

## СТАНОВИЩЕ

за научните трудове, представени по конкурса за заемане на академична длъжност  
„професор“

област „Хуманитарни науки“, професионално направление „Филология“,

(български език – морфология, синтаксис и корпусна лингвистика)

кандидат: Петя Начева Осенова, доцент в Катедрата по български език, Факултет по славянски филологии, СУ „Св. Климент Охридски“

Доц. д-р П. Осенова участва в конкурса със значителна по обем, съдържателност и многостранност научна продукция – 44 труда, от които една монография и 43 статии, като 10 от статиите са самостоятелни, а 33 – в съавторство (като приносят на кандидатката във всяка една от тях е ясно разграничен). Представените трудове са показателни за извънредно плодотворната научна дейност на П. Осенова, очертават облика ѝ на сериозен, задълбочен и талантлив изследовател с голям капацитет и разнообразни научни интереси, който сполучливо съчетава разработването на проблеми от областта на морфологията, синтаксиса и формалната лингвистика със създаването на езикови ресурси и езикови технологии за българския език. Участието на кандидатката в значителен брой съвместни публикации е показателно за нейното умение да работи в екип, което е особено важно за разработките в областта на корпусната и компютърната лингвистика, които по необходимост имат интердисциплинарен характер и изискват участие на по-големи колективи от специалисти. Свидетелство за научното равнище на представените публикации е високата им оценка от международната научна общност, за което говори фактът, че значителна част от тях са публикувани в най-авторитетни международни издания и сборници от конференции в областта на компютърната и корпусната лингвистика.

П. Осенова съчетава успешно научноизследователската и преподавателската си дейност, повече от 17 години е преподавател в СУ „Св. Климент Охридски“, като основните дисциплини, на които са посветени лекционният ѝ курсове, са морфология, синтаксис, съвременен български език и формални граматика.

Публикациите, с които П. Осенова участва в конкурса, са впечатляващи както с разнообразната си тематика, така и с подчертаната си актуалност. Основна особеност на нейното научно творчество е тясното преплитане на теоретични и научноприложни проблеми от областта на морфологията и синтаксиса – от една страна, и корпусната лингвистика, от друга. Значителна част от публикациите на П. Осенова са посветени на проблеми, възникнали при разработването на различни езикови ресурси и езикови технологии за българския език, в което кандидатката активно участва.

Свидетелство за високата квалификация на П. Осенова и за успешното ѝ интегриране в международната научна общност е нейното активно участие в

международни проекти (в периода от 2009 досега – 6 на брой) и инициативи, насочени към прилагане на езикови технологии в различни сфери (електронно обучение, машинен превод и др.) и към разработване на ресурси за синтактично анотирани корпуси. П. Осенова има участие и в национални проекти, като е ръководител на проекта „Развитие и функциониране на научноизследователски център за анализ на политическата и журналистическата реч“ (2009-2013 г.), финансиран от МОН, резултат от който е създаденият корпус на политическата и журналистическата реч.

Основният труд, с който П. Осенова участва в конкурса, е монографията *Граматическо моделиране на българския език (с оглед на обработката на естествен език)*. В труда са представени два подхода за граматично моделиране на морфосинтактичните явления в българския език: чрез разработване на ресурсна граматика и чрез трансформация на синтактичната анотация на корпуса Бултрибанк (първоначално анотиран въз основа на Опорната фразова граматика) в друга формална теория.

Един от най-сериозните приноси на авторката в този труд (част от проблематиката на който е разгледана и в публикации № 8, 27 и 32) е разработването на ресурсна формална граматика за българския език въз основа на Опорната фразова граматика (ОФГ). Теоретичен принос представлява прилагането за първи път върху български езиков материал на тази теория така системно и последователно. Както отбелязва П. Осенова, задачата за моделирането на морфосинтактичните явления е сложна и трудоемка по принцип поради ред причини. Значителна сложност при граматичното моделиране поражда и необходимостта от представяне в граматичната матрица на ОФГ на специфични за българския език явления. Авторката се е справила много успешно с тази трудна задача. При кодирането на специфични за българския език морфосинтактични явления, които липсват в първоначалната граматична матрица, тя прилага творчески решения, което обуславя високото равнище на създадения от нея ресурс. Постигане на авторката е моделирането на специфични или носещи специфични особености единици като клитичните местоимения, удвоеното допълнение, местоименията, глаголите. Оригинален принос на П. Осенова е адаптираните на известните от традиционната ни граматика особености на частите на речта към модела на моделиране, използван в ОФГ – например конструкциите с *щях* и *да-форми* са моделирани като конституционални – съставени от глагол *щях* и да-форма с функция на комплемент (с. 75).

Създадената от авторката ресурсна граматика е много полезен езиков ресурс, който дава възможност за обективна съпоставка на българския език с други езици и въз основа на това – за разкриване по обективен начин както на универсалните явления в него, така и на специфичните му характеристики. Трябва да се подчертае и това, че чрез кодирането на специфични за българския език явления и особености в граматичната матрица създадената от П. Осенова ресурсна граматика допринася и за разширяване и обогатяване на самата матрица.

Много ценни са разработените с участието на П. Осенова два нови модела на синтактично описание на българския език, които са резултат от трансформацията на анотационната схема на Бултрибанк (изградена първоначално върху конституентна теория) в анотационната схема на депendentната теория (в нейните разновидности *Граматика на зависимостите* и *Универсален модел на Граматиката на зависимостите*). Като важен принос трябва да се отбележи моделирането на синтактичните явления в езика ни в рамките на депendentната теория, което се извършва за първи път по отношение на българския език. В резултат на трансформацията на лингвистичното знание в Бултрибанк (проведена в рамките на обща международна инициатива, в която са включени и редица други езици) се създава един много полезен синтактично анотиран ресурс за българския език, който осигурява по-универсална рамка на моделиране на езика ни и улеснява съпоставката му с други езици.

Извършената трансформация на лингвистичното знание от теоретичната рамка на една формална теория в друга е сложен процес, основан на определени правила при прехвърляне на кодировките на частите на речта и на пренос на релации от едната схема на анотация в другата. Участието на П. Осенова в тези дейности е значително. Проблемите при прехвърлянето на синтактичната анотация от единия теоретичен модел в другите два са разгледани подробно в монографията (също и в публикации № 5, 42), като особено детайлно са анализирани различията в начина на представяне на частите на речта при различните модели.

Научните разработки на П. Осенова съдържат и многобройни други важни приноси, които е невъзможно да бъдат обхванати при ограничения обем на становището, поради което тук ще отбележим обобщено и в съкратен вид само най-съществените от тях.

1. С приносен характер е разработеният от П. Осенова унифициран начин за представяне в лексикона и анализиране в текстовете в Бултрибанк на различни типове многокомпонентни конструкции (идиоми, устойчиви съчетания и др.). За тази цел е използван определен формален инструмент – катена, чрез който те се кодират в лексикона като едно цяло с определена семантика. Предложеният начин за кодиране на тези единици дава възможност за най-адекватно отразяване на особеностите в тяхната семантика и в синтактичното им поведение (свързани с възможни разкъсвания от други елементи и др.) (срв. № 2, 6, 9, 11). Набелязан е и начин за представяне чрез катена на връзката между сложни думи и съответстващи им синтактични словосъчетания пак с оглед на моделирането на тази връзка в лексикона (срв. № 11).

2. Особена значимост имат изследванията на П. Осенова, насочени към проучване на определени езикови явления чрез използване на корпуси и на интегрираните в тях средства. Тези изследвания илюстрират възможностите на новите подходи за получаване на обективни данни за езиковите явления. Чрез прилагането на тези подходи П. Осенова достига до важни наблюдения за честотността на различните

валентни типове глаголи в българския (срв. № 20) и за разпространението в българските диалекти на редица явления (напр. ятовия преглас и др.) (срв. № 44).

Особено интересни и перспективни са изследванията на П. Осенова на някои особености на публичната политическа реч въз основа на данни от създадения с нейно участие специализиран корпус с политическа реч (включващ два подкорпуса: с парламентарна реч и с журналистически интервюта). Кандидатката достига до важни изводи относно използването в речта на политиците на модалните глаголи *мога*, *искам* и *трябва*, като установява, че най-често срещаният от тях е *трябва* (т.е. изразяването на необходимост). Направени изводи са получени на обективна основа – чрез съпоставка на фреквентността и на контекстите на употреба на модалните глаголи в политическия корпус и в общ медиен корпус (срв. № 14, 18). П. Осенова разглежда и друга перспективна възможност за анализ на речта на политиците чрез корпусни подходи и средства, свързана с установяване на тяхното отношение и мнение по определена тема (срв. № 37).

3. Много сериозен е приносът на П. Осенова в разработването на значителен брой езикови ресурси и езикови технологии за български език. Участието ѝ в тези дейности е значително и се основава на високото равнище на нейната лингвистична компетентност и теоретична подготовка.

П. Осенова участва в създаване на цялостен модел за автоматична обработка на българския език, който включва сегментация, разделяне на изречения, морфологичен анализ, лематизация и синтактичен анализ (срв. № 19). Значителен е нейният принос за усъвършенстване на автоматичните средства за морфосинтактичен анализ на текстове в корпус и за автоматично снемане на морфологична многозначност. Извършеното от П. Осенова описание на типовете многозначности в българския и създадените правила за снемане на многозначността (срв. № 19 и 26) са теоретична основа за автоматично снемане на многозначността, без която този тип обработка на текстове в корпуса не би могла да се осъществи ефективно. Проучването на типовете морфологична многозначност представлява и сериозен теоретичен принос по отношение на изучаването на това явление в българския език.

Много сериозно е участието на П. Осенова и в семантичното аотиране на Бултрибанк със значения от лексикалната база WordNet и от DBPedia (срв. № 10). Важна стъпка в обогатяване на този корпус със семантична информация представлява и създаденият с участието на П. Осенова онтологично базиран речник за българския език, който съдържа информация за отношенията между лексикалните единици в речника и понятията в онтологията (срв. № 40).

П. Осенова участва и в разработването на многобройни други полезни и важни езикови ресурси и средства за обработка на българския език. Тя взема участие в създаването на автоматичен модул за разпознаване и категоризиране на собствени имена в корпуса (срв. № 43); в създаване на уеб базиран морфосинтактичен модул за българския (срв. № 31); в прилагането на конституентен парсер върху Бултрибанк –

подход, който се прилага за първи път по отношение на български език (срв. № 12); в създаването на българско-английски синтактичен паралелен корпус (като участва в разработването на неговата анотационна схема и в правилата за подравняване – срв. № 23, 34); в създаването на статистически модел за машинен превод (срв. № 3, 17 и 22) и др.

С важно значение са и публикациите на П. Осенова, насочени към обучението по български език и в частност – към преподаването на знания от областта на синтаксиса (срв. № 15, 28). В тези разработки се съдържат редица ценни идеи, насочени към повишаване равнището на обучението, включително и чрез използване на корпуси. Много полезна е идеята за по-оптимално отношение между изучаването на нормативните норми и езиковата практика при чуждоезиковото обучение по български, при което се залага на поставянето на обучавания в реална езикова среда.

Трудовете на кандидатката имат внушителен брой цитирания – в базата *Web of Science (Thomson Reuters)* са отбелязани 183 цитирания на 113 нейни публикации (в периода 1985-2016 г.), а в Springer – 109 цитата. П. Осенова е посочила 69 цитирания на свои публикации за периода от 2009 г. досега въз основа на данни от личния си архив. Показателен за международното признание на научноизследователската ѝ дейност е фактът, че всички тези цитирания се съдържат в публикации в международни издания.

Заклучение. Направеният обзор на научноизследователската дейност на доц. д-р П. Осенова я очертава като много сериозен и задълбочен изследовател с извънредно широки интереси, който съчетава плодотворно изследвания на високо теоретично равнище с впечатляващи по обхвата си приложни разработки, насочени към създаване и усъвършенстване на важни езикови ресурси за българския език и средства за тяхната обработка. С оглед на това с пълна убеденост препоръчвам на почитаемото научно жури да присъди академичната длъжност професор на доц. д-р Петя Начева Осенова.

18.05.2016

проф. д-р С. Колковска

(Институт за български език – БАН)