

РЕЦЕНЗИЯ

по конкурс за професор
по научната специалност 01.01.12 “Информатика“
с кандидат доц. д-р Людмила Петрова Димитрова-Рашкова

Рецензент: доц. д-р Данаил Антонов Дочев,
Институт по информационни и комуникационни технологии
- БАН

За участие в конкурса за професор по специалност 01.01.12 “Информатика (компютърна лингвистика – средства и системи за обработка на лингвистични знания)“, обявен от Института по математика и информатика – БАН /ИМИ – БАН/ за нуждите на секция „Компютърна лингвистика“ (ДВ брой 58/29.07.2011 г.) е подал документи единствен кандидат - доц. д-р Людмила Петрова Димитрова-Рашкова, доцент в ИМИ – БАН. Людмила Димитрова-Рашкова е завършила висше образование в Софийския университет “Св. Климент Охридски” - Математически факултет през 1969 г. като магистър по математика. отличен успех по специалност "Математика - производствен профил". В периода 1969-1973 г. е работила като програмист и научен сътрудник. В периода 1973 – 1977 г. е редовен докторант в Московския държавен университет “М.В.Ломоносов” – Факултет по изчислителна математика и кибернетика и през 1977 г. успешно защитава дисертация за научната и образователна степен "доктор". От 1978 г. работи като научен сътрудник II/I ст., а от 1996 е доцент в секция „Компютърна лингвистика” на ИМИ – БАН.

Людмила Димитрова-Рашкова има общо 64 публикации, а за конкурса е представила 29 публикации. От тях 2 са на български език и 27 - на английски език. 3 публикации са самостоятелни, 8 - в съавторство с един автор, 9 - в съавторство с двама автори, 5 - в съавторство с трима автори и 4 - в съавторство с петима автори. 23 от трудовете са публикувани след 1996 г., т.е. след получаването на научната и образователна степен “доктор“.

Представените трудове са отпечатани в следните видове издания:

1 – монография,

4 - в български издания,

10 – в международни списания,

1 - в трудове на национална конференция,

13 - в трудове на международни конференции.

Тази класификация на изданията се отличава малко от представената от кандидата класификация (/5/ и /6/ са публикации в български издания - серия годишници на Югозападния университет – Благоевград „Дискретна математика и приложения”, а не в трудове на международни конференции). Всички трудове напълно съответствуват на научната специалност, по която е обявен конкурса.

През цялата си кариера на изследовател Людмила Димитрова-Рашкова работи в областта Компютърна лингвистика – изследване, създаване и експериментиране със средства за компютърна обработка на естествено-езикови текстове. В началото на изследователската си дейност тя е работила по теми на Математическата лингвистика – изследване и моделиране на формални граматически схеми за представяне на естественоезикови структури. Дългосрочните изследователски интереси на Людмила Димитрова-Рашкова са в областта на методи и средства за компютърно представяне и обработка на лингвистични знания и особено в тяхното приложение при създаване на компютърни езикови ресурси – анотирани корпуси от български и многоезикови текстове, лексически бази от данни, двуезични компютърни речници. Напоследък Л. Димитрова-Рашкова работи по моделиране на електронна изследователска инфраструктура за езикови ресурси на славянски езици. При мнозинството изследователски резултати по създаване на компютърни езикови ресурси и изследователска Е-инфраструктура за тях е явна плодотворната връзка с активното и последователно участие на Л. Димитрова-Рашкова в разработката на международни проекти.

Първото направление на представените за конкурса изследвания е свързано с анализ и експерименти с модели от теорията на формалните граматика за представяне на граматически конструкции от българския език - трудове /4/, /5/ и /6/. Интересът е съсредоточен върху формализма на логическите граматика, успешно използвани за описване на свободен словоред. Изследвани са и съот-

ветствия между определени типове формални езици и дискретната структура „магазинен автомат”.

Второто направление на представените за конкурса изследвания е посветено на изследователски и реализационни въпроси на компютърната лингвистика. Трудове /10/, /12/ и /13/ разглеждат въпроси на представянето и обработката на лингвистични знания при създаване на многоезични компютърни ресурси (лексикони, описание на общи и специфични морфосинтактични характеристики, анотирани подравнени корпуси). Описани са морфосинтактични характеристики на частите на речта в българския език и начин за автоматично снемане на морфосинтактична многозначност. Резултатите от горните изследвания определят единна методологична и стандартизационна рамка (/7/) за следващите конкретни разработки на цифрови многоезикови корпуси: българо-полски корпус (/14/, /16/), българо-словашки подравнен корпус (/27/), българо-полски-литовски паралелен корпус (/21/, /22/, /24/).

Третото направление на изследванията на Л. Димитрова-Рашкова засяга разработване на средства и системи за обработка и управление на езикови ресурси. Трудове /1/-/3/ представят първи български разработки по автоматична обработка на български език – алгоритъм и система за автоматично сегментиране на български словоформи, както и програмна система за автоматичен анализ на български текст при автоматично съставяне на речници и речникови архиви. Трудове /8/, /17/ и /20/ разглеждат модели и реализации на лексически бази данни, като са определени лексикографски спецификации за българския език съгласно международните стандарти за електронни речници. В /9/ са обсъдени въпроси на хармонизацията на класификаторите в цифровите многоезични речници. Изследванията по лексически бази данни са материализирани в разработката на двуезичен полско-български компютърен речник, разработен като Интернет-приложение (/18/, /19/, /29/). В /15/ са разгледани специфики на представянето на български глаголи в речника.

Трудове /23/, /26/ и /28/ разглеждат структури и стандарти за изследователски електронни инфраструктури, поддържащи цифрови компютърни ресурси: корпуси, речници, бази от данни, лексикони, тезауруси, онтологии.

Людмила Димитрова-Рашкова е оформен и активен изследовател в областта на информатиката. Проявява добра научна ерудиция и широки интере-

си. Успешно се занимава както с теоретични задачи - създаване и изследване на информационни модели, така и със задачи с приложен характер - разработка на практически значими цифрови ресурси и програмни системи. Със своите интереси и разработки тя се утвърди като водещ изследовател в модерното и актуално направление “езиково инженерство“, обединяващо прилагането на съвременните информационни технологии към разнообразните задачи за обработка на естествено-езикови текстове с практическа значимост. Л. Димитрова–Рашкова развива и активна учебно-педагогическа дейност в 3 български университета. Особено активно е участието ѝ в магистърската програма “Езикови и мултимедийни технологии” на ВТУ “Св. Св. Кирил и Методий” – 8 курса, от които 2 разработени за първи път в България.

Основните научни и научно-приложни приноси на Л. Димитрова – Рашкова в рецензираните трудове са в следните направления:

- Прилагане на логически граматика за описание на свободния словоред в българския език, както и на синтактичните взаимодействия на глаголни аспекти, трудно формализирани при българския език /5/, /6/. Това са първите изследвания за български език в тази област, материализирани и в програмна система /4/. Резултатите от тази група са постигнати преди избора за доцент.

- Научните и научно-приложни приноси в цикъла от изследвания по компютърна лингвистика, отразяващи директно и изследвания по представяне на лингвистични знания, са свързани със следните резултати:

- Разработен е български лексикон с практически значим обем и специфични езикови ресурси за сегментиране на аотирани текстове, отговарящи на международни стандарти за описване на морфосинтактични характеристики /10/. В този контекст са представени морфосинтактични спецификации на български глаголни форми /12/ и са съпоставени български и словашки морфосинтактични спецификации /13/.

- В рамките на международен проект е разработен модел на многоезичен подравнен аотирани ресурс /7/.

- На основата на горния модел са реализирани първите български цифрови аотирани корпуси (/11/, /25/), както и също за първи път българо-полски (/14/, /16/), българо-словашки /27/ и българско-полско-литовски (/21/, /22/, /24/) паралелни корпуси със значителен обем. Многоезичните цифрови корпуси на-

мират многообразни приложения: изследвания по съпоставителна лингвистика, автоматично класифициране и резюмиране на текстове, машинен превод, автоматизирано семантично аотиране на обекти в цифрови библиотеки и т.н.

- Приносите в областта на средствата за обработка и управление на езикови ресурси се състоят в:

- Разработен е алгоритъм и система за автоматично сегментиране на български словоформи без използване на речник от лексеми /1/, /2/. Тези резултати са постигнати преди избора за доцент.

- Разработени са лексикографски спецификации за българския език съгласно международните стандарти за кодиране на електронни речници /8/.

- Предложен е модел на двуезична лексическа база данни, определящ съвкупност от структурни маркери и маркери за съдържание, съвместима с TEI /17/.

- Предложени са нови класификатори за български глаголи в контекста на създаване на двуезични компютърни речници /15/, /18/.

- Разработени са полско-български и българо-полски компютърни речници, реализирани като Интернет-приложение /19/, /20/, /29/.

- В рамките на международен проект в сътрудничество с чуждестранни изследователи е предложен концептуален модел на електронна изследователска инфраструктура за езикови ресурси на славянски езици, представен в монографията /26/. Моделът поддържа съвременните езикови технологии и всички познати типове цифрови езикови ресурси, и се опира върху Грид-технологията като технологична платформа.

Основната част от приносите могат да бъдат квалифицирани като обогатяване на съществуващите знания.

В публикациите със съавтори Л. Димитрова-Рашкова е на първо място в 24 публикации, на второ място - в 2 публикации и на трето място – в 1 публикация. Това показва, че нейните приноси в съвместните работи са съществени.

Към документите на кандидатката е приложен списък с цитирания, съдържащ общо 89 открити цитирания на 14 от публикациите на Л. Димитрова-Рашкова. Освен това са открити 13 цитирания на електронни материали с нейно съавторство. 5 цитирания са на български автори, а останалите - на чуждестранни автори. Прави впечатление, че трудовете са цитирани в

престижни издания: Springer LNCS - 5, Prentice Hall - 2, Kluwer, Routledge, John Benjamins и др., в трудове на престижни международни конференции, както и в 5 чуждестранни дисертации.

Към представените трудове могат да се отправят следните критични бележки:

- понякога преобладава описателната част на резултатите за сметка на по-подробно мотивиране на приетите решения;

- има частично припокриване на съдържанието в трудовете, описващи българо-полско-литовския корпус; напр. статия /22/ е доста близък вариант на статия /21/ с малко повече примери.

По препоръчителните изисквания (Чл. 2., т. 6 от „Правилника на ИМИ-БАН за приложение на ЗРАСРБ“) към кандидати за академичната длъжност „професор“ в областта на информатиката се установява следното:

- представени са 1 монография и 28 други публикации (при минимално изискване за 15 други публикации);

- от представените публикации 23 (повече от 5) не са представяни за заемане на академичната длъжност „доцент“;

- от представените публикации 23 (повече от 10) са в списания или в сборници от международни конференции;

- приведени са общо 102 цитирания без автоцитати (при минимално изискване за 20 цитирания);

- кандидатката няма поне двама защитили докторанти;

- представен е списък от 18 труда, поместени в престижни издания, които се позовават на идеи и/или съществено използват резултати на кандидатката;

- представени са коментари за научната дейност на Л. Димитрова –Рашкова от: Института по лингвистика на Словашката АН за успешна съвместна работа по европейски проект; от Института по славистика на Полската АН - две покани за участие в редакционната колегия на международното списание *Cognitive Studies/Études Cognitives* за 2010 и 2011-2014 г.

По съществените незадължителни изисквания към кандидатите (чл.3 от „Правилника на ИМИ-БАН за приложение на ЗРАСРБ“):

- кандидатката е координатор, водещ участник и участник в общо 20 международни и национални научноизследователски проекти;

- през последните 5 години е участник в 7 програмни комитети и е изнесла 25 доклада на международни научни мероприятия;
- участник е в редколегията на международно списание и е редактор на 5 международни научни издания;
- членува в авторитетните международни професионални организации Association for Computational Linguistics и European Woman in Mathematics;
- участвала е в реализацията на изследванията си при разработката на 7 приложни информационни продукти;
- води аудиторни занятия в СУ “Св. Кл. Охридски“ (4 курса във ФМИ и ФКНФ), НБУ (1 курс), ВТУ “Св. Св. Кирил и Методий”(8 курса);
- разработила е за пръв път в България 3 университетски курса;
- ръководи 1 докторант и 10 успешно защитили дипломанти.

Личните ми впечатления от работата на Людмила Димитрова-Рашкова в Института по математика и информатика – БАН напълно потвърждават създаващото се от нейните трудове и многообразна дейност мнение за енергичен и утвърден изследовател в областта на информатиката с безспорен авторитет сред специалистите по компютърна лингвистика. От дълго време тя работи успешно в съвместни колективи с чуждестранни учени и поддържа широки международни контакти.

ЗАКЛЮЧЕНИЕ

Въз основа на анализа на научните постижения и изявите на кандидатката смятам, че тя напълно отговаря на изискванията за професор на Закона за развитието на академичния състав в Република България и Правилника за приложението му в ИМИ – БАН и убедено предлагам на Уважаваното научно жури да препоръча на Научния съвет на Института по математика и информатика – БАН да избере Людмила Петрова Димитрова-Рашкова за професор по информатика.

София, 31 октомври 2011г.

РЕЦЕНЗЕНТ: