



РЕЦЕНЗИЯ

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 (= официальный опонент): проф. д-р Красимира Славчева Алексова,
Кафедра болгарского языка, Софийский университет им. Св. Климента Охридского,
член научного жюри, определенного приказом Директора Института болгарского
языка, БАН, проф. Светлы Коевой РД-09-10 от 07.02.2020 г.

1. Актуальность разрабатываемой в диссертационном труде проблемы в научном и
научно-прикладном отношении.

Просмотр документов по процедуре показывает, что соблюдены Минимальные
национальные требования к приобретению ученой и образовательной степени „доктора“, а
также и требования Правильника о докторантах в БАН.

Тема диссертационного труда бесспорно диссертабельна и актуальна, что
подкрепляется анализом достигнутых результатов в исполнении поставленных целей и
задач. Болгарско-английские параллели времен и как система, и как употребление сами по
себе представляют собой самостоятельную диссертабельную тему, а в рецензируемом
труде эти теоретические наблюдения последовательно и целенаправленно подчиняются
актуальной для компьютерной лингвистики задаче – улучшению машинного перевода.
Результаты разработок диссертанта имеют неоспоримое практическое применение, так как
могут непосредственно послужить улучшению машинного перевод с болгарского на
английский язык.

2. Знает ли докторант состояние проблемы и научную литературу по вопросу?

Труд показывает очень хорошее знание теоретических взглядов (как болгарских ученых, так и утвержденных в лингвистике иностранных исследователей), касающихся грамматических глагольных категорий, которые являются значимыми для разработки. При этом углубленные познания Т. Лазарова относятся к двум областям, взаимопроникающих в диссертационном труде: это теоретическая грамматика болгарского и на английском языков с акцентом на грамматические категории глагола, с одной стороны, и с другой стороны – машинного перевода, который является частью широкой области компьютерной лингвистики. В работе производит впечатление факт, что концепции отдельных ученых представлены не самоцельно, а через сравнительный анализ, при этом в работе очерчиваются значимые тенденции в развитии машинного перевода.

Библиография охватывает более 150 цитированных единиц, которже являются достаточным основанием представить ясную картину научных исследований, связанных с глагольной системой в двух сопоставляемых языках и с машинным переводом.

3. Выбранная методика исследования может ли дать ответ на поставленные в диссертационном труде цели и задачи?

В исполнении целей и задач диссертационного труда Т. Лазаров применил вполне адекватную методику. На базе воспринятых теоретических позиций о семантике и составе темпоральных систем болгарского и английского языков проводится формально-семантическое описание темпоральных значений обоих языков. Это становится базой для создания гипотез о переводческих соответствиях. Они проверяются с помощью эмпирических данни из созданного на основе сформулированных критериев Болгарско-английского корпуса, содержащего соотнесенные глагольные формы. Это означает, что корпусный подход используется для верификации исследовательских гипотез докторанта. Особенно значим и статистический анализ темпоральных форм в обоих языках и сопоставительный анализ по отношению к трансферу информации о морфологических категориях, дефинированных докторантом как время, вид действия и таксис в болгарском языке по отношению к английским морфологическим категориям времени и вида действия, а также и по отношению к трансферу информации о категориях лица, числа и рода болгарского глагола при переводе на английский язык.

4. Краткая аналитическая характеристика естества и достоверности материала, на основе которого строится вклад диссертационного труда.

Так как применяется корпусный анализ, можно утверждать, что достоверность материала гарантирована. Конечно, анализ и выводы базируются на созданном Болгарско-английском корпусе соотнесенных глагольных форм, который имеет определенный (крайний) объем, вне которого могут остаться некоторые случаи редких употреблений, напр.. будущего предварительного и будущего предварительного времени в прошедшем в

болгарском языке. Именно этого попытался избежать диссертант, включив в Корпус примеры Грамматики Р. Ницоловой, изданной на английском языке.

5. В чем заключается научный и научно-прикладной вклад диссертационного труда, и какова конкретнее значимость вклада для науки и практики?

Диссертационный труд отличается логической последовательностью изложения, с ясной структурой и наличием связей между отдельными частями. Обособленные две макрочасти, посвященные лингвистической теории машинного перевода, так целенаправленно связаны, что гарантируется максимальная реализация основной цели исследования – обнаружить конкретные грамматические параллели между темпоральными отношениями на основе параллельного болгарско-английского корпуса и создать статистические модели перевода глагольных форм, чтобы улучшить машинный перевод. Пятая глава полностью убеждает читателей, что результаты диссертационного труда имеют прямое практическое применение – они позволяют усовершенствовать машинный перевод. С полной уверенностью могу сказать, что теория и практика в исследовании максимально взаимосвязаны.

По-моему, успех реализации поставленных целей зависит в большой степени от выбора теоретической позиции по отношению к организации категории *времени*. Это осознает и сам докторант, что видно из записанной основной цели диссертации на с. 3, и, точнее, первой части этой цели – „провести на теоретическом уровне сопоставление между семантикой темпоральных значений в двух языках, что поможет теоретической обосновке возможных переводческих эквивалентов ...“ (в переводе с болгарского на русский язык). Докторант опирается на теоретические взгляды Ивана Куцарова о грамматических категориях болгарского глагола, с одной стороны, а, с другой, разделяет видение о 9 временах в болгарском языке, что на первый взгляд кажется некоторым противоречием. Тодор Лазаров преодолевает его, воспринимая идею, что категория времени - это гиперкатегория, и представляет собой совокупность трех морфологических категорий в болгарском языке по классификации глагольных грамматических категорий Ив. Куцарова: времени, вида действия и таксиса. И несмотря на то, что можно спорить по поводу взгляда Ив. Куцарова о трех отдельных категориях, сам выбор Т. Лазарова опираться на видение, что грамматическое время - гиперкатегория, дает общую основу, которая позволяет докторанту искать параллели между двумя по-разному организованными с типологической точки зрения темпоральными системами – болгарской и английской.

Несомненно, выбор теоретической лингвистической позиции, с которой нужно исходить в таком сопоставительном по своей сущности исследовании, совсем не простая задача, при наличии разнородных взглядов не только в грамматических традициях двух языков, но и в типологии. При этом обязательно как особую трудность нужно учитывать и различия в наборе грамматических глагольных категорий в двух языках и не одинаковый состав и организация глагольных категорий. Докторант показал возможность извлечь

положительное из неоднородных позиций с целью обосновать свою теоретическую концепцию и однозначно дефинировать использованный терминологический аппарат.

Положительной чертой избранного докторантом подхода является учет взаимодействия между грамматическими и лексическими элементами при выражении темпоральности, потому что без этого невозможно искать оптимизирования переводческих соответствий. Все-таки грамматические категории функционируют в тексте через свое употребление в конкретном контексте во взаимодействии с лексическими элементами. Реализация болгарских времен в связанном тексте подчеркнута контекстуальна зависима, особенно в сложном предложении, в отличие от согласования времен в английском языке. Этот факт предопределяет разнородность переводческих соответствий болгарских глагольных времен в английских текстах. В диссертационном труде рассматривается этот существенный вопрос, который непосредственно касается реализации одной из исследовательских целей, связанных с переводными соответствиями темпоральных форм – это правило согласования времен (с. 61) в английском языке.

Очень важными для реализации целей диссертационного труда являются две гипотезы, которые автор проверяет через извлечение статистической информации из параллельного болгарско-английского корпуса, а именно: существует ли какое-либо соответствие между видом глагола в болгарском и продолжительные глагольные формы в английском (так от употребления можно идти к теоретическому обобщению, возможно ли в конечном счете говорить о виде в широком смысле, который грамматикализуется по-разному в разных языках – с помощью вида в славянских языках и с помощью продолжительных глагольных форм, – при этом в принципе это разные семантические явления более общего в семантическом плане поля), и вторая гипотеза – существует ли соответствие между аористом и простым прошедшим в английском, и между имперфектом и продолжительным прошедшим временем при учете видовых различий в болгарском языке (т.е. эмпирическим путем искать доказательства грамматикализования схожих семантических признаков разным способом) (с. 81). Эти гипотезы оформлены как рабочие вопросы в параграфе 6 на с. 84.

Докторант исчерпывающе представил историческое развитие идей и подходов в машинном переводе (как традиционные статистические системы машинного перевода, так и машинный перевод, основанный на нейронных сетях). Необходимая часть диссертационного труда – рассмотрение характеристик языковых корпусов в качестве ресурсов машинного перевода, так как достаточное количество эмпирического материала и применение значимых критериев к корпусу гарантируют адекватность исследования.

Я считаю, что в диссертации корпус, с которым работает докторант – Болгарско-английский корпус с соотнесенными глагольными формами - достаточно ясно представлен, выяснены его параметры. Дискутируются проблемы относительно отношений между критериями балансированности и иллюстративности используемого корпуса из-за наличия глагольных форм с низкой частотностью. Автор сам подчеркивает, что вопреки представительности и иллюстративности языковых данных в Корпусе,

остается проблема грамматического разнообразия. В отдельной части представлены и аннотационные конвенции (два аннотационных слоя) при создании использованного корпуса.

Статистические данные о распределении болгарских глагольных времен в использованном корпусе, абсолютные стоимости и процентные соотношения в четырех типах текстов (а это художественная литература, публицистические тексты, научные тексты и субтитры фильмов) дают бесспорные основания для формулировки утверждений, которые доказывают высказанные в научной литературе гипотезы, базирующейся на исследовательской интуиции. Подобные утверждения касаются частотности употребления времен в зависимости от степени их семантической маркированности, замещения более маркированных времен их немаркированными коррелятами, зависимости частотности употребления времен от типа текста и т.д. Сопоставление между распределением болгарских и английских времен в двух частях корпуса (болгарской и английской) подкрепляет мнение, что частотность употребления следует семантической нагрузке членов одной категории, а их распространение зависит и от характерных стилистических особенностей текста.

С точки зрения моих научных интересов, Таблица 15 дает распределение переводческих соответствий форм болгарского настоящего времени в английских текстах. Оно показательное, особенно что касается положения в художественных текстах и субтитрах фильмов, потому что причины заключаются не только в том, что представлены больше всего транспозиционные употребления, как указывает автор, но и в факте, что в болгарском языке грамматические категории отличаются повышенной степенью контекстуальной зависимости – если контекст выразил соответствующее значение, в случае – темпоральный признак, то маркированное этим признаком время может быть заменено его немаркированным коррелятом. Это относится не только к 1, но и к 2 и к 3 признакам, что приводит к замене даже будущего предварительного времени настоящим временем в болгарском языке. Таблица 15 – непосредственное доказательство этого факта. Подобные доказательства содержатся и в Таблице 20, которая показывает, что формам настоящего времени в болгарском языке в 32.4% из случаев соответствуют прошедшие времена в английском и 5.92% будущие времена.

При употреблении прошедшего времени сов. вида, несмотря на наличие нескольких переводческих соответствий, вариантность меньше, что обусловлено его семантической нагрузкой признаком предшествования по отношению к акту коммуникации. Мне кажется, что случаи единичных переводческих соответствий обусловлены семантическими и/или видовыми особенностями болгарского глагола. Данные Таблицы 19 с распределением переводческих соответствий болгарского прошедшего времени несов. вида, на мой взгляд, являются прямым доказательством, что оба времени имеют различное место в организации темпоральных систем двух языков и не имеют одну и ту же семантику. Данные о переводческих соответствиях болгарского перфекта в английских текстах были бы достоверны, если бы употребления конклюдивного аориста не были бы кодированы как

перфект. Наибольшее количество переводческих соответствий с помощью простого прошедшего времени в английском языке доказывает различия не столько в организации категории времени, сколько в функционировании времен, в т.ч. в сложном предложении.

Вывод, что категория таксис в болгарском не имеет строго определенных переводческих эквивалентов в английском, хотя почти 60% составляют переводческие соответствия с простыми временами в английском, имеет свою причину, которая коренится, по-моему, в факте, что в английском языке функционирование времен управляется правилами согласования, в то время как реализация индиректных времен (названных зависимым таксисом) подчиняется контекстуальной зависимости, а не облигаторности.

Особенно значимым для языковедческой теории и типологических исследований является тезис, доказанный с помощью корпусного подхода, что нет непосредственной связи на формальном уровне между категорией вида глагола в болгарском языке и продолжительными и перфектными временами в английском, а это значимо для передаваемой семантики. Это важный вывод, который непосредственно касается теоретической классификации этих грамматических признаков, часто относящихся к более широкому семантическому полю.

Конкретные данные при обработке использованного корпуса говорят о том, что в 70% случаев темпоральная форма в болгарском глаголе однозначно может предопределить темпоральную форму в английском языке (с. 172), что имеет существенное значение для машинного перевода. Важны также анализы докторанта о причинах несоответствия между формально-семантическим анализом и эмпирическими данными Корпуса, потому что они касаются самых важных различий между семантикой и функционированием в тексте темпоральных систем в двух языках.

Сделанные наблюдения на материале конкретных параллелей в Корпусе и представленные статистические данные являются особо ценными доказательствами, которые верифицируют исследовательские гипотезы и превращают выводы в доказанные научные факты.

6. Оценка публикаций по теме диссертационного труда: количество; характер изданий, в которых они опубликованы (международные, национальные). Каково их отражение в науке – использование и цитирование другими авторами.

7 научных публикаций по теме диссертационного труда опубликованы. Две из них опубликованы в сборниках одного из самых значимых международных научных форумов по компьютерной лингвистике у нас. Эти труды, как и опубликованные статьи в журналах „Български език“, „Съпоставително езиковедие“, в электронном журнале *Littera et Lingua* и в сборнике Международной юбилейной конференции Института болгарского языка в 2017 г. являются достаточным доказательством, что научная общность имела возможность

ознакомиться со значимыми научными поисками и достижениями в диссертационном труде.

7. Изготовлен ли автореферат согласно требованиям, правильно ли отражает основные положения и научный вклад диссертационного труда?

Автореферат отвечает структуре диссертационного труда и верно отражает достижения автора. Список с пунктами научного вклада адекватно представляет положительные результаты исследования.

8. Заключение с ясным утверждением, присудить ли ученую степень кандидату, или нет.

В результате анализа научного и научно-прикладного вклада в диссертационном труде, его значимости для языковедческой науки и для компьютерной лингвистики я убежденно предлагаю Тодору Димитрову Лазарову **присудить** ученую и образовательную степень „доктора“ (=кандидата наук) за его диссертационный труд „Българско-английски граматични паралели с оглед на машинния превод“ („Болгарско-английские грамматические параллели в перспективе машинного перевода“) и голосую **положительно** за это решение.

24.04.2020 г.

Рецензент:
(проф. д-р Красимира Алексова)

