

Zápis z obhajoby disertační práce Mgr. Anny Nedoluzhko

konané dne 9. 9. 2010

Téma práce: Rozšířená textová koreference a asociační anafora
(Koncepce anotace českých dat v Pražském závislostním korpusu)

Přítomni: Prof. PhDr. Petr Mareš, CSc. (předseda)
Prof. PhDr. Oldřich Uličný, DrSc. (školitel)
Prof. PhDr. Jarmila Panevová, DrSc.
PhDr. Paed.Dr. Miloslav Vondráček, Ph.D.
Doc. RNDr. Vladimír Petkevič, CSc.
Prof. PhDr. Eva Hajičová, DrSc. (oponentka)

Omluveni: Doc. RNDr. Karel Oliva (oponent)

1. Předseda komise prof. Mareš zahájil obhajobu a představil přítomným kandidátku.
2. Školitel prof. Uličný představil kol. Nedoluzhko a seznámil komisi s uchazečkou a její disertací.
3. Kandidátka představila přítomným svou disertační práci; šlo zejména o následující myšlenky (podrobněji viz příložená prezentace č. 1):
 - a. Textová koreference a anafora dosud nebyly na češtině v celku zpracovány.
 - b. Systematizace a klasifikace dat z Pražského závislostního korpusu – praktický cíl: příprava anotace; zároveň kladen důraz na teoretické vysvětlení jednotlivých praktických rozhodnutí.
 - c. Teoretický základ:
 - i. Ruské a české teorie reference; práce německých slavistů (založeny na korpusech)
 - ii. Projekty světové počítačové lingvistiky
 - d. Schéma anotace – představena navržená klasifikace a obecné zásady anotace
 - e. Teoretické řešení vybraných problematických případů; pokud je vyloučeno jednoznačné teoretické řešení, je stanovena jednotná konvence (např. při rozlišování jednotlivých typů v klasifikaci či u nejednoznačných antecedentů)
 - f. Představení aplikace navržené klasifikace a evaluace této anotace.
4. Poté oponentka prof. Hajičová přednesla závěr svého posudku.
Předseda prof. Mareš přečetl posudek druhého oponenta, doc. Olivy (omluven).
5. Odpověď Mgr. Nedoluzhko na připomínky posudků – výběrově (podrobně viz příložená prezentace č. 2):
 - a. Rozlišení termínů *anafor* a *anafora* je provedeno důsledně.
 - b. Termín *pokleslý referent* – jde o terminologii E. V. Padučevové.
 - c. Jazyková stránka – práce prošla věcnými a jazykovými korekturami, zjevně byly nedostatečné, před tiskem musí práce znovu projít další kontrolou.
 - d. Termín *odděleně reprezentovaný příznak negace* byl přejat z Manuálu Mikulové et al. 2005, stejně tak termín *kontejner*.
 - e. Překlady z ruštiny – vysvětlení přesného ruského a českého významu, obhájení původních překladů.

- f. Atribut sempos – adj. denot. – bylo třeba ukázat, u kterých adjektiv byla a nebyla provedena anotace koreference a jak jsou značena v Pražském závislostním korpusu.
- g. Práce Hajičová – Oliva – Sgall se týká jiného tématu (gramatická koreference, pronominální textová koreference; v práci kol. Nedoluzhko jde o další typy textové koreference).
- h. K výtce nevědeckosti práce: Neexistuje souhrnná studie o koreferenci v češtině, její vytvoření je příliš rozsáhlý úkol. Tato práce se zabývá prvními kroky tohoto rozsáhlého výzkumu (sběr, analýza, klasifikace materiálu). Další kroky, běžné ve vědeckém postupu, jsou obsaženy porůznu v jednotlivých kapitolách.
Práce je uspořádána podle vědeckého postupu běžného v počítačové lingvistice (obsahuje např. evaluaci).
- i. K výtce, že práce není o češtině: byla vytvořena na základě českých dat, na základě veškeré dostupné bohemistické literatury; jiné než cizí korpusy dosud nebyly k dispozici.
6. Zhodnocení vystoupení Mgr. Nedoluzhko ze strany oponentky prof. Hajičové:
- Prezentace k námitkám je vynikající, dotazy byly v úplnosti zodpovězeny.
 - Vyjádřila lítost nad nepřítomností doc. Olivy.
7. Diskuse
- Prof. Uličný: procedurálně – posudek doc. Olivy se vzdává výsledného hodnocení, obsahově je negativní. Jak k němu přistupovat?
Komise: komise se shodla, že posudek přijímá jako neurčitý co do závěru.
 - Doc. Petkevič: jak je možné, že práce obsahuje tolik jazykových nedostatků, co na to školitel?
Prof. Uličný: četl předběžnou verzi.
Mgr. Nedoluzhko: ukázalo se, že je nutná rozsáhlejší korektura.
Prof. Hajičová, prof. Uličný: až půjde práce do tisku, řádně korigovat.
Prof. Hajičová: nebýt u Mgr. Nedoluzhko přísnější než u anglicky psaných prací českých studentů, nejde o takový problém. Porozumění při čtení není narušeno.
Prof. Uličný: i česky psané práce od českých vědců mívají řadu jazykových nedostatků.
 - Prof. Panevová: co vedlo doc. Olivu k závěru, že práce není o češtině? (Otázka pro nepřítomného doc. Olivu)
Mgr. Nedoluzhko: poznámky doc. Olivy se věnují hlavně první třetině práce, která obsahuje přehled zahraniční literatury.
 - Prof. Panevová: lze na základě této práce najít oblast koreference, kterou by bylo možné zpracovávat automaticky?
Mgr. Nedoluzhko: jde o oblast pojmenovaných entit.
 - Dr. Cinková: existují syntaktická kritéria pro rozlišení specifické a nespécifické/generické reference? Existují jazyky s formálním vyjádřením reference?
Mgr. Nedoluzhko: pro češtinu jde o kombinaci s anotací mezivýpovědních významových vztahů (diskurz) a s anotací aktuálního členění – tato kombinace může zjednodušit automatické vyhledávání. U jazyků s členem (angl., něm.) je automatické řešení koreference podstatně jednodušší.

- f. Doc. Petkevič: podejte vysvětlení k Olivově výtce ke str. 66 – „subjekty se používají predikativně“.
Mgr. Nedoluzhko: (podrobnější vysvětlení viz příložená prezentace č. 2).
V projektu MATE se označuje koreference mezi subjektem a jmennou částí přísudku; problematické jsou struktury typu *Řešení je takové. Takové je řešení. To je řešení.*
Doc. Petkevič: původní formulace v disertaci byla nejasná.

8. Vyhlášení výsledku tajného hlasování:
Počet členů komise: 6
Přítomno členů komise: 5
Kladných hlasů: 5
Komise navrhla udělit titul doktor (Ph.D.)

Přílohy: 1. vytištěná prezentace Mgr. Nedoluzhko představující její disertaci
2. vytištěná prezentace Mgr. Nedoluzhko s odpověďmi na připomínky oponentů

V Praze dne 9. 9. 2010 zapsala PhDr. Šárka Zikánová, Ph.D.


Předseda komise prof. PhDr. Petr Mareš, CSc.