Posudek oponenta disertační práce Any Adamovičové „Partikule v mluveném jazyce“ předkládané v roce 2018 v Ústavu českého národního korpusu FF UK

I. Stručná charakteristika práce

Předkládaná práce se zabývá distribucí a užitím partikulí v Pražském mluveném korpusu (PMK) Českého národního korpusu (ČNK). Na základě autorčiny mnoholeté zkušenosti ve tvoření korpusu jsou mapovány distinktivní rysy partikulí v běžné mluvě v Čechách.

II. Stručné celkové zhodnocení práce

Autorka si klade za cíl podávat vyčerpávající přehled používání partikulí v mluvené češtině a tento úkol splnila. Co se týče délky i obsahu, jde o úplnou a pečlivě zpracovanou disertaci, která nabízí čtenářům detailní pohled do mnoha aspektů PMK, nejen do partikulí.

I přes objemnost a detailnost je několik lakun, které by do budoucnosti zasloužily pozornost. Ty jsou zmíněny v III.1 a III.3. níže: definice partikulí, rozbor problematiky partikulí v srovnávacím kontextu a využívání statistických metod.

III. Podrobné zhodnocení práce a jejích jednotlivých aspektů

1. Struktura argumentace.

Výklad je celkově srozumitelný a na každém místě je cítit autorčina hluboká znalost PMK. Práce je členěna přehledně. Ke struktuře mám čtyři poznámky:

a) Obecný cíl je velmi široký: podávat úplný přehled výskytu partikulí v mluveném korpusu. V tom autorka uspěla, nicméně mi často chyběla tradiční struktura výzkumné práce počínaje výzkumnou otázkou přes hypotézu, analýzu k závěrům, abych se v detailech mohl orientovat a přijít na to, proč se autorka zabývá zrovna tím či oním bodem.

b) Stěžejní strukturní mezerou je, že se autorka ani v kapitole o různých postupech ke zkoumání partikulí nesnaží podat definici partikule. Vícekrát upozorňuje na to, že jde o problematický slovní druh, pro čtenáře by bylo ale podstatnější a důležitější , aby od autorky zazněla pracovní definice, která by se mohla na závěr podrobit detailnějšímu přezkoumání.

c) Místy odbočuje od tématu a zabývá se tématy, která nejsou zcela nezbytná pro linii argumentace. Autorka by mohla několik postranních podkapitol vypustit a důsledněji směřovat čtenáře k zamýšlenému cíli.

d) V některých částech by autorka mohla rozložit strukturu argumentace přehlednějším způsobem. Např. v 6. kapitole analýza začíná definicemi různých pozic a pokračuje přes četné tabulky s minimálním komentářem, či v 8. kapitole na s. 381-382 se nabízí 6 tabulek po sobě bez vysvětlení, proč je co v které tabulce a bez následného komentáře, co je v těchto informacích relevantní. Bylo by žádoucí tu poskytnout výchozí hypotézu a vysvětlení metody, aby čtenář nemusel sám odhadnout, kam práce míří a proč je zajímavé či poučné se zabývat tříděním partikulí podle pozic. Práci by prospěl selektivnější výběr dat a výkladu.

2. Formální úroveň práce

Na formální úrovni se mi práce zdála adekvátní: kompetentně pracuje se zdroji, posuzuje na vysoké odborné úrovni. Kvalitu češtiny nechávám na rodilých mluvčích, mně se četla dobře.

Tabulky (i když je jích místy až mnoho) jsou většinou srozumitelné, i když je každá jiná a čtenář musí každou prostudovat jednotlivě, aby rozluštil zkratky a množství informací do ní zakomponované.

Označování příkladů je různorodé: většina není číslována, místy má každý svůj řádek, jinde jsou rozděleny středníkem, další číslovány jsou.

Obecně by se mělo přihlížet důsledněji k relevantnosti předkládaných informací. Na více místech jsou uváděny dlouhé tabulky, ale v následné diskusi nejsou zmíněny (Tabulka 3.2, s. 93-95), anebo se připouští, že se z nich nedá nic relevantního vyčíst (Tabulka 3.4, s. 97-98).

3. Práce s prameny či s materiálem

Autorka prokazuje širokou znalost české literatury o partikulích, korpusech a mluvnicích češtiny, dokáže z nich čerpat a hodnotit je nestranně a věcně. Jasně a přesvědčivě zdůvodňuje sběr dat z PMK a vysvětluje své závěry. Svědomitě a pečlivě používá základní deskriptivní statistiky.

Výše jsem se zmínil o několika místech v literatuře a metodologii a rád bych na ně obrátil pozornost:

Za prvé, uvítal bych snahu dát „české“ definice a „českou“ literaturu o partikulích do širšího lingvistického kontextu. Jde o značně problematický slovní druh: problémy s jeho vymezením jsou probírány v prvních kapitolách, i když v disertaci se autorka vyhýbá určení vlastní definice. Typologický kontext by možná vrhl na vymezení tohoto slovního druhu užitečný pohled. V seznamu literatury je zmíněno několik prací o partikulích v angličtině, francouzštině atd., které ale nepřispívají významným způsobem k předkládanému popisu oboru.

Za druhé, analýza překypuje tabulkami a základními deskriptivními statistikami, ale chybí v ní další logický krok, což jsou statistiky inferenční. Bez nich je užitečnost analýzy omezena, je možné jen konstatovat, že je něčeho mnoho nebo málo (viz např. s. 106, kde padne zmínka o rozdílech „ve výskytech sice velké, i když už ne tak obrovské“; s. 134, kde se jedná o sociolingvistické rozdíly; 7. kapitola, kde se rozdíly mezi slovními druhy opírají o fakt, že nějaké spojení je „častější“ u partikulí než u příslovcí nebo naopak; či s. 386, kde se bez definice rozlišuje mezi „podobnými“ a „srovnatelnými“ pozicemi ve frekvenčních tabulkách).

Zvlášť vzhledem k *sociolingvistickým proměnným*, kde se autorka pokouší o hierarchizaci důležitosti faktorů, by inferenčně statistická analýza byla vhodná; bez ní není interpetace dat opodstatněna tak, jak byla vyložena . Zvlášť v analýzách méně frekventovaných partikulí hrozí např., že jde o kovariaci s proměnnou „mluvčí“, tj. že převahu toho či onoho tvaru způsobuje jeho nadměrné užití u pouze několika mluvčích.

Důsledkem toho je, že se nabízí popis situace v jednom korpusu bez možnosti odvodit z něj pravděpodobnost, že jde o spolehlivý výsledek platný i pro mluvenou češtinu obecně. Na opravu tohoto nedostatku by většinou stačily nesložité inferenční testy (ve kterých se testuje PMK oproti korpusu psaného jazyka a hodnotí se skutečný výsledek oproti očekávanému výsledku čerpanému z většího korpusu, jak se to např. dělá v 8. kapitole). Tím by analýza nabyla na významnosti a nabízela definitivnější závěry o stavu partikulí v mluveném jazyce.

Za třetí, autorka se opírá důsledně o korpusové značkování textového stylu, i když je možné, že díky subjektivnosti značkování referuje spíše o rozhodnutích kóderů než o lingvistických parametrech textu. Jako příklad bych uvedl snahu klasifikovat styl partikulí v odd. 3.3. Získáváme tím nové poznatky o stylu partikulí, anebo spíše o tom, jak kódeři pocítili formalitu daného tvaru (*dovopravdy* nemůžeme značkovat jako neutrální styl)?

4. Vlastní přínos

Autorka přináší mnoho nových poznatků o užití partikulí v PMK. Popis struktury, kompilace a anotace korpusu je detailní. Cílem práce je úplný popis užití partikulí v PMK a tento úkol je bezesporu splněn. Se znalostí těchto jevů ve smyslu jejich detailního popisu jsme tedy teď o krok dále. Autorka do teoretických otázek však zasahuje jen velmi opatrně: využívá poznatky, demonstruje znalost a pochopení teorie, ale vyhýbá se snahám o vytříbení, přeinterpretování či přetvaření teorií a zaměřuje se spíše na poznámky ohledně jejich výstupů (např. návrhy na značkování mluvených korpusů).

IV. Dotazy k obhajobě

1. Jaká by byla podle autorky uspokojivá definice „partikule“?
2. Americký lingvista Arnold Zwicky konstatoval, že „partikule jsou slova, která zbyla po zařazení ostatních slov do syntaktických kategorií“ (1985:292) a argumentoval pro zrušení „partikulí“ jako samostatného slovního druhu. Koherentnost této kategorie se dá zpochybnit na základě korpusových dat, jak to např. dělají Endresenová et al. 2016. Jak by autorka obhájila existenci partikulí jako slovního druhu, zvlášť vzhledem k rozmanitosti a nejednotnosti kategorie, které ve svém výzkumu odhalila?
3. Do jaké míry si autorka myslí, že je označení *stylů* v PMK konzistentní a nepovede nás do kruhu? Mě na tom znepokojil nejdříve nepoměr mezi styly 1/2 (označují celé úseky textu) a 3/4 (označují spíše jen lexémy) – že „styl“ je buď vlastnost textu anebo vlastnost lexému, ale ne pro jeden styl tak a pro druhý styl onak. Jaký to má pak vliv na tvrzení, že styly 1 a 2 jsou základní a že styly 3 a 4 jsou marginální?
4. Říká se, že vědecká studie věrně reprodukuje předsevzetí vědců, kteří se na ní podíleli. V tomto ohledu přináší ruční anotace korpusu celou řadu otázek, včetně rozhodnutí o seskupování fonologických slov do víceslovných lemmat; o zařazení do stylů; o určení slovního druhu lexému. (Autorka sama uvádí na s. 59 v pozn. 50 podobné pochybnosti.) Je ruční značkování v mnoha případech příliš idiosynkratické a individuální, aby bylo vědcům nápomocno?
5. Třídění partikulí do klasifikačních kategorií (kap. 5) je při sémantickém kódování korpusu nezbytné. Nicméně stojí za otázku, jestli takto s partikulemi operuje lidská mysl – myslím tím, jak lidé přiřazují k partikulím význam a jak si vybírají vhodnou partikuli v dané situaci. Jak by vypadal kognitivně adekvátní popis sémantiky partikulí (bez ohledu na kódování)?
6. Jaký je vědecký přínos výčtů kumulací partikulí evidovaných v kap. 6 a statistiky o frekvenci jednotlivých partikulí v kap. 7 – resp. jak to posouvá dále naše pochopení statusu partikulí a jejich roli v diskurzu?

V. Závěr

Předložená disertační práce splňuje požadavky kladené na disertační práci, a proto ji doporučuji k obhajobě a v předběžně ji klasifikuji jako *prospěla*.

V Sheffieldu dne 1.6.2018 Prof. Neil Bermel