



Ústav anglického jazyka a didaktiky

V Praze dne 23. 8. 2016

Oponentský posudek diplomové práce Bc. Simony Zvěřinové:

„N-gramy v mluveném projevu českých a rodilých mluvčích angličtiny“

Práce Simony Zvěřinové srovnává užití víceslovných spojení v mluvené angličtině rodilých mluvčích češtiny a rodilých mluvčích angličtiny. Srovnání založené na datech dvou srovnatelných mluvených korpusů – studentském a rodilém – se zaměřuje především na čtyřslovná spojení, která jsou analyzována jak kvantitativně, tak na menším vzorku kvalitativně.

V teoretické části se autorka zabývá vztahem mezi užíváním víceslovných spojení a úrovní jazykových dovedností/plynulostí vyjadřování (proficiency, fluency), přičemž naráží i na obecnější problematický koncept gramatiky mluvené řeči. Pojednává také relevantní výzkum na poli víceslovných spojení a to jak z pohledu frazeologie, tak z pohledu čistě frekvenčního. Na tomto základě podává přehled terminologie a funkčních kategorií, do kterých lze víceslovná spojení zařazovat.

V části věnované materiálu a metodologii popisuje autorka oba použité korpusy (LINDSEI a LOCNEC), způsob výběru vzorků a samotnou metodu extrakce dat určených k analýze. V závěru této části pak formuluje výzkumné otázky pro následující část analytickou.

Předposlední část je rozdělená na dvě podkapitoly zaměřené na kvantitativní a kvalitativní analýzu získaných dat. Kvantitativní analýza vychází z dat pro všechny mluvčí obou korpusů a srovnává především rozdíl v celkovém počtu víceslovných výrazů (n-gramů) mezi oběma korpusy. Po odstranění opakování stejných slov autorka konstatuje signifikantní, ač ne výrazný rozdíl: nerodilí mluvčí používají méně víceslovných výrazů než rodilí. Nejnáročnější kvalitativní část analýzy je založena na menším vzorku 15 mluvčích z každého korpusu, přičemž výběr mluvčích byl dán ohledem na parametr rychlosti mluvy. Jednotlivé výskyty víceslovných výrazů pak byly rozděleny do čtyř funkčních kategorií (referential, interactional, discourse-organising a propositional) a v rámci nich bylo opět srovnáno užití rodilými a nerodilými mluvčími.

V závěru práce autorka odpovídá na všechny stanovené výzkumné otázky a mimo jiné konstatuje i zajímavý poznatek, že kromě kvantitativních rozdílů v produkci víceslovných výrazů naznačuje kvalitativní analýza rozdíl v míře formálnosti výrazů používanými mluvčími zkoumaných skupin. Pokud lze vyloučit signifikantní rozdíl ve způsobu sběru dat, potvrzuje se tak obecně známý problém pokročilých studentů jazyka s formálností projevu.

Autorku bych u obhajoby poprosil o odpovědi na následující otázky:

1. Co myslí výrokem „In the period spanning 1960s and 1970s, the common understanding of language was reached...“ na str. 9?



Ústav anglického jazyka a didaktiky

2. Autorka opakovaně řeší problém překryvu n-gramů. Jak by problém řešila, kdyby s nimi pracovala znovu. Je to skutečně problém (nechybí v první tabulce na s. 35 logicky také „would like to talk“)?
3. Autorka často naráží na omezení daná fixní délkou sledovaných n-gramů, jak by si s tímto problémem do budoucna poradila?
4. Mohla by na zkoumaný rozdíl ve fluency/proficiency/nativeness vhrnout světlo také samotná délka zkoumaných n-gramů, případně délka v kombinaci s frekvencí?
5. Proč/jak jsou relevantní opakující se slova nalezená mimo rámeček 4-gramů?

Jazykově, formálně i technicky je práce dobře zvládnuta, drobné chyby především formulačního rázu (např. „to reflect the findings back on research“ str. 4) a překlepy jsem vyznačil do výtisku práce.¹ Je třeba podotknout, že závěr a především úvod práce působí jazykově o poznání hůře než většina práce – jako by byly napsány ve spěchu a bez náležité kontroly. Dvě tabulky na s. 35 nejsou číslovány.

Práce je srozumitelně uspořádána, drobné výtky směřují jednak k umístění výzkumných otázek – vhodnější je formulovat otázky před nebo na začátku popisu materiálu a metodologie, jednak ke kvalitativní části analýzy – z pohledu čtenáře by bylo vhodnější krátce uvést a definovat jednotlivé funkční kategorie dříve, než jsou po jedné detailně rozebrány, jelikož smysluplnost jejich rozdělení je možné posoudit až v celku.

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

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

¹ Časté jsou také chyby snad způsobené automatickou kontrolou pravopisu: inform/infirm (s. 37), „any club that would have be as a member“ (s. 28) a především záhadný chobotnatec: ve větě „...word combinations from the opaque, semantically opaque and elephant fixed, to the most transparent...“ na s. 15 (?!).