Posudek oponenta disertační práce PhDr. Any Adamovičové

“Partikule v mluveném jazyce”

předkládané v roce 2018 na Ústavu Českého národního korpusu

427 stran

Vedoucí práce: prof. PhDr. František Čermák, DrCs.

Disertační práce přináší detailní popis partikulí v mluvené češtině, který vychází z konkrétní materiálové základny i teoretického popisu použitého pro první korpus mluvené češtiny: Pražský mluvený korpus (PMK). Autorka se sama přípravy korpusu i manuálního značkování zúčastnila, uvádí tedy velmi zasvěcené komentáře ke zkoumaným datům. Na základě korpusu se věnuje podrobné kvalitativní i kvantitativní analýze skutečně všech tvarů partikulí, všímá si i jejich variantnosti, kombinatorických vlastností a zejména funkce, kterou partikule plní v mluveném jazyce. Jde také o první představení toho, jak jsou partikule v jazyce užívány z pohledu sociolingvistických parametrů. Nechybí ani srovnání s psaným jazykem.

Autorka však nevychází pouze ze zmíněného manuálního značkování PMK, ale podrobuje tehdejší rozhodnutí týmu, kterého byla součástí, kritické analýze. Jedná se o velice detailní popis materiálu, v němž se vyskytuje řada nejasných a sporných užití. Je třeba ocenit, že autorka se nevyhýbá jejich řešení, zmiňuje nejednoznačnosti, nabízí různé pohledy, představuje výhody i nevýhody zvolených řešení. Při tom všem vychází vždy z konkrétního kontextu dané promluvy. Velkým přínosem je i část věnovaná partikulím jako homonymním slovním druhům a ověření jejich skutečného zastoupení v konkrétních promluvách. Přehledně jsou zde zachyceny a materiálově zdokumentovány procesy partikularizace. Práce je velkým přínosem pro popis užívání partikulí v češtině a jejich důležitosti v mluveném projevu, neboť se jedná o poměrně frekventovaný slovní druh. Jejím nemalým vedlejším efektem je i jasné uvedení do některých problémů manuální anotace PMK. Otevírá další možnosti komparace s novějšími mluvenými korpusy a může přispět i ke zlepšení jejich anotace.

Při takto detailní práci s korpusovým materiálem, jako je tomu v případě předkládané disertace, badateli vždy komplikují výzkum omezení, která jednotlivé typy zpracování materiálu přinášejí. V tomto případě je to na prvním místě nedostupnost zvukového záznamu, která dle mého soudu ztěžuje, až znemožňuje některé detailní analýzy, jako je např. rozhodnutí, zda se jedná o víceslovnou partikuli, frazém či pouze kumulaci několika partikulí nebo rozdíly uváděné v užívání partikulí na s. 166. Dalšími takovými faktory bránícími do budoucna srovnání partikulí v jiných i mluvených korpusech je stylová charakteristika lemmat a užívání tzv. stylového lemmatu.

S některými problémovými kategoriemi se autorka sama úspěšně vypořádala: patří k nim přehodnocení charakteristiky mluvní situace a její rozčlenění na spíše monologickou a dialogickou (místo původní formální a neformální), což je právě faktor, který se ukazuje jako důležitý nejen v případě užívání partikulí. Podobně velmi funkční pro popis partikulí je změna označení primární a sekundární partikule na slovnědruhově homonymní a nehomonymní v mluvené češtině a stanovení výčtu takovýchto partikulí.

Přes veškeré klady PMK nelze ani při výkladu zapomínat na to, že shromažďuje materiál pouze ze specifického byť rozsáhlého regionu a nikoliv ze všech nářečních oblastí naší republiky a tudíž nelze na jeho základě tvrdit, že přirozený mluvený jazyk „je ze své podstaty převážně a většinově nespisovný a obecněčeský“ (str. 56).

Autorčin přístup k materiálu a jeho popisu je velice fundovaný a výklad je jasný a srozumitelný. Nicméně při uvádění takového množství frekvenčních údajů a výčtu by práci prospěla větší část

věnovaná interpretaci předkládaných výsledků. Některá uváděná srovnání, zejména mezi psaným a mluvených jazykem, jsou zbytečně podrobná až nadbytečná vzhledem k účelu, který práce sleduje (např. srovnání nejčastějších substantiv a adjektiv v psaném a mluveném korpusu v 2. kapitole, tabulka 2.11 na str. 73). Z terminologického hlediska by bylo vhodné podrobněji vysvětlit pojetí kolokace a valence (6. kapitola), neboť se v současné lingvistice jedná o velmi různě užívané termíny.
Při srovnávání korpusových dat tak odlišného rozsahu jako je tomu u mluvených a psaných korpusů by velmi prospělo uvádění nějaké formy normalizované frekvence (např. přepočet na milion korpusových pozic, který umožňuje rozhraní Kontext na rozdíl od rozhraní Bonito, ale který lze doplnit i manuálně). U mnoha údajů, zejména u těch málo doložených, se nabízí samozřejmě srovnání s dalšími korpusy mluvené češtiny, které autorka ale nemohla použít vzhledem k nedostupnosti jejich značkování v době vzniku této práce. Její výsledky ale jistě budou dobrým východiskem pro budoucí srovnání.
Po formální stránce by takto rozsáhlé práci prospělo více odkazů na literaturu s přesnou citací stran, např. na s. 96 je citován 3. svazek SČFI (správně patří 2. svazek) a jelikož víceslovné částice nejsou jeho hlavní náplní, bylo by vhodné doplnit přesnou citaci strany. Některé údaje by bylo vhodnější prezetovat formou poznámkového aparátu. Vzhledem k rozsahu práce je množství překlepů a nepřesností velmi malé (např. s. 260 navíc předložka v u spojení úzus a frekvence, 390 *okodovína* místo okódována, 421 *nivě* místo nově, na s. 17 nepřesná informace o tom, že korpusy ORAL a ORTOFON nebyly dosud anotovány: oba korpusy byly v červnu 2017 zveřejněny již s lemmatizací a morfologickým značkováním). Jinak je práce přehledná a úpravná.
Předností této práce je vyčerpávající popis partikulí, který vychází ze skutečného materiálu mluvené češtiny a nevyhýbá se žádným těžkostem, který tento materiál přináší. Výsledkem je velmi zajímavý výčet partikulí užívaných v mluvené češtině, který doplňje popis mluveného jazyka J. Šonkové, zaměřený především na autosémantika. Důležitým přínosem je také stanovení míry slovnědruhové homonymie pro jednotlivé výrazy a zejména pak ty případy, kdy v mluveném jazyce k této homonymii nedochází. Tyto výsledky by mohly vést ke zlepšení anotace dat. Další dílčí výsledky se díky podrobnosti zpracování stávají výzvou pro budoucí ověřování co do vlivu socialingvistických parametrů pro data získaná z jiného časového období. V neposlední řadě jde o ověření jedné z naznačených tezí, a totiž, zda užívání partikule podléhá určité jazykové módě. Předností práce jsou také přehledně vyložené a na konkrétním materiálu ilustrované procesy partikularizace.

Dotazy k obhajobě
Při obhajobě bych ráda položila autorce následující otázky:

1) Při srovnání s psaným jazykem byl zvolen korpus SYN2005, ač časově bližší svým složením by byl korpus SYN2000? Proč tomu tak bylo? Kterou verzi značkování jste použila? Korpusy jsou anotovány vylepšenou verzí anotačních nástrojů vždy v poslední souhrnné verzi korpusu SYN, při čemž data zůstávají neměnná, lze tedy sledovat, zda a jak se anotace v dílčích bodech zlepšila.

2) Vysvětlila byste detailněji rozdíl mezi víceslovnou partikulí a frazémem? Teoretické pojetí vychází ze samostatnosti frazému vůči promluvě a uvádíte je na několika místech práce (např. str. 101 „co je v těsném sepjetí s promluvou a je těsně zakomponováno do její struktury, by mělo být považováno za víceslovnou partikuli, na rozdíl od frazémů, které se vyskytují samostatně“; nebo str. 193 „frazémy stojí samostatně, víceslovná partikule se vždy váže k určitému úseku věty či ke slovu“, avšak uváděné příklady (zejména na str. 97, např.

dejme tomu, *tak ňák,* a str. 193, např. *jen jestli*) jsou dle mého názoru frazémy a při jejich určení je důležitá intonace. Může mít frazém funkci víceslovné partikule?

3) Na str. 208 uvádíte pod tabulkou 6.7 představující 20 partikulí s blíže neurčenou valencí (např. *jako, teda, tak, prostě* atd.): **„**Tyto partikule se vyskytují v kontextu narušené větné struktury, nedokončených vět, elize, různých vybočení, digrese, tápání při hledání lepší formulace a adekvátnějšího uchopení myšlenky. Partikule při takovýchto sděleních nemá tedy svou běžnou funkci…“ Podle mé zkušenosti se jedná o typickou součást mluvené komunikace, se kterou by měl zvolený teoretický popis počítat. Jedná se především o expletiva, která byla v práci zmíněna. Podobné problémy nastávají také např. u opakování slov a hezitací. Jaký je váš názor na zpracování těchto jevů jiným způsobem než v rámci tradiční slovnědruhové klasifikace (např. zmiňované discourse/pragmatic marker apod.)? Domníváte se, že s tradičním slovnědruhovým dělením při popisu mluveného jazyka vystačíme?

4) Pokoušela jste se na základě zjištěných údajů o nějakou formální kontrolu manuálního značkování (např. vyhledáním bigramů na začátku a konci promluv, pomocí kolokačních měr), které by pomohlo manuální kódování verifikovat, případně rozšířit? Pokud ano, jaké to přineslo výsledky?

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

24. 5. 2018 Marie Kopřivová