

Věcné připomínky a problémy k vlastnímu textu disertace:

V kapitole 2.2 (Spojovací konstrukce) se používá pojem „závislost“, aniž by bylo předtím jasně řečeno, zda jde o pojem závislosti „technické“ (vyjádřené hranou stromu) nebo „lingvistické“ (tj. – byť ne zcela formálně – ve smyslu FGP), nebo něco „mezi tím“. Zejména pokud si čtenář představí reprezentaci na analytické rovině, kde je pojem závislosti pohřbíván technický (neboť „lingvisticky nezávislostních“ hran je velké množství, ne-li většina), může se divit, jak a proč byl vůbec pojem „lingvistické“ závislosti na analytické rovině definován (ve skutečnosti se s ním operovalo jen na velmi intuitivní úrovni, a ani v dokumentaci na analytické rovině není vztah mezi FGP závislostí a jednotlivými anotacími pravidly a pokyny nikde explicitně uveden). Teprve z textu vyplývá, že jde o závislost „lingvistickou“, nicméně čtenář si ani potom není zcela jist, jak a kam až se přitom má zajít; jistě může být překvapen např. tím, že se při řešení spojovacích konstrukcí na analytické rovině bez dalšího uvažuje i o „přechodných“ uzlech (předložky, spojky), ačkoli (a) ty existují i mimo spojovací konstrukce a (b) po úvodu (kap. 1) si pečlivý čtenář pamatuje (nebo snadno dovodí, viz též výše), že jsou i jiné „nezávislostní“ hrany – chybí zde zdůvodnění obého (situace (a) i (b)). I když to jistě přispělo ke kratšímu textu, je potom diskutabilní, zda popis řešení spojovacích konstrukcí na analytické a tektogramatické rovině je vhodné takto sloučit.


Tabulka 3.1. patří spíše do závěru práce (kde autor přesně specifikuje vlastní přínos k práci v kontextu celého projektu PZK).

Vzhledem k tomu, že autor dosud relativně málo publikoval (viz výše), a k tomu, že práce je psaná česky, doporučuji vhodně upravenou verzi publikovat i na mezinárodním fóru (zejména vhodně strukturovaný závěr s typologizací problémů a doporučeními pro budoucí korpusy, podpořený čísly, příklady a dalšími údaji z PZK, jak jsou uvedeny v kapitole třetí), a věci spíše „lingvistické“ (spojovací konstrukce v češtině, složené slovesné tvary a jiné lingvistické problémy, které z detekce chyb v PZK vyplynuly) pak i v kvalitním českém odborném lingvistickém tisku.

Závěrečná kapitola ve své druhé části správně shrnuje základní poznatky z celého procesu detekce, opravy a kontroly korpusu; přesto by jí svědčilo (snad s větším časovým odstupem) ještě výrazněji typy problémů klasifikovat a strukturovat, aby případný následník a organizátor podobného procesu při anotaci dalších korpusů měl jasnou představu, co je z toho na takový korpus aplikovatelné.

Překlepy a formální nedostatky: v Úvodu, str. 1 „nepřesný ... interpretace“ má být „nepřesná“. Na obr. 2.1 na str. 24 se překrývá šipka vedoucí od společného rozvíjení „horní“ spojovací konstrukce s levým členem „dolní“ spojovací konstrukce, čímž vzniká dojem, že tento uzel rozvíjí pravý člen „dolní“ spojovací konstrukce.

Závěr: celkově práce prokazuje, že autor dostatečně prokázal schopnost k samostatné, houževnaté a tvořivé (ale přitom ve prospěch velkého projektu orientované) práci, a jeho disertaci tedy doporučuji k obhajobě s kladným výsledkem.

  
Praha, 11. 8. 2006, Jan Hajič, ÚFAL MFF UK