

**Nguy Gang Linh: Návrh souboru pravidel pro analýzu anafor v českém jazyce**

Předložená diplomová práce si klade za úkol náročný cíl, totiž vytvořit systém, který by umožnil – alespoň v základních rysech – automaticky identifikovat koreferenční vztahy v českém textu. Po stručném (a zasvěceném) uvedení do základního pojmosloví této oblasti, které zdaleka není jednoduché, autorka věnuje pozornost některým systémům popsaným v literatuře, které se věnují tomuto problému v jiných jazycích, především pro angličtinu, a v další části pak stručně probírá návrh pravidel publikovaný Kučovou a Žabokrtským pro češtinu. V hlavní části práce popisuje svůj návrh analyzátoru anafory v češtině. Materiálem, ze kterého vychází, je Pražský závislostní korpus, a to především s anotacemi na podkladové (tektogramatické) rovině. Svůj návrh pravidel opírá jak o informaci o typu anaforických (odkazovacích) výrazů, tak o informaci o druhu závislostního vztahu a o aktuálně-členské platnosti daného uzlu. V závěru autorka předkládá vyhodnocení úspěšnosti svého analyzátoru podle jednotlivých druhů anafory a navrhuje, jak by bylo možné úspěšnost analyzátoru v navazujícím vývoji systému vylepšit. Práci doplňují dvě přílohy – v první z nich, zřejmě pro jisté osvěžení čtenářových vědomostí – uvádí příklady anafory v některých jiných jazycích než je čeština a ve druhé pak předkládá seznam sloves a substantiv prokazujících vlastnost „kontroly“ a reciprocity, na nichž jeden z postupů analýzy zakládá. K práci je přiložen cd-rom s počítačovým systémem.

Především je třeba vyzdvihnout, jak autorka, která není rodilou mluvčí češtiny, zvládla tak lingvisticky náročný úkol jako jsou otázky koreference. V úvodní části se zmiňuje o rozdílu mezi koreferencí gramatickou a textovou, ve vlastním systému však tento rozdíl nezmiňuje – přitom by bylo zajímavé, jak se autorce tento rozdíl jeví v jejích výsledcích (dá se předpokládat, že určení gramatické koreference bude dávat výsledky lepší, protože je možné se gramatickou informací „naučit“ z PDT). Také by bylo zajímavé vidět, jak se ve výsledcích projevuje rys shoda – neshoda v přiřazení hodnoty aktuálního členění (hodnoty atributu TFA). Opět by se na první pohled dalo předpokládat, že shoda hodnoty v tomto atributu mezi odkazujícím výrazem a výrazem, k němuž se odkazuje, není důležitá, respektive, že odkazující výraz bude mít (převážně, nebo téměř vždy) hodnotu „kontextově zapojený“, zatím co výraz, k němuž se odkazuje, může být jako kontextově zapojený, tak kontextově nezapojený.

Pokud jde o jazykovou stránku práce, je obdivuhodné, jak autorka zvládla češtinu, která není jejím mateřským jazykem. Některé vazby jsou sice neobvyklé, někde jsou pravopisné chyby, ale to nijak nesnižuje vypovídací hodnotu práce. Pokud jde o terminologii, mám několik poznámek:

- str. 2: „odkazovaný subjekt“ snad lépe „referent“
- str. 4: „jmenno-frázová anafora“ – asi lépe „odkazování substantivní frází“ (autorka používá doslovného překladu angl. noun phrase anaphora)
- str. 8: nejde o složité druhy anafory, ale jen o další příklady anafory
- str. 19: rozlišujeme uzly kontextově zapojené, kontrastivní uzly kontextově zapojené (nikoli „opačně zapojené“) a uzly kontextově nezapojené

Také by bylo vhodné v příkladu tektogramatického zápisu na str. 24, obr. 3.1 uvést hodnoty TFA, protože se jich pak dále využívá.

V textu práce se hojně využívá odkazů na literaturu, ale bohužel ne vždy vhodně: u mnoha odkazů je uvedeno jen jméno autorů nikoli rok publikace (to je uvedeno až v seznamu literatury). Někde se odkazuje jen na jméno, ale tento údaj v literatuře není (např. odkaz na Kučovou na str. 24, ale i u jiných autorů jinde). Ovšem je třeba zdůraznit, že autorka literaturu studované oblasti zná a vhodně ji využívá.

Závěrem ráda zdůrazňuji, že moje dotazy a připomínky jsou jen okrajového rázu a nijak nesnižují hodnotu předkládané práce. Navrhuji proto přijmout diplomovou práci **Nguy GangLinh: Návrh souboru pravidel pro analýzu anafor v českém jazyce** k obhajobě.

V Praze 26. 8. 2006



Prof. PhDr. Eva Hajičová, DrSc.  
ÚFAL MFF UK