протогерманской* основы ***taikijan** «to show» (показывать). Первое упоминание слова появляется в XIV веке [ADWO; EDO; MW; ODOWH]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. слово «*teacher*» имеет одно современное значение «someone whose job is to teach in a school or college» (человек, который преподает в школе или колледже) [CALD], которое появилось в 1375 году в форме «one who teaches or instructs at school» [OED];

2. слово «*teacher*» как *существительное* имеет также ряд частных случаев употребления: первоначальное *устаревшее* значение «that which shows or points out, an indicator, the index-finger» [OED], обозначающее указательный палец или предмет, который показывает или указывает на что-то, восходит к этимологии слова и зафиксировано в текстах с 1290 по 1350 год, и *специализированное* значение «one of several officers appointed to teach» [OED], которое использовалось в конгрегационалистских церквях Новой Англии как должность священнослужителя в период 1834-1850 гг.

Слово «**university**» в английском языке употребляется как существительное и является лексемой для сопоставления со словом «*университет*» в русском языке. Этимологически слово пришло в *среднеанглийский* язык из *французского* как адаптация слова **université**, произошедшего от слова **universitas** в *средневековой латыни* со значением «society, community» (общество, сообщество), которое в свою очередь являлось сокращением фразы «*universitas magistrorum et scholarium*» «community of masters and scholars» (сообщество мастеров и ученых). *Латинская* основа слова **universus** обозначала «combined into one» (объединенный в один). Первое упоминание слова появляется в XIV веке [AD; ADWO; EDO; MW; ODOWH]. Изучая изменения в семантике лексемы по материалам OED и CALD можно выделить следующие особенности:

1. слово «*university*» имеет одно современное значение «a place of higher education where people study for an undergraduate (= first) or postgraduate (= higher level) degree» [CALD] для обозначения высшего учебного заведения, которое появилось в 1300 году в форме «the whole body of teachers and scholars engaged, at a particular place, in giving and receiving instruction in the higher branches of learning» [OED] и сохранилось до настоящего момента;

2. слово «*university*» как *существительное* имеет также ряд устаревших значений, которые восходят к этимологии слова: значение «*the universe, universality*» (вселенная, универсальность) [OED], зафиксированное в текстах в 1494-1642 гг., и значение «*the whole of something, all things, etc. A body or company of persons associated together for some purpose*» (объединение чего-либо, компания людей, объединенных вместе для какой-то цели) [OED], которое использовалось в период с 1374 по 1843 гг.

Завершая этимологический и семантико-дефиниционный анализ выбранных ключевых слов тематического поля «Education» в английском языке в сравнительно-историческом аспекте, можно отметить, что исследуемые лексемы в большей степени являются многозначными и употребляются в широком ряде контекстов помимо сферы образования. При изучении семантических изменений в среднем выделяется не менее трех дифференцируемых значений у каждой лексемы в списке ключевых слов. С этимологической точки зрения основная часть лексики вошла в обиход в период с XIV по XVI век и была заимствована из латинского и греческого языков. Такая особенность слов делает их словами-когнатами для русского языка, например «*examination / экзамен*», «*professional / профессиональный*», «*university / университет*», «*school / школа*», «*student / студент*». В качестве частичных когнатов можно выделить слова «*academic*» и «*manager*». Начиная с XIX века у ряда слов появляются дополнительные коннотации и частные случаи употребления. В целом ключевая лексика тематического поля «Education» в английском языке подверглась значительным изменениям в семантике и сфере употребления слов, что вполне объясняется особенностями истории развития и эволюции языка. В следующем параграфе мы проведем сопоставительный анализ семантических особенностей русскоязычных ключевых слов и их эквивалентов в английском языке с целью выявления схожих и различных черт в сферах употребления лексем. На основе обработки полученных данных можно будет сделать предположения о культурных трендах в сфере образования России и Великобритании.

4.2.6 Сопоставительный анализ тематических полей

«Образование / Education»: сходства и различия

Настоящий параграф описывает результаты сопоставительного анализа семантических особенностей русскоязычных ключевых слов тематического поля «Образование» и их эквивалентов в тематическом поле «Education» английского языка. Таким образом, составленный параллельный зеркальный список ключевых слов, которые выступают заголовочными единицами в моделируемом словаре, позволяет выявить основные сходства и различия семантики лексем в двух сопоставляемых языках.

Для дальнейшего анализа важно первоначально определить понятие *слов - когнатов*, или *межъязыковых омонимов*, в лингвистике. В большей степени вопросами изучения когнатов занимаются исследователи в области теории и методики преподавания иностранных языков, процесса усвоения языка (*language acquisition*) и билингвизма, однако этот аспект также может быть интересен для изучения исторических семантических изменений при сопоставлении лексики пары языков. **Когнаты** (*от латинского cognati – «родственные»*) понимаются как слова разных языков, имеющие общее происхождение и восходящие к одной и той же основе в языке. При этом семантические особенности такой лексики могут обладать индивидуальными характеристиками в разных языках и полностью не совпадать друг с другом, в таком случае принято выделять *полные* и *частичные* когнаты. Слова-когнаты возникают в процессе исторического взаимодействия двух и более языков, например в результате языковых заимствований из неблизкородственных языков или как языковые дублеты для близкородственных [Акуленко, 1972; Безукладников, Штиглуз, Мансурова, 2016].

Кроме этого, для английского языка свойственен процесс *омонимии частей речи*, то есть трансформации слов из одной части речи в другую. При этом такие слова имеют единое звуковое и графическое оформление [Соколова, 2014]. В настоящем исследовании такой процесс наблюдается у ряда лексем английского языка, где происходит омонимия пар прилагательное / существительное или

существительное / глагол. В связи с тем, что подобное явление не присутствует в анализируемой лексике русского языка, мы приняли решение не дифференцировать значения слов английского языка по частям речи, с целью проведения более полного и обобщенного анализа семантических характеристик.

Далее мы переходим к сопоставительному анализу ключевой лексики тематического поля «Образование» / «Education» в русском и английском языках. На данном этапе мы рассматриваем и сопоставляем *современные значения* и контексты употребления слов, которые актуальны при изучении иностранного языка. Первая пара лексем «**наука / science**» с этимологической точки зрения имеют разное происхождение. Существительное русского языка «наука» имеет праславянские корни, тогда как английская лексема «*science*» была заимствована из латинского языка через старофранцузский. Первое упоминание слова «*science*» появляется в XIV веке. Данные о времени появления слова «наука» отсутствуют. Семантически лексема «наука» имеет большее количество значений, чем лексема «*science*». Общими для обоих языков является значение «*система знаний о закономерностях развития природы и общества и способах воздействия на окружающий мир*», а также устаревшее в обоих языках значение «*обучение какому-либо ремеслу*». Дополнительно в русском языке выделяется разговорное значение «*то, что поучает, даёт опыт, урок*».

Пара лексем «**образование / education**» с этимологической точки зрения имеют разное происхождение. Существительное русского языка «образование» происходит от русского церковнославянского языка, тогда как английская лексема «*education*» была заимствована из латинского языка. Первое упоминание слова «*education*» появляется в 1531 году, начало использования слова «образование» датируется XIX веком. Семантически обе лексемы имеют по два значения. Общим для обоих языков является значение «*процесс усвоения знаний, обучение, просвещение*». Дополнительно в американском варианте английского языка выделяется значение «*the study of methods and theories of teaching*» (изучение методов и теорий обучения), которое используется по отношению к направлениям и программам подготовки педагогов в университетах и колледжах США. В русском

языке у слова «образование» существует значение «*то, что образовалось, возникло в результате какого-либо процесса*», тематически не связанное с исследуемой группой лексики.

Пара лексем «**преподаватель / teacher**» с этимологической точки зрения также имеют разное происхождение. Существительное русского языка «*преподаватель*» имеет праславянские корни, а английская лексема «*teacher*» восходит к протогерманской основе. Первое упоминание слова «*teacher*» появляется в XIV веке., данные о времени появления слова «*преподаватель*» отсутствуют. Семантически лексемы имеют одно общее значение «*человек, который занимается преподаванием чего-либо*», при этом в русском языке слово «*преподаватель*» преимущественно используется для обозначения сотрудников в среднем специальном или высшем учебном заведении, а в школах используется слово «*учитель*», тогда как в английском языке слово «*teacher*» используется в широком ряде контекстов и строго не привязано к уровню учебного заведения.

Пара лексем «**профессиональный / professional**» являются словами-когнатами и этимологически восходят к латинскому языку. Первое упоминание слова «*professional*» появляется в 1616 году, слово «*профессиональный*» и его производные пришли в русский язык в XVIII – XIX вв. Семантически английская лексема «*professional*» имеет три значения в качестве прилагательного и три в качестве существительного, которые могут быть соотнесены не только с лексемой «*профессиональный*», но и с однокоренным словом, существительным «*профессионал*». Два значения прилагательных являются общими для обоих языков: «*относящийся или связанный с профессией*» и «*относящийся к профессионализму*». Дополнительно в английском языке выделяется значение прилагательного и существительного «*used to describe someone who does a job that people usually do as a hobby*» (человек, который выполняет работу, которую люди обычно делают в качестве хобби), которое может быть соотнесено со значением слова в «*профессионал*» в русском языке как заимствование и когнат. На продвинутом уровне владения (C1) в английском языке появляется разговорный контекст употребления слова «*professional*» в значении «*someone who has worked*

hard in the same type of job for a long time and has become skilled at dealing with any problem that might happen» (человек, который долгое время упорно работал на одном и том же типе работы и научился решать любые проблемы, которые могут возникнуть).

Пара лексем «**руководитель / manager**» с этимологической точки зрения имеют разное происхождение. Существительное русского языка «*руководитель*» имеет общеславянские и греческие корни, тогда как английская лексема «*manager*» была заимствована из итальянского языка и восходит к латинской основе. Однако обе основы (**roka** и **manus**) имели значение «рука», что, несмотря на разное историческое происхождение слов, делает их близкими по семантике. Первое упоминание основы слова «*руководитель*» датируется XI веком, слово «*manager*» появляется позднее в 1588 году. Семантически лексема «*manager*» имеет два значения, а лексема «*руководитель*» одно. Общим для обоих языков является значение «человек, который руководит или направляет деятельность кого-, чего-либо». Дополнительно в английском языке выделяется технически специализированное значение «*computer program or system that helps you to control or organize a particular type of activity*» (компьютерная программа или система, которая помогает вам контролировать или организовывать определенный вид деятельности). Также в русском языке существует прямое заимствование из английского языка – слово «менеджер», но эта лексема не используется в контексте образовательной деятельности в исследуемых текстах СМИ.

Пара лексем «**студент / student**» являются словами-когнатами, этимологически они были заимствованы из латинского языка. Первое упоминание слова «*student*» появляется в XV веке, а слово «*студент*» впервые было использовано в 1721 году. Семантически лексемы имеют одно общее значение «*учащийся*». При этом в русском языке слово «*студент*» преимущественно используется для обозначения учащихся в высших учебных заведениях, а в английском языке слово «*student*» может обозначать также и старших школьников.

Пара лексем «**университет / university**» являются словами-когнатами, этимологически восходят к латинскому языку. Первое упоминание слова «*university*» появляется в XIV веке, слово «*университет*» получило широкое распространение с 1755 года. Семантически лексемы имеют одно общее значение «*высшее учебно-научное заведение*».

Пара лексем «**учебный / academic**» с этимологической точки зрения имеют разное происхождение. Прилагательное русского языка «*учебный*» имеет праславянские корни, тогда как английская лексема «*academic*», употребляемая в качестве прилагательного и существительного, была заимствована из латинского языка и восходит к греческой основе. Первое упоминание слова «*academic*» появляется в 1580-х гг., данные о времени появления слова «*учебный*» отсутствуют. Семантически лексема «*academic*» имеет большее количество значений, чем лексема «*учебный*». Единственным общим для обоих языков является значение «*связанный с образовательным процессом, учебой*». Дополнительно в английском языке выделяется значение «*someone who teaches at a college, or who studies as part of their job*» (кто-то, кто преподает в колледже или учится в рамках своей работы), которое может быть соотнесено со значением слова «*академик*» в русском языке как частичный когнат. На уровне носителя языка (C2) в английском языке появляются два других значения: «*used to describe someone who is clever and enjoys studying*» (умный человек, которому нравится учиться) и «*based on ideas and theories and not related to practical effects in real life*» (основан на идеях и теориях и не связан с практическими эффектами в реальной жизни), сопоставимое со словом «*академический*» в русском языке.

Пара лексем «**школа / school**» являются словами-когнатами, этимологически были заимствованы из латинского языка и восходят к греческой основе. Первое упоминание слова «*school*» появляется ранее XII века, а слово «*школа*» получило широкое распространение в XVI–XVII вв. Английская лексема «*school*» употребляется в качестве существительного в трех значениях и имеет одно глагольное значение. Лексема «*школа*» имеет четыре значения существительного. Общими для обоих языков являются значения «*учебное заведение, которое*

осуществляет общее образование и воспитание молодого поколения, или где даются специальные знания, профессиональные навыки» и «направление, течение в науке, искусстве, литературе и т.п.». Кроме этого, современное в русском языке специализированное в области сельского хозяйства значение «*питомник, где выращивают растения до посадки на постоянное место*» может быть соотнесено с устаревшим в английском языке значением глагола «*to rear (a plant) in a nursery*». Дополнительно в английском языке выделяется значение глагола «*to train a person or animal to do something*» (обучать человека или животное делать что-либо), а в русском языке значение «*приобретение знаний, выучки, опыта и система приёмов, обязательных упражнений, видов работ при изучении чего-либо, при овладении каким-либо мастерством, искусством*».

Пара лексем «**экзамен / examination**» являются словами-когнатами и этимологически восходят к латинскому языку. Первое упоминание слова «*examination*» появляется в XIV веке, в русском языке слово «экзамен» известно с начала XVIII века. Семантически лексема «*examination*» имеет два значения, а лексема «экзамен» одно. Общим для обоих языков является значение «*проверочное испытание чьих-либо знаний и навыков*». Дополнительно в английском языке выделяется значение «*the act of looking at or considering something carefully in order to discover something*» (процесс рассмотрения или внимательного изучения чего-либо), которое тематически не связано со сферой образования.

Таким образом, можно выделить ряд сходств и различий в ключевой лексике тематического поля «Образование / Education» в русском и английском языках. Во-первых, часть лексем являются словами-когнатами и имеют общую этимологию: «*школа / school*», «*студент / student*», «*экзамен / examination*», «*университет / university*», «*профессиональный / professional*». Во-вторых, анализируемые ключевые слова русского языка по большей части вошли в употребление в XVIII-XIX вв., что значительно позднее в сравнении с их англоязычными эквивалентами, которые появляются в языке с XIV по XVI век. В-третьих, ключевая лексика русского языка обладает меньшим количеством

значений в сравнении с многозначными словами английского языка, в ряду которых также присутствует частеречная омонимия. Проведя анализ параллельного списка выбранных ключевых лексем, мы полагаем, что описанные выше слова могут стать подходящей базой данных для моделирования компьютерного учебного мультимодального словаря историко-культурологического типа, который бы смог выступить в качестве модели знаний о лексике тематического поля «Образование / Education» в русском и английском языках и инструмента визуализации культурных трендов в сфере образования двух стран, России и Великобритании. В следующем параграфе мы перейдем к описанию процесса разработки компьютерного приложения модели словаря и особенностей ее пользовательского интерфейса.

4.3 Проектирование компьютерного приложения модели учебного мультимодального словаря историко-культурологического типа

Проектирование компьютерного приложения модели учебного мультимодального словаря историко-культурологического типа включает в себя решение как теоретических вопросов, так и практических задач. Настоящий параграф описывает этапы работы по выбору языка программирования, подходящего для технологии разработки словаря, проектированию структуры хранения данных, поиску и анализу оптимальных дополнительных справочных источников и мультимодальных элементов, разработку программной платформы и графического пользовательского интерфейса, вводу содержимого словарных статей и апробации результатов.

Для реализации графического интерфейса словаря был выбран формат интерактивной платформы **Shiny Dashboard**²² на языке программирования R. *Shiny* – это пакет R с открытым исходным кодом, разработанный для создания интерактивных веб-приложений в формате *информационных панелей* (dashboard), которые состоят из трех частей (рис. 16): *заголовок* (header), *боковая панель*

²² Shiny Dashboard, URL: <https://shiny.rstudio.com/>

(sidebar) и *тело сайта* (body). Такие приложения отличаются разнообразием средств для визуализации данных и сравнительной простотой создания кода веб-страницы [Shinydashboard]. Основные концепции, принятые при проектировании графической оболочки, включают: *однооконный интерфейс* без использования дополнительных диалоговых окон, организацию переходов по разделам словаря при помощи *гиперссылок*. Интерфейс словаря построен максимально удобно для обычных пользователей и специалистов, не имеет лишних «окон» и «модулей», что обеспечивает высокую продуктивность работы.

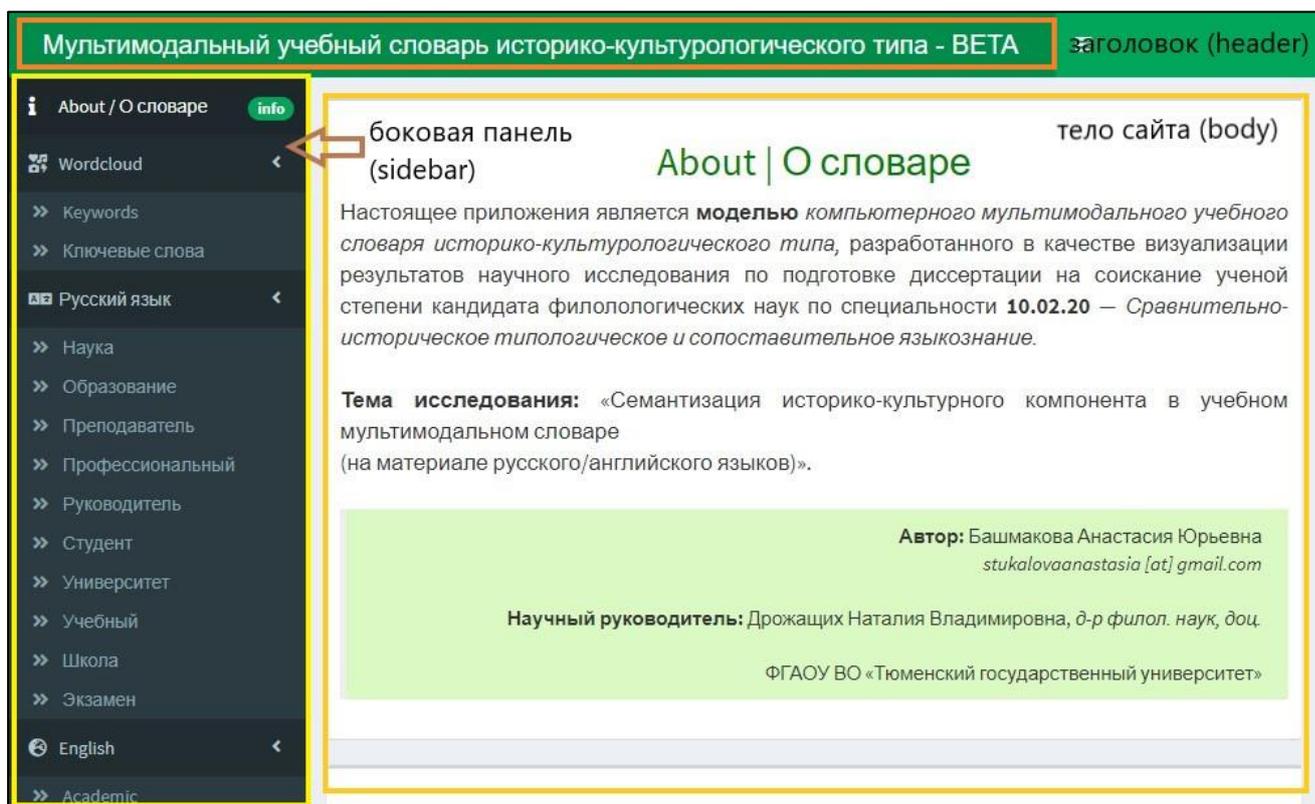


Рисунок 16. Графический пользовательский интерфейс словаря

Макроструктура словаря обеспечивает *одноканальный вход* в словарную статью по алфавитному указателю списка слов в боковой панели и включает в себя:

- *вводную часть: раздел «About | О словаре»* в боковой панели словаря;
- *облака слов: раздел «Wordcloud»* (рис. 17), находящийся в боковой панели и состоящий из двух элементов выпадающего списка, на страницах которых представлены облака слов ключевой лексики тематического поля «Образование / Education» в русском и английском языках, составленные по материалам частотного анализа корпуса текстов образовательных СМИ. При

Рисунок 18. Словарная статья лексемы «Наука»

Рисунок 19. Словарная статья лексемы «Academic»

4) *аудиопрезентация слова* (запись женского или мужского голоса; требуется воспроизведение аудио файла вручную);

5) *современные дефиниции слова с примерами употребления* (текст статичен; дефиниция оформлена обычным начертанием, выравнивание по левой стороне; пример употребления слова выделен в отдельный блок светло-зеленого цвета, текст примера отображается курсивом, выравнивание по правой стороне);

- 6) *этимологическая справка* (вкладка со статичным текстом, включает также дополнительные интересные факты об истории слова);
- 7) *историко-культурный комментарий* (вкладка со статичным текстом, культурологическая справка и информация об истории развития понятия в стране изучаемого языка);
- 8) *информация о частотности употреблений лексемы в диахронии* (гиперссылка на лексикографический ресурс Google Books Ngram Viewer);
- 9) *примеры употребления слова в корпусе* (гиперссылка на НКРЯ для русского языка или на онлайн-сервис Linguazza.com для английского языка);
- 10) *иллюстрации* (статичное изображение; опционально);
- 11) *видео-контент* (встроенный виджет на видео с платформы YouTube; опционально);
- 12) *раздел «Морфемы»* (статичное изображение разбора слова по составу и гиперссылка на онлайн-словарь «Карта слов» в левом нижнем углу (рис. 20); только для русского языка);
- 13) *CEFR Levels* (уровень владения языком, на котором студент должен овладеть тем или иным значением слова по материалам онлайн-словаря CALD; только для английского языка);

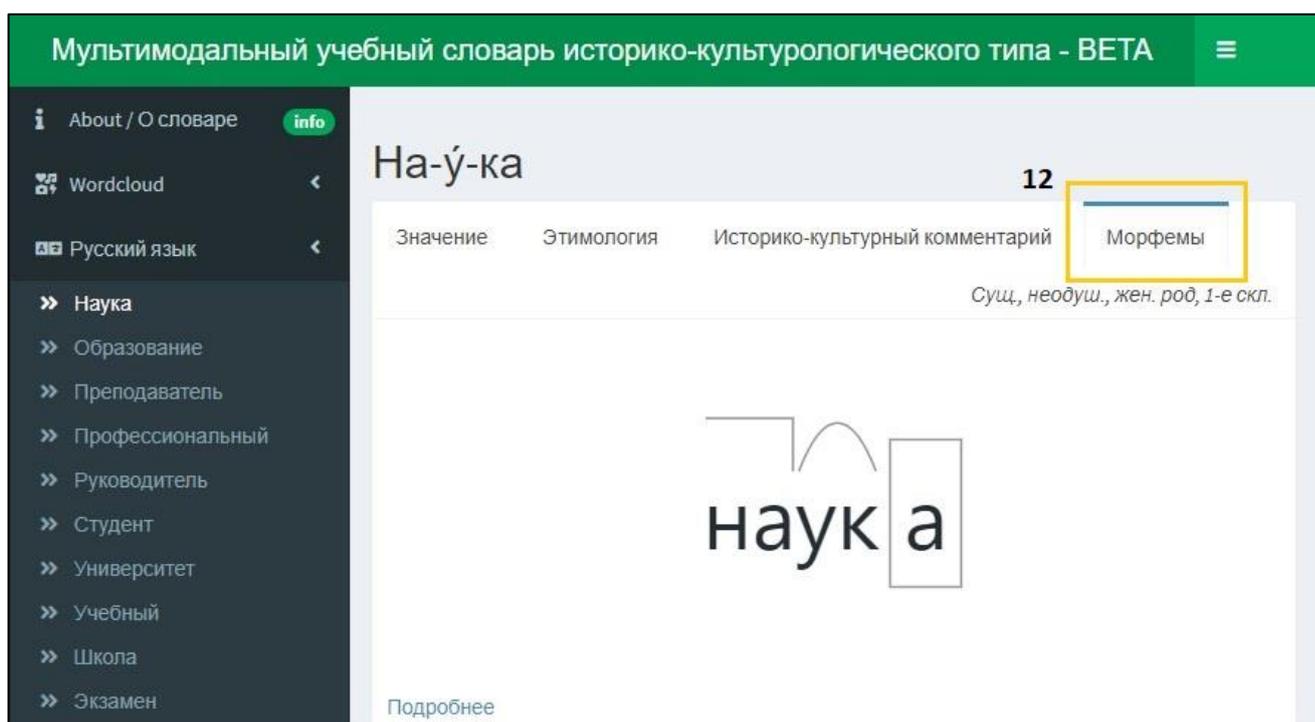


Рисунок 20. Словарная статья лексемы «Наука», раздел «Морфемы»

14) блок «Timeline» (текст и визуализация; только для английского языка).

Отдельно также необходимо описать блок «Timeline», состоящий из *текстовой части*, которая описывает хронологию развития семантических изменений в значениях слов (рис. 21) и ее *визуализации* в виде интерактивной панели в формате таймлайн, расположенной ниже под текстовой частью (рис. 22). Информация в текстовой части вкладки «Timeline» выделяется разными цветами для каждой части речи (зеленый и синий), даты выделены жирным шрифтом, стилистические пометы при наличии отмечены курсивом.

Мультимодальный учебный словарь историко-культурологического типа - BETA

About / О словаре info

Wordcloud <

Русский язык <

English <

Academic >>

Education >>

Examination >>

Manager >>

Professional >>

School >>

Science >>

Student >>

Teacher >>

University >>

English

Professional

Modern definition Etymology Timeline adjective

Adjective

- 1. 1420 - 1881** *Obs.* pertaining to or marking entrance into a religious order
- 2. 1747 - modern** engaged in one of the learned or skilled professions, or in a calling considered socially superior to a trade or handicraft
- 3. 1879 - 1978** disparagingly applied to one who pursues relentlessly an activity or belief that is regarded with disfavour; inveterate, habitual, ruthless

Noun

- 1. 1848 - modern** one who belongs to one of the learned or skilled professions
- 2. 1811 - modern** one who makes a profession or business of any occupation, art, or sport, otherwise usually or often engaged in by amateurs as a pastime
- 3. 1861 - 1977** a prostitute

Рисунок 21. Словарная статья лексемы «Professional», текстовая часть раздела «Timeline»

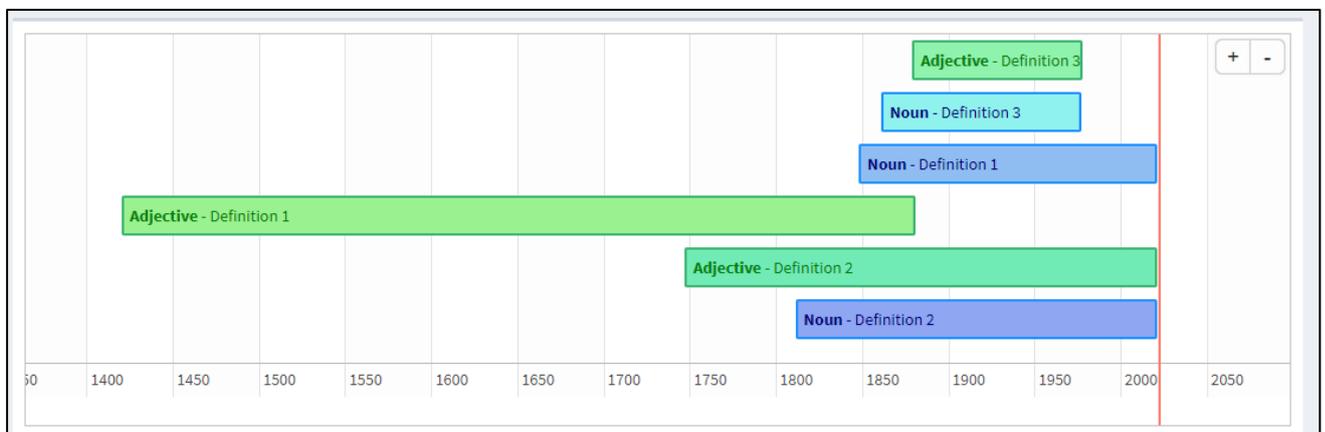


Рисунок 22. Словарная статья лексемы «Professional», визуализация раздела «Timeline»

Номера значений в текстовой части соотносятся с номерами блоков, расположенных на временной шкале «Timeline», в которых значения слова также окрашены в разные цвета для каждой части речи соответственно.

Важно отметить, что удобной характеристикой пользовательского интерфейса моделируемого словаря является его *интерактивность*, внедренная в различные элементы словаря: *переключение по вкладкам* значения, этимологии и комментариев (при выборе любой из вкладок словарной статьи она подсвечивается голубой полоской сверху); *анимация* появления лексем и дополнительной информации об их частотности на карте слов; *блок таймлайн* для визуализации семантических изменений в лексике английского языка (пользователь может увеличивать или уменьшать масштаб отображения и передвигаться по шкале удерживая правую кнопку мыши); *гиперссылки* на сайты других компьютерных лексикографических ресурсов; *проигрыватель* видео и аудио, встроенный в страницу словарной статьи.

Далее мы переходим к рассмотрению содержательных элементов модели словаря, таких как база данных словарных статей, мультимодальные элементы и дополнительные справочные ресурсы. Список включенных в словарь лексем, данные об их частотности, историко-культурологический комментарий, информация об этимологии и другие метаданные представлены в виде файлов-таблиц формата .csv:

- *список современных значений лексем русского языка*, состоящий из столбцов: слово, часть речи, транскрипция, современное значение, пример;
- *историко-культурологические комментарии и этимология лексем русского языка*, столбцы: слово, дата первого упоминания использования слова, этимология, культурный комментарий;
- *список современных значений лексем английского языка*, столбцы: слово, часть речи, транскрипция, уровень CEFR, современное значение, пример;
- *список исторических значений лексем английского языка*, столбцы: слово, часть речи, дата первого упоминания использования слова, дата последнего использования слова, лексическая помета, значение. Данные этой таблицы легли

в основу визуализации семантических изменений в английском языке в формате *таймлайн* (timeline), в котором дата вхождения каждого значения фиксируется как дата первого примера употребления слова из словаря OED, а дата окончания его использования как дата последнего примера употребления в указанном значении в случае, если такое значение не совпадает с современным толкованием слова по материалам онлайн-словаря CALD;

– *историко-культурологические комментарии и этимология лексем английского языка*, столбцы: слово, дата первого упоминания использования слова, этимология, культурный комментарий;

– *частотные списки ключевых слов* русского и английского языков отдельными файлами, столбцы: слово, частотность.

Кроме этого, в статьях представлены различные мультимодальные элементы (аудио- и видеозаписи, иллюстрации) и дополнительные справочные материалы, которые способствуют семантизации историко-культурного компонента значения лексем, представленных в словаре, и повышают привлекательность интерфейса моделируемого приложения. Изображения были подобраны индивидуально для статей на русском и английском языках, чтобы проиллюстрировать реалии каждой из стран (здания школы и университета, схемы организации систем образования России и Великобритании). Технические особенности компьютерного приложения моделируемого мультимодального словаря позволяют включить в словарную статью все элементы, описанные в ее микроструктуре, но не каждая лексема имеет достаточный набор справочных данных или может быть проиллюстрирована изображением или схемой. В связи с этим не все словарные статьи обладают полным списком мультимодальных элементов.

Словарные статьи иллюстрированы цифровыми растровыми изображениями в формате .jpeg или .png и содержат аудиопрезентацию произношения слов на английском и русском языках в формате .mp3. Аудиофайлы и изображения были взяты с портала *Wikimedia Commons (Викисклад)*²³, который является хранилищем

²³ Wikimedia Commons, URL: https://commons.wikimedia.org/wiki/Main_Page

свободных изображений, видео, звуков и других мультимедийных файлов. Дополнительные видеоматериалы оформлены в виде встроенных виджетов с прямой ссылкой на платформу *YouTube*²⁴. Для обеспечения интегрированного доступа к текстовым корпусам и другим ресурсам, наличие которых является одним из основных требований к созданию современного компьютерного словаря, были добавлены гиперссылки на результаты поисковых запросов лексем на сайтах следующих компьютерных лексикографических ресурсов:

- *Google Books Ngram Viewer* для изучения диахронических изменений в частотности употреблений лексем;
- *Национальный корпус русского языка*²⁵ (НКРЯ) для демонстрации выборки примеров употребления слов в русском языке в формате KWIC («key word in context» ключевое слово в контексте);
- *Онлайн-сервис Linguazza.com*²⁶ для демонстрации примеров употребления слов в английском языке. *Linguazza.com*, как ресурс, реализованный командой онлайн-тезауруса «Карта слов», является инструментом для изучения английского языка и позволяет проверить правильность построения фразы на иностранном языке благодаря контекстным примерам употребления. Данный ресурс собирает английские фразы, относящиеся к разным темам, и помещает их в единое дерево, по иерархии которого пользователь может перемещаться и видеть примеры использования интересующих его фраз. Выбор данного ресурса обосновывается его тематической структурированностью, открытым доступом и постоянно обновляющейся базой данных. Официальные корпуса английского языка, такие, как *British National Corpus*²⁷ (BNC) или *Corpus of Contemporary American English*²⁸ (COCA), требуют предварительной регистрации пользователя и не позволяют сохранить прямую ссылку на поисковый запрос для размещения ее в словарной статье, что является значительным недостатком для пользователя.

²⁴ YouTube, URL: <https://www.youtube.com/>

²⁵ Национальный корпус русского языка, URL: <https://ruscorpora.ru/new/index.html>

²⁶ Linguazza.com, URL: <https://linguazza.com/>

²⁷ British National Corpus, URL: <https://www.english-corpora.org/bnc/>

²⁸ Corpus of Contemporary American English, URL: <https://www.english-corpora.org/coca/>

При создании компьютерного веб-приложения модели учебного мультимодального словаря историко-культурологического типа использовались современные технологии и программные продукты (Python 3, R, Shiny Dashboards, HTML и CSS). Бета-версия словаря была протестирована на базе кафедры английской филологии и перевода и кафедры английского языка института социально-гуманитарных наук Тюменского государственного университета, а также в рамках научно-исследовательской стажировки по грантовой программе У. Фулбрайта для выпускников вузов и аспирантов (Fulbright Foreign Student Program) в департаменте лингвистики Государственного университета Монклер (г. Монклер, США), и различных научно-практических конференциях.

Подводя итог, мы считаем, что основным достоинством моделируемого лексикографического продукта является возможность *мультимодальной презентации* материала, *ориентированность словаря на пользователя-иностранца* и *интерактивность* пользовательского интерфейса, которые повышают мотивацию изучающих иностранный язык и способствуют успешному развитию коммуникативной компетенции. При этом общая структура настоящего словаря соответствует всем общепринятым стандартам и технологиям лексикографического моделирования, пользовательский интерфейс прост и интуитивно понятен. Таким образом, моделируемый компьютерный учебный мультимодальный словарь историко-культурологического типа может быть использован для семантизации и изучения новой лексики иностранного языка. В следующем параграфе мы подведем итоги четвертой главы.

Выводы по Главе 4

Результатом заключительной главы настоящего исследования является разработка модели компьютерного учебного мультимодального словаря историко-культурологического типа и анализ языкового материала – основы базы данных словаря для пары языков русский-английский, в процессе работы над которыми мы пришли к следующим выводам:

1. Согласно нашей концепции *учебный мультимодальный словарь историко-культурологического типа* выступает как системный многофункциональный образовательный ресурс, ориентированный на человека, изучающего иностранный язык. Такой словарь не только представляет контекст употребления той или иной лексемы, но и обладает дополнительной экстралингвистической информацией, исторической и культурологической справкой, содержит гипертекстуальный доступ, соединяющий словарь с различными языковыми корпусами и базами данных. Он интерактивный и мультимодальный, что позволяет наглядно визуализировать динамику культуры с помощью инфографики, иллюстраций, видео- и аудиофайлов, отрывков культурно значимых текстов.

2. В качестве исследовательского языкового материала и языковой основы базы данных моделируемого словаря был автоматически собран корпус текстов новостных статей, состоящий из двух частей: *русскоязычная часть корпуса* представлена текстами с сайта образовательного портала «*EDU-Inform*», *англоязычная часть корпуса* – статьи онлайн-издания «*Education Today Magazine*». Языковой материал данных источников прошел обработку методами компьютерной лингвистики с целью выявления ключевых слов, которые способны наиболее репрезентативно продемонстрировать основные темы, характеризующие тематическое поле «Образование / Education» в русском и английском языках.

3. Тексты исследуемых образовательных СМИ в русском и английском языках по материалам анализируемых онлайн-изданий обладают рядом различий как в количественном, так и в качественном аспектах. Так временной период публикаций для русского языка составляет 10 лет (2011 – 2021 гг.), а для английского языка 7 лет (2014 – 2021 гг.). В обоих случаях были использованы все имеющиеся на момент сбора корпуса материалы. Несмотря на небольшую разницу во временном отношении количество публикаций русскоязычного издания (1119 статей) в два раза больше, чем количество англоязычных публикаций (534 статьи). Данный показатель также отразился и на объеме двух частей исследовательского корпуса: 770793 словоупотребления для русского языка

и 285320 словоупотреблений для английского языка. Такие статистические данные показывают, что англоязычная часть корпуса практически в три раза меньше русскоязычной, что не может быть соотнесено в процентном соотношении с различием во временном периоде публикаций и считаться погрешностью. Таким образом, можно сделать вывод, что русскоязычные образовательные СМИ, по данным исследовательского материала настоящего исследования, публикуют большее количество и объем новостных статей, чем их коллеги в Великобритании.

4. Ключевые лексемы представляют собой репрезентативное обобщение основных актуальных тем, связанных со сферой образования России и Великобритании, в которых можно выделить следующие характеристики:

– с лингвистической стороны тексты новостных статей про образование *на русском языке* подвержены значительному влиянию со стороны государства и демонстрируют классические характеристики официально-делового стиля письма. Русскоязычные публикации изобилуют канцеляризмами, именами собственными (названия городов, фамилии, названия министерств и органов власти), сложными прилагательными и нейтральной лексикой. Тематически ключевую лексику публикаций об образовании в России можно разделить на следующие группы: *слова, описывающие участников, типы и процесс образования; слова, связанные с органами государственной власти; слова, описывающие эпидемиологическую ситуацию в условиях коронавируса;*

– тексты публикаций про образование *на английском языке* демонстрируют более типичную для данной сферы ключевую лексику, которую можно поделить на следующие тематические группы: *участники образовательного процесса; учебные дисциплины; процесс обучения; образовательные учреждения; благотворительность; образовательный менеджмент; пандемия.* Тематика благотворительности является культурно-специфической для Великобритании, в то время как вопросы образовательного менеджмента являются общими и для материалов русскоязычного источника;

– интересной культурной особенностью также можно выделить преобладание ключевой лексики, тематически связанной с освещением

организации образовательного процесса в условиях пандемии коронавируса, в новостных статьях на обоих языках. Несмотря на небольшой временной промежуток публикаций, полтора года с начала дистанционного формата обучения, ключевая лексика, связанная с этой темой, имеет высокую частотность, что демонстрируют данные визуализации облаков слов.

5. Несмотря на имеющиеся различия в русскоязычной и англоязычной частях корпуса, собранного для настоящего исследования, для дальнейшего «ручного» этимологического, семантико-дефиниционного и сопоставительного анализа были выделены лексемы, которые составили зеркальный список ключевой лексики двух языков на основе их тематической принадлежности к сфере современного образования и частотности: «*профессиональный / professional*», «*наука / science*», «*образование / education*», «*преподаватель / teacher*», «*учебный / academic*», «*университет / university*», «*руководитель / manager*», «*студент / student*», «*школа / school*», «*экзамен / examination*». Задачей дальнейшего анализа было проследить как исторически менялась семантика ключевых лексем, выявленных по материалам современных новостных статей в сфере образования, который показал следующие особенности:

- ключевая лексика английского языка начала входить в употребление в XIV-XVI вв., что значительно раньше слов русского языка, появившихся в XVIII-XIX вв. Данную особенность можно объяснить историческим развитием системы образования Европы и России. Первые университеты в Европе и Великобритании начали появляться как раз в указанный период, с XIII-XIV вв., что возможно способствовало формированию и развитию ключевой лексики тематического поля «Education» в английском языке. Первые университеты в России, в свою очередь, были открыты только в XVIII веке. Это может также стать причиной того, что половина слов из анализируемого списка ключевой лексики были заимствованы в русский язык и поэтому являются словами-когнатами, имеющими общую этимологию с англоязычными лексемами;

- немногие ключевые слова английского языка сохранили стабильное значение на протяжении всей своей истории. Анализ показал, что исследуемая

лексика английского языка в большей степени является многозначной и имеет в среднем по три дифференцируемых значения. Кроме этого, ряд лексем имеют частеречную омонимию, что также расширяет список семантических характеристик слова;

– ключевые слова русского языка в большей степени имеют только одно значение, которое не было подвержено значительным изменениям во временной период от вхождения слова в употребление до настоящего момента.

Здесь также важно отметить, что многозначность слов может стать сложностью при изучении новой лексики иностранного языка учащимися, но при этом она отражает ценную информацию об историческом развитии семантики слов и может стать основой для анализа культурных трендов в сфере образования России и Великобритании.

Таким образом, проанализированный языковой материал, описывающий ключевую лексику тематического поля «Образование / Education» в русском и английском языках, с подробной семантической, историко-культурной и экстралингвистической информацией стал лингвистической базой данных для модели компьютерного *учебного мультимодального словаря историко-культурологического типа*. На наш взгляд, такой словарь может помочь решению актуальных проблем традиционной и компьютерной лексикографии, так как его образовательная ценность состоит в том, что пользователь имеет возможность изменять параметры поиска по различным критериям: хронология, этимология, контекст, частотность, семантика, примеры употребления, и нелинейно изучать как язык, так и историю и культуру его носителей. Завершая основную часть настоящего исследования, в следующем параграфе мы перейдем к заключению диссертации.

ЗАКЛЮЧЕНИЕ

Настоящее диссертационное исследование было направлено на изучение особенностей и выявление принципов семантизации историко-культурного компонента в учебном мультимодальном словаре с целью разработки модели компьютерного учебного мультимодального словаря, способного визуализировать динамику семантических трансформаций лексики в диахронии на материале русского и английского языков. В заключении сформулируем основные выводы, полученные в результате исследования.

В первой главе мы рассмотрели теоретические предпосылки настоящей работы, изучив исследования в области *компьютерной лексикографии* и *мультимодальности*. В рамках анализа была описана история развития компьютерной лексикографии, а также основные характеристики и типология компьютерных лексикографических ресурсов. На примере существующих технологических решений для русского и английского языков (*мезаурис PyТез, RuWordNet, мезаурис «Карта слов», WordNet, FrameNet, The Visual Thesaurus, Fine Dictionary*) были проанализированы особенности использования в словаре элементов различных модальностей, таких как *текст, изображения, аудио- и видеоматериалы*, которые способны облегчить процесс изучения иностранного языка, а также повысить интерес и вовлеченность обучающихся.

Вторая глава была посвящена изучению актуальных направлений и результатов исследований в области *компьютерной лингводидактики* и *учебной лексикографии* в рамках *мультимодального подхода*. Теоретический анализ литературы позволил описать основные тенденции в разработке компьютерных образовательных ресурсов для изучения иностранного языка, а также обозначить роль компьютерных словарей в процессе обучения языку и знакомства с новой иноязычной лексикой. В рамках обзора были комплексно описаны вопросы создания, задачи и этапы разработки учебных словарей для изучающих иностранный язык на примере методологии различных научных лексикографических школ.

В третьей главе была определена значимость вопроса *семантизации историко-культурного компонента лексемы* при создании компьютерного учебного мультимодального словаря для изучающих иностранный язык как нового типа словаря. В главе были описаны основные направления исследований в области *лингвокультурной лексикографии, семасиологии, цифровой гуманитаристики, культурологии*. Теоретический анализ литературы позволил дать определение **историко-культурного (диахронического) компонента** значения лексемы как *совокупности информации об истории семантического развития слова, национально-культурных коннотациях, контексте и особенностях его употребления в определенную историческую эпоху, которая описывает слово как достояние определенной лингвокультуры*. В дальнейшем были проанализированы способы переводной и беспереводной семантизации, а также сформулированы **принципы и особенности семантизации историко-культурного компонента**:

- *описательное вербальное толкование слова на иностранном языке;*
- *краткая этимологическая справка;*
- *экстралингвистический комментарий;*
- *мультимодальные элементы;*
- *наглядная визуализация языкового окружения лексемы;*
- *гипертекстовый доступ.*

По итогам главы была представлена схема-инфографика, отражающая *содержательные и структурные компоненты процесса моделирования компьютерного учебного мультимодального словаря для изучающих иностранный язык*.

Четвертая глава была посвящена описанию этапов и технологии разработки модели компьютерного учебного мультимодального словаря для визуализации результатов диссертационного исследования и изучению практических аспектов состава тематического поля «Образование / Education» в русском и английском языках. В качестве финального продукта была представлена пилотная модель компьютерного учебного мультимодального словаря историко-культурологического типа для пары языков русский-английский,

проанализирована его макро- и микроструктура и языковой материал тематического поля «Образование / Education».

Моделируемый словарь является средством обучения и обширной базой данных, системным многофункциональным справочным пособием. Он соответствует характеристикам, которыми должен обладать современный компьютерный учебный словарь для изучающих иностранный язык, определенным в теоретической части исследования, а также в нем реализованы сформулированные принципы и особенности семантизации историко-культурного компонента. Отличительной особенностью учебного мультимодального словаря историко-культурологического типа является учет исторического развития семантики слова, визуально представленной на *интерактивной хронологической шкале таймлайн (timeline)*. *Наглядность, интерактивность, нелинейность, динамичность, историчность, антропоцентричность, культурологическая ориентированность* представляют важнейшие параметры создания и функционирования разрабатываемого словаря, который, по большому счету, можно охарактеризовать как словарь культурных трендов эпохи. Слово в таком словаре выступает как проводник в мир вечно меняющихся значений культурной парадигмы.

В качестве лингвистической базы данных для моделируемого словаря был создан корпус текстов тематического поля «Образование / Education» в русском и английском языках посредством автоматизированного веб-скрейпинга новостных публикаций образовательных онлайн-порталов «*EDU-inform*» и «*Education Today Magazine*». С использованием моделей нейросети BERT по материалам исследовательского корпуса были выявлены ключевые слова в сфере современного образования и описаны их семантические характеристики. По результатам этимологического, семантико-дефиниционного и сопоставительного анализа ключевой лексики был выявлен ряд сходств и различий между языками:

– статистические показатели исследовательского корпуса новостных текстов на русском и английском языках подтверждают, что русскоязычные

образовательные СМИ публикуют большее количество и объем новостных материалов, чем англоязычные;

– исследуемые тексты образовательных СМИ на русском и английском языках отражают культурные особенности сферы образования двух стран. Так, тексты новостных статей на русском языке подвержены значительному влиянию со стороны государства, что представлено в обилии канцеляризмов, имен собственных, сложных прилагательных и нейтральной лексике. В публикациях на английском языке частотной становится тематика благотворительности, которая является культурно-специфической для Великобритании. В том числе в последние годы в обеих странах становятся актуальными вопросы образовательного менеджмента и отмечается преобладание лексики, тематически связанной с освещением организации образовательного процесса в условиях пандемии коронавируса;

– анализируемая ключевая лексика английского языка начала входить в употребление значительно раньше (XIV-XVI вв.), чем лексика русского языка (XVIII-XIX вв.). Данную особенность можно объяснить историческим развитием системы образования Европы и России, что также может стать причиной заимствований и наличия слов-когнатов, имеющих общую этимологию в русском и английском языках;

– сопоставительный анализ зеркального списка ключевых слов тематического поля «Образование / Education» в русском и английском языках показал, что немногие слова сохранили стабильное значение на протяжении всей своей истории. Кроме того, исследуемая лексика английского языка более многозначна, чем слова русского языка. Такая особенность может стать сложностью при изучении новой лексики иностранного языка учащимися, но при этом она отражает ценную информацию об историческом развитии семантики слов и может стать основой для анализа культурных трендов в сфере образования России и Великобритании.

Резюмируя итоги проведенного исследования, отметим, что гипотеза настоящего исследования, согласно которой принципы и технологии

моделирования в современной мультимодальной лексикографии могут позволить создать учебный компьютерный словарь нового типа, способный визуализировать диахронические изменения в языке и семантизировать историко-культурные особенности лексики типологически разных языков, может считаться подтвержденной.

В заключение также отметим, что настоящее исследование затронуло лишь некоторые аспекты изучения семантизации новой лексики иностранного языка и лексикографического моделирования в контексте компьютерной мультимодальной учебной лексикографии и цифровой гуманитаристики, которые являются актуальными направлениями в современной лингвистике. На наш взгляд, в перспективе представленный в настоящем исследовании подход к созданию модели компьютерного учебного мультимодального словаря историко-культурологического типа может найти применение при создании аналогичных и более совершенных лексикологических ресурсов как для пары русский / английский язык, так и для других языков.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Авдеева, М. В. Компьютерная лексикография в системе практической и теоретической лексикографии / М. В. Авдеева // Актуальные вопросы теории и практики лингвострановедческой лексикографии : материалы международной научно-практической конференции. – 2012. – С. 26-28.
2. Агапова Н. А. О принципах создания электронного словаря лингвокультурологического типа: к постановке проблемы / Н. А. Агапова, Н. Ф. Картофелева // Вестник Томского государственного университета. – 2014. – № 382.
3. Айсакова, Е. А. Проблема семантизации слов тематической группы «Образование» / Е. А. Айсакова, Т. А. Филипская // Мир науки, культуры, образования. – 2018. – № 6 (73).
4. Акуленко, В. В. Вопросы интернационализации словарного состава языка / В. В. Акуленко. – Харьков : Издательство Харьковского университета, 1972. – 213 с.
5. Аникина, М. Н. Русский словарь : учебный словарь русского языка для иностранцев / М. Н. Аникина. – 2-е изд., стереотип. – Москва : Дрофа, 2014. – 974, [2] с.
6. Ансимова, О. К. Лингвокультурология как отдельная филиация общей лексикографии / О. К. Ансимова // Сборник конференций НИЦ Социосфера. – 2012. – № 44. – С. 12-18.
7. Ансимова, О. К. Семантизация лингвокультурных единиц в толковых словарях русского языка / О. К. Ансимова // Вестник Вятского государственного университета. – 2017. – № 7.
8. Ансимова, О. К. Словари лингвокультуры в системе средств обучения русскому языку как иностранному / О. К. Ансимова // Актуальные проблемы русского языка и методики его преподавания в иностранной аудитории. – 2019. – С. 71-75.

9. Антоненко, К. Б. Количественный анализ культурного следа технического прогресса / К. Б. Антоненко // Дослідження з історії техніки: збірник наукових праць. – 2011. – Вип. 14. – С. 6-22.
10. Арнольд, И. В. Семантическая структура слова в современном английском языке и методика ее исследования / И. В. Арнольд. – Ленинград : Просвещение, 1966. – Т. 191. – С. 5.
11. Атрашевская, О. К. Семантизация англоязычного слова путем анализа его морфемных составляющих / О. К. Атрашевская // Карповские научные чтения : сб. науч. ст. : в 2-х ч. – Минск : Белорусский Дом печати, 2012. – Вып. 6, Ч. 1. – С. 63–67.
12. Ахманова, О. С. Очерки по общей и русской лексикологии / О. С. Ахманова. – Москва : Государственное учебно-педагогическое Издательство Министерства Просвещения РСФСР. – 1957.
13. Башмакова, А. Ю. Культуромика в корпусном исследовании / А. Ю. Башмакова // Множественность интерпретаций : верификация гуманитарного знания : материалы VII молодежной научно-практической конференции (Тюмень, 19 декабря 2018 г.) / Тюменский государственный университет, Институт социально-гуманитарных наук ; научный редактор Е. В. Михалькова. – Тюмень : Издательство Тюменского государственного университета, 2019. – С. 17-22. – URL: <https://elib.utmn.ru/jspui/handle/ru-tsu/3387> (дата обращения: 25.02.2022).
14. Башмакова, А. Ю. Выявление ключевых слов тематического поля "Образование/Education" / А. Ю. Башмакова // Филологические науки. Вопросы теории и практики. – 2021. – Т. 14. – № 12. – С. 3986-3990. – DOI 10.30853/phil20210598
15. Башмакова, А. Ю. Классические и современные подходы к изучению семантических изменений / А. Ю. Башмакова // Успехи гуманитарных наук. – 2022. – № 1. – С. 80-85.
16. Безукладников, К. Э. О необходимости обучения распознаванию немецко-английских когнатов при обучении немецкому языку как второму иностранному / К. Э. Безукладников, Л. Б. Штиглуз, Е. Ф. Мансурова //

- Проблемы романо-германской филологии, педагогики и методики преподавания иностранных языков. – 2016. – № 12.
17. Беликов, В. И. Что и как может получить лингвист из оцифрованных текстов / В. И. Беликов // Сибирский филологический журнал. – 2016. – № 3. – С. 17-34.
 18. Белозерова, Н. Н. Использование электронных библиотек и корпусная лингвистика / Н. Н. Белозерова // Современные библиотеки в инновационном образовательном процессе : материалы Всероссийской научно-практической конференции, Тюмень, 16–17 октября 2004 года. – Тюмень: Тюменский государственный университет, 2004. – С. 50-55.
 19. Белозерова, Н. Н. Функции когнитивных доминант "смежность" и "замещение (селекция)" при усвоении иностранных языков в ходе учебного процесса / Н. Н. Белозерова, С. Н. Антонова // Инновационные проекты и программы в образовании. – 2017. – № 1. – С. 41-47.
 20. Беляев, Б. В. Очерки по психологии обучения иностранным языкам / Б. В. Беляев. – Москва : Просвещение, 1965. – Т. 228. – С. 219.
 21. Беляева, Л. Н. Лингвистические технологии в современном сетевом пространстве. Language worker в индустрии локализации / Л. Н. Беляева. – Книжный дом, 2016. – 134 с.
 22. Беляевская, Е. Г. Семантика слова : учебное пособие для институтов и факультетов иностранных языков / Е. Г. Беляевская. – Москва, Высшая школа. – 1987.
 23. Березовская, Е. А. Современная лексикография: возможности электронных словарей / Е. А. Березовская, Е. В. Сухова // Русский язык : человек, культура, коммуникация – IV : сборник материалов Международной научной конференции. – Екатеринбург, 2014. – С. 179-183.
 24. Берков, В. П. Словарь и культура народа / В. П. Берков // Мастерство перевода. – Москва : Советский писатель, 1975. – С. 402-420.
 25. Бовтенко, М. А. Компьютерные средства обучения языку: современные возможности / М. А. Бовтенко // Компьютерные инструменты в образовании. – 2000. – № 6.

26. Богачёва, Г. Ф. Электронный идеографический словарь : теоретический и прикладной аспекты (на материале «Русского тематического словаря») / Г. Ф. Богачёва, А. И. Ольховская, М. К. Парамонова // Вопросы лексикографии. – 2017. – № 12. – С. 39-60.
27. Большакова, М. А. Использование семантической сети при создании интеллектуального электронного словаря для немецко-русского отраслевого перевода / М. А. Большакова // Известия Волгоградского государственного педагогического университета. – 2013. – № 4 (79). – С. 49-53.
28. Бонч-Осмоловская, А. А. Культуромика в национальном корпусе русского языка, к постановке задачи: три века русских дорог / А. А. Бонч-Осмоловская // Труды института русского языка им. В. В. Виноградова. – 2015. – Т. 6. – С. 605-641.
29. Бонч-Осмоловская, А. А. Имена времени: Эпитеты десятилетий в национальном корпусе русского языка как проекция культурной памяти / А. А. Бонч-Осмоловская // Шаги/Steps. – 2018. – Т. 4, № 3-4.
30. Бочегова, Н. Н. Диалог культур в аспекте интерлингвокультурологии / Н. Н. Бочегова // Вестник Курганского государственного университета. – 2018. – № 2(49). – С. 11-14.
31. Боярская, Е. Л. Qualia-структура Дж. Пустейовского и ее значение для проведения лингвистических исследований / Е. Л. Боярская // Вестник Балтийского федерального университета им. И. Канта. Серия: Филология, педагогика, психология. – 2012. – № 2.
32. Будагов, Р. А. Категория значения в разных направлениях современного языкознания / Р. А. Будагов // Вопросы языкознания. – 1974. – Т. 4.
33. Бутакова, Л. О. Языковая способность, речевая компетенция, речевая деятельность представителей поколения Z / Л. О. Бутакова // Русский язык в школе. – 2021. – Т. 82. – № 4. – С. 61-74. – DOI 10.30515/0131-6141-2021-82-4-61-74.
34. Ваулина, Е. Ю. Толковый словарь. Возможности электронной версии / Е. Ю. Ваулина // Структурная и прикладная лингвистика. – 2007. – С. 187-196.

35. Верещагин, Е. М. Лингвострановедческая проблематика страноведения в преподавании русского языка иностранцам / Е. М. Верещагин, В. Г. Костомаров. – Москва : МГУ, 1971.
36. Виноградов, В. В. Слово и значение как предмет историко-лексикологического исследования / В. В. Виноградов // Вопросы языкознания. – 1999. – № 1. – С. 5-36.
37. Волошина, И. С. Модель электронного переводного словаря-справочника естественнонаучных омонимичных терминов : специальность 10.02.21 «Прикладная и математическая лингвистика» : диссертация на соискание ученой степени кандидата филологических наук / И. С. Волошина. – Тюмень, 2015. – 239 с.
38. Габдулхаков, В. Ф. О лингводидактике поликультурного образования / В. Ф. Габдулхаков, А. М. Гареева, А. М. Гарифуллина // Образование и саморазвитие. – 2012. – № 5. – С. 89-95.
39. Гак, В. Г. Проблема создания универсального словаря (энциклопедический, культурно-исторический и этнолингвистический аспекты) / В. Г. Гак // Национальная специфика языка и её отражение в нормативном словаре : сборник научных статей. – Москва, 1988. – С. 119-125.
40. Герд, А. С. Прикладная лингвистика / А. С. Герд. – Санкт-Петербург : Издательство Санкт-Петербургского университета, 2005. – 268 с.
41. Глобина, Л. В. Лексико-семантическое поле партитивной лексики в современном русском языке : специальность 10.02.01 «Русский язык» : автореферат диссертации на соискание ученой степени кандидата филологических наук / Л. В. Глобина ; Воронежский государственный университет. – Воронеж, 1995. – 18 с.
42. Григоренко, О. В. Словарное освоение новых слов русского языка средствами компьютерной лексикографии / О. В. Григоренко // В мире науки и искусства : вопросы филологии, искусствоведения и культурологии. – 2015. – № 54. – С. 95-99.

43. Гринев, С. В. Основы лексикографического описания терминосистем : специальность 10.02.04 «Германские языки» : диссертация на соискание ученой степени доктора филологических наук / С. В. Гринев. – Москва, 1990. – 436 с.
44. Голубкова, Е. Е. Новый облик современной лексикографии : будет ли существовать словарь через 20 лет? / Е. Е. Голубкова // Вестник Московского государственного лингвистического университета. – 2014. – № 20 (706).
45. Гулевская, А. Е. Мультиmodalный словарь фитонимов нового типа / А. Е. Гулевская, С. Л. Мишланова // Редакционный совет. – 2017. – С. 58.
46. Дакукина, Т. А. Семантизация лексических единиц при обучении чтению оригинальных иноязычных текстов / Т. А. Дакукина // Научно-педагогическое обозрение. Pedagogical review. – 2015. – № 1 (7).
47. Дорофеева, И. В. Общенаучный метод моделирования и специфика его применения в лингвистике / И. В. Дорофеева // Вестник Московской международной академии. – 2015. – № 2.
48. Дрожащих, Н. В. Лингвистические структуры в диахроническом образовательном дискурсе Англии / Н. В. Дрожащих // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. – 2012. – № 1. – С. 43-48.
49. Дрожащих, Н. В. Введение в динамическую синергетику языка / Н. В. Дрожащих, Т. Ю. Белоногова. – 2-е издание, переработанное и расширенное. – Тюмень : Тюменский государственный университет, 2018. – 212 с.
50. Дрожащих, Н. В. Учебный мультиmodalный словарь историко-культурологического типа в преподавании иностранного языка / Н. В. Дрожащих, А. Ю. Башмакова. – DOI: 10.21684/2411-197X-2021-7-1-43-61 // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. – 2021. – Том 7, 1 (25). – С. 43-61.
51. Дубичинский, В. В. Теоретическая и практическая лексикография / В. В. Дубичинский. – Вена ; Харьков, 1998. – 160 с.

52. Дуринова, Г. В. Когнитивные методы в исторической семантике / Г. В. Дуринова // Вестник Московского университета. Серия 9: Филология. – 2015. – № 6. – С. 124-133.
53. Зализняк, А. Семантическая деривация в синхронии и диахронии : проект «Каталога семантических переходов» / А. Зализняк // Вопросы языкознания. – 2001. – Т. 2. – С. 13-25.
54. Зализняк, А. А. Семантический переход как объект типологии // Вопросы языкознания. – 2013. – № 2. – С. 32-51.
55. Зарипова, Р. Р. Компьютерные технологии в инновационном обучении иностранным языкам : конспект лекций / Р. Р. Зарипова. – 2014. – 106 с.
56. Звегинцев, В. А. История языкознания XIX и XX веков в очерках и извлечениях / В. А. Звегинцев. – 2-е изд., доп. – Москва : Учпедгиз, 1960.
57. Ибрагимова, В. Л. Лексика орудийности: семантическое и историко-мотивационное содержание (по данным словарей) / В. Л. Ибрагимова // Динамика языковых и культурных процессов в современной России. – 2018. – № 6. – С. 237-242.
58. Ибрагимова, В. Л. Проблема мотивированности языкового знака в контексте сопоставительной лингвокультурологии / В. Л. Ибрагимова, В. Чуньму // Вестник Башкирского университета. – 2019. – Т. 24. – № 2. – С. 390-395.
59. Иванищева, О. Н. Лексикографирование культуры в двуязычном словаре : специальность 10.02.19 «Теория языка» : диссертация на соискание ученой степени доктора филологических наук / О. Н. Иванищева ; Санкт-Петербургский государственный университет. – Санкт-Петербург, 2005.
60. Иванова, Е. П. Мультимедийные и электронные словари французского языка в их отношении к традиционной печатной лексикографии / Е. П. Иванова // Вестник Челябинского государственного университета. – 2007. – № 1. – С. 44-50.
61. Ильин, Д. Н. Развитие лексической семантики: процессы позитивации / негативации значения слова в русском языке / Д. Н. Ильин ; Южный

- федеральный университет. – Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2018. – 150 с.
62. Ильина, Е. В. Полиинтерпретируемость текста и сходство читательских интерпретаций : проблема методики верификации в филологической герменевтике / Е. В. Ильина // Понимание и рефлексия в культуре, науке и образовании : международная научно-практическая Интернет-конференция и круглый стол, посвященная 85-й годовщине со дня рождения заслуженного деятеля образования, профессора, доктора филологических наук Георгия Исаевича Богина: сборник научных трудов / Н. Ф. Крюкова (отв. редактор) : Тверской государственный университет, 2015. – С. 41-45.
63. Кантышева, Н. Г. Моделирование терминосистемы «экологический аудит» и её описание в систематизирующей глоссарии : специальность 10.02.21 «Прикладная и математическая лингвистика» : диссертация на соискание ученой степени кандидата филологических наук / Кантышева Надежда Геннадьевна. – Тюмень, 2011. – 198 с.
64. Кантышева, Н. Г. Гипермедиальные структурные элементы представления специальных знаний в компьютерной лексикографии / Н. Г. Кантышева // Вестник Челябинского государственного университета. – 2012. – № 32(286). – С. 47-50.
65. Кашкин, Е. В. Типы информации о лексических конструкциях в системе ФреймБанк / Е. В. Кашкин, О. Н. Ляшевская // Труды института русского языка им. В. В. Виноградова. – 2015. – Т. 6. – С. 464-556.
66. Кибрик, А. А. Мультимодальная лингвистика / А. А. Кибрик // Когнитивные исследования. – 2010. – Т. 4. – С. 134-152.
67. Клейман, Э. И. Формирование профессионально-ориентированного тезауруса-лексикона средствами многоуровневой мультимедийной презентации в рамках обучения иностранных студентов на довузовском этапе / Э. И. Клейман, Е. Ю. Мамонова // Межкультурная коммуникация: лингвистические аспекты : сборник материалов VIII международной научно-практической конференции,

- Новосибирск, 28–29 марта 2019 года. – Новосибирск: Новосибирский государственный технический университет, 2019. – С. 370-376.
68. Ковязина, М. А. Функциональная модель двуязычного экологического словаря-тезауруса : специальность 10.02.21 «Прикладная и математическая лингвистика» : диссертация на соискание ученой степени кандидата филологических наук / Ковязина Марина Анатольевна. – Тюмень, 2006. – 373 с.
69. Колокольникова, М. Ю. Историческое развитие принципов и методов семасиологических исследований / М. Ю. Колокольникова ; Саратовский государственный университет имени Н. Г. Чернышевского. – 2012.
70. Конопелько, И. П. Сопоставительный анализ в исследовании и преподавании языка / И. П. Конопелько. – Воронеж : ООО «РИТМ», 2019. – 226 с.
71. Крамаренко, О. Л. Проблема лексикографирования культурно-маркированных лексических единиц в учебном словаре / О. Л. Крамаренко, О. Ю. Богданова // Верхневолжский филологический вестник. – 2020. – № 3(22). – С. 164-170. – DOI 10.20323/2499-9679-2020-3-22-164-170.
72. Крамаренко, О. Л. К вопросу о выборе методики анализа культурологической ценности лексикографических произведений / О. Л. Крамаренко, О. Ю. Богданова, Ю. В. Мошкина // Казанская наука. – 2021. – № 3. – С. 114-116.
73. Кузьмина, Е. О. Способы семантизации новой лексики на начальном этапе обучения РКИ / Е. О. Кузьмина // Проблемы современного образования. – 2019. – № 1.
74. Лаврищева, Е. М. Интерфейс в программировании / Е. М. Лаврищева // Проблемы програмування. – 2007. – № 2 – С. 126-139.
75. Левенкова, А. Ю. Структурно-функциональная модель терминологического словосочетания солнечной энергетики и её лексикографическое описание : специальность 10.02.21 «Прикладная и математическая лингвистика» : диссертация на соискание ученой степени кандидата филологических наук / Анна Юрьевна Левенкова. – Тюмень, 2016. – 274 с.

76. Левенталь, И. В. Семантизация лексических единиц в учебных словарях для иностранцев : поиск нового подхода / И. В. Левенталь // Филологические науки. Вопросы теории и практики. – 2014. – № 6-2. – С. 120-125.
77. Левковская, К. А. Теория слова, принципы её построения и аспекты изучения лексического материала / К. А. Левковская. – Москва, 1962. – 296 с.
78. Лексико-семантическое поле как диагностическая модель и средство методической коррекции развития лексикона учащихся / Н. А. Аксарина [и др.] // Филологические науки. Вопросы теории и практики. – 2015. – № 3-1. – С. 19-22.
79. Лингвистическое моделирование / С. Антонова, И. Афонина, А. Дрожащих [и др.]. – Тюмень : Издательство «ВекторБук», 2009. – 188 с.
80. Линдгрэн, Н. Органы чувств животных и их электронные аналоги / Н. Линдгрэн // Электроника. – 1962. – № 7. – С. 22–27.
81. Лукашевич, Н. В. Тезаурусы в задачах информационного поиска / Н. В. Лукашевич. – Москва, 2010. – 396 с.
82. Лукьянова, Н. А. Лингвокультурология в системе современной российской лексикографии (90-е гг. XX-начало XXI в.) / Н. А. Лукьянова // *Techn grammatike* (Искусство грамматики) : сборник памяти профессора К. А. Тимофеева. – Новосибирск : Новосибирский государственный университет. – 2005. – № 2. – С. 624.
83. Лыкова, Н. Н. Генезис языка права: начальный этап (на материале французских и русских документов X-XV веков) : специальность 10.02.20 «Сравнительно-историческое, типологическое и сопоставительное языкознание» : диссертация на соискание ученой степени доктора филологических наук / Лыкова Надежда Николаевна. – Екатеринбург, 2005. – 411 с.
84. Лыкова, Н. Н. Региональные и миноритарные языки во франкофонном культурном ареале : историко-культурный аспект / Н. Н. Лыкова // Вестник Тюменского государственного университета. Гуманитарные исследования. *Humanitates*. – 2015. – Т. 1, № 3. – С. 168-175.

85. Лысякова, М. В. Лексико-семантические парадигмы : лингвистический статус, критерии разграничения / М. В. Лысякова // *Russian Journal of Linguistics*. – 2005. – № 7.
86. Мальцев, В. В. Семантизация нового иноязычного материала как один из компонентов процесса усвоения иностранного языка в учебных условиях / В. В. Мальцев // *Альфа-2018 : сборник научных статей*. – Гродно : Гродненский государственный университет, 2018. – С. 741-744.
87. Мамонова, Е. Ю. Принципы составления англо-русского переводческого лексикона-тезауруса по тематике "Добыча нефти" / Е. Ю. Мамонова, А. Т. Фатуллаева // *Филологические науки. Вопросы теории и практики*. – 2020. – Т. 13. – № 9. – С. 269-275. – DOI 10.30853/filnauki.2020.9.50.
88. Марти, А. О понятии и методе всеобщей грамматики и философии языка / А. Марти // *История языкознания XIX и XX вв. в очерках и извлечениях* / В. А. Звягинцев. – Москва, 1960. – Ч. 2. – С. 7-12.
89. Мартынюк, А. Я. Русская лексикография : традиции и перспективы / А. Я. Мартынюк // *Культура народов Причерноморья*. – 2004. – № 49-1. – С. 180-182.
90. Миронов, В. В. Трансформация культуры в пространстве глобальной коммуникации / В. В. Миронов // *Гуманитарий Юга России*. – 2012. – № 1. – С. 101-120.
91. Михалькова, Е. В. Применение лингвистических методов при анализе поведения пользователей в социальных сетях / Е. В. Михалькова, Ю. Е. Карякин // *Математическое и информационное моделирование*. – Тюмень : Тюменский государственный университет, 2015. – С. 162-173.
92. Моргачёва, Е. Н. Сравнительно-исторический метод : обзор подходов к классификации в общественных науках / Е. Н. Моргачёва // *Вестник РГГУ. Серия: Психология. Педагогика. Образование*. – 2016. – № 2 (4).
93. Морковкин, В. В. Базовый учебный словарь русского языка для учащихся определенной национальности: научные основы и способы использования / В. В. Морковкин, Б. Брандт // *Русский язык за рубежом*. – 1985. – № 6. – С. 61-68.

94. Моченов, С. В. Векторная модель представления текстовой информации / С. В. Моченов, А. М. Бледнов, Ю. А. Луговских // Материалы международной научной конференции. – 2006. – С. 133-139.
95. Мошкина, Ю. В. Обучающие эффекты инфографики в преподавании иностранных языков в неязыковом вузе / Ю. В. Мошкина, О. Л. Крамаренко, О. Ю. Богданова // Глобальный научный потенциал. – 2021. – № 4(121). – С. 72-75.
96. Нелюбин, Л. Л. Перевод и прикладная лингвистика / Л. Л. Нелюбин. – Москва : Высшая школа, 1983. – 208 с.
97. Основы педагогики высшей школы / Л. Л. Товажнянский, О. Г. Романовский, В. В. Бондаренко, О. С. Пономарев, З. О. Черванева. – Харьков : НТУ «ХПИ», 2005.
98. Палкова, А. В. Основные понятия электронной лексикографии / А. В. Палкова // Вестник Тверского государственного университета. Серия: Филология. – 2015. – № 4. – С. 88-93.
99. Пауль, Г. Принципы истории языка : пер. с нем. / Г. Пауль ; под ред. А. А. Холодовича ; вст. ст. С. Д. Кацнельсона. Москва : Иностранная литература, 1960. - 500 с.
100. Переводный и беспереvodный способы семантизации лексики при изучении татарского языка / С. Х. Айдарова [и др.] // Вестник Казанского технологического университета. – 2013. – Т. 16, № 1. – С. 319-322.
101. Песина, С. А. Исследование особенностей перевода аудиовизуального текста / С. А. Песина, Т. Ю. Баклыкова // Современные исследования социальных проблем. – 2020. – Т. 12. – № 1. – С. 218-231. – DOI 10.12731/2077-1770-2020-1-218-231.
102. Песина, С. А. Многоуровневая структура полисеманта: инвариантно-кластерный анализ / С. А. Песина // Вопросы когнитивной лингвистики. – 2021. – № 1. – С. 60-69. – DOI 10.20916/1812-3228-2021-1-60-69.
103. Петрушова, О. Л. Учебные словари в межкультурной коммуникации : специальность 10.02.19 «Теория языка» : автореферат диссертации на

- соискание ученой степени кандидата филологических наук / О. Л. Петрушова. – 2009. – 23 с.
104. Пиоттух, К. В. Из истории развития английской лексикографии (лексикографический анализ словаря Chambers's Twentieth Century Dictionary of the English Language printed in 1901, reprinted in 1911. Edinburgh: W. & R. Chambers) / К. В. Пиоттух // Вестник Воронежского государственного университета. Серия: Лингвистика и межкультурная коммуникация. – 2017. – № 4. – С. 181-186.
105. Пиотровская, К. Р. Современная компьютерная лингводидактика / К. Р. Пиотровская // Научно-техническая информация. Серия 2: Информационные процессы и системы. – 1991. – № 4. – С. 26-29.
106. Питина, С. А. Реализация концепта УЧИТЕЛЬ в английском и русском языках / С. А. Питина // Вестник Челябинского государственного университета. – 2019. – № 10(432). – С. 118-123. – DOI 10.24411/1994-2796-2019-11017.
107. Питина, С. А. Об использовании визуализации при обучении английскому языку в вузе / С. А. Питина, А. Д. Щербов // Современная высшая школа: инновационный аспект. – 2020. – Т. 12. – № 2(48). – С. 70-77. – DOI 10.7442/2071-9620-2020-12-2-70-77.
108. Покровский, М. М. О методах семасиологии / М. М. Покровский // Избранные работы по языкознанию. Семасиологические исследования в области древних языков. – 1959. – С. 63-170.
109. Полонникова, Е. Г. Условия развития умений языковой догадки при обучении иностранцев семантизации русской фразеологии / Е. Г. Полонникова // Русский язык как иностранный и методика его преподавания. – 2020. – № 31. – С. 87-92.
110. Порождение тезауруса типа WordNet для русского языка / Н. В. Лукашевич, Г. Э. Лашевич, А. А. Герасимова, В. В. Иванов, Б. В. Добров // Пятнадцатая национальная конференция по искусственному интеллекту с международным участием : труды конференции : в 3-х т. – Смоленск : Универсум, 2016. – Т. 2. – С. 89-97.

111. Потапенко, А. А. Семантические векторные представления текста на основе вероятностного тематического моделирования : специальность 05.13.17 «Теоретические основы информатики» : диссертация на соискание ученой степени кандидата физико-математических наук / А. А. Потапенко. – Москва, 2018. – 147 с.
112. Потебня, А. А. О связи некоторых представлений в языке / А. А. Потебня. – Воронеж : Тип. В. Гольдштейна, 1864. – 32 с.
113. Помаролли, Д. Концепция сопоставительного лингвокультурологического словаря пищевой метафоры на материале русского и итальянского языков / Д. Помаролли // Вестник Томского государственного университета. – 2018. – № 434.
114. Пономарева, О. Б. Субстандартная семантическая деривация : когнитивно-коммуникативный аспект / О. Б. Пономарева, Н. Ю. Шабалина. – Тюмень : Вектор Бук, 2011. – 266 с.
115. Прошина, А. А. Моделирование двуязычного словаря-тезауруса по экономике : специальность 10.02.20 «Сравнительно-историческое, типологическое и сопоставительное языкознание» : диссертация на соискание ученой степени кандидата филологических наук / А. А. Прошина. – Екатеринбург, 2008. – 276 с.
116. Процик, И. Презентация новой лексики в лингводидактики для иностранцев / И. Процик // Теория и практика преподавания украинского языка как иностранного : сборник научных трудов. – Львов, 2007. – С. 24-30.
117. Рублёва, О. С. Слово в электронном словаре : с позиций пользователя электронными ресурсами : специальность 10.02.19 «Теория языка» : диссертация на соискание ученой степени кандидата филологических наук / О. С. Рублёва. – Киров, 2010. – 160 с.
118. Русакова, А. В. Лингвистическая модель двуязычного электронного текстоориентированного словаря юридических терминов (на материале Конституции Российской Федерации, Конституции Французской Республики и Конвенции о защите Прав человека и основных свобод) : специальность

- 10.02.21 «Прикладная и математическая лингвистика» : автореферат диссертации на соискание ученой степени кандидата филологических наук / А. В. Русакова ; Тюменский государственный университет. – Тюмень, 2008. – 25 с.
119. Рябцев, Е. В. Способы выдвижения графического уровня в поликодовом тексте: новаторство романа "Жутко громко и запредельно близко" Дж.С. Фоера / Е. В. Рябцев, Н. Н. Бочегова // Вестник Курганского государственного университета. – 2018. – № 1(48). – С. 17-20.
120. Савина, О. Ю. Модель генерации текстоориентированного лексического минимума : специальность 10.02.21 «Прикладная и математическая лингвистика» : диссертация на соискание ученой степени кандидата филологических наук / Савина Ольга Юрьевна. – Тюмень, 2017. – 197 с.
121. Сафиуллина, Г. Р. Сравнительная лексикография : практика составления русско-татарско-англо-испанского иллюстративного словаря для детей / Г. Р. Сафиуллина // Известия Волгоградского государственного педагогического университета. – 2018. – № 6. – С. 149-155.
122. Сафиуллина, Г. Р. Принципы построения татарско-английского онлайн-словаря проекта Glosbe на уровне макро-и микроструктуры / Г. Р. Сафиуллина // Филологические науки. Вопросы теории и практики. – 2021. – Т. 14, № 4. – С. 1120-1124.
123. Селегей, В. П. Электронные словари и компьютерная лексикография / В. П. Селегей. – URL: <https://docplayer.com/36764723-Elektronnye-slovani-i-kompyuternaya-leksikografiya.html> (дата обращения: 28.12.2018).
124. Серебренников, Б. А. Языковая номинация: общие вопросы / Б. А. Серебренников, А. А. Уфимцева. – Наука, 1977. – Т. 1.
125. Серова, Т. С. Функции учебного терминологического словаря как средства формирования иноязычной речевой лексической компетенции / Т. С. Серова, Г. Р. Чайникова // Язык и культура. – 2015. – № 1 (29).
126. Сивакова, Н. А. Лексикографическое описание английских и русских фитонимов в электронном глоссарии : специальность 10.02.21 «Прикладная и

- математическая лингвистика» : диссертация на соискание ученой степени кандидата филологических наук / Н. А. Сивакова. – Тюмень, 2004. – 162 с.
127. Смирницкий, А. И. К вопросу о слове (проблема отдельности слова) / А. И. Смирницкий // Вопросы теории и истории языка. – Москва, 1952.
128. Соболева, А. В. Когнитивная готовность к межкультурному общению как компонент межкультурной компетенции / А. В. Соболева, О. А. Обдалова // Язык и культура. – 2015. – № 1 (29). – С. 146-155.
129. Соколова, Е. П. Переходные явления в системе неполнозначных частей речи русского языка : лексикографический аспект / Е. П. Соколова // Проблемы истории, филологии, культуры. – 2014. – № 3 (45).
130. Сорока, О. Г. Визуализация учебной информации / О. Г. Сорока, И. Н. Васильева // Университет педагогического самообразования. – 2015. – № 12. – С. 12.
131. Сорокина, Ю. В. Понятие мультимодальности и вопросы анализа мультимодального лекционного дискурса / Ю. В. Сорокина // Филологические науки. Вопросы теории и практики. – 2017. – № 10-1 (76).
132. Стёпин, В. С. Моделирование / В. С. Стёпин, Г. Б. Гутнер, Ф. Н. Голдберг // Гуманитарная энциклопедия / Центр гуманитарных технологий – 2010. – URL: <https://gtmarket.ru/concepts/7025> (дата обращения: 05.02.2022).
133. Сысоев, П. В. Обучение иноязычному письменному юридическому дискурсу студентов направления подготовки «юриспруденция» / П. В. Сысоев, В. В. Завьялов // Язык и культура. – 2018. – № 41. – С. 308-326.
134. Сьянова, Е. И. К проблеме создания электронного сводного диалектного словаря / Е. И. Сьянова // Вестник Пермского университета. Российская и зарубежная филология. – 2020. – Т. 12, № 1.
135. Табанакова, В. Д. Идеографическое описание научной терминологии в специальных словарях : специальность 10.02.21 «Прикладная и математическая лингвистика» : диссертация на соискание ученой степени доктора филологических наук / Табанакова Вера Дмитриевна. – Тюмень, 2001. – 288 с.

136. Табанакова, В. Д. Концептуальный фон лексемы или совмещение структурного, когнитивного и функционального подхода к описанию семантики слова / В. Д. Табанакова // Вестник Тюменского государственного университета. Гуманитарные исследования. Humanitates. – 2015. – Т. 1, № 4. – С. 31-42.
137. Табанакова, В. Д. Моделирование семантики, синтактики и прагматики термина в переводном электронном словаре / В. Д. Табанакова // Современные проблемы лексикографии : материалы конференции, Санкт-Петербург, 27–30 апреля 2015 года. – Санкт-Петербург : Общество с ограниченной ответственностью «Нестор-История», 2015. – С. 172-173.
138. Теоретический контекст интегрированного предметно-языкового обучения в вузе / С. К. Гураль, Э. П. Комарова, С. А. Бакленева, А. С. Фетисов // Язык и культура. – 2020. – № 49. – С. 138-147. – DOI 10.17223/19996195/49/8.
139. Трофимова, У. М. О некоторых особенностях семантизаций тематически связанных слов в психолингвистическом эксперименте / У. М. Трофимова // Вопросы филологии. – 2005. – № 1. – С. 33-40.
140. Ульман, С. Семантические универсалии / С. Ульман // Новое в лингвистике. – 1970. – № 5. – С. 250-299.
141. Уфимцева, А. А. Опыт изучения лексики как системы (на материале английского языка) / А. А. Ульман. – Москва : Изд-во Академии наук СССР, 1962. – 288 с.
142. Филин, Ф. П. О лексико-семантических группах слов / Ф. П. Филин // Езиковедски изследвания в чест на академик Стефан Младенов. – 1957. – С. 530.
143. Филиппович, Ю. Историческая компьютерная лексикография-terra incognita в компьютерном мире / Ю. Филиппович, М. Чернышева // Компьютерра. – 1999. – № 45.
144. Халина, Н. В. Терминологическое сопровождение "визуального поворота" в современной культуре познания / Н. В. Халина, Л. М. Комиссарова, Т. Н.

- Внучкова // Филология и человек. – 2018. – № 3. – С. 185-196. – DOI 10.14258/filichel(2018)3-15.
145. Халина, Н. В. Культурный интеллект как мотивировочная база овладения языком иной культуры / Н. В. Халина, А. В. Жеребненко // Современные исследования в области преподавания иностранных языков в неязыковом вузе. – 2018. – № 7. – С. 149-160.
146. Хомутова, Т. Н. Энциклопедический мультимедийный словарь как средство развития интеллектуального туризма / Т. Н. Хомутова, М. Г. Мамонтова // Герценовские чтения. Иностранные языки : сборник научных статей, Санкт-Петербург, 12-13 апреля 2018 года. – Санкт-Петербург : Российский государственный педагогический университет им. А. И. Герцена, 2018. – С. 192-194.
147. Хомутова, Т. Н. Вербализация культурных ценностей в англоязычном научном дискурсе / Т. Н. Хомутова // Лингвокультурные ценности в языковом сознании и коммуникативной практике : материалы Международной научной конференции, Тяньцзинь, 16 ноября 2018 года / научные редакторы В. И. Карасик, Ван Минъюй. – Тяньцзинь : ООО «ПринТерра-Дизайн», 2019. – С. 271-279.
148. Хутыз, И. П. Мультимодальность академического дискурса как условие его коммуникативной успешности / И. П. Хутыз // Вестник Адыгейского государственного университета. Серия 2: Филология и искусствоведение. – 2016. – № 1 (172).
149. Чикнаверова, К. Г. Использование семантизации для пропедевтики лексических ошибок студентов при обучении юридическому английскому языку в вузе / К. Г. Чикнаверова // Язык и культура. – 2019. – № 47.
150. Чумарина, Г. Р. Теория и практика составления электронных словарей : специальность 10.02.02 «Языки народов Российской Федерации» : автореферат диссертации кандидата филологических наук / Г. Р. Чумарина. – Казань, 2010. – 28 с.

151. Шаклеин, В. М. Русская лингводидактика: история и современность / В. М. Шаклеин. – Москва : РУДН, 2008.
152. Шамов, А. Н. Методика преподавания иностранных языков / А. Н. Шамов. – 2008.
153. Шанский, Н. М. О некоторых актуальных вопросах методики русского языка как науки / Н. М. Шанский // Проблемы совершенствования содержания и методов обучения русскому языку. – 1969. – № 4-8. – С. 3-12.
154. Шаповал, В. В. Теория и практика верификации словарных данных на основе источников : специальность 10.02.01 «Русский язык» : специальность 10.02.19 «Теория языка» : диссертация на соискание ученой степени доктора филологических наук / В. В. Шаповал. – Москва, 2016. – 686 с.
155. Шафтельская, Н. В. Мотивационно-сопоставительный словарь предметно-бытовой лексики русского и английского языков / Н. В. Шафтельская // Вопросы лексикографии. – 2013. – № 1 (3).
156. Шереметьева, С. О. Платформа для концептуального аннотирования многоязычных текстов / С. О. Шереметьева, О. И. Бабина // Вестник Южно-Уральского государственного университета. Серия: Лингвистика. – 2020. – Т. 17. – № 4. – С. 53-60. – DOI 10.14529/ling200409.
157. Шляхова, А. В. Электронный словарь и его специфика / А. В. Шляхова // Вестник Тюменского государственного университета. Социально-экономические и правовые исследования. – 2008. – № 1. – С.181-185.
158. Шухардт, Г. Этимология и исследование истории слова / Г. Шухардт // Избранные статьи по языкознанию : пер. с нем. / ред., предисл. и примеч. Р. А. Будагова. – 1950. – С. 210-220.
159. Щерба, Л. В. Опыт общей теории лексикографии / Л. В. Щерба. – 1974.
160. Щипицина, Л. Ю. Информационные технологии в лингвистике : учебное пособие / Л. Ю. Щипицина. – Москва : Флинта : Наука, 2013. – 128 с.
161. Щипицина, Л. Ю. Возможности геолокационных приложений на занятиях по иностранному языку / Л. Ю. Щипицина // Открытое и дистанционное образование. – 2020. – № 1(77). – С. 18-24. – DOI 10.17223/16095944/77/3.

162. Юсупова, А. М. Роль способов семантизации лексики в методике преподавания русского языка в национальной школе / А. М. Юсупова // Синергия. – 2017. – № 1.
163. Ярославская, И. И. Лингводидактические термины в русском языке: социокультурный и сопоставительный аспекты / И. И. Ярославская, Е. В. Астремская // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. – 2014. – № 2.
164. Ятаева, Е. В. Учебный глоссарий как средство развития учебной иноязычно-лексической компетенции в профессиональном языковом образовании : специальность 13.00.08 «Теория и методика профессионального образования» : диссертация на соискание ученой степени кандидата педагогических наук / Е. В. Ятаева. – Екатеринбург, 2006.
165. Amsler, R. A. The structure of the Merriam-Webster pocket dictionary [Diss] / R. A. Amsler. – Austin : University of Texas. – 1980.
166. Anandarajan, M. Practical Text Analytics. Maximizing the Value of Text Data / M. Anandarajan, C. Hill, T. Nolan (Advances in Analytics and Data Science. Vol. 2). – 2019.
167. Anderman, G. M. Words, words, words: the translator and the language learner / G. M. Anderman, M. Rogers (ed.). – Multilingual Matters, 1996. – Vol. 7.
168. Automating the lexicon: research and practice in a multilingual environment / D. E. Walker [et al.] (ed.). – Oxford University Press, 1995.
169. Baxter, J. The dictionary and vocabulary behavior : A single word or a handful? / J. Baxter // Tesol Quarterly. – 1980. – P. 325-336.
170. Beheydt, L. The semantization of vocabulary in foreign language learning / L. Beheydt // System. – 1987. – Vol. 15, № 1. – P. 55-67.
171. Béjoint, H. The foreign student's use of monolingual English dictionaries: A study of language needs and reference skills / H. Béjoint // Applied linguistics. – 1981. – Vol. 2, № 3. – P. 207-222.

172. Bert: Pre-training of deep bidirectional transformers for language understanding / J. Devlin [et al.] // Cornell University. – 2018. – URL: <https://arxiv.org/abs/1810.04805> (дата обращения: 25.02.2022).
173. Bezemer, J. Multimodality, learning and communication : A social semiotic frame / J. Bezemer, G. Kress. – Routledge, 2015.
174. Biesaga, M. Dictionary tradition vs. pictorial corpora: which vocabulary thematic fields should be illustrated? / M. Biesaga // Lexikos. – 2017. – Vol. 27. – P. 132-151.
175. Bochegova, N. N. Semantic prototypes and national ethos / N. N. Bochegova // Studia Linguistica (Санкт-Петербург). – 2019. – No 28. – P. 18-21.
176. Bondar, O. A. English-Russian cultural contacts in XVI-XIX centuries : Пособие предназначено для студентов-бакалавров, обучающихся по направлениям подготовки «Педагогическое образование (Иностранный язык)» и «Лингвистика» / O. A. Bondar, N. A. Plastinina, E. M. Ryanskaya. – Ufa : Общество с ограниченной ответственностью "Аэтерна", 2020. – 59 p. – ISBN 978-5-00177-068-8.
177. Breal, M. Essai de semantique: science des significations / Michel Breal. – Paris : Hachette, 1897. – 349 p.
178. Campbell, L. Historical Linguistics : An Introduction / L. Campbell. – 2004.
179. Calzolari, N. A lexical workstation: from textual data to structured database // N. Calzolari, E. Picchi // Computational approaches to the lexicon / B. T. Sue Atkins and Antonio Zampolli, eds. – 1994. – P. 439-467.
180. Carter, R. Developments in the teaching of vocabulary / R. Carter, M. McCarthy // Vocabulary and language teaching. – 1988. – P. 39-59.
181. Cowie, A. P. English Dictionaries for the Foreign Learner / A. P. Cowie // Lexicography : Principles and Practice / R. R. K. Hartmann. – 1983. – P. 135-152.
182. Cowie, A. P. English dictionaries for foreign learners: A history / A. P. Cowie. – Oxford University Press on Demand, 1999.
183. Decoo, W. The Didactics of Semantization in Historical Perspective / W. Decoo // Deseret Language and Linguistic Society Symposium. – 1985. – Vol. 11, № 1. – P. 8.

184. Diachronic word embeddings and semantic shifts: a survey / A. Kutuzov [et al.] // Cornell University. – 2007. – URL: <https://arxiv.org/abs/1806.03537> (дата обращения: 25.02.2022).
185. Fellbaum, C. WordNet / C. Fellbaum // Theory and applications of ontology: computer applications. – Springer, Dordrecht, 2010. – P. 231-243.
186. Fillmore, C. J. Toward a frame-based lexicon: The semantics of RISK and its neighbors / C. J. Fillmore, B. T. Atkins // Frames, fields and contrasts: New essays in semantic and lexical organization. – 1992. – Vol. 75. – P. 102.
187. GDEX: Automatically finding good dictionary examples in a corpus / A. Kilgarriff [et al.] // Proceedings of the XIII EURALEX international congress. – Barcelona, Spain : Documenta Universitaria, 2008. – P. 425-432.
188. Gerd Neuner. Wortschatzarbeit und Bedeutungsvermittlung. Fernstudieneinheit 08. / Gerd Neuner // Die Unterrichtspraxis\teaching German. – 1994. – 119 p.
189. Grootendorst, M. Keybert: Minimal keyword extraction with bert / M. Grootendorst. – 2020. – URL: <https://github.com/MaartenGr/KeyBERT> (дата обращения: 25.02.2022).
190. Grus, J. Data science from scratch: first principles with python / J. Grus. – O'Reilly Media, 2019.
191. Grygiel, M. The methodology of analysing semantic change in historical perspective / M. Grygiel // Studia Anglica Resoviensia. – 2005. – Vol. 3. – P. 25-47.
192. Hamilton, W. L. Diachronic word embeddings reveal statistical laws of semantic change / W. L. Hamilton, J. Leskovec, D. Jurafsky // Cornell University. – 2016. – URL: <https://arxiv.org/abs/1605.09096> (дата обращения: 25.02.2022).
193. Happy or not: Generating topic-based emotional heatmaps for Culturomics using CyberGIS / Shook E. [et al.] // 2012 IEEE 8th International Conference on E-Science. – IEEE, 2012. – P. 1-6.
194. Hughes, G. A history of English words / G. Hughes. – Wiley-Blackwell, 2000. – Vol. 1.
195. Jackson, H. Words and their Meaning / H. Jackson. – Longman, 1988.

196. Jones, K. S. A statistical interpretation of term specificity and its application in retrieval / K. S. Jones // *Journal of documentation*. – 2004.
197. Kaser, O. Tag-cloud drawing: Algorithms for cloud visualization / O. Kaser, D. Lemire // Cornell University. – 2007. – URL: <https://arxiv.org/abs/cs/0703109> (дата обращения: 25.02.2022).
198. Kegl, J. Machine-readable dictionaries and education / J. Kegl // *Automating the lexicon: research and practice in a multilingual environment*. – Oxford, 1995. – P. 249-284.
199. Kipfer, B. A. Dictionaries and the intermediate student: Communicative needs and the development of user reference skills / B. A. Kipfer ; Published by Max Niemeyer Verlag 1987 // *The dictionary and the language learner* / In Cowie, A. P. (ed.). Papers from the EURALEX Seminar at the University of Leeds, 1-3 Apr. 1985 (Lexicographica Series Maior 17). Tübingen: Niemeyer. – P. 44-54.
200. Kress, G. Multimodality: Challenges to thinking about language / G. Kress // *TESOL quarterly*. – 2000. – Vol. 34, № 2. – P. 337-340.
201. Kress, G. Reading images: Multimodality, representation and new media / G. Kress // *Information Design Journal*. – 2004. – Vol. 12, № 2. – P. 110-119.
202. Kress, G. Multimodality: A social semiotic approach to contemporary communication / G. Kress. – Routledge, 2010.
203. Kutuzov, A. Distributional word embeddings in modeling diachronic semantic change / A. Kutuzov ; University of Oslo. – 2020.
204. Landau, S. I. Dictionary: The art and craft of lexicography / S. I. Landau // New York Charles Scriber s Sons. – 1984. – Vol. 211.
205. Lawson, M. J. The vocabulary-learning strategies of foreign-language students / M. J. Lawson, D. Hogben // *Language learning*. – 1996. – Vol. 46, № 1. – P. 101-135.
206. Levy, M. Computer-assisted language learning: Context and conceptualization / M. Levy. – Oxford University Press, 1997.
207. Lew, R. Multimodal lexicography : The representation of meaning in electronic dictionaries / R. Lew // *Lexikos*. – 2010. – Vol. 20. – P. 290-306.

208. Maybury, M. T. Multimedia Interaction for the New Millennium / M. T. Maybury // Sixth European Conference on Speech Communication and Technology. – 1999.
209. Measuring diachronic evolution of evaluative adjectives with word embeddings: the case for English, Norwegian, and Russian / J. Rodina [et al.]. – Association for Computational Linguistics, 2019.
210. Meringer, R. Wörter und Sachen / R. Meringer // Indogermanische Forschungen. – 1904. – Vol. 16, № 1. – P. 101-196.
211. Michel, J. B. Quantitative analysis of culture using millions of digitized books / J. B. Michel // Science. – 2011. – Vol. 331, № 6014. – P. 176-182.
212. Mining semantics for culturomics: towards a knowledge-based approach / Borin L. [et al.] // Proceedings of the 2013 international workshop on Mining unstructured big data using natural language processing. – 2013. – P. 3-10.
213. Nagao, M. A methodology for the construction of a terminology dictionary / M. Nagao // Computational approaches to the lexicon / B. T. Sue Atkins and Antonio Zampolli, eds. – 1994. – P. 398-411.
214. Nattinger, J. Some current trends in vocabulary teaching / J. Nattinger // Vocabulary and language teaching. – 1988. – Vol. 1. – P. 62-82.
215. Phani, S. Culturomics on a bengali newspaper corpus / S. Phani, S. Lahiri, A. Biswas // 2012 International Conference on Asian Language Processing. – IEEE, 2012. – P. 237-240.
216. Poole, R. Concordance-based glosses for facilitating semantization and enhancing productive knowledge of academic vocabulary / R. Poole. – The University of Alabama, 2011.
217. Prensky, M. Digital Natives, Digital Immigrants. Part 1 / M. Prensky // On the Horizon. – 2001. – Vol. 9, № 5. – P. 1-6.
218. Professor K. Reisig's Vorlesungen über lateinische Sprachwissenschaft. – Verlag der Lehnhold'schen Buchhandlung, 1839.
219. Pustejovsky, J. The generative lexicon / J. Pustejovsky. – MIT press, 1998.
220. Quantifying the evolutionary dynamics of language / E. Lieberman [et al.] // Nature. – 2007. – Vol. 449, № 7163. – P. 713.

221. Rogers, A. A primer in bertology : What we know about how bert works / A. Rogers, O. Kovaleva, A. Rumshisky // Transactions of the Association for Computational Linguistics. – 2020. – Vol. 8. – P. 842-866.
222. Sato, T. Could a Multimodal Dictionary Serve as a Learning Tool? An Examination of the Impact of Technologically Enhanced Visual Glosses on L2 Text Comprehension / T. Sato // The EuroCALL Review. – 2016. – Vol. 24, № 2. – P. 3-12.
223. Sharifian, F. Applied cultural linguistics: Implications for second language learning and intercultural communication / F. Sharifian, G. B. Palmer (ed.). – John Benjamins Publishing, 2007. – Vol. 7.
224. Sharma, P. Self-Supervised Contextual Keyword and Keyphrase Retrieval with Self-Labeling / P. Sharma, Y. Li // Computer Science, Economics, Education. – 2019.
225. Shkurina, M. V. Сравнительный анализ извлекающих методов автоматического аннотирования для текстов на русском языке / M. V. Shkurina, O. Y. Sabinin // Theoretical & Applied Science. – 2019. – № 6. – P. 164-169.
226. Sosińska-Kalata, B. Semantization and standardization—cooperative or conflicting trends in knowledge organization? / B. Sosińska-Kalata // Knowledge organization in the 21st century, 2014. – P. 580-587.
227. Stowell-Ruzicka, L. Acquiring vocabulary in a foreign language: some effects of the nature of instruction on the knowledge and use of words / L. Stowell-Ruzicka. – Louisiana State University and Agricultural & Mechanical College, 1991.
228. Suchanek, F. M. Semantic culturomics / F. M. Suchanek, N. Preda // Proceedings of the VLDB Endowment. – 2014. – Vol. 7, № 12. – P. 1215-1218.
229. Summers, D. The role of dictionaries in language learning / D. Summers // Vocabulary and language teaching. – 1988. – P. 111-125.
230. Tf-idf // Encyclopedia of machine learning / C. Sammut, G. I. Webb. – Springer, 2011. – P. 986-987.
231. The Oxford handbook of computational linguistics / R. Mitkov (ed.). – Oxford University Press, 2004.

232. Tracing cultural diachronic semantic shifts in Russian using word embeddings: test sets and baselines / V. Fomin [et al.] // Cornell University. – 2019. – URL: <https://arxiv.org/abs/1905.06837> (дата обращения: 25.02.2022).
233. Visions and open challenges for a knowledge-based culturomics / Tahmasebi N. [et al.] // International Journal on Digital Libraries. – 2015. – Vol. 15, № 2-4. – P. 169-187.
234. Wilks, Y. Electric words: dictionaries, computers, and meanings / Y. Wilks, B. M. Slator, L. M. Guthrie. – MIT Press, 1996.
235. Xue, M. Representing the cultural dimension of meaning in learner's dictionaries- from the perspective of chinese EFL learners in l2 reception / M. Xue // Lexikos. – 2017. – Vol. 27, № 1. – P. 578-596.

СЛОВАРИ И ИХ СОКРАЩЕНИЯ

236. Академик. Словари и энциклопедии. – URL: <https://academic.ru/> (дата обращения: 06.10.2021).
237. БТСРЯ – Кузнецов С. А. Большой толковый словарь русского языка / С. А. Кузнецов. – Норинт, 2004.
238. КАРТАСЛОВ.РУ – Карта слов и выражений русского языка – URL: <https://kartaslov.ru> (дата обращения: 23.05.2021).
239. Лингвистическая онтология «Тезаурус РуТез». – URL: <http://www.labinform.ru/pub/ruthes/> (дата обращения: 13.06.2021).
240. ЛЭС – Лингвистический энциклопедический словарь / ред. В. Н. Ярцева. – Москва : Большая Российская энциклопедия, 2002.
241. МОС – Тихонов, А. Н. Морфемно-орфографический словарь / А. Н. Тихонов. – Москва : Школа-Пресс, 1996. – 701, [1] с.
242. НСМТП – Азимов, Э. Г. Новый словарь методических терминов и понятий (теория и практика обучения языкам) / Э. Г. Азимов, А. Н. Щукин. – Москва : Издательство ИКАР, 2009.

243. ГРАМОТА.РУ : [справочно-информационный портал]. – URL: <http://www.gramota.ru> (дата обращения: 06.10.2021).
244. ЭРЦ – Платонов, О. А. Святая Русь : энциклопедический словарь русской цивилизации / О. А. Платонов. – Москва : Православное издательство «Энциклопедия русской цивилизации», 2000.
245. ЭСРЯ – Этимологический словарь русского языка. – Москва : Русский язык от А до Я : Издательство «ЮНВЕС», 2003.
246. ЭСРЯ(К) – Крылов, Г. А. Этимологический словарь русского языка / Г. А. Крылов. – Санкт-Петербург : ООО «Виктория плюс», 2004.
247. ЭССРЯ – Фасмер, М. Этимологический словарь современного русского языка : в 4-х т. / М. Фасмер. – Москва : Астрель : АСТ, 2009. – Т. 1: А-Д. – 588 с.
248. ЭССРЯ(Ш) – Этимологический словарь современного русского языка : в 2-х т. / сост. А. К. Шапошников. – Москва : Флинта ; Наука, 2010.
249. AD – Dr. Goodword's AlphaDictionary. – URL: <http://www.alphaDictionary.com/> (дата обращения: 28.09.2021).
250. ADWO – Ayto, J. Arcade dictionary of word origins / J. Ayto // The histories of more than. – 1990. – Vol. 8.
251. CALD – Cambridge Advanced Learner's Dictionary / Cambridge University Press, 2020. – URL: <https://dictionary.cambridge.org/ru/> (дата обращения: 28.09.2021).
252. COBUILD – Cobuild, C. Dictionary English Language / C. Cobuild // London: William Collins Sons And Col, Ltd. – 1987.
253. EDO – Online Etymology Dictionary. – URL: <http://www.etymonline.com/> (дата обращения: 28.09.2021).
254. FineDictionary. – URL: <http://www.finedictionary.com/> (дата обращения: 06.06.2021).
255. FrameNet. – URL: https://framenet.icsi.berkeley.edu/fndrupal/framenets_in_other_languages/ (дата обращения: 28.06.2021).

256. LDOCE – Longman, P. Longman Dictionary of Contemporary English / P. Longman. – Harlow, 1978.
257. MWNBWH – Merriam-Webster, Inc. The Merriam-Webster new book of word histories. – Merriam-Webster, 1991.
258. MW – Merriam-Webster: Dictionary and Thesaurus. – URL: <http://www.merriam-webster.com> (дата обращения: 28.09.2021).
259. OALDCE – Hornby, A. S. Oxford Advanced Learners Dictionary of Current English: Revised and Updated / A. S. Hornby. – 1974.
260. ODE – Wallace, S. A dictionary of education / S. Wallace. – (2 ed.). – OUP Oxford, 2015.
261. ODOWH – The Oxford dictionary of word histories / G. Chantrell (ed.). – Oxford University Press, 2002.
262. ODOWO – Cresswell, J. Oxford dictionary of word origins / J. Cresswell. – (2 ed.). – Oxford University Press, 2010.
263. OED – Simpson, J. A. The Oxford English Dictionary / J. A. Simpson, E. S. C. Weiner. – Oxford : Clarendon Press ; Oxford ; New York : Oxford University Press, 1989.
264. RuWordNet. – URL: <https://ruwordnet.ru/> (дата обращения: 31.05.2021).
265. The Visual Thesaurus. – URL: <https://www.visualthesaurus.com/> (дата обращения: 06.05.2021).
266. WordNet. – URL: <http://wordnetweb.princeton.edu/> (дата обращения: 23.05.2021).

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

267. Облако слов. – URL: <https://облакослов.рф/> (дата обращения: 26.01.2021).
268. Education Today Magazine. - URL: <https://www.education-today.co.uk/> (дата обращения: 13.08.2021).
269. EDU-inform : образовательный портал. – URL: <http://edu-inform.ru/> (дата обращения: 14.08.2021).

270. KeyBERT. – URL: <https://github.com/MaartenGr/KeyBERT> (дата обращения: 15.05.2021).
271. Shinydashboard. – URL: <https://rstudio.github.io/shinydashboard/> (дата обращения: 17.11.2021).
272. Studiorum. Образовательный портал Национального корпуса русского языка. – URL: <https://studiorum-ruscorpora.ru/forum/> (дата обращения: 26.01.2021).
273. Wikimedia Commons (Викисклад). – URL: <https://commons.wikimedia.org/> (дата обращения: 15.10.2021).