

Posudek vedoucí diplomové práce

Ondřej Kučera: Pražský závislostní korpus jako cvičebnice jazyka českého

Předložená diplomová práce je prací aplikačně-implementačního charakteru. Cílem práce bylo sestavit elektronickou cvičebnici češtiny s větami vybranými z Pražského závislostního korpusu. Cvičení měla pokrývat procvičování jak tvarosloví, tak i větné stavby spolu s určováním větných členů.

Diplomová práce je velice přehledně členěna do osmi kapitol, do uživatelské a programátorské příručky a je doprovázena CD. První kapitola je kapitolou motivační. Druhá kapitola je charakteru rešeršního – autor podává přehled několika systémů elektronických cvičebnic a výukových programů češtiny, které jsou k dispozici na trhu. Autor zdůrazňuje, že ani jeden z uvedených systémů nenabízí cvičení pro komplexní zpracování (tvarosloví a větný rozbor) českých vět. Tento fakt je ovlivněn skutečností, že systémy vznikaly „ručně“, tedy že autoři nejdříve shromáždili věty z textů, s kterými se denně potkávali a následně tyto věty zpracovali dle hledisek, která si sami stanovili. Originalita přístupu, kterou předložená dp zpracovává, spočívá ve využití materiálu, který již vybraná hlediska obsahuje. Tímto materiálem je anotovaný korpus češtiny, Pražský závislostní korpus (PZK). Sestavení cvičebnice pak spočívá v převedení akademických anotací do zápisů srozumitelných žákům základních a středních škol, jak je konstatováno v kap. 3. Popisu logické a fyzické struktury PZK je věnována čtvrtá kapitola.

Z pohledu lingvisticko-aplikačního jsou stěžejní kap. 5 a 6, z pohledu implementačního kap. 7. V rámci převedení akademických anotací se vyberou takové věty, zejména co do délky a pokrytí jevů, na které jsou žáci „zvyklí“ – proběhne fáze tzv. *filtrování* vět. Následně proběhne *transformace* syntaktických stromů, tedy převedení akademických rozborů do rozborů školských. Jednotlivá transformační pravidla jsou formulována systematicky v souladu s rozbohem jevů v anotačním manuálu PZK. Nemohou být zcela přímočará vzhledem k tomu, že v některých jevech se akademický přístup od rámcových osnov odlišuje. Pro implementaci systému bylo zvoleno prostředí Java. Systém cvičebnice tvoří dva uživatelské programy, *Charon* a *Styx - Charon* je administračním nástrojem pro výběr vět pro procvičování, *Styx* je samotnou cvičebnicí. Cvičebnice svým objemem (12 tis. vět) převyšuje všechny ostatní, nejen české cvičebnice. Oba programy se podařilo dle instrukcí v uživatelské příručce nainstalovat bez problémů.

V kap. 8 se autor zamýšlí nad dalšími kroky, které se neobejdou bez zpětné odezvy od těch, kterým je cvičebnice určena, od učitelů a žáků.

Předloženou dp hodnotím vysoce kladně. Cíl byl bez výhrad splněn. Autor postupoval systematicky, musel během řešení zvládnout lingvistickou teorii syntakticko-analytických rozborů zachycených v PZK, osvěžit si školské větné rozборы a implementovat systém, který bude zájmat (doufejme) nejen premianty tříd.

V Praze 31. ledna 2006

Barbora Vidová Hladká
ÚFAL MFF UK