

ОТЗЫВ

члена научного жюри доцента д -ра Людмилы Ивановой Кировой

Кафедра болгарского языка и специализированного обучения иностранцев,
Департамент языкового обучения, Софийский университет им. Св. Климента Оридского

на диссертационную работу на соискание образовательной и ученой степени доктора

в области высшего образования - 2. Гуманитарные науки; Профессиональное
направление: 2.1. Филология; **Специальность:** Общее и сравнительное языкознание
(математическая лингвистика)

Автор диссертации: Виктория Красимилова Петрова-Любенова, Сектор компьютерной лингвистики, Институт болгарского языка им. проф. Любомира Андрейчина, Болгарская академия наук

Тема диссертации: ПОЛУАВТОМАТИЧЕСКАЯ РАЗРАБОТКА МНОГОЯЗЫЧНЫХ
ТЕРМИНОЛОГИЧЕСКИХ РЕСУРСОВ

Научный руководитель: проф. д -р Светла Коева, Сектора компьютерной лингвистики, Институт болгарского языка им. проф. Любомира Андрейчина,, Болгарская академия наук

Тема диссертации актуальна, важна и имеет практическую направленность. Она касается быстро развивающейся области с широкой перспективой развития. Такая тема требует знаний и умений во многих областях: в области языка, перевода, конкретных технологических инструментов, которые помогают переводчикам в процессе их работы (программы, терминологические базы данных и другие терминологические ресурсы).

Диссертантка знает хорошо состояние проблемы и научную литературу по этому вопросу, что видно из подробного изложения в первой и второй главах и приложенной библиографии. Разделы этих глав анализируют возможности программ перевода с помощью компьютера для создания терминологических баз, программ для извлечения терминов, а также доступные в настоящее время терминологические ресурсы для болгарского языка. Анализируя ситуацию в настоящее время, в условиях нашей страны, диссертантка выполняет первую из своих поставленных целей и обосновывает свою вторую цель - предложить способ заполнить отсутствие необходимых терминологических ресурсов.

Диссертантка предлагает методику создания терминологических ресурсов для специализированной области, в частности, представляет терминологический ресурс в области компьютерных технологий. В **третьей главе** она описывает технологии, подходы

и виды методов для автоматического извлечения терминов и терминологических ресурсов в области компьютерной терминологии. Сравнивая инструменты для автоматического извлечения терминов, существующих до настоящего времени анализа, В. Петрова приходит к выводу, что свободно доступных инструментов не так много, а те, которые извлекают термины на болгарском языке, еще ограничены. Чтобы преодолеть эти ограничения, диссертантка описывает и аргументирует предложенную ей самой методологию полуавтоматического создания терминологической базы в области компьютерных технологий на основе двуязычного корпуса (английско-болгарского). В. Петрова перечисляет требования, которым должен соответствовать конкретный инструмент для извлечения терминов из корпуса, а также учитывает положительные стороны и недостатки двуязычного корпуса, который возможно использовать. Исходя из этого, она предпочитает работать с комбинацией программ, с помощью которых можно извлекать термины на обоих языках. С помощью применяемой методологии диссертантка расширяет возможности обработки корпуса. К функциональным возможностям разработки программы, направленной на болгарский язык, она добавляет возможность одновременного извлечения двуязычной терминологии. Английские термины были идентифицированы, потом в болгарской части используемого корпуса их переводы были найдены, и результаты визуализированы в программе MateCat. К английскому термину, который служит ключевым словом, болгарские эквиваленты расположены рядом в контексте предложения, в котором они употреблены. Затем диссертантка выбирает Excel как наиболее доступный формат, который позволяет прямой импорт в любой программе перевода с помощью компьютера (или т.наз. автоматизированного перевода) и располагает результаты в таблице. Наконец, она применяет проверку вручную и техническое оформление терминологической базы.

В. Петрова разработала такую комбинацию методов, которая позволяет форматам терминологической базы варьироваться в зависимости от технических возможностей программы перевода при помощи компьютера или от конкретных требований клиента. Диссертантка также указывает на ряд других, самостоятельных средств, посредством которых извлеченная терминологическая база может быть преобразована из одного формата в другой (где это необходимо) для импорта в заданный проект перевода.

В. Петрова также добавила два эксперимента, в которых с помощью автоматизированного перевода измеряет точность перевода в своем корпусе по сравнению с другим двуязычным корпусом (из регламентированных ЕС), в котором используется готовая база и является нормой для сравнения. Проверка доказывает, что термины визуализируются правильно, и результат в отношении искусственно введенных ошибок одинаков. Таким образом, диссертантка дополняет методологию создания терминологической базы в области компьютерных технологий с проверкой для измерения точности перевода терминологической базы.

В. Петрова также сравнивает эффективность наиболее распространенных специализированных программ оценки качества перевода, которые имеют гораздо большую функциональность, чем инструменты для этой цели, встроенные в компьютерные программы автоматизированного перевода. Она анализирует ошибки и классифицирует их, чтобы определить их влияние на качество перевода. Все эти процедуры доказывают, как процесс перевода может быть завершен использованием программ автоматической проверки качества перевода. Чтобы аргументировать наблюдения, в какой степени можно полагаться на них и где необходимы персонализированные пользовательские настройки и оценка человека, диссертантка комментирует международные стандарты, которые определяют качество перевода, заявляя, какие требования они предъявляют к процессам, ресурсам и всем другим аспектам для предоставления качественной службы перевода.

Вклад диссертантки

Таким образом, предложенная и реализованная методология является ядром труда и ее основным достоинством. Методология также может быть применена для извлечения терминов из других терминологических областей, как один из возможных путей, чтобы компенсировать отсутствие достаточных ресурсов и программ для обработки текста на болгарском языке, содержащего терминологию.

Анализ технологий и средств, используемых для создания терминологических ресурсов в контексте компьютерных автоматизированных систем перевода, является многомерным, точным и высоко значимым. Проанализированы системы компьютерного автоматизированного перевода, их технические характеристики, их компоненты, а также изменения, которые эти инструменты накладывают на рабочий процесс, образование и подготовку переводчиков.

Публикации, представленные диссертанткой, соответствуют требованиям по количеству, связаны с темой диссертации и опубликованы в болгарских изданиях – сборниках с докладами конференций. Две из трех публикаций на английском языке.

Мои критические замечания незначительны. При сравнительной проверке типа ошибок с помощью программ Xbench, QA Distiller и Trados есть примеры из корпуса, который служит нормой, но нет примеров из корпуса с компьютерной терминологией. В заключении было бы хорошей идеей привести каждый из многочисленных конкретных вкладов труда, если не с названиями, по крайней мере, с нумерацией. В автореферате есть некоторые повторения. Тем не менее, автореферат правильно, четко и кратко отражает основные положения и научный вклад диссертации.

Заключение

Диссертационный труд является продуктом очень хорошо проведенного исследования, результаты которого имеют научно-прикладное и прикладное значение.