

РЕЦЕНЗИЯ

на дисертационен труд за придобиване на образователната и научната степен

доктор

Автор на дисертационния труд: Христина Кукова

Секция по компютърна лингвистика

Институт за български език „Проф. Любомир Андрейчин“

Тема на дисертационния труд: „Формално описание на перцептивните и менталните предикати в българския език (снемане на семантична многозначност)“

Рецензент проф. д-р Йовка Тишева, СУ „Св. Климент Охридски“

1. Данни за автора на дисертационния труд и за докторантурата

Христина Кукова е завършила бакалавърската програма „Руска филология“ във Факултета по славянски филологии на СУ „Св. Климент Охридски“ през 2001 г., а през 2002 г. – бакалавърската програма „Българска филология“ в същия факултет. В периода 2003 – 2004 г. се обучава в магистърската програма „Компютърна лингвистика. Интернет технологии в хуманитаристиката“. Тя разширява своята подготовка в областта на компютърната лингвистика и като специалист по проекта “Интегриране на нови практики и знания в обучението по компютърна лингвистика” (2012 – 2013 г.). В същия период работи и върху Българския национален корпус. През 2010 г. Хр. Кукова е назначена като проучвател в Секцията по компютърна лингвистика на ИБЕ, по-късно продължава работата си като асистент. В момента е филолог към същата секция и се занимава с разработката и тестването на езикови ресурси.

В периода от 1 март 2004 г. до 1 март 2008 г. Христина Кукова е зачислена като докторант по професионално направление 2.1. Филология, докторска програма „Общо и сравнително езикознание (математическа лингвистика)“ в Секцията по компютърна лингвистика на Института за български език „Проф. Л. Андрейчин“. Обучението ѝ е в редовна форма с научен ръководител проф. д-р Радка Влахова. От 1 септември 2008 г. е отчислена с право на защита. По време на докторантурата Хр. Кукова повишава своята научна и професионална подготовка чрез участие в курсовете, организирани от Центъра за обучение на БАН, Учебна програма за докторанти, както и в специализирани обучения по вероятностни модели и формална семантика. С решение на Научния съвет на ИБЕ от 2 декември 2021 г. изпитите, положени от докторантката по време на обучението ѝ, се признават при процедурата за защита на дисертационния труд.

Христина Кукова е взела участие в един национален проект на Института за български език – „Написаното остава! Пиши правилно!“. Тя е част от екипа на два текущи научни проекта: „Онтология на ситуациите за състояние – лингвистично моделиране. Съпоставително изследване за български и руски“ и „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“.

Представеният по процедурата текст е с общ обем от 379 стр., от които 280 стр. са основен текст. Разработката съдържа 38 таблици с обобщения и 4 фигури. Включени са и две приложения. Първото представлява таблица със синонимни множества (ID №, литерали и дефиниция) от Уърднет, приписан фрейм от Фреймнет и тип според предложената в Глава 4 типология, а второто – съпоставка на речниковите значения на

глаголите за перцепция в Речника на българския език и в Уърднет. В списъка с цитираните източници фигурират 147 заглавия.

2. Актуалност на разработвания в дисертационния труд проблем

Дисертационният труд на Христина Кукова „Формално описание на перцептивните и менталните предикати в българския език (снемане на семантична многозначност)“ е разработен в актуална и важна в теоретикоприложна област на компютърната лингвистика. В теоретично отношение темата е пресечна точка на изследователски въпроси, разглеждани в редица когнитивни науки и в езикознанието, тъй като отразяват връзката между семантичните и синтактичните структури, между механизмите на пораждање и експлициране на лексикалните и на синтактичните значения. Работата представлява цялостно и задълбочено изследване на семантичните особености на един клас предикати в българския език, които в определени контексти се реализират със значения, отнасящи ги към друг лексикалносемантичен клас. Многозначността, означавана още като полисемия, омонимия, размитост, синкретизъм или конверсия, е изследвана с оглед на процесите, които я пораждат, на конкретни класове думи, при които се отбелязва този феномен, напр. при съществителните, или на определено езиково равнище, напр. морфологичната многозначност. Трудът на Хр. Кукова представлява цялостно проучване на връзката между глаголите за възприятие в българския език, представящи предикатите за перцепция, и семантичния клас на менталните предикати.

Освен в собственолингвистичен план, актуалността на проблематиката се свързва и с продължаващите в компютърната лингвистика опити за изработване на модели и алгоритми за автоматичното снемане на семантична многозначност. Автоматичното отстраняване на семантичната многозначност е съществен проблем при компютърната обработка на естествения език. За разлика от снемането на граматичната многозначност, което в една или друга степен е преодоляно за повечето европейски езици, разрешаването на семантичната многозначност е значително по-сложен и все още нерешен проблем. Автоматичното снемане на този тип многозначност е свързано с избора на подходящото значение за дадена дума от множество от предварително определени възможни значения. В българския език предикатите от разглеждания клас се характеризират с множество лексикални значения, както и с голямо структурно разнообразие при синтактичната реализация. Както авторката отбелязва, за целите на автоматичното снемане на семантична многозначност следва да бъдат описани всички възможни комбинации, тоест синтактични конструкции, в които разглежданите предикати се реализират с основното си значение възприятие или като ментални предикати.

Обектът на проучването е ясно и точно определен. Авторката си поставя три научни цели: допълване на досегашните описания и типологизиране на предикатите за перцепция; съпоставителен анализ на аргументните и комплементните характеристики на предикатите с акцент върху онези представители на глаголите за перцепция, които биха могли да се срещнат в двете групи (перцепция и менталност) с цел подпомагане на автоматичното снемане на семантична многозначност; типологизирано представяне на фреймовете и фреймовите елементи на перцептивните глаголи. Те са реалистични, постижими и отговарят на темата на дисертацията. Деветте научни задачи произтичат от основната цел и точно представят етапите, през които преминава проучването. Смятам,

че първата от формулираните три хипотези отразява един общ постулат за връзката между лексикалната семантика на глаголите (не само тези за перцепция) и синтактичната им реализация, доказан в редица предходни изследвания.

3. Познаване на състоянието на проблема и научната литература по темата

Теоретичната рамка на изследването е разработена след проучването на 147 библиографски източника на български, руски и английски език. Христина Кукова познава общотеоретичните модели за интерпретация на явлениято многозначност, както и на лексикалната многозначност (полисемия и омонимия). Тя умело интерпретира, обобщава и творчески прилага в аналитичните части от дисертацията критерии за определяне на явлениято и класификации на видовете многозначност, предложени от Дж. Лейкъф, Дж. Катц и Дж. Фодор, Д. Крус, Дж. Пустейовски и др. За българския език е коментирано проучването на Н. Михайлова-Сталянова, в което се прилага теорията за компонентния анализ. Прегледът на мненията по въпроса позволява да се прецизира предметът на изследване във връзка с многозначността. Хр. Кукова отбелязва, че ще разглежда „три вида нееднозначност, а именно многозначност, омонимия (макар и да се среща по-рядко при глаголи) и метафорични употреби. Перцептивните предикати демонстрират висока продуктивност по отношение на семантичното обогатяване с помощта на езиковата метафора“ (стр. 42).

Докторантката познава теоретичните основи и практическия опит, свързан с автоматично отстраняване на семантичната многозначност. В основата на процеса е определянето на значението, с което е използвана думата в даден контекст. От речник или ресурс, представящ възможните значения, които се приписват на лексемата, с помощта на алгоритми трябва да се определи онова значение, което е актуализирано в конкретния случай. Във втора глава от дисертацията Хр. Кукова прави преглед на различните подходи за отстраняване на семантичната многозначност, като посочва, че те се различават по типа и количеството информация, която се обработва. „В по-ранните и основополагащи научни търсения в тази посока, подходите използват един или два типа информация. По-новите и надграждащи подходи най-често комбинират различни типове информация и статистически данни.“ (стр. 48) Авторката коментира резултатите, постигнати чрез три групи подходи: подходи, използващи синтагматична и статистическа информация (работещи обикновено върху корпусни данни); подходи, използващи концептуална и статистическа информация (отчитащи връзката между темата/сферата на употреба на конкретната дума и значението); подходи, използващи парадигматична, синтагматична и концептуална информация (работещи с речниковите дефиниции на думите, за да се открие семантичната близост на думи от дефинициите с думи от контекстуалното обкръжение). Вариант на последния е подходът, използващ Уърднет, при който се разширява моделът на субкатегоризационната информация. Чрез синтактичен анализ на повърхнинно ниво се поручава информация за предикатно-аргументната структура на изреченията, като се отчитат случаите/значенията, които не нарушават селективните ограничения на предиката.

Аналитичният преглед позволява на Хр. Кукова да дефинира централната за пета глава задача: да бъде предложено описание на предикатите за перцепция, основано на Фреймнет, което да подпомага процеса на разрешаване на лексикалната полисемия. Това

включва систематизиране на значенията, с които глаголите са представени в Уърднет, дефиниране на единиците и фреймовите им елементи, както и онагледяване с примери от Българския национален корпус. След очертаване на възможните семантични ограничения за фреймовите елементи и на синтактичните категории, с които се представят аргументите, може да се достигне до „най-подходящата комбинация от лингвистични данни, които да доведат до коректно автоматично разрешаване на многозначността при глаголите за възприятие в българския език“ (стр. 57).

Изследването на Христина Кукова е свързано с конкретна група предикати, което предполага познаване и на теоретичните принципи за лингвистична класификация на тези единици. семантична класификация на предикатите за различни езици. Авторката се позовава на класическите изследвания на З. Вендлер, У. Чейф, Дж. Лайънс, Б. Левин, Р. Ван Валин и Р. ЛаПола, а с оглед на славянските езици – Л. Шчерба, О. Селивърстова, Ю. Апресян, Е. Падучева. Направени са прегледи на конкретни проучвания, отнасящи се към семантичните класификации (деление на семантични класове) или към аспектиалните класификации на предикатите (действия, процеси, състояния или качества). За представяне на аргументите към анализирания предикати се използва основно типологичната система, разработена в рамките на Фреймнет. Тя отчита семантичните типове на комплементите (обекти, събития и пропозиции по класификацията на С. Шуле) и структурните и семантичните характеристики на фреймовите елементи, с които те се изразяват. В четвъртата глава на дисертацията е направен опит за типологизиране на комплементите към перцептивните предикати в българския език. Хр. Кукова дефинира фреймовите елементи, разглежда „сечението“ между перцептивните и менталните предикати и определя структурните и семантичните особености на комплементите при реализиране на различните значения. Така се извеждат съответствията и несъвпаденията в структурните и семантичните модели, които перцептивните и менталните предикати реализират. Тези изводи позволяват конкретизиране на подходите в пета глава при типологията на аргументите и решенията за снемане на семантичната многозначност.

Акцентът в прегледа на проучванията, свързани конкретно с перцептивните предикати, не е върху тематичната класификация по типа сетиво. Във връзка с темата на дисертацията Христина Кукова коментира основно изследвания, в които се анализира трансформацията от конкретно сетивното към когнитивното и перцептивните предикати имат епистемична, концептуална, индиректна или абстрактна интерпретация, която ги свързва с глаголите за ментална дейност. При неепистемичните значение глаголите назовават чистата перцепция, без наличието на когнитивен процес, за разлика от епистемичните значения. В двата случая предикатите се реализират и с различни комплементи.

4. Методика на изследването

В хода на научното изследване Христина Кукова е приложила комплексна методика, включваща общонаучни логически методи (анализ, синтез, обобщение, класификация), емпирични методи (проучване на източници, наблюдение, оценка, описание и сравнение, качествени и количествени анализи на съдържание). Методиката е адекватна на поставените научни задачи. Прилагането ѝ води до точно и непротиворечиво представяне

на моделите при реализиране на неепистемичните и епистемичните значения на перцептивните предикати и избора на аргументи при всеки конкретен тип значение.

Методиката на работа включват няколко стъпки, водещи от емпиричните данни към типологичните обобщения. В началния етап е ексцерпиран е езиков материал от Уърднет (синонимни множества, съдържащи перцептивни предикати), който е съпоставен и разширен чрез данни от Речника на българския език. Предложена е систематизация на типовете перцептивни предикати според четири диференциални признака: времева локализация; ориентация (субект експериенцер, обект стимул); активно или пасивно възприятие; реално или потенциално възприятие.

За определянето на случаите, в които се наблюдава преход от класа на перцептивните към този на менталните предикати авторката търси сечението на синонимните множества от изследвания клас със синонимни множества от други семантични класове според характеристиките в Уърднет (напр. за социални отношения и дейности (verb.social), когнитивни (verb.cognition), глаголи за създаване (verb.creation) и др.). Чрез анализ на синтактичните фреймове от Фреймнет и примери от Българския национален корпус са направени изводи за синтактичната реализация на глаголите в зависимост от конкретното значение (за перцепция или за когнитивна дейност). Формулирани са правила на изграждане на глаголните фрази в различните значения и субкатегоризация. След анализ на типовете аргументи, компленти и комплементизатори е направен семантичен и синтактичен анализ на фреймовите елементи (ядрени, периферни и екстратематични). Предложенията за подходи при автоматично снемане на семантичната многозначност с оглед на разглеждания клас предикати се базират на типологията на комплементите им.

5. Научните и научно-приложните приноси на дисертационния труд. Значимост на приносите за науката и практиката

Приносите в работата са ясно и точно представени. Те произтичат от проведения анализ на емпиричните данни. Изследователските хипотези, формулирани в началото на изследването, са потвърдени. Успешно са изпълнени научните цели на изследването.

В теоретичен план, разширени са знанията за семантичните особености и типологията на предикатите за перцепция в българския език. Съпоставителният анализ на синтактичната реализация на глаголите за перцепция, които биха могли да се употребят и със значение на ментални предикати, доказва връзката между синтактичната структура и семантичните характеристики на опората. Тези зависимости са систематизирани с цел подпомагане на автоматичното снемане на семантична многозначност при глаголите за възприятие. Промените в значението се наблюдават както по отношение на назоваването на други типове възприятия, така и за назоваване на когнитивни дейности или социалните отношения. Хр. Кукова отделя прототипни и непрототипни значения на разглежданите глаголи. Най-податливи на семантична деривация са спонтанните перцептивни глаголи, като на първо място този процес се наблюдава при глаголите за зрително и за слухово възприятие. Най-много когнитивни разширения могат да се припишат на глаголите за зрително възприятие. При анализа на прототипните употреби се очертават две групи глаголи: за визуални, слухови и тактилни възприятия – за обонятелни и вкусови възприятия. Комплементите към глаголи от първите три групи може да бъдат въведени с

че или да, докато при останалите два се срещат основно с че-изречения. Глаголите от типа *виждам, чувам, усещам* могат да бъдат употребявани с комплемементи с когнитивно съдържание, докато това е по-малко характерно за останалите два типа глаголи.

Непрототипните значения на глаголите за зрително възприятие се отнасят към когнитивни дейности, означавани с менталните предикати *разбирам, предвиждам, представям си, преценявам*, към социалната сфера, напр. *срещам се, посещавам* и др., към оценъчна група, напр. *откривам, грижа се* и отделни значения като *ставам свидетел и отнасям се*. Като непрототипни значения на глаголите за слухово възприятие са определени тези, които се представят чрез *обръщам внимание, подчинявам се, информиран съм и разбирам*. За тактилните предикати са характерни значенията *вземам си (консумирам), влияя, достигам, занимавам се с*. Към семантичните разширения на глаголите за обоняние се отнасят употребите, при които те реализират значения като *проследявам, предполагам, подозирам, разследвам*. Глаголите за вкусово възприятие могат да реализират и значения като *преживявам, наслаждавам се, предвиждам*.

Разработката на Хр. Кукова разширява знанията за семантико-синтактично особености на предикатите за перцепция и техните аргументите чрез изводи за типовете пояснения, които тези глаголи присъединяват, когато изразяват физическо възприятие или ментални процеси. На базата на типологията на предикатно-аргументните отношения във Фреймнет и направените анализи за групите предикати са изведени зависимости между семантичните типове и фреймовите елементи.

Представеният труд показва постиженията на Христина Кукова при решаването на научно-приложен проблем чрез методи, съответстващи на съвременните постижения на науката и практиката. Седемте приноса, посочени в дисертацията, могат да бъдат обобщени в две групи. Значим научно-приложен принос е представеното оригинално решение на актуален проблем, свързан с извеждане на семантичните особености при реализация на разглежданите глаголи. Дисертацията разширява полето на синтактичните анализи, тъй като проучва зависимости при реализация на многозначни думи. Принос в областта на синтактичните описания на българския език са систематизираните модели за представяне на семантичната структура на предикатите за перцепция и на синтактичните модели при реализиране на прототипни и на непрототипни значения. Тези модели са с практикоприложна стойност за компютърната лингвистика, при разработването на програми за автоматично снемане на семантична многозначност и автоматичен превод, а за преводаческата практика и за обучението по български език като роден или като чужд.

6. Преценка на публикациите по дисертационния труд и на автореферата

Христина Кукова е представила списък с осем публикации по темата на дисертацията. Една от тях е на английски, а останалите – на български. Две от статиите са в издания, реферирани и индексирани в световноизвестни бази данни с научна информация. Две статии са публикувани в сп. „Български език“, а три – в сборници с доклади от научни форуми. Количествената и качествената оценка на представените публикации показва, че те отговарят на минималните изисквания за придобиване на образователната и научна степен „доктор“.

Авторефератът на дисертацията точно и ясно представя структурата на труда, направените наблюдения и анализи и резултатите от тях, както и научните приноси на дисертационния труд.

7. Заключение

Текстът, представен от Христина Кукова, отговаря на формалните и съдържателните изисквания за дисертационен труд. Това е цялостно задълбочено изследване по актуален научен проблем, в резултат на което авторката достига до важни изводи в полето на синтаксиса и формалното описание на българския език, лексикалната семантика, компютърната обработка и моделиране.

Оценявайки актуалността на темата на дисертационния труд „Формално описание на перцептивните и менталните предикати в българския език (снемане на семантична многозначност)“, постигнатите резултати и направените изводи, предлагам на уважаваното Научно жури да присъди на Христина Кукова образователната и научна степен „доктор“ в областта на висшето образование 2. Хуманитарни науки, професионално направление 2.1. Филология, научна специалност „Общо и сравнително езикознание“.

15.05.2022 г.

проф. Йовка Тишева