

МИНОБРНАУКИ РОССИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра английского языка  
и методики его преподавания

**КОМПЬЮТЕРНАЯ ЛЕКСИКОГРАФИЯ: ТЕРМИНОЛОГИЧЕСКИЕ БАЗЫ  
ДАНЫХ И ЭЛЕКТРОННЫЕ СЛОВАРИ**

АВТОРЕФЕРАТ ВЫПУСКНОЙ КВАЛИФИКАЦИОННОЙ РАБОТЫ БАКАЛАВРА

Студентки 4 курса 416 группы  
направления 44.03.01 Педагогическое образование  
профиль – «Иностранный язык»  
факультета иностранных языков и лингводидактики

Байджаева Ашыргелди  
Алланазарова Арслана

Научный руководитель  
профессор кафедры английского  
языка и методики его преподавания,  
д.ф.н., доцент

\_\_\_\_\_ Е.Ю. Балашова

дата, подпись

Зав. кафедрой  
английского языка  
и методики его преподавания  
канд. пед. наук, доцент

\_\_\_\_\_ Г.А. Никитина

дата, подпись

Саратов 2022 год

**Введение.** *Актуальность* данной работы определяется тем, что эта проблема имеет не только теоретическое, но и практическое значение, поскольку ее решение позволяет прогнозировать направления лексикографической деятельности.

*Объектом* данного исследования является лексикография, а предметом исследования выступают словари иностранных языков.

*Цель* настоящего исследования заключается в комплексном изучении процесса создания и развития словарей.

*Задачами* данной выпускной квалификационной работы являются:

1. ознакомиться с историей развития лексикографии английского языка;
2. рассмотреть место и роль словаря в системе современной английской лексикографии;
3. ознакомиться с многообразием словарей различных типов.

*Структура работы* определена задачами исследования. Работа состоит из введения, двух глав, разделенных на параграфы, заключения и списка использованной литературы.

*Во введении* рассматривается актуальность изучения данного вопроса, а также цель, задачи.

*Первая глава* посвящена описанию теоретической стороны вопроса, то есть дается информация, начиная от истории компьютерной лексикографии и до отличия от бумажных словарей.

*Вторая глава* посвящена описанию практической стороны вопроса.

*В заключении* даются основные выводы по работе.

Работа завершается списком *использованных источников*.

**Основное содержание.** В первой главе «Основные понятия компьютерной лексикографии» отмечено, что существует определенное противоречие между новыми языковыми компьютерными технологиями и старым традиционным словарем, заключающееся в содержании, позволяющем полностью использовать эти технологии. Другими словами, в электронном словаре представлена речевая модель, построенная на других принципах, нежели традиционные формальные модели лексики, а это краеугольный камень новых технологий.

Проблема составления такого словарного содержания, которое предполагает анализ отдельной лексической смысловой единицы, а не морфологической лексемы, представляется наиболее перспективным направлением в компьютерной лексикографии.

Понятно, что это решение требует специальной синхронизации словарных описаний и используемых формальных моделей по технологиям анализа. В итоге это должно быть единое целостное лексико-синтаксически-семантическое описание лексики в электронном словаре.

Широко распространенное программное обеспечение, в том числе и электронные словари, характеризуется частой сменой версий и наличием постоянной обратной связи от тысяч пользователей. Поэтому компьютерная лексикография неизбежно является наиболее актуальной лексикографией.

В свете постоянно расширяющихся межъязыковых контактов и межкультурной коммуникации на рубеже второго и третьего тысячелетий проблемы лексикографии приобретают большое значение. Все расширяющиеся технические возможности способствуют переходу лексикографической теории и практики на новый продвинутый уровень.

Автоматизация работы над словарем и широкий доступ к глобальным сетям приводят к новым задачам составления и использования словарей, отвечающих современным требованиям

Во второй главе «Терминологические базы данных (ТБД) и их применение в различных областях знаний» отмечено, что База данных ISO предоставляет согласованную платформу для поиска, разработки и поддержки

концептуально-терминологического содержания для всех стандартов ИСО и позволяет осуществлять поиск по трём основным категориям: графические символы; термины и определения; коды (стран, валют, языков и т.д.).

IATE (Inter-Active Terminology for Europe – интерактивная терминология для Европы) – многоязычная база данных терминов, используемых в документах Европейского союза.

EuroVoc (2012 г.) – многоязычный тезаурус, состоящий из более чем 6700 иерархически организованных предметных областей, используемых европейскими институтами и многими органами в государствах-членах Европейского Союза (ЕС) для классификации и поиска официальных документов.

Многоязычная терминологическая база данных ООН (UNTERM) – это лингвистический инструмент, который переводит терминологию и номенклатуру, используемые в Организации Объединенных Наций (ООН), на шесть официальных языков ООН (арабский, китайский, английский, французский, русский и испанский).

Теория и практика проектирования и построения терминологических баз данных чаще всего основывается на анализе больших объёмов информации, в связи с чем можно говорить об интеграции терминоведения с компьютерной лексикографией и корпусной лингвистикой. Научные труды в этих областях также учитывались нами в ходе проведения исследования: В. А. Плунгян (2008), А. Б. Кутузов (2006), Г. Г. Бабалова (2009), А. А. Поликарпов (2011), Е. А. Сидорова (2013), В. П. Захаров (2014).

В данной области выделяется ряд нерешенных проблем:

Среди учёных нет четкого понимания того, что включает в себя когнитивная природа термина, какие характеристики термина можно причислить к когнитивным и на каком основании.

Большинство современных терминологических баз отражают статичный характер термина и, по сути, являют собой электронные переводные словари, основная функция которых – справочная. В то же время терминологическая база нового поколения должна отражать динамическую

природу терминологических систем, быть многогранной, служить средством узнавания.

Хотя некоторые ученые указывают на необходимость разработки новых методов подготовки переводчиков для значительного повышения качества перевода, исследований по дидактическому техническому переводу не проводилось.

Большинство современных терминологических баз отражают статичность термина, и по сути основная функция – это электронные словари-переводы со справочной информацией. Но они также необходимы и полезны.

**Заключение.** Исходя из вышеизложенного, можно сделать вывод, что компьютерная лексикография – это прикладная научная дисциплина в лингвистике, изучающая методы использования компьютерных технологий для составления словарей и их применения, а также достаточно эффективно внедряет компьютерную лексикографию в образовательных и других областях. Также использование компьютерной лексикографии способствует повышению качества работы, лучшего понимания текстов или терминов.

Существует два основных подхода к машинному переводу в переводческой практике и в информационных технологиях. С одной стороны, результаты машинного перевода можно использовать для быстрого ознакомления с содержанием документа на незнакомом языке. В этом случае он не требует тщательного редактирования. Другой подход предполагает использование машинного перевода вместо обычного "человеческого". Это включает в себя тщательное редактирование и адаптацию системы перевода к конкретной предметной области. Здесь играют роль полнота словаря, эффективность способов устранения лексической двусмысленности, поиск переводческих соответствий и алгоритмы синтеза.

Переводчик потратит 20% времени на перевод, 40% на поиск незнакомых слов и 40% на регистрацию.

Компьютер потратит 95% непосредственно на перевод и 5% на обновление словаря. Количественные характеристики указывают на

преимущество машинного перевода. Однако качество такого перевода значительно уступает тому, который сделан человеком.

Первая демонстрация машинного перевода состоялась в 1954 году в США, затем в СССР. Система была примитивной (словарь из 250 слов, грамматика из 6 правил, перевод только простых фраз). Однако уже в 70-80-е годы началось широкое использование систем машинного перевода.

Говоря о проблемах машинного перевода, следует отметить, что независимо от того, какие пары языков используются в таких языках, полного семантического равенства не существует: как правило, несколько слов языка, что переводится отвечают. То есть лингвистически точный перевод невозможен. Можно сказать, что процесс перевода текстов носит творческий характер. Стоит отметить, что перевод-это задача, которая до сих пор не имеет однозначного алгоритма решения, что создает трудности при построении алгоритмов машинного перевода.

В ходе этой работы мы изучили историю развития компьютерной лексикографии и продемонстрировали примеры работы компьютерной лексикографии. Исходя из этого, мы можем сказать, что цели и задачи этой работы были достигнуты.