



Ústav anglického jazyka a didaktiky

V Praze dne 5. 9. 2014

Oponentský posudek diplomové práce Bc. Lucie Gillové:

„Tagging a spoken learner corpus“

Diplomová práce Lucie Gillové se zabývá značkováním mluveného žákovského korpusu angličtiny a návrhem vhodného značkovacího schématu na základě existujícího značkování pro korpus LINDSEI.

V teoretické části práce autorka shrnuje základní poznatky o typických rysech jazyka studentů cizích jazyků a o historii výzkumu v této oblasti. Podobně se věnuje i úzce souvisejícímu tématu chybové analýzy a definici chyby obecně. Následně popisuje historii žákovských korpusů a detailně se věnuje chybovému značkování a korpusu LINDSEI, který hraje zásadní roli ve výzkumné části práce. Popisuje také specifika mluveného jazyka z pohledu korpusové lingvistiky.

Teoretická část je zpracována důkladně a kriticky s ohledem na relevantní literaturu.

Ve výzkumné části práce vychází autorka z nahrávek a jejich přepisů, které byly pořízeny pro LINDSEI na ÚAJD. Výhodou je právě dostupnost původních nahrávek, protože přepisy se ukazují pro identifikaci chyb někdy ne zcela dostatečné. Autorka označovala 20 přepisů, z nichž vybrala 200 příkladů s ohledem na co nejširší pokrytí chyb specifických pro mluvený jazyk. Tyto příklady dále analyzuje a na základě analýzy navrhuje změny ve značkovacím schématu LINDSEI.

Výsledkem autorčiny analýzy jsou skutečně návrhy doplnění značkovacího schématu a navíc i změny ve schématu existujících. Za zajímavé považuji také prefixy (rys charakteristický pro mluvený jazyk) a sufixy obecně modifikující stávající značky (nejistota značkování). Jedinou obecnější výtka k práci mám k uváděné šíři kontextu u analyzovaných příkladů, často se mi zdálo, že není dostatečná k posouzení daného jevu jako chybného (u konkrétních pro mě v tomto smyslu problematických příkladů se můžeme zastavit během obhajoby, jsou to příklady 4, 5, 39, 40, 65, 67). Zároveň by bylo i vzhledem k dalšímu využití výsledků práce šikovné, kdyby autorka do závěru zařadila i přehled (např. formou tabulky) navrhovaných změn.

Dále uvádím několik jednotlivostí, u nichž se lze v průběhu obhajoby také zastavit:

- Na str. 33 autorka označuje za největší nevýhodu Free Text Project fakt, že značí každou kategorii samostatnou značkou. Je to skutečně nevýhoda? Na jiném místě se autorka potýká (a je to častý značkovací problém) s překryvem značkových znaků (např. pravopisná chyba v rámci širší chybné konstrukce) – nemůže podobný způsob značkování naopak pomoci s tímto problémem?
- Diachronních mluvených korpusů je více (str. 39), jedním z nich je např. *The Newcastle Electronic Corpus of Tyneside English*.



Ústav anglického jazyka a didaktiky

- Nenapadají autorku možnosti, které by umožnily i výzkum výslovnosti, aniž by byl nutný náročná fonetická transkripce (jak o ni píše na str. 43)?
- Definice chyby je, jak autorka sama naznačuje, poměrně problematické. V některých případech se mi navíc zdá, že by bylo dobré jasněji definovat rozdíl mezi konkrétními typy chyb. Jak se má třeba příklad 43 označený jako jev typický pro mluvenou řeč k příkladům chyb v časové souslednosti resp. nepřímé řeči? Nelze příklad 71 interpretovat také jako opravu (self-correction), nebo příklad 78 a možná častěji příklady s „really“ jako hesitační výplň?
- K těmto nejednoznačnostem bych ještě podtrhl problém správného protikladu značkové chyby, kterého se autorka jen letmo dotýká na str. 52-3 (př. 43).
- V obecnější rovině: domnívá se autorka, že vzhledem k jí nově zavedené značce rysů specifických pro mluvený jazyk by mohlo být užitečné značkování ještě rozšířit, aby pokrylo více jevů a stejným způsobem pak značkovat i srovnávací korpus rodilých mluvčích?

Závěrem bych vyzdvihl především praktický přínos práce – nejde jen o relativně skromné (což není nutně špatné) navrhované úpravy značkovacího schématu, ale i o samotné označování velmi aktuálních dat získaných na mateřském ústavu (a tedy o příspěvek k jejich rozvoji) a o systematickost, na jejímž základě lze o tomto konkrétním problému značkování poučeněji mluvit.

Jazykově i technicky je práce na vysoké úrovni, některé zřídka chyby jsem označil do výtisku práce.

Práce tak zcela splňuje požadavky kladené na diplomové práce a proto ji doporučuji k obhajobě s navrženým hodnocením: výborně.

Mgr. Ondřej Tichý, Ph.D.
Ústav anglického jazyka a didaktiky