СТАНОВИЩЕ

от доц. д-р Боряна Колева Братанова

катедра „Англицистика и американистика“, Филологически факултет,

ВТУ „Св. св. Кирил и Методий“, професионално направление 2.1. Филология. Съвременен английски език – граматика и семантика

за придобиване на научната степен „доктор на науките“ по професионално направление 2.1. Филология. Теория и практика на превода

Нов български университет, Магистърски факултет, Департамент чужди езици и култури

от гл. ас. д-р Велислава Райкова Стойкова

Тема на дисертационния труд: Подходи за машинен превод на български език

1. Значимост на изследвания проблем в научно и научно-приложно отношение

Представеният дисертационен труд систематизира различни подходи в областта на статистическия машинен превод и тяхното приложение при превод на български език. Темата е особено значима, тъй като разработването на езикови технологии е насочено основно към езици с голям брой говорещи, като най-разработени са те за глобалния английски език. Дисертантката си поставя за цел да разгледа подходите за машинен превод на български език, който не попада в групата на големите световни езици, поради което технологии за обработка за българския език се разработват основно в България от носители на езика. Като славянски език с богата морфология българският език е особено благодатен за компютърна обработка. В този контекст дисертационният труд на д-р Велислава Стойкова представлява значима научна разработка, в която са интегрирани както теоретични аспекти, така и научно-приложни изследвания. В този смисъл дисертационният труд напълно се вписва в научната област теория и практика на превода.

2. Обоснованост на целите и задачите в дисертационния труд

Целите на изследването са ясно формулирани единствено в автореферата. Като основна цел на дисертационния труд се определя разглеждането на възможностите за прилагането на формалните езикови теории за компютърен анализ на българския език като цяло и в частност приложението им с оглед на машинния превод. Изследването се осъществява чрез установяване приложимостта на три подхода за машинен превод на

български език – DATR език за представяне на лексикална информация, универсалния мрежов език (УМЕ) за представяне на граматическо знание за съответния език и Sketch Engine (SE) (Скетч Енджин) за статистическа обработка на двуезични електронни корпуси. Изследователските задачи, които е необходимо да намерят своето решение за постигане на поставените цели не са ясно дефинирани, а по-скоро са преплетени с целите на изследването.

3. Съответствие между методологията на изследване и целите на дисертационния труд

Водеща методология на изследването е контрастивният анализ, който се основава на извеждането на Tertium Comparationis и прилагането му при разглеждане на езикови явления в два или повече езика с цел установяване на сходства и различия между тях. Ясно е представено мястото на машинния превод в по-широката област на преводознанието, както и спецификата на текстовете, които са подходящи за машинен превод. Набляга се на традициите, които изследванията в областта на машинния превод имат в българското езикознание и преводознание. Очертават се основните различия между професионалния превод, направен от човек и машинния превод от гледна точка на богатство на използваната лексика и на езиковите структури, образност на превода, боравене с културноспецифична лексика и др. Представя се и спецификата на двата основни подхода при превод с помощта на компютър – машинен превод и компютърно подпомогнат превод.

В изследването са представени етапите на функциониране на софтуерите за машинен превод във връзка с кодирането и декодирането на информация при автоматичен превод от изходния език на целевия език. При извличане на значението от изходния език се минава през междинния етап на символна междуезикова репрезентация преди значението да се предаде като преводно съответствие чрез структурите на езика цел. В машинния превод се използват основно формални подходи при обработката на езика. В дисертационния труд се представят трите основните подхода при машинния превод – машинен превод, основан на правила, статистически машинен превод и невронен машинен превод. Въпреки че с навлизането на невронните мрежи се разширява обхватът на текстове, които се превеждат автоматично, в дисертационния труд ясно са разграничени областите, които засега са постижими единствено за професионалните преводачи.

Дисертационният труд представлява съпоставително изследване на приложенията на машинния превод за анализ на семантичните и граматическите особености на българския език. Избраната методологията на изследване съответства на поставените цели, тъй като включва подробно представяне на трите подхода за машинен превод на български език – DATR, универсалния мрежов език (УМЕ) и Sketch Engine (SE) (Скетч Енджин) в теоретичен и практико-приложен план. Основните подходи на съпоставителното езикознание и машинния превод са умело интегрирани в анализа и са приложени при изследване на езикови явления в различни езици, напр. флективната морфология в българския, руския и полския език, прономиналната система в българския, английския и руския език и др. Дисертационният труд цели да представи приложимостта на методи за езикова обработка при осъществяване на съпоставителни изследвания, което се онагледява чрез прилагане на модел на машинен превод между двойки езици като български и словашки, български и сръбски, български и английски и др. Практико-приложните аспекти на изследването са онагледени чрез анализ на машинния превод на някои единици като напр. екологични термини и времеви изрази в българския и словашкия и др. От съществено значение за изследването е представянето на компютърните подходи към терминологията. Поради своята еднозначност, термините повишават успеваемостта на машинния превод. Като цяло дисертационният труд показва умело интегриране на методологията на съпоставителните изследвания с подходите на машинния превод за целите на изследването. Проблематиката, разглеждана в дисертационния труд, е представена изчерпателно и в автореферата, чийто текст е логически структуриран с плавен преход между отделните раздели и следва точно съдържанието на дисертационния труд.

4. Научни и научно-приложни приноси на дисертационния труд

Дисертационният труд съдържа редица приноси както в теоретичен план, така и в практико-приложен аспект с оглед на технологиите за обработка и машинен превод на българския език. Като по-важни приносни моменти може да се посочат: приложение на съвременните формални езикови теории към различните подходи за машинен превод на български език; извеждане на начин за компютърно измерване на качеството на резултата от машинния превод на български език; очертаване на съвременните достижения в областта на машинния превод за български език; установяване приложимостта на машинния превод като подпомагащ дейността на професионалните преводачи. Дисертационният труд поставя

акцент върху т.нар. малки езици с фокус върху българския език и по този начин онагледява значимостта и приложимостта на технологиите за машинен превод без обвързване с броя говорещи съответния език. Дисертацията е написана на английски език, което също може да се посочи като принос на изследването, тъй като дава възможност трудът да стане достояние на глобалната академична и изследователска общност, която работи в областта на технологиите за превод, и да популяризира системите за машинен превод за български език.

5. Преценка на публикациите по дисертационния труд

Д-р Велислава Стойкова е автор на множество научни публикации основно в областта на езиковите технологии и машинния превод, които са публикувани в специализирани издания в страната и в чужбина. Разработките са написани на английски и на български език. За последните десет години д-р Стойкова има над четиридесет публикации, някои от които са реферирани в Scopus. Към дисертацията са представени дванадесет публикации, сред които също има реферирани в Scopus и Web of Science.

6. Цитирания

Д-р Стойкова има над 40 цитирания. Значителна част от тях са в чужбина, което показва международното признание на нейната научна дейност.

7. Мнения, препоръки и бележки

Представените подходи към машинен превод за българския език може да бъдат тествани върху специализирани текстове, които съдържат терминология, напр. икономика, финанси, здравеопазване и т.н. Изследванията върху машинния превод биха могли да се обвържат в по-голяма степен и с областта на терминологията. В дисертационния труд има някои незначителни езикови грешки, които е препоръчително да бъдат отстранени, напр. *the first* (стр. 8) → *the former*, *later* (стр. 8) → *the latter*, *underlay* (стр. 9, 10, 17) → *underlie*, *can regarded*, *can derived* (стр. 9) → *can be regarded*, *can be derived*, *can be occurred* (стр. 13) → *can occur*, *sentense* (стр. 25) → *sentence*, *is represent* (стр. 38) → *are represented*.

8. Заключение и оценка на дисертационния труд

Дисертационният труд, научните публикации и цитиранията на д-р Велислава Стойкова удовлетворяват наукометричните изисквания на Закона за развитието на академичния състав в Република България и съпътстващите го правилници. Като цяло дисертацията представлява съществен принос в развитието на технологиите за компютърна обработка и машинен превод на български език, за който все още няма достатъчно

разработени технологични ресурси. Давам своята положителна оценка за дисертационния труд и препоръчвам на членовете на уважаемото научно жури да гласуват положително за присъждането на научната степен „доктор на науките“ на д-р Велислава Стойкова в област на висше образование 2. Хуманитарни науки, професионално направление 2.1. Филология. Теория и практика на превода.

10.09.2022 г. Изготвил становището:

гр. Велико Търново /доц. д-р Боряна Братанова/