

СТАНОВИЩЕ

От члена на НЖ Максим Иванов Стаменов, д.ф.н., проф. в Института за български език „Проф. Любомир Андрейчин“ към БАН

за дисертационен труд за придобиване на (образователната и) научна степен .доктор.

Автор на дисертационния труд: **Виктория Красиминова Петрова-Любенова**, редовен докторант в Института за български език „Проф. Любомир Андрейчин“ при БАН.

Тема на дисертационния труд: **„Полуавтоматично разработване на многоезикови терминологични ресурси“.**

Научен ръководител проф. д-р Светла Коева от Института за български език „Проф. Любомир Андрейчин“.

Общият обем на дисертацията е 237 страници, което включва 176 страници основен текст и 61 страници приложения и библиография. Съдържа увод, четири глави, заключение, приноси, 3 приложения и списък на термини. Трудът включва 23 фигури и 6 таблици.

1. Актуалност на разработвания в дисертационния труд проблем в научно и научно-приложно отношение и постигнати резултати

Проблемът, на който е посветена предлаганата дисертация, е твърде актуален в контекста на най-новите насоки на развитие на информационните технологии с оглед на разширяване на възможностите на компютърноподпомогнатия превод (на базата на системи за преводна памет, машинен превод и терминологични бази). Трудът представлява обстойно изследване, чиято цел е да предложи методика за създаване на терминологични ресурси за дадена специализирана област. За да бъде тя постигната, са анализирани възможностите на програмите за компютърноподпомогнат превод за създаването на терминологични бази, на програмите за извличане на термини, както и на наличните към момента терминологични ресурси за български език.

В Увода към дисертацията се посочва актуалността на изследването, мотивацията му, неговият предмет, цели и задачи, методи, използвани при работата и се очертава структурата му. Като основна цел се посочва ориентацията да предложи методика за създаване на терминологични ресурси за дадена специализирана област. Предмет на изследването стават терминологичните бази и тяхната употреба в системите за компютърноподпомогнат превод, както и начините, по които съвременните информационни технологии позволяват и улесняват създаването и употребата им.

В първа глава се прави преглед на съвременните системи за компютърноподпомогнат превод, като се акцентира върху техните основни компоненти (преводна памет, машинен превод, терминологични бази). Тук се описват по подходящ начин новите познания и умения, които се изискват от преводачите във връзка с използването на технологиите за компютърноподпомогнат превод. За тази цел се очертават съвременното състояние на използване на системите за компютърноподпомогнат превод, техните компоненти и нови функционалности. За подходяща ориентация на актуалните практики към днешна дата, се провежда проучване

сред преводачите в България (73 професионални преводачи) относно нагласите при използването на системите за компютърноподпомогнат превод.

Във втора глава са представени видовете терминологични речници и тяхната структурата в съпоставка със структурата на терминологичните бази. Целта е да се очертаят приликите и разликите, както и преимуществата при използването на терминологични бази в процеса на работа със системи за компютърноподпомогнат превод. Т също се обсъждат структурата и особеностите на терминологичните бази, като за целта се обсъждат световноизвестни терминологични бази, които могат да се използват във или извън системите за компютърноподпомогнат превод. Специално внимание се обръща на създаването и управлението на терминологичните бази в системите за компютърноподпомогнат превод. Изтъкнати са основните начини за създаване и управление на терминологични бази в системи за такъв тип превод. В допълнение се разглеждат приложения, които могат да се използват заедно със системите за компютърноподпомогнат превод, както и самостоятелно функциониращи системи за работа с терминология, които могат да бъдат използвани паралелно. Демонстрира се използване на терминологичните бази в зависимост от различните типове текст за превод и приложението на контролирани езици. Основно внимание в това отношение се обръща на изискванията за последователност и непротиворечивост при употребата на термини и възможността да се наложат специфични ограничения.

В трета глава се тълкуват особености на компютърната терминология, като ориентацията е към обобщение от гледна точка на включването на компютърните термини в терминологични ресурси. Описват се начините за създаване на терминологични ресурси, като се започне от традиционните отпечатани терминологични речници, премине се през онлайн терминологичните речници и се достигне до автоматичното извличане на термини и представянето на съществуващите инструменти за тази цел. Описанието на инструментите се съпътства от сравнение на техните функционалности в рамките на определен период от време на тестване. Създадена е също така методика за полуавтоматично създаване на терминологична база в областта на компютърната терминология и се описват конкретните етапи от работата по разработването на двуезична терминологична база за английски и български език. Представени са характеристиките на така създадената двуезична терминологична база. Предлага се унифициран начин на работа при създаването на терминологични бази, който се основава се от най-добрите практики и взема под внимание възможностите и начина на работа на преводачите.

В четвърта глава се предлага и обсъжда оценка на точността на превода със системите за компютърноподпомогнат превод без или с използването на терминологични бази. Обсъждат се проблемите, свързани с измерването на точността на превода с функционалности, вградени във или външни за инструментите за компютърноподпомогнат превод. Предложена е класификация на грешките при превод със системи за компютърноподпомогнат превод с оглед на тяхната значимост на съдържанието на текста. Направено е сравнение на ефективността на програмите за оценка на превода по предварително зададена методика.

Като обобщение на постигнатите резултати може да се посочи, че в дисертацията се предлага методика за полуавтоматично създаване на двуезична терминологична база с посока на превода от английски на български, а областта, за която се отнася, е компютърната терминология. Разгледани са също някои възможности за измерване на качеството на превода в системите за компютърноподпомогнат превод, които имат вградени функционалности за това. За тази цел в дисертацията се представят резултатите от два експеримента, с които се цели проверка на поведението на програмите за качество при превод от английски на български език.

2. Ориентация на дисертантката за състоянието на проблема и научната литература

Авторката показва добро познаване на литературата по темата на работата. В библиографията към дисертацията са посочени общо 134 публикации на наши и чуждестранни учени, с които дисертантката е работила при реализацията на дейностите, свързани с изследването и написването на своя труд, и 18 източника от интернет.

3. Характер на избраната методика

При разработването на дисертацията авторката използва както съвременни методи (статистически, за компютърна обработка, корпусен), така и някои традиционни (описателен, аналитичен, експериментален). Като основни за постигане на изследователските цели се възприемат анализът и описанието. Анализ се прилага при подбора на подходящи езикови ресурси и средства за тяхната обработка. Хипотезите се верифицират експериментално посредством използването на такива инструменти като анкета, проведена онлайн, работа с програми за автоматична обработка на текстове, за автоматично извличане на термини, за автоматично подравняване на паралелни документи по изречения и за автоматично създаване на двуезикови терминологични бази. При работата с всички програми за компютърноподпомогнат превод се използват съвременни средства за компютърна обработка на текстове, на терминологични бази, използване на машинен превод, за създаване на преводна памет. Прилагат се и количествени методи за обобщение на събраната информация, както това е налице при анкетата сред преводачите и при автоматичното извличане на термини. Корпусният подход се прилага при аотирането на преводни текстове с грешки с цел да се верифицират функционалностите на програми за проверка на качеството на превода.

4. Кратка аналитична характеристика на научните и/или научно-приложните приноси на дисертационния труд

Потвърждаваме наличието на следните приноси в труда. Те са ориентирани в две основни посоки – теоретични и с приложен потенциал.

Сред теоретичните подходящо е да бъдат посочени следните. На първо място, представени и анализирани са в обстоен преглед възможностите на системи за компютърноподпомогнат превод и промените, които те носят със себе си. Това става на базата на преглед на актуалните проучвания в областта на компютърноподпомогнатия превод. На второ място се проучват и обсъждат наличните начини за създаване на терминологични бази за системите за такъв тип превод. На трето място, предлага се описание и анализ на системите за проверка на качеството.

Сред приносите с приложна насоченост следва да се изтъкнат следните. Извършена е анкета, с която изследователката се обръща към преводачите в България, които работят със системи за компютърноподпомогнат превод, за да изследва аспекти на техните практики в това отношение. Разработена е методология за полуавтоматично създаване на терминологични ресурси за български език за системите за компютърноподпомогнат превод, като се вземат под внимание някои ограничения (безплатен софтуер, поддръжка на български и в областта на компютърната терминология). Не на последно място се предлага унифицираща класификация на грешките при превод.

5. Преценка на публикациите по дисертационния труд: брой, характер на изданията, в които са отпечатани, цитирания

Дисертантката отчита като публикации, свързани с темата на дисертацията, общо три статии. Две от тях са били представени на годишните конференции на ИБЕ и публикувани в съответните Сборници с доклади, а третата е публикувана в материалите от уъркшоп на компютърните технологии за превод. Описаните публикации отговарят на минималните изисквания за придобиване на научната степен „доктор по филология“. Приложен е и списък с две цитации на нейни публикации.

6. Критични бележки

Наличието на Заключение (стр. 170-175) не е отбелязано в Съдържанието на дисертацията (стр. 3-5), а така също и в описанието на структурата на дисертацията в Автореферата (стр. 6-7).

7. Авторефератът правилно ли отразява съдържанието на дисертационния труд

Авторефератът е подробен (61 стр.) и правилно отразява основните положения и научните приноси на дисертационния труд. Той отговаря на съдържанието му.

8. Заключение

Имайки предвид гореизложеното, предлагам на Научното жури да присъди на дисертантката **Виктория Красимирова Петрова-Любенова** научната и образователна степен „доктор“ по професионално направление 2.1 Филология, научна специалност Общо и сравнително езикознание (компютърна лингвистика) за дисертацията ѝ **„Полуавтоматично разработване на многоезикови терминологични ресурси“**.

Дата 10.12.2022 г.

Изготвил рецензията:

(подпис)