

Bc. Simona Zvěřinová, N-grams in the speech of Czech and native speakers of English

Diplomová práce

Ústav anglického jazyka a didaktiky, FFUK, září 2016

Posudek vedoucího práce

Autorka měla pro svou diplomovou práci stanoven cíl prozkoumat užívání n-gramů v mluveném projevu českých pokročilých mluvčích angličtiny a zjistit, zda se v této oblasti odlišují od projevu mluvčích rodilých. Práce se tak nachází na poli kontrastivní analýzy žakovského jazyka a přispívá k rostoucímu korpusu literatury srovnávající řeč rodilých a nerodilých mluvčích především z hlediska plynulosti. Snaží se totiž zjistit, zda užívání víceslovných spojení k plynulosti přispívá.

Práce má standardní strukturu. Její teoretická část začíná pojednáním o charakteristikách mluveného jazyka, popisuje některé principy konverzační gramatiky, užívání víceslovných spojení v žakovské angličtině a problematice plynulosti a jejích komponent. Vychází tak z modelu, který pokládá víceslovná spojení za složku, která se podílí na vytváření plynulého projevu. Poté podrobněji pojednává o víceslovných spojeních a srovnává frazeologický přístup s přístupem frekvenčním a zmiňuje i různé přístupy k vytvoření jejich typologie. V celé teoretické části výborně pracuje s literaturou, srovnává různé přístupy a studie, syntetizuje zjištění a případně s nimi neváhá polemizovat, kde je to nutné.

Metodologická část je přehledně napsána: je zřejmé, že autorka rozumí postupům, které získala studiem korpusové lingvistiky. Daří se jí dobře vysvětlit, proč volí tu kterou techniku a výborně upozorňuje na četné metodologické problémy, na které narazila. Jako vedoucí práce musím říct, že tyto problémy autorka pečlivě konzultovala a vždy přišla s připravenými možnostmi řešení, která dokázala dobře zdůvodnit. Rozhoduje se čerpat data z korpusů LINDSEI_CZ a LOCNEC a volí kvantitativní průzkum a následnou kvalitativní sondu do zúženého vzorku mluvčích. I v této části byly v průběhu práce nutné změny, protože se ukázalo, že původně stanovenou velikost vzorku bylo nutno rozšířit a studentka tak musela velkou část analýz provádět znovu. Oceňuji, že to v krátkém časovém úseku zvládla a že se nesnažila pracovat s menším vzorkem a výsledky "kamuflovat". Mohla snad jen lépe vysvětlit, proč z n-gramů zvolila právě 4-gramy.

V kvantitativní analýze především srovnává frekvenci jednotlivých typů v obou korpusech, výsledky jasně prezentuje, užívá statistický test, ale nezmiňuje jaký. Uvedená hodnota "p" je sama o sobě jen vyjádřením pravděpodobnosti toho, že výsledek je náhodný, neříká nám však nic o velikosti účinku. Pokud studentka užívala log-likelihood test, měla uvést ještě alespoň hodnotu G2, v případě chí-kvadrátu pak tuto hodnotu. V kvalitativní analýze vybírá z každého korpusu po 15 mluvčích a vytváří funkční typologii 4-gramů, které užívají. Zde nachází poměrně zajímavé výsledky, které však není nikterak snadné interpretovat a nabízí se otázka, do jaké míry se dají zobecnit. Analýza je výborně popsána a shrnuta, celkově jsem se však v textu místy trochu ztrácel; zřejmě by bývalo stálo za to text této části přehledněji strukturovat.

Velice přehledně a s nadhledem je napsáno závěrečné shrnutí, kde se autorce daří systematicky prezentovat výsledky, srovnání i vše, co z nich vyplývá. Zodpovídá na všechny výzkumné otázky. Při srovnání se studii De Cockové a Larsson Aasové mluví o rozdílech ve statistické významnosti, což však nemusí ještě nutně vypovídat o velikosti tohoto rozdílu, resp. velikosti účinku. Oceňuji vysvětlení zjištěných rozdílů, když autorka upozorňuje na ekologickou validitu korpusu LINDSEI_CZ; ta dozajista má vliv na styl projevu, který čeští mluvčí zvolili pro nahrávání korpusu. Touto pasáží autorka přechází k výčtu a diskusi možných nedostatků své studie; jde o velmi informované zhodnocení.

Práce je napsána výbornou angličtinou a dodržuje zásady akademického stylu. Je též dobře redigována, ač o některých částech toto platí o něco méně. Například trochu narychlo byl zjevně napsán abstrakt, který nedodržuje zcela zásady pro jeho sepsání (např. nemluví o výsledcích). Pro svou budoucí akademickou

práci by se autorka měla s kritérii pro psaní abstraktů seznámit, protože mohou výrazně ovlivnit přijetí její práce odborným publikem. Autorka výborně pracuje s literaturou, měla by si však být vědoma, že i jazykové korpusy je třeba řádně očitovat, což se týká jak LINDSEI_CZ, tak LOCNEC.

Celkově hodnotím práci jako velmi zralou a kvalitní. Jako vedoucí práce konstatuji, že autorka pracovala iniciativně a s velkou samostatností. Byla vždy výborně připravena na konzultace a dotazovala se na skutečně podstatné problémy, ke kterým vždy nabízela možná řešení. Celkově práce splňuje zadání, je jen škoda, že autorka hlouběji nevysvětluje, proč nebylo možné přistoupit k hledání korelace mezi četností n-gramů a tempem mluvy (v poslední konzultaci nad prací to přitom důkladně vysvětlila).

Závěr

Domnívám se, že práce Bc. Simony Zvěřinové je velmi kvalitní a přináší nové poznatky. S ohledem na výše uvedené práci doporučuji k obhajobě a navrhuji hodnocení **výborná**.

PhDr. Tomáš Gráf, Ph.D.
30.8.2016