

СТАНОВИЩЕ

за дисертационния труд на Мария Андреева Тодорова
Типология и свойства на устойчиви словосъчетания в българския език.

Глаголни фразеологизми

за присъждане на образователната и научна степен „доктор”

в област на висшето образование 2. Хуманитарни науки,

професионално направление: 2.1. Филология

(компютърна лингвистика)

Дисертационният труд на Мария Андреева Тодорова *Типология и свойства на устойчиви словосъчетания в българския език. Глаголни фразеологизми* се състои от 184 страници – увод, четири глави, заключение, библиография и три приложения. Целите на дисертацията, дефинирани в увода на работата: да се характеризират устойчивите словосъчетания (класификация, дефиниции, отличителни признаци) на базата на съществуващите изследвания в традиционната и модерната лингвистика и да се опишат морфологичните и синтактичните свойства на глаголните фразеологизми с оглед на компютърната обработка на езика, са изпълнени успешно.

Както традиционната, така и модерната лингвистика не предлагат описания, приложими при компютърната обработка на езика – наблюденията и анализите са или твърде общи, или ако са детайлни, засягат единични езикови явления. По тази причина разработките, свързани с компютърна лингвистика, предполагат на първо място задълбочен и прецизен лингвистичен анализ, който е предпоставка за непротиворечиво формализирано описание. Това означава, че лингвистичният анализ трябва да включва формулировки, достатъчни за еднозначната компютърна обработка на дадена лингвистична единица при възможните ѝ обкръжения. Следователно за изследване, свързано с лингвистично мотивирана компютърна обработка на езика (базирана на речници и правила), се очаква да има преди всичко теоретичен езиковедски принос, който без съмнение ще намери и практическо приложение. Това се отнася и до дисертационния труд на Мария Андреева Тодорова. Задачата на докторантката – да създаде модел за описание на глаголните фразеологизми, приложим за компютърна обработка, изисква ориентиране в съществуващото теоретично и терминологично многообразие и решаване на редица неизяснени досега граматични въпроси: за основната форма на глаголните фразеологизми; за парадигмата на отделните граматични типове; за възможните перифрази и т.н. Като научен ръководител си позволявам да дам цялостна висока оценка на постигнатите от Мария Тодорова резултати.

Първа глава, озаглавена *Устойчиви словосъчетания – теоретична рамка*, съдържа синтезиран преглед на голяма част от изследванията, чийто обект са устойчивите словосъчетания (библиографията съдържа 206 заглавия). В Таблица 1. се обобщават пресечните точки между различни термини и дефиниции, а в Таблица 2. са сравнени класификациите на Дж. Наги, Ю. Балтова, Ш. Бали, Н. Д. Шмельов, В. Виноградов, А. Саг, Т. Болдуин, Ч. Филмор и др., К. Ничева, В. Кювлиева, И. Мелчук – представители на различни направления и подходи. Трябва да подчертая, че показаното равнище на обобщение се постига изключително трудно, и авторката демонстрира завидна прецизност при откриването и показването на приликите и разликите между

изследваните теоретични постановки. Високата степен на генерализация, обаче, в някои случаи може да бъде следвана само от читатели, които до голяма степен са запознати с проблематиката. Не бих казала, че това е търсен ефект, а по-скоро е следствие от факта, че информацията, която трябва да бъде анализирана и категоризирана, е огромна като количество и многообразна по отношение на теории, подходи, класификации и равнище на формализация.

Във втора глава – *Типология на устойчивите словосъчетания в български*, се изследват основните проблеми на формата, значението и комуникативните функции на устойчивите словосъчетания с оглед на извеждането на система от признаци за тяхното разпознаване, описание и типологизиране. Устойчивите словосъчетания и техните видове са описани с помощта на формалния апарат на теорията Смисъл <=> Текст на И. Мелчук и допустимите за даден тип комбинации от признаци: [+/-мотивираност; +/-анализируемост; +/-продуктивност; +/-буквалност; +/-преносност]. Според приложените дефиниции, ако е налице признакът [-анализируемост], то винаги трябва да е налице признакът [-мотивираност], ако е налице [+анализируемост], то трябва да е налице и [+буквалност] и т.н. Характеризирането посредством признаци може в бъдеще да се развие в две посоки – да се потърсят непротиворечиви критерии за дефиниране на проявите на всеки един признак и да се открият групите от фразеологизми, при които комбинациите от дадени признаци се проявяват или само с положителните, или само с отрицателните си стойности. Също така признаците, които характеризират надреден тип, трябва да обединяват признаците на типовете, които ги съставят, което може да се прецизира.

Третата глава – *Морфо-синтактични свойства на глаголните фразеологизми в български*, е централна за дисертацията. При категоризирането на глаголните фразеологизми се отчитат: пълнотата или ограничението на парадигмата (до един или определено множество от членове) както на глаголната опора, така и на останалите конституенти; допустимата съчетаемост между лексикалните единици; словоредните и съгласувателните ограничения; възможностите за добавяне, изпускане или замяна на конституенти; възможностите за перифразиране. С други думи, изследването цели (и постига) пълно морфологично и синтактично описание на глаголните фразеологизми в български. Мария Тодорова следва критериите за минималност и неутралност при дефинирането на абстрактна основна форма и на практика прилага разбирането ни, че лексикалните единици (думи, съставни думи, фразеологизми) трябва да имат унифицирано описание по отношение на категориалната и парадигматичната информация (характеризиращи основните форми и показващи групирането на думите по части на речта и парадигматични класове) и граматичната информация (характеризираща образуването на словоформите и показваща групирането на думите с еднакво формообразуване, което обхваща не само синтетичните, но и аналитичните форми). Теоретично може да се очаква, че всички обособени в предходни наши разработки двамадесет групи лични глаголи, десет – третолични, и десет – безлични, могат да участват при образуването на глаголни фразеологизми (в момента са регистрирани само за някои от тях – стр. 102 и следващи).

Към специфичните дейности, които съпътстват подобен тип изследвания, трябва да се отбележи и ексцерпцията на данни – не може просто да се наблюдава какво е регистрирано в речниците или корпусите и да се привеждат примери. Наличният материал трябва да се извлече на базата на критерии, свързани с целта на разработката,

да се унифицира, защото в различните речници и корпуси е представен по различен начин, да се отстранят неточностите и да се вземе решение как да се подходи към различните типове многозначност, които неминуемо съпътстват ръчната обработка на езикови данни. Справяйки се с тези задачи, Мария Тодорова е съставила колекция от над 70 000 устойчиви словосъчетания, сред които за нуждите на изследването е извадила 27 902 глаголни фразеологизма.

В четвърта глава, озаглавена *Формализиран морфо-синтактичен речник на глаголни фразеологизми в български*, се представя начинът за създаване на компютърен речник и система от правила, които описват парадигмата, словореда, съгласуването и перифразирането на глаголните фразеологизми. Лингвистичната информация е организирана според формализма за представяне на компютърни морфологични речници и компютърни синтактични правила, отразяващи лексикални и граматични зависимости, които се свързват с името на Морис Грос. И двата начина на представяне съответстват на крайните преобразуватели, математически модели на устройства за разпознаване на регулярни езици и приписване на допълнителна информация (в случая граматични характеристики). Компютърните морфологични речници и синтактичните правила за комбиниране позволяват автоматичното разпознаване на дадена дума (в случая глаголен фразеологизъм) и съотнасянето ѝ с всички правилни лингвистични интерпретации, ако са повече от една. Мария Тодорова описва 28 глаголни типа и 44 именни типа (на съществителни, прилагателни и местоимения, компоненти на глаголните фразеологизми). Подходът, който авторката е приела, е възможен, но един различен подход, при който типовете са само глаголни и характеризират конкретната съчетаемост на класове за даден фразеологизъм, би бил в съответствие с концепцията за унифицирано описание на лексикалните единици.

Работещите в областта на компютърната лингвистика се сблъскват и преодоляват още един проблем, който не е познат за останалите езиковедски разработки. Обобщеното лингвистично знание трябва да бъде представено във формата и ограниченията на дадено компютърно приложение (в конкретния случай най-доброто, известно към момента), което в общия случай изисква значителна комбинативност, за да се запази пълнотата и непротиворечивостта на езиковите данни.

Мария Андреева Тодорова е завършила магистърската програма *Компютърна лингвистика. Интернет технологии в хуманитаристиката* към Софийския университет „Св. Климент Охридски“ през 2003 година, а в периода 2006 – 2009 година е редовна докторантка по компютърна лингвистика в Института за български език. Собственият ми опит с докторанти (трима успешно защитили и един редовен докторант в момента), както и наблюденията ми върху работата на другите докторанти на Секцията по компютърна лингвистика показва, че дисертацията на Мария Тодорова е претърпяла най-многобройни редакции, което се дължи на относително продължителния период за нейното създаване (поради прекъсване по майчинство) и на високата и самокритичността на докторантката. Колкото и парадоксално да звучи, многобройните редакции са причина за част от техническите несъвършенства, които високата читател трудно може да открие в текста (например неточно употребено местоимение на стр. 50, употреба на термина *ограниченост* вместо *рестриktivност* на стр. 68, неточно цитиране на стр. 71 и др.). Докторантката беше изключително коректна към предложенията на научния ръководител, но отстояваше и собствените си различни виждания, което заслужава уважение. Например, според мене

читателят би вникнал по-добре в теорията Смисъл<=>Текст, ако теорията е изложена в целостта си на едно място. Представянето на различни части от теорията на няколко места отговаря на логическата структура на текста, но води до повторения или пропуски. Например, от изброените на стр. 50 и 51 видове фразеологизми: прагматемите, идиоми, полуидиоми и колокации, само прагматемите са описани в самостоятелно обособена част на стр. 76, а за останалите три вида читателят трябва да има допълнителни знания и да прояви съобразителност, за да ги намери в текста. Друг пример – изброените към съставните наименования подтипове на стр. 73 са всъщност само именуванни обекти, които могат да се класифицират като собствени имена и каскадни наименования.

Авторефератът отразява основните наблюдения и изводи, направени в дисертацията. Мария Тодорова има значителен брой публикации, свързани с темата на дисертацията. Приносите на дисертационния труд са обективно, но доста скромно описани от авторката. Дисертационният труд има сериозни теоретични постижения при формулирането на критериите за минималност и неутралност на абстрактната основна форма на глаголните фразеологизми; при описанието на парадигмата им, което отчита морфологичните, синтактичните, словоредните и съгласувателните особености; както и при категоризирането на глаголните фразеологизми в структурни и флективни типове. Правилните теоретични заключения и обобщения на авторката за наблюдаваните закономерности на различно езиково равнище правят възможно създаването на компютърния морфологичен речник и формулирането на компютърните синтактични правила, които имат приложение при автоматичната обработка на български език.

В заключение на изложеното дотук убедено предлагам почитаемото жури да вземе решение за присъждането на образователната и научна степен „доктор” на Мария Андреева Тодорова.

25 февруари 2015 г.

Проф. д-р Светла Коева

