



ОТЗЫВ

на диссертационный труд на соискание образовательной и ученой степени доктора (= кандидата наук, PhD), в области:

2. Гуманитарные науки, профессиональное направление 2.1. Филология, докторская программа „Общее и сравнительное языкознание“.

Автор кандидатского диссертационного труда: **Тодор Димитров Лазаров**, докторант (= аспирант) в Секции по компьютерной лингвистике, в Институте болгарского языка им. Проф. Л. Андрейчина - БАН.

Тема диссертационного труда: „**Болгарско-английские грамматические параллели в перспективе машинного перевода**“ (на болгарском языке).

Рецензент (на правах отзыва): проф. д.ф.н. Максим Иванов Стаменов, Институт болгарского языка им. проф. Любомира Андрейчина, член научного жюри, определенного приказом Директора Института болгарского языка, БАН, проф. Светлы Коевой РД-09-10 от 07.02.2020 г.

1. Актуальность разрабатываемого в диссертационном труде проблема в научном и научно-прикладном отношении

Проблема, которой посвящена настоящая диссертация, весьма актуальна в контексте новейших направлений развития компьютерной лингвистики. Работа ориентирована на исследование болгарско-английских грамматических параллелей в области машинного перевода и, более конкретно, объект исследования ограничен анализом и моделированием темпоральных систем двух языков и выражением темпоральных отношений. Труд посвящен исследованию болгарско-английских грамматических параллелей с ориентацией к машинному переводу и, более конкретно, объект исследования ограничен до анализа и моделирования темпоральных систем на двух языках и выражения на темпоральных отношений в болгарском и английском языках и способами передачи грамматического значения на темпоральных форм из болгарского в английский язык. С одной стороны, исследование этой проблематики с помощью подходов и методологии на машинного перевода мотивировано возможностями, которые научное поле предлагает для систематического и формального анализа комплексных языковых явлений, какими являются темпоральные системы и глагольное формообразование. С другой стороны, слабая разработанность проблемы и отсутствие достаточных языковых корпусов, которые имеют аннотации и информацию о способах соотношения глагольных форм, предполагает необходимость теоретико-прикладного аспекта исследования.

Специально необходимо подчеркнуть сочетание в диссертации теоретического и теоретико-прикладного аспектов. Основная ее цель – реализовать сопоставление на теоретическом уровне между семантикой темпоральных значений в двух языках, что создаст возможности теоретического обоснования возможных переводческих эквивалентов и статистического анализа данных, извлеченных из параллельного корпуса с соотношенными глагольными формами на болгарском и на английском языках. Результаты исследования ориентированы на их приложение при усовершенствовании работы над системами для компьютерно ассистированного перевода, а также и при селекции одноязычных и параллельных данных для обучения систем для машинного перевода.

Цели работы сформулированы имея ввиду описанную ориентацию: представить и проанализировать болгарско-английские грамматические параллели при переводе темпоральных форм с болгарского на английский и представить статистические данные о грамматикализованной информации в темпоральных формах при переводе с болгарского на английский язык.

2. Ориентация диссертанта в состоянии проблема и научной литературы

Автор показывает хорошую ориентацию в литературе по теме диссертации. В библиографическом списке к диссертации включены в общей сложности 149 публикаций болгарских и иностранных ученых, с которыми диссертант работал при осуществлении деятельности, связанных с реализацией исследования и написанием текста диссертационного труда.

3. Характер выбранной методик

Метод исследования основывается на возможностях формально-семантического описания свойств языка и к приложению этого типа описания для прикладных целей. Обсуждение гиперкатегории время в болгарском и английском языках предоставляет возможности для параллельного формально-семантического анализа значений темпоральных категорий. Проблема о грамматических параллелях между темпоральными системами на двух языках рассматривается с помощью методов семантического анализа, корпусного анализа и статистического моделирования. Грамматические параллели между болгарским и английским с точки зрения формально-семантического описания сходств и различий их темпоральных систем проходит эмпирическую верификацию на основании статистических данных, извлеченных из „Болгарско-английского параллельного корпуса с соотнесенными глагольными формами“. Для достижения этой цели корпусный анализ сочетается с статистическим моделированием.

4. Краткая аналитическая характеристика научных и научно-прикладных вкладов диссертационного труда

Подтверждаем наличие следующих вкладов в обсуждаемой работе. На первом месте это относится к представлению единого в теоретическом отношении описания темпоральных систем в болгарском и английском языке – как в по отдельности, так и в сопоставительном плане. Механизм конструирования грамматического темпорального значения для болгарских и английских глагольных форм исследована в сопоставительном плане и формулированы основные закономерности взаимодействия между темпоральными морфологическими категориями в обоих языках и концептуальная структура гиперкатегории время.

В диссертации конструирован и аннотирован Болгарско-английский параллельный корпус с соотнесенными глагольными формами. Корпус представляет собой языковой ресурс, который предоставляет возможности исследования грамматических параллелей между болгарским и английским языками. Новое в нем состоит в том, что он выполняет функцию средства формального описания грамматических параллелей между темпоральными системами болгарского и английского языков и источником извлечения статистических моделей отношений переводной реализации глагольных форм в обоих языках. При этом способе конструкции и употребления языковые данные могут послужить для обогащения

переводных моделей для нужд машинного перевода, а также для теоретических изысканий. Представленная процедура структурирования и аннотирования Корпуса может в дальнейшем быть использована при разработке подобных лингвистических ресурсов для исследования сходных языковых явлений вне обхвата актуального исследования.

Существенны обобщения в облике количественной и дистрибутивной информации о глагольных системах болгарского и английского языков, которая подходит для различных типов научного анализ и вне обхвата настоящего исследования. Таким образом обеспечены достоверные лингвистические данные для создания альтернативных языковых приложений для компьютерной обработки текстов на естественном языке как статистическая модель для аннотации и статистическая модель для перевода глагольных форм.

Дополнительным вкладом является интегрирование лингвистических данных из Болгарско-английского параллельного корпуса со соотнесенными глагольными формами в софтвер для компьютерно-ассистированного перевода с целью его практического приложения как дополнительное средство при переводе. Таким способом статистические данные могут быть как использованы при процессе перевода, так и быть скорректированы и обогащены дополнительным языковым материалом.

Нельзя однако принять как вклад результаты положенного кандидатом труда для представления процесса селекции ресурсов, для предварительной обработки, аннотации и использования софтверных приложений для работы с корпусом. Если речь идет об оформлении самой процедуры по структурированию и аннотации Корпуса и возможностях ее использовать для других целей, то она уже была упомянута при обсуждении вклада самого Болгарско-английского корпуса с соотнесенными глагольными формами.

5. Оценка публикаций, связанных с диссертационным трудом: количество, характер изданий, в которых они были отпечатаны, цитирования

Диссертант опубликовал по теме диссертации в общей сложности семь публикаций. Все они под его авторством. Три из его публикаций отпечатаны в реферированных изданиях – две в сп. „Български език“ и одна в сп. „Съпоставително езикознание“. Он имеет и одну публикацию в эл. журнале „Littera et Lingua“. Две от его публикаций на английском языке и опубликованы в материалах от международных конференций по компьютерной лингвистике, проведенных в Болгарии (CLIB 2016 CLIB 2018). Одна публикация отпечатана в материалах юбилейной конференции Института болгарского языка, проведенной в 2017 г. Отсутствуют сведения о наличии цитаций публикаций диссертанта другими авторами. Описанные публикации значительно превышают минимальные требования в этом отношении для приобретения научной степени „доктор по филологии“.

6. Критические замечания

При презентации особенностей болгарской и английской систем глагола обсуждается и история вопроса, что не всегда выглядит целесообразным, имея ввиду выбранную тему. Так например при обсуждении английской системы глагола делается обзор взглядов лингвистов от XVI века до нашего времени (с. 44, 46-48).

Как бы не была похвальна теоретическая ориентация в рамках исследования в русле компьютерной лингвистики, есть и случаи, в которых она проявляет обратную сторону в облике спекулятивных утверждений: „Совместное присутствие всех этих

значений в тексте требует формального анализа инвариантного значения, что обосновывает рассмотрение позиции формы в системе языка на экстралингвистическом уровне (Чергова 2014: 187)“ (с. 52).

Проблематично утверждение, что „когда действие в подчиненном предложении является будущим по отношению к действию в главном, употребляется будущем времени в прошлом (Ранкова и Иванова 1980: 226). Оно действует при структурировании сложных предложений и последовательных высказываний и представляет собой скорее синтаксическую особенность построения текста.“ (с. 61). Sequence of tenses однако приложимо на уровне предложения, а не текста. Как нетривиальный в данном случае можно рассматривать вопрос насколько в разговорной речи одно предложение можно оформить как два высказывания, которые дополняют друг друга. И в этом случае, по нашему мнению, определяющим условием для правильного оформления с точки зрения синтаксиса предложения.

7. **Правильно ли отражает автореферат содержание диссертационного труда**

Автореферат правильно отражает основные положения научных вкладов диссертационного труда. Он отвечает его содержанию.

8. **Заключение**

Имея ввиду вышесказанное, делаю предложение Научному жюри, чтобы оно **присудило** диссертанту **Тодору Димитрову Лазарову** научную и образовательную степень „доктора“ по профессиональному направлению 2.1 Филология, научная специальность Общее и сравнительное языкознание (компьютерная лингвистика) за его диссертацию „**Болгарско-английские грамматические параллели в перспективе машинного перевода**“.

Дата 25.04.2020 г.
(подпись)

Автор отзыва:

