

RNDr. Eduard Bejček
Automatické propojování lexikografických zdrojů a korpusových dat
posudek vedoucí disertační práce

Předkládaná disertační práce se zabývá vzájemným propojováním a obohacováním slovníkových a korpusových dat. Zaměřuje se především na dva velice aktuální úkoly: na vzájemné propojení valenčních slovníků a na propojení slovníku víceslovných jednotek s textem.

Po úvodních kapitolách autor představuje jazykové zdroje, se kterými pracuje, podrobně přibližuje i jejich formáty. Za přínosnou považuje jeho poznámku k tzv. univerzálním formátům: výstižně konstatuje, že pro provázání slovníků není často diskutovaná problematika formátu klíčová – mnohem důležitější je podobná granularita a podobnost informací, které lze k porovnání využít.

Klíčové jsou kapitoly 5 a 6. První z nich je věnována vzájemnému propojení lexikografických dat – jejím přínosem je podrobná analýza problematiky a návrh a implementace automatického systému pro propojování valenčních slovníků. Jejím praktickým výstupem potvrzujícím možnosti a úspěšnost zvolených metod (i jejich omezení) jsou pak propojené valenční slovníky VALLEX a PDT-Vallex vedoucí k doplnění desítek tisíc anotovaných vět z treebanků PDT a PCEDT do dat slovníku VALLEX.

Kapitola 6 potom představuje metody, implementaci a vyhodnocení automatického systému. Zajímavá je především diskuse o podobě tektogramatické reprezentace víceslovných jednotek a vlastnostech takové reprezentace.

Disertační práce svým způsobem dobře charakterizuje svého autora – představuje dvě základní oblasti, kterým se v době doktorského studia věnoval, a to projekt budování valenčního slovníku VALLEX a projekt Lexemann zastřešující anotaci víceslovných výrazů v PDT. Tato témata jsou spjata též se dvěma granty GAUK, jejichž byl hlavním řešitelem – s grantem *Obohacování a rozvoj lexikografických zdrojů* a s grantem s titulem *Anotace víceslovných lexémů a pojmenovaných entit v PDT*.

Práce dokládá pečlivost autora, jeho snahu do hloubky pochopit lingvistické aspekty úkolů a adekvátně je modelovat (i zde možná pramení tendence volit pravidlové metody?); precizně a podrobně vysvětluje zvolená řešení a je nanejvýš poctivý při hodnocení jejich přiměřenosti i další využitelnosti.

Práce dále dokládá, že se Eduard Bejček úspěšně zapojil do řady projektů školícího pracoviště a stal se tak důležitým členem v řadě projektových týmů; mimo již výše zmíněných projektů VALLEX a Lexemann je nutno zmínit zejména projekt Pražského závislostního korpusu. Tomu odpovídá i bohatá publikační činnost, typicky ve větších autorských týmech.

Shrnutí: Průběh studia RNDr. Eduarda Bejčka, jeho zapojení do výzkumných týmů, publikační aktivita i předkládaná disertace svědčí o jeho způsobilosti práci obhájit a doktorské studium úspěšně ukončit.

V Praze 9.9.2015

doc. RNDr. Markéta Lopatková, Ph.D.
školitelka