Univerzita Karlova

Filozofická fakulta

Ústav českého národního korpusu

Matematická lingvistika

Autoreferát

PhDr. Ana Adamovičová

Partikule v mluveném jazyce

vedoucí práce prof. PhDr. František Čermák, DrCs.

2018

Tato práce si klade za cíl zmapovat téma, kterému v české lingvistice zatím nebyla systematicky věnována dostatečná pozornost. Vychází z materiálu *Pražského mluveného korpusu* (PMK) obsahujícího řadu nahrávek autentického mluveného jazyka. Cílem práce je v prvé řadě podat vyčerpávající popis partikulí vyskytujících se v PMK (2. a 3. kapitola), dále na základě jejich skutečného výskytu v přirozeném kontextu promluvy představit sémantické třídy partikulí (5. kapitola), blíže se podívat na jejich kolokabilitu a syntagmatiku včetně valence (6. kapitola) a tím lépe vysvětlit jejich funkci v běžně mluveném jazyce. Pozornost je též věnována jejich frekvenci a oblibě v rámci jednotlivých kombinací sociolingvistických parametrů, respektive čtyř proměnných (pohlaví, věk, vzdělání a typ promluvy), zahrnutých do výběru mluvčích ve snaze vytvořit jejich co nejvyváženější kombinace při pořizování nahrávek (4. kapitola). Řada partikulí je homonymní s dalšími slovními druhy: míra homonymie a její kontextová podmíněnost je analyzována v 7. kapitole. Frekvenční výsledky, k nimž jsme dospěli, jsou porovnány s frekvencí výskytu partikulí v psaném korpusu SYN2005 (zveřejněném v *Statistikách češtiny*) a též byla zmapována prezentace partikulí ve slovnících českého jazyka (8. kapitola).

Takovýto důkladný rozbor by měl vést k lepšímu formálnímu a funkčnímu vymezení partikulí. Taky by mohl přispět k jejich snadnějšímu vymezení při sestavování trénovacích souborů pro automatické značkování mluvených korpusů. Jedním z dílčích cílů práce je též poukázat na nedostatky popisu partikulí v lexikografické praxi. V neposlední řadě by důkladný kontextový popis partikulí měl vést i k lepšímu uchopení této problematiky při výuce češtiny pro cizince. Partikule nejsou doménou jen mluveného jazyka, i když je jejich frekvence v něm mnohem vyšší, mají též řadu pragmatických funkcí podílejících se velkou měrou i na uspořádání psaného diskursu. Zatímco rodilí mluvčí používají partikule přesně a spontánně a pomocí nich různými způsoby modifikují svou výpověď či zaujímají stanovisko k adresátovi či obsahu sdělení, anebo odkazují na segmenty své výpovědi jak při interakci, tak při strukturaci textu, nerodilí mluvčí je užívají mnohem méně, někdy vůbec ne, často ovšem nesprávně, což v úhrnu přispívá k prohloubení jejich jazykové nejistoty a oddaluje zvládnutí kýženého cíle – přiblížit svůj jazykový projev projevu rodilých mluvčích.

Na základě *Pražského mluveného korpusu* vyšel v r. 2007 *Frekvenční slovník mluvené češtiny* a zatím vznikly pouze jedna větší monografie – morfologie mluveného jazyka (Šonková, J.: *Morfologie mluvené češtiny: frekvenční analýza,* NLN, Praha 2008) a ukázky nejzajímavějších nahrávek tvořících *Pražský mluvený korpus* – monologů a dialogů – zahrnutých do *Čítanky mluvené češtiny* (Adamovičová, A., Čermák, F., Pešička J., Karolinum, Praha 2008).

Cílem projektu PMK – FSMČ bylo vůbec poprvé souvisle a v úplnosti zmapovat autentickou mluvenou češtinu, a tak nezkresleně představit její jádro – obecnou češtinu. Korpus má přídomek „pražský“, protože se vycházelo z představy, že Praha, vzhledem ke svému centrálnímu postavení a velké fluktuaci obyvatelstva, svým způsobem reprezentuje běžnou českou mluvu v Čechách a zčásti i na Moravě. PMK díky této strategii – systematickému pokrytí jazykové oblasti obohaceného o intenzivní anotaci – , skýtá zcela ojedinělý obraz mluvené češtiny na konci 20. století, přičemž odráží jak jazyk předchozího společenského uspořádání, tak případné změny na začátku uspořádání nového. Přestože lze namítat, že je daný rozsah pokrytí omezený, což je ovšem pro mluvené korpusy celkem běžné, slovník nicméně představuje cenný materiálový základ pro studium mluvené češtiny.

PMK svým rozsahem patří spíše k menším korpusům. Korpusoví lingvisté nejednou vyzdvihovali fakt, že skutečně nezkreslené údajů o jazyce lze získat na základě analýzy miliónových korpusů: čím větší korpus, tím jistější data. Tento fakt nikterak nelze zpochybnit, otázkou ale zůstává, co svým průzkumem chceme ukázat a čeho chceme docílit. Pro popis některých jazykových jevů se rozsáhlé korpusy nemusí hodit: i sebezkušenější lingvista se v záplavě dat může ztratit. Proč tedy při popisu partikulí vycházíme právě z PMK, i když dnes máme k dispozici rozsáhlejší korpusy mluveného jazyka (např. řada ORAL či ORTOFON)? Jednak tyto korpusy ještě nebyly lemmatizovány a morfologicky označkovány, byť se s tím do budoucna počítá, jednak je obecně známým faktem, že jsou partikule, vzhledem ke své povaze a velké míře homonymie s jinými slovními druhy, špatně dohledatelné ve velkých automaticky značkovaných korpusech. Materiál opírající se o PMK, jakkoli omezený svým rozsahem, poskytuje autentickou bázi pro studium partikulí v mluveném jazyce a též umožňuje jejich všestrannou analýzu ve skutečném úzu a kontextu, což je nesmírně cenné. Korpus je celkem vyvážený co do počtu mluvčích, jejich věku, dosaženého vzdělání (sociolingvistické parametry) při zohlednění formálních a neformálních situací promluvy.

Pozornost soustředěně a cíleně věnovaná českým partikulím, poprvé na objektivním korpusovém základu, vyžaduje pro nejasnou a někdy i mlhavou představu, kterou tento slovní druh často vzbuzuje, nutně komplexní a systematický přístup. Musí vycházet jak z neoddiskutovatelného potenciálu, který korpusy nabízejí, tj. z analýzy forem a jejích funkcí v pokud možno rozsáhlém mluveném korpusu s jeho bohatstvím kontextů, tak metodologicky se věnovat cizelování výchozích předpokladů a hypotéz, které si analýza dat vynucuje, a jež v nemalé míře zasahují i do málo probádaných aspektů pragmatických.

1. KORPUSOVÁ LINGVISTIKA. PARTIKULE. MLUVENÝ JAZYK

V této kapitole je blíže osvětleno jak širší metodologické východisko práce – korpusová lingvistika, tak samotný obsah práce, a sice slovní druh partikule a mluvený jazyk:

- korpusy poskytují nenahraditelnou bázi pro nový a nezkreslený popis jazykových dat. Týká se to především studia jazykových dat v širokém kontextu a opravdovém úzu, což je pro zkoumání funkcí a významu partikulí přístup naprosto klíčový;
- shrnutí dosavadních přístupů českých lingvistů k partikulím (Trávníček, Havránek, Jedlička, Šmilauer, Uličný, Grepl, Nekula, Čermák a další) potažmo jejich zpracování v základních příručkách o českém jazyce (Česká mluvnice, AMČ, Nauka o českém jazyku, N/ESČ, PMČ, MSoČ a další);
    - tento přehled mj. i na diskusi příkladů a ukázkách z PMK odhalil dvě věci: jednak naprosto nejednotný, nekonzistentní, výběrový a často vágní přístup lingvistů k témuž zkoumanému jevu, jednak to, že partikule zatím nebyly popsány systematicky a ve světle nových poznatků, jež umožňují korpusová data a rozsáhlé kontexty;
- vymezení pojmu mluvený jazyk. Zde se přikláníme hlavně k pojetí mluvenosti, jak jej navrhují Sgall, Hronek, Čermák, Cvrček a někteří další lingvisté. Stručně je zmapována neutěšená situace týkající se místa obecné češtiny a běžně mluveného jazyka v diskusích o stratifikaci češtiny, stylových funkcích jazyka, jazykové kultuře a v souvislosti s tím často vyzdvihovaném nedotknutelném a výsadním postavení spisovného jazyka.

2. METODICKÉ VÝCHODISKO ZPRACOVÁNÍ DAT V PRAŽSKÉM MLUVENÉM KORPUSU

Tato kapitola se věnuje metodickému východisku práce v užším slova smyslu – popisuje vznik a budování PMK (2.1), metody zpracování dat v PMK a sociolingvistické parametry, pomocí nichž lze provádět jemnější analýzu materiálu (2.2) a *Kódovník PMK* (2.3), na jehož základě proběhla ruční anotace celého materiálu. Další podkapitola (2.4) je věnována stylové charakteristice lemmat v PMK. Partikulím a dalším slovním druhům v PMK, hlavně jejich frekvenční analýze i porovnání s psaným korpusem SYN2005, se věnuje předposlední podkapitola (2.5). Poslední podkapitola (2.6) je přímo zaměřena na podobu a tvary partikulí vyskytujících se v PMK-FSMČ.

Budování PMK a metodická východiska

- **první fáze projektu** spočívala v pořízení 304 nahrávek obsahujících přes 15 015 promluv, které vznikaly v letech 1988-1996. Při nahrávkách byly pořízeny:
    - **formální rozhovory** (F, též M) – 193 nahrávky (monology vycházející z dotazníku zaměřeného na zachycení názorů respondentů na otázky týkající se běžného života: práce, škola, vzdělání, rodina, výchova dětí, mezilidské vztahy, život ve městě vs. život na venkově, šťastný a úspěšný život; v druhé vlně po převratu 1989 přibyly otázky o podnikání a cestovním ruchu);
    - **rozhovory neformální** (N, též D) – 103 nahrávky (dialogy jsou volným rozhovorem dvou respondentů na jakékoli téma). Respondentům (v úhrnu jich bylo 504: 287 žen a 217 mužů) byla zaručena anonymita a nebyl jim sdělen účel (tj. ryze jazykový výzkum), za jakým se nahrávky pořizují;
    - při výběru mluvčích se maximálně dbalo o to, aby proporce čtyř sociolingvistických parametrů byly pokud možno co nejvyváženější:
        - **pohlaví** (ženy a muži v poměru 56 % 44 %), **věk** mluvčích (mladší – od 20 do 34 let a starší – nad 35 let v poměru 55 % : 45 %), **vzdělání** (nižší – základní a střední, včetně nedokončeného vysokoškolského a vyšší – vysokoškolské v poměru 55 % : 45 %) a **typ promluvy** (tj. formální a neformální v poměru 60 % : 40 %);
        - vzhledem k oscilacím v zastoupení jednotlivých sociolingvistických parametrů všechny statistické údaje, které zveřejňujeme, zohledňují propočty poměrných hodnot distribuce toho kterého parametru;
- **v druhé fázi projektu** se všechny pořízené nahrávky přepsaly podle zásad správného přepisu do počítače maximálně věrně (pár jich muselo být vyřazeno pro špatnou kvalitu). Takto vlastně vznikl PMK o rozsahu tři čtvrtě miliónů jednotek;
- **ve třetí fázi projektu** se přistoupilo k anotaci takto získaného materiálu na základě *Kódovníku PMK*, který pro tyto účely sestavil F. Čermák. Šlo o ruční označení každého slova až jedenáctimístnými kódy, poskytujícími především úplnou morfologickou informaci včetně valence, částečně také syntaktickou informaci a údaj o stylové charakteristice promluvy. K syntaktické informaci nutno podotknout, že ačkoli se s její analýzou při sestavování kódovníku počítalo, v praxi se od toho z větší části upustilo, neboť se rozdíly v syntaxi mluveného a psaného jazyka ukázaly být tak velké, že se syntax psaného jazyka pro analýzu jazyka mluveného stala jen stěží použitelnou. Vznikaly různé problémy související nejen s označením hranice věty v souvislém proudu vyprávění, ale také s jejím vnitřním členěním a odlišnou větnou perspektivou mající často eliptický ráz.

PMK je první český korpus, který nabízí plné, systematické a pečlivě ověřené tagování všech svých forem, kde se u každého slovního druhu hlouběji rozlišuje až desítka dalších kategorií. Pro časovou náročnost celé práce část korpusu byla okódována jen redukovaně, tj. označena

jen tří až čtyřmístnými kódy zohledňujícími jevy systémové, nikoli parolové. Vedle deseti slovních druhů byly také označeny a blíže zařazeny i citátové výrazy jednoslovné i víceslovné, zkratky a propria. Byly také poprvé označeny všechny víceslovné analytické gramatické formy (*abych šel, kdyby přišel* atd.), jakož i další víceslovné výrazy, včetně partikulí, víceslovných spojek a frazémů (mimochodem, česká frazeologie je z větší částí obecněčeská).

V PMK, potažmo FSMČ, najdeme množství informací týkajících se především morfologie, lexikonu, ale ovšem také frekvence slov. Lze z něj vycházet při studiu obecněčeské morfologie, korpus nabízí možnost studia variantních podob a frekvence mnoha českých slov a tvarů. Pokud jde o lexikon, je nutno poznamenat, že četnost výskytů některých slov, obzvlášť substantiv a některých frází, může být poněkud zkreslená, neboť respondenti automaticky opakovali slova obsažená v otázkách dotazníku a navíc se vyjadřovali k předem určeným tématům. Týká se to hlavně slov a obratů jako *škola, učitelka, rodiče, děti, trestat, brát pásek, vysoká rozvodovost, postavení žen na pracovišti, mít k sobě blíž, vidět si do talíře, zaměstnání, doplňovat si klasifikaci, šťastný a úspěšný život* ap.):

- porovnáme-li 100 nejfrekventovanějších autosémantik v PMK a SYN2005, jen **čtvrtina substantiv** se vyskytuje mezi první stovkou nejfrekventovanějších slov v obou korpusech: v SYN2005 se nevyskytují jednak slova, která jsou podmíněna typem nahrávek a tematickými okruhy dotazníku (škola – *školství, třída, učitel*; pracovní podmínky – *povolání, zaměstnání, prostředí*; výchova dětí – *manželství, výchova, trest, mládež*; život ve městě vs. na venkově: *vesnice, venkov*; vzdělávání – *obor*), jednak slova typická pro **mluvený jazyk**: *lidi, kluk, holka, ženská*, *mužskej, barák*, ovšem také abstrakta *pocit* a *zkušenost,* odrážející subjektivní hodnocení při neformální konverzaci. Naproti tomu, v PMK nenajdeme mezi první stovkou slov jednak části těla, jednak slova z politicko-společenského spektra (*země, voda, zákon, právo, firma, část, skupina, řada, století, číslo* ap.);

- zhruba **tři čtvrtiny adjektiv** a **adverbií** se vyskytuje mezi první stovkou nejfrekventovanějších adjektiv a adverbií v obou korpusech: v PMK nenajdeme mezi první stovkou adjektiva mající publicisticko-politický nádech (*státní, politický, sociální, světový, veřejný* apod.), což je vzhledem ke složení textů naprosto předvídatelné, a naopak adjektiva hodnotící, intenzifikační či „identifikační“, mající osobní neformální ráz, se vyskytují pouze v mluveném korpusu (např. *normální, úspěšnej, spokojenej, náročnej, strašnej, hroznej, mužskej, ženskej* apod.). Absence některých adverbií souvisí se stylem: v psaném korpusu najdeme slova jako *již* a *téměř*, která se v mluveném jazyce nepoužívají, a naopak v mluveném korpusu se vyskytují varianty či synonyma typické pro mluvený jazyk jako *dneska, ňák, eště, trošku, teďka, hrozně, dycky, furt* ap.;

- **více než 90 %** nejfrekventovanějších **sloves** v PMK se vyskytuje mezi nejčetnějšími 100 slovesy i v psaném korpusu. Jen 7 sloves z PMK nenajdeme v psaném korpusu: *líbit se, jezdit, číst, bydlet, učit se, věnovat se* a *trestat* – pouze poslední sloveso přímo souvisí

s tematickými okruhy dotazníku (téma výchova děti), ostatní jsou spjatá s typickými tématy neformální konverzace: záliby, cestování a bydlení; v PMK se nevyskytují mezi první stovkou slovesa *patřit, vést, uvést, získat, znamenat, rozhodnout, objevit, cítit a ukázat*.

Výskyt partikulí, na rozdíl od zmíněných autosémantik a některých frazémů, není však ovlivněn žádnými vnějšími faktory, a tudíž jejich frekvenci v PMK lze považovat za plně vypovídající a nosnou.

Další kvantitativní údaje

Na základě materiálu PMK jsme poprvé získali cenné kvantitativní údaje:

- partikule jsou po slovesech a zájmenech 3. nejfrekventovanějším slovním druhem v mluveném jazyce;
- frekvence partikulí v psaných korpusech je však mnohem nižší – v SYN2005 je najdeme na předposledním místě před citoslovci;
- nejvyšší rozdíly v distribuci slovních druhů ve dvou typech korpusů kromě partikulí vykazují též zájmena – která jsou mnohem četnější v mluveném korpusu, a substantiva a adjektiva – která jsou několikanásobně četnější v korpusu psaném. Distribuce ostatních slovních druhů nevykazuje v obou korpusech tak markantní rozdíly;
- na vysoké frekvenci tvarů partikulí se však podílí poměrně malý **počet lemmat**: v PMK se vyskytlo 400 lemmat partikulí – těsně před partikulemi se umístily číslovky, za partikulemi interjekce, konjunkce, zájmena a prepozice. Jen 20 % lemmat partikulí má však frekvenci vyšší než 100 výskytů (připomeňme si prvních 20 nejfrekventovanějších stylových lemmat partikulí v PMK: *no, jako, tak, teda, prostě, třeba, a, že_jo, asi, jo, taky, to, i, ale, vlastně, zase, no_tak, už, eště, vůbec* – poslední slovo má kolem 1000 výskytů). Až polovina lemmat partikulí má tedy i zde, v mluveném jazyce, velmi nízkou frekvenci;
- ze zmíněných 400 lemmat je jich jen 270 „původních“ a zhruba 1/3 **variant**. Varianty lze rozdělit na ustálené (např. *teprve* a *teprv*) a neustálené (*tepre* – většinou jde o nedbalou výslovnost), fonetické (*dyž* a *když*) a lexikální (*teď* a *teďka* atd.). Při rozboru materiálu jsme však zjistili, že výchozí přístup k variantám nebyl vždy opodstatněný jejich skutečným územ a funkcí: vyhraněné stylové varianty byly většinou anotovány jako 2 samostatná lemmata (např. *aspoň* a *alespoň*, *taky* a *také*). K několika dalším variantním tvarům v PMK je však potřeba přistoupit stejně: např. *spíš* a *spíše* (první dubleta v PMK má až 43krát vyšší četnost), ovšem hlavně *teda* a *tedy*, kde kromě stylového rozdílu roli hraje i jiná funkce slova: *teda* (v PMK až 91 % výskytů) hlavně v kombinaci se slovem anebo v koncové pozici, má mnohem blíže k expresivní partikuli *ale* než k faktuálnímu *tedy*, a tak ji nelze automaticky považovat za variantu *tedy*;

- na **frekvenci výskytu slovních druhů** měla vliv hlavně **mluvní situace**: v PMK v monologu (souvislém vyprávění) častěji byly použity substantiva, adjektiva, spojky a frazémy, zatímco se v dialogu (spontánním rozhovoru) častěji vyskytly zájmena, číslovky, slovesa, adverbia, partikule, citoslovce a kategorie *jiné* (převážně propria). Celkově se na výskytu partikulí výrazněji podílel i socioindex vzdělání – na rozdíl od většiny dalších slovních druhů partikule v úhrnu byly poměrně častěji použity méně vzdělanými respondenty. Z dalších socioindexů se výrazně častěji podílel pouze socioindex pohlaví na užití adjektiv: adjektiva v průměru častěji byla užita muži.

Stylová charakteristika lemmat

- základní styl neformálních ústních projevů, tj. **styl běžně mluvený** se vyskytl u 89, 5 % tvarů PMK, nejvíce u **interjekcí, partikulí** a u kategorie **jiné** (propria), z autosémentik nejvíce u verb, nejméně u adjektiv a substantiv; pokud jde o sociolingvistické parametry, styl běžně mluvený se o něco častěji vyskytl u žen (+2,2 %), mladších respondentů (+6,6 %), respondentů s dosaženým nižším vzděláním (+10 %) a v dialogu (+11,6 %)

- styl mluveného projevu výrazně se blížící stylu projevů psaných, tj. **styl mluvený = styl psaný** se vyskytl u 10, 3 % tvarů PMK, nejvíce u **adjektiv, substantiv** a **prepozic**; z hlediska sociolingvistických parametrů styl běžně mluvený se o něco častěji vyskytl u mužů (+2,1 %), starších respondentů (+6,5 %), respondentů s dosaženým vyšším vzděláním (+8,1 %) a v monologu (+11,6 %);

- styl **knižní** se v PMK vyskytl jen u 0,6 % tvarů, nejčastěji u **spojek** (např. *jelikož, leč, pakliže, tudíž, aniž* apod.), **sloves** (pasívum) a **adverbií** (např. *dosti, posléze, již* apod.); častěji se vyskytl u mužů, mladších respondentů, respondentů s dosaženým vyšším vzděláním a v monologu;

- styl **vulgární** se vyskytl jen ojediněle u **frazémů** (např. *stát za hovno*) a **substantiv** (např. *hajzl*); zatímco první dva sociolingvistické parametry jsou v obou příznakových stylech stejně četné, styl vulgární se mnohem častěji vyskytl u respondentů s dosaženým nižším vzděláním a téměř výlučně v dialogu.

Frekvence slov a tvarů v PMK

- mezi 100 nejfrekventovanějšími slovy v PMK je pouze třetina autosémantických, dvě třetiny připadají na synsémantická slova – zájmena, partikule, spojky a prepozice. Z autosémantik se nejméně, jak už víme, vyskytují adjektiva – pouze 2, a to až na samém konci seznamu (93. *dobrej* a 96. *jinej*). Dalo by se říci, že hodnotící funkci adjektiv v mluveném jazyce přebírají partikule;

- podíl partikulí na frekvenci 100 nejfrekventovanějších slov v PMK je více než třetinový; více než polovinu prvních 20 nejfrekventovanějších slovních tvarů v PMK tvoří partikule (*to, a, že, je, no, na, jako, v, tak, ale, já, tam, ty, teda, prostě, nebo, třeba, už, jo, protože*).

3. PARTIKULE PODLE FORMY A FREKVENCE – DRUH PARTIKULÍ A JEJICH STYLOVÁ CHARAKTERISTIKA

V této kapitole jsou představeny všechny typy partikulí vyskytujících se v PMK, potažmo ve FSMČ (3.1 a 3.2). V prvních dvou podkapitolách se hlavně zabýváme druhým číslem kódovníku, tj. druhem partikulí (stejném při plném a redukovaném, čili krátkém kódování). Další podkapitola (3.3) je věnována stylové charakteristice partikulí, tj. poslednímu číslu kódovníku, jež je taktéž stejné pro celý korpusový materiál. V závěru kapitoly (3.4) je seznam partikulí, jež se v PMK nevyskytly.

Skutečný **počet partikulí** je celkově mnohem vyšší než je zmíněné číslo udávající lemmata doložená v PMK (400) – v podkapitole 3.4 jsme představili dalších 45 partikulí, které se v PMK buď nevyskytly (např. *kéž, lautr, nesporně, nuže, též, zajisté* apod. – některé výrazy z tohoto seznamu lze označit za zastaralé či zastarávající, některé za knižní), anebo nebyly jako partikule anotovány (tato slova většinou byla anotována jako adverbia, přičemž zpětná analýza kontextů zřetelně poukázala na funkční homonymii těchto výrazů, např. *bezesporu, bezpochyby, jednoduše, nepochybně, počítám, škoda, údajně* apod.). Započítáme-li tedy tyto potenciálně homonymní výrazy a všechny uzuální varianty, lze říci, že je ve skutečnosti v češtině kolem 440 lemmat partikulí (jednoslovných a víceslovných).

Nízká frekvence partikulí v SYN 2005 neodráží pouze jejich nižší frekvenční podíl v psaném diskursu, souvisí též i s velice nespolehlivým automatickým značkováním tohoto slovního druhu vykazujícího vysokou míru homonymie:

- téměř polovina lemmat partikulí (48 %) je **homonymní** s dalšími slovními druhy, z čehož je nejvíce homonymních lemmat s adverbii (46 %). Homonymie tvarů v kontextu promluvy je však ještě vyšší (64 %), z čehož je jich zase nejvíce s adverbii (36 %) a přibližně stejně se spojkami (27 %) a citoslovci (26 %); homonymie tvarů je nejtvrdším oříškem při automatickém rozpoznání partikulí v kontextu. V dalším přehledu uvádíme poměrnou četnost tvarů homonymních partikulí vs. další slovní druhy:

PART : INTER	79 % : 21 %	PART : tvar V	22 % : 78 %
PART : ADV	55 % : 45 %	PART : KONJ	21 % : 79 %
PART : IF	42 % : 58 %	PART : PRON	9 % : 91 %

- při mapování druhu partikulí došlo k značnému, až pětinásobnému nárůstu primárních, tj. **nehomonymních partikulí** v kontextu promluvy, pro něž jsme ve shodě s *Kódovníkem PMK* používali ne příliš vhodného termínu partikule vlastní. Tato skupina při sestavování kódovníku počítala především s původními partikulemi a několika

dalšími výrazy, jež „nálepku“ partikule již „obhájily“, např. *asi, aspoň, dyť, snad, spíš, taky, teprve* ap. Mezi nehomonymní partikule byla pochopitelně zařazena všechna další slova, jež se v kontextu promluvy v PMK vyskytla **pouze ve funkci partikule**: např. *celkem, dokonce, dovopravdy, hlavně, každopádně, nicméně, ovšem, vlastně, teda, tedy, však* atd. Původní termín „vlastní“, obdobně jako u prepozic, příliš implikuje původ výrazu, a proto jsme zde navrhli termíny primární (=nehomonymní v kontextu) a sekundární (=funkčně homonymní). Celkově se v PMK vyskytlo až 52% lemmat nehomonymních partikulí a 36 % tvarů (jednoslovných a víceslovných); jednoslovných partikulí (včetně homonymních) je zhruba třikrát více než víceslovných, přičemž víceslovné partikule mají navíc velmi nízkou frekvenci tvarů.

Vlastní víceslovné partikule

Poprvé byly soustavně popsány v SČFI, 3. svazek. Podílí se 20 % na celkovém výskytu lemmat partikulí a jen 7 % na celkové frekvenci tvarů partikulí. Pouze 3 víceslovné partikule mají vysokou frekvenci: *že jo, no tak* a *přece jenom* (10krát méně četné než *že jo*). Nízká frekvence většiny těchto výrazů může být způsobena i metodologickými nejasnostmi ohledně zařazení těchto slov, což na druhou stranu může být dáno i velkou mírou jejich splývavosti s jinými slovními druhy, především frazémy či víceslovnými spojkami. K tomu, abychom jistou jednotku začali vnímat jako víceslovnou, musíme mít oporu v úzu. Tyto jednotky sice vznikají častým opakováním téže kombinace, pouhá kumulace jednoslovných výrazů však nemusí být dostatečným důvodem k tomu, abychom ji přiřadili statut víceslovnosti, respektive ustálenosti, většina z nich je textovými realizacemi kumulace.

Při mapování **kumulací partikulí** (od poměrně frekventované dvojnásobné až po vzácnou pětinásobnou) jsme však vystopovali řadu dalších výrazů, jež by vzhledem k svému vysokému indexu opakování v podobných kontextech a poměrně vyhraněnému významu měly být též označeny za ustálené, čili systémové **víceslovné jednotky**: např. *tak jako*, *tak asi*, *asi tak*, *to teda*, *ba přímo*, *eště navíc*, *zase naopak*, *takže vlastně*, *vůbec ne, ani ne, potom/pak eště*, *ty jo*. Ze všech faktuálních partikulí je z hlediska kombinatoriky nejflexibilnější partikule *asi* (zřejmě také pro svou velmi krátkou formu): stejně často se vyskytuje jak na první (246), tak na druhé pozici (307) kumulativního řetězce. Téměř všechny kolokáty partikulí *jako, teda* a *prostě*, vyskytnou-li se tyto partikule na prvním místě řetězce, mají jen jeden výskyt – tím, že tyto partikule nevytvářejí žádnou pevnější vazbu s dalšími slovy se jen potvrzuje jejich převážně výplňková funkce. Automatická identifikace víceslovných partikulí je však značně obtížná: řada těchto výrazů svou definitivní slovnědruhovou platnost či funkci získává až větněčlenským zapojením, případně též intonací: např. *to teda nevim; to teda sme byli uplně z toho šokovaný*; *to teda*! (zájmeno + partikule; víceslovná partikule; frazém). Naprostá většina

víceslovných partikulí představuje kontaktní kombinace, čili jejich komponenty se vyskytují těsně vedle sebe, na rozdíl od víceslovných spojek, jejíchž kombinace jsou často distanční (srov. *dyž tak*). Zde je snad jedinou výjimkou víceslovná partikule *čím-tím* a její různé variantní obměny. Je však potřeba si uvědomit, že jde o otevřenou třídu lexikalizovaných gramatických slov majících oproti jazyku psanému výraznou převahu v mluveném jazyce, o čemž svědčí i téměř 100 % podíl **stylu základního, tj. běžně mluveného**, při jejich anotaci.

Stylová charakteristika partikulí

- **základní styl** u partikulí je zastoupen o 6 % víc, než je průměr výskytu tohoto stylu zahrnující všechny slovní druhy v PMK. Z hlediska jejich formy jsou to podle frekvence: 1. vlastní víceslovné partikule (téměř 100 %), 2. partikule homonymní s adverbiem (téměř 50 %), 3. partikule homonymní s dalšími slovními druhy vč. tvarů sloves (39 %) a 4. vlastní jednoslovné partikule (37 %). Zde vidíme, že zatímco u partikulí homonymních s jinými slovními druhy a u vlastních partikulí nejsou výraznější rozdíly v distribuci základního stylu, u lemmat partikulí homonymních s adverbii se základní styl vyskytuje podstatně více, kdežto u víceslovných vlastních partikulí je základní styl zcela dominantní. Mohlo by to svědčit o tom, že u vlastních jednoslovných partikulí, na rozdíl od víceslovných, mnohem více záleží jednak na jejich formě (*dyť* vs. *vždyť*, *aspoň* vs. *alespoň*, *taky* vs. *také* apod.), která se z hlediska případného procesu partikulizace již poměrně ustálila a má často neutrální příznak, jednak zřejmě ovšem také i na kontextu promluvy, kdežto u víceslovných vlastních partikulí, které jsou většinou kombinací dvou synsémantik, jde především o lexikalizované, výrazně mluvené prostředky vyjádření. Na druhou stranu mnohem větší zastoupení neutrálního stylu u partikulí homonymních s adverbii by nás mohlo vést k další úvaze, a sice že tato skutečnost nejspíš může souviset s tím, že pragmatické funkce nově vznikajících partikulí přehodnocením funkce původně autosémantických slov v kontextu promluvy mají mnohem vyšší expresivitu, a tudíž i výraznější mluvený raz (srov. *čistě, hezky, klidně, teď* apod.);

- **stylem neutrálním** bylo označeno pouze 4 070 tvarů partikulí (151 lemmat), což představuje jen 5, 4 % celkového materiálu (např. *opravdu* 108/108, vopravdu – 609; také 63/65; *ono* 45/45, vono – 478; zejména 39/39; ano 31/31; že ano 19/19; nicméně 17/19; *spíše* 14/14, spíš – 595; alespoň 11/11; *vždyť* 7/7, dyť – 377; *nejméně* 6/6, nejmíň – 14; *prý* 6/6, prej – 152; rovněž 5/5; kupříkladu 5/6; patrně 4/5; *on* 4/4, von – 52; *oni* 3/3, voni 37; *prosím* 2/2; *zvláš* 2/3, zvlášť – 51; ostatně 2/2; *ona* 2/2, vona – 43; raději 2/2; takřka 2/2).

4. PARTIKULE VE SVĚTLE SOCIOLINGVISTICKÝCH PARAMETRŮ A DOBY VZNIKU NAHRÁVEK

Zde jsme zkoumali distribuci 72 partikulí (zhruba čtvrtiny lemmat bez variant) s četností vyšší než 100 výskytů v PMK seřazených podle frekvence z hlediska jednotlivých sociolingvistických parametrů. Zajímala nás distribuce konkrétních partikulí, pokud jde o pohlaví (Z – ženy, M – muži), věk (I – mladší, V – starší) a dosažené vzdělání mluvčích (B – nižší, A – vyšší), ovšem také preference jednotlivých partikulí v konkrétní mluvní situaci – monologu (M), nebo dialogu (D). Partikule jsme rozdělili do tří velkých skupin podle frekvenčního pásma: a) první skupina obsahuje 20 partikulí a 6 variant s frekvencí vyšší než 1000 výskytů, b) druhá skupina obsahuje dalších 29 partikulí a 8 variant s frekvencí od 200 až 1000, c) třetí skupina obsahuje 23 partikulí a 2 varianty s frekvencí od 100 do 200 výskytů (viz Tabulka 4.1 - 4.3):

- v úhrnu na 27 % partikulí socioindexy neměly žádný vliv (na užití partikule *už,* jakožto na jedinou partikulí v PMK, neměl vliv žádný z parametrů; tři parametry – pohlaví, věk a vzdělání – neovlivnily výskyt partikulí *tak, zase, i, jinak, pak, naprosto*; na *vlastně, eště, aspoň, normálně, vyloženě, ani, vono, rozhodně, bohužel, totiž, akorát, jo, zrovna, přece, co*, *nebo tak* neměly vliv dva parametry v různých kombinacích – viz více Tab. 4.8);

- na 55 % sledovaných partikulí sociolingvistické parametry měly výrazný vliv: na distribuci partikulí se jednoznačně nejvíce podílela **mluvní situace** – až 93,1 %, další tři parametry jsou zastoupeny poměrně stejně: **věk** – 69,4 %, **pohlaví** – 66,6 % a **vzdělání** – 62,5 %. Ačkoli jsou tyto tři parametry zastoupeny poměrně podobnými čísly, lze nicméně vyvodit závěr, že na užití partikulí má o něco větší vliv věk mluvčích a jejich pohlaví než dosažené vzdělání. Vzhledem k tomu, že partikule mají převážně mluvený ráz, tento závěr jistě podporuje tezi o společném jádru běžně mluveného jazyka majícím univerzální uplatnění bez ohledu na profesní zázemí mluvčích;

- v **monologu** se mnohem častěji vyskytly výrazy, kterými mluvčí podkresluje a vybarvuje svou výpověď, anebo různým způsobem strukturuje svou promluvu (např. podle frekvence: *teda, prostě, třeba, že jo, asi, vlastně, zase, samozřejmě, určitě, vopravdu, spíš, takže* ap.);

- na druhou stranu v **dialogu** se vyskytla především krátká slova, typická pro začátek výpovědi – *no, tak*, anebo konec výpovědi – *jo*, expresivní – *taky, vono, docela*, kontaktní výrazy – *viď, víš*, evaluativní – *stejně, dyť* apod. (viz Tab. 4.6);

- pokud jde o socioindex pohlaví, zdá se, že partikule, které v PMK byly častěji použity **muži**, jsou více spjaté s kognitivním myšlením, vyšší mírou přesnosti či snahou o lepší uspořádání struktury promluvy (např. *myslim, přímo, celkem, naopak, takže, zřejmě*). Jistota u mužů byla podtržena slovy *samozřejmě*, *ovšem*, nejistota či neurčitost *asi, možná, prostě*, *jaksi*, důraz slovy *dokonce, právě*. Na druhou stranu **ženy** upřednostnily spíše pocitová, více

subjektivní a méně přesná hodnocení (*docela, nebo tak*), či informace „z druhé ruky“ (*prej*), kontaktní slova (*viď, víš*), jistotu více vyjadřovaly slovy *vopravdu, určitě, fakt, skutečně,* nejistotu *jako, snad,* při důrazu často použily partikuli *hlavně*, *taky*, *zrovna*, *dyť, přece, úplně,* při rezignaci *stejně;*

- v rámci socioindexu věku **mladší** respondenti jednoznačně nejvíc použili partikule *jako, prostě, třeba* a *vlastně*. Jistotní modalitu vyjadřovali nejčastěji partikulemi *asi, určitě* a *možná*. U nich taky nejčastěji zazněly partikule *například* a *dejme tomu*. Velké oblibě se v této věkové skupině těšily módní výrazy *fakt, normálně* a *vyloženě*. Pro **starší** respondenty jsou typické úplně jiné partikule: *že jo, taky, viď, víš, ne, snad, skutečně, teda, přece jenom, prej, no tak* apod;

- rozdíly v užití partikulí v rámci socioindexu stupně dosaženého vzdělání jsou taktéž předvídatelné: respondenti s dosaženým **nižším** vzděláním použili častěji partikule *no, a, ale, vono, viď, snad, dyť, víš, teď, jako, prostě* apod. U respondentů s dosaženým **vyšším** vzděláním se více vyskytly partikule *tedy, asi, samozřejmě, spíš, právě, možná, skutečně, rozhodně, jaksi, totiž, bohužel, zřejmě, že jo* apod.

Vzhledem k tomu, že sběr materiálu proběhl ve dvou vlnách (1988-1992 a 1994-1996) nabízela se možnost ještě jedné sondy: vystopovat případné rozdíly v četnosti užití partikulí v obou vlnách nahrávání, jež by mohly odrážet oblibu partikulí ve dvou významných obdobích, tj. v období těsně před změnou a těsně po změně společenského uspořádání. Byť se nejedná o velké rozdíly v distribuci jednotlivých slov, což při rozsahu PMK ani není možné, některá čísla svědčí o jistých dobových trendech. Nejvyšší rozdíly ve frekvenci slov **v první vlně**, respektive jejich mnohem vyšší frekvence na přelomu 80. a 90 let, jsou zaznamenány u slov *dost, eště, hlavně, i, ňák, ovšem, pravděpodobně, také, teda, řekněme, vopravdu, zřejmě*. **Ve druhé vlně** jsou to *dejme tomu, docela, fakt, jaksi, naprosto, prej, prostě, přece, samozřejmě, skutečně, spíš, tedy, víceméně, vlastně, vono, vůbec*.

Je zajímavé sledovat distribuci synonym, z nichž se některá výrazně častěji objevila buď v první, nebo v druhé vlně:

První vlna	**Druhá vlna**
vopravdu třeba / řekněme ňák pravděpodobně teda	samozřejmě / skutečně / pochopitelně / fakt dejme tomu jaksi možná tedy

Nejvyšší frekvenční rozdíly ve dvou sledovaných časových úsecích (zde si nebudeme všímat frekvence nižší než 100 výskytů, neboť tyto údaje nejsou dostačující pro věrohodnější závěry):

- **v první vlně** je to jednoznačně ***teda*** – jeho výskyt na přelomu 80. a 90 let byl v PMK víc než 2x vyšší než v polovině 90. let (1. vlna 8,04 % : 2. vlna 3,82 %);
- **v druhé vlně** se nejvyšší rozdíly vyskytly u slov ***docela, naprosto***, jejichž frekvence je 2,5x vyšší než v první vlně (obě slova jsou však ve třetím nejnižším frekvenčním pásmu – mají méně než 200 výskytů). O 2 % vyšší frekvence byla zaznamenána u slova ***prostě***, o 0,5 % u slova ***vlastně*** a o čtvrt procenta u slova ***vůbec***.

A konečně je zajímavé, že se varianty s kvantitou *jó, nó, že jó* mnohem častěji vyskytly v druhé vlně nahrávání a všechny u mladších respondentů.

5. SÉMANTICKÉ TŘÍDY A FUNKCE PARTIKULÍ

Základní typy partikulí, jež popsal F. Čermák (2008:64-65; 2010:179-181) se promítly i do značkování partikulí podle *Kódovníku PMK*. Čermák navrhl pro partikule dvě základní třídy: 1. **primární** zahrnující partikule **postojové, tj. atitudinální** – odráží kognitivní funkce buď postojové, nebo emocionální, jež naznačují dvě různé linie či vrstvy v promluvě, v širším slova smyslu postojů mluvčího (se třemi podskupinami: **faktuální**, **voluntativní** a **expresivní**) a **emocionální**; 2. **sekundární** – fakultativní, doprovodná, jež se různě s typem základním kříží, prolíná, doplňuje ho a dotváří, tj. modifikuje. Ze své podstaty (hodnocení a intenzita) je spíše subjektivní (se dvěma podskupinami: **evaluační** a **intenzifikační**). Klasifikace obsahuje i tzv. skupinu/třídu **jinou**, která měla být naplněna nejasnými příklady – v současné době se ukazuje, že je značně nesourodá, při revizi materiálu z větší části byla zrušena.

Kódovník PMK nabídl 9 tříd partikulí:

- **faktuální** (signalizuje nezaujatý vztah mluvčího k faktu, ke skutečnosti, popř. hodnotí **pravděpodobnost**): *tak, jinak, jako, teda, jo; možná, asi, spíš;* **2. faktuálně-evaluativní** (s podílem hodnocení podaného **kvaziobjektivně**): *stejně, prakticky, dyt', čím dál tím;* **3. faktuálně-intenzifikační** (s podílem hodnocení **míry, intenzity** ap.): (to) *vůbec* (není pravda), *rozhodně* (to neuznávám), (dost se vo tom píše a) *zvlášt'* (ve Vlastě), (todlecto) *ani* (neni dobrý); **4. voluntativní**: *at', řekněme;* **5. voluntativně-evaluativní** (spojení **subjektivního hodnocení s projevem vůle/lhostejnosti**): (když člověk pomyslí) *třeba* (na letce), *klidně* (by mohla), (trochu si toho mám vážit) *když už* (je jednou mám); **6. expresivně-evaluativní / intenzifikační**: *no tak, docela, tak ňák/nějak, a vůbec, jen* (koukni); **7. emocionálně-evaluativní / intenzifikační**: *bohužel, vyloženě;* **8. faktuálně-expresivní (+ evaluativní)**: *prostě, halt, vono* (je to dost jako špatný); **9. jiné/nejasné:** např. kombinace *no/nó, ale*, výčtové *pak* (má malej plat).

Třídy zachycené kódovníkem se snaží reflektovat v „čisté“ podobě jen dvě atitudinální základní třídy – faktuální (1) a voluntativní (4), další základní atitudinální třída – třída expresivní (6), ovšem také třída emocionální (7) se vyskytují jako smíšené typy s modifikujícími sekundárními třídami zachycujícími evaluaci a intenzifikaci. Kromě toho další čtyři třídy mají též smíšený charakter – faktuálně-evaluativní (2), faktuálně-intenzifikační (3), voluntativně-evaluativní (5) a faktuálně-expresivní (8). Tato poslední třída je zároveň i nejkomplexnější třídou, neboť je kombinací dvou základních typů, přičemž zahrnuje i jeden příznak modifikační, tj. evaluativnost. Poslední třída – „jiné“ (9) byla pak ponechána pro případy, které by se ani do tak detailně rozpracované klasifikace nemusely vejít. Na začátku práce na anotaci korpusu v ní byly uvedeny jen dva příklady (viz výše), během procesu anotace se stala předmětem dlouhých diskusí: její výsledná podoba v PMK svědčí o potížích, na něž anotátoři naráželi, stejně jako o různě vybočujících kontextových aktualizacích slov vyskytujících se na jejím seznamu, což je jistě důležitým svědectvím. Byla nicméně podrobena revizi, jednak proto, že je příliš rozsáhlá (cca 19 % všech příkladů), jednak proto, že obsahuje velice nesourodé partikule. Většina partikulí, jež při anotaci byla zařazena do této třídy, při revizi materiálu s přihlédnutím ke kontextům byla přestěhována do faktuálně-expresivní (+ evaluativní) třídy, k jejímuž popisku byla přidána též intenzifikace, anebo do emocionálně-evaluativní +intenzifikační třídy. Jedná se např. o partikule *fakt, naprosto, přímo, tudle, úplně*, *div, eště že, náhodou* apod.

Frekvenční přehled tvarů vyskytujících se v jednotlivých třídách:

- nejfrekventovanější třídou je třída čistě **faktuální** (47 %). Za čistě faktuální třídou je co do frekvence svých tvarů nejvíce zastoupena třída faktuálně-expresivní (14,4 %) a hned po ní faktuálně-intenzifikační (13,6 %);

- faktuální třída dohromady s dalšími třídami nesoucími v názvu jakousi faktuálnost v kombinaci s doprovodnými sémantickými odstíny – evaluativností (2. třída), intenzifikací (3. třída) či expresivitou (8. třída) – tvoří téměř 80 % materiálu. Dalších 20 % výskytu partikulí je rozloženo mezi třídou jiné (10, 5 % – v níž po revizi materiálu zbyly jen partikule *no* a *nó*) a další půlka tvoří souhrn různých tříd: voluntativní (třídy 4 a 5) – 6 %, expresivní a emocionální (v obou případech s nádechem evaluativnosti a intenzifikace – třídy 6 a 7) – 4,6 % výskytu všech partikulí;

- souhrnná frekvence všech partikulí s příznakem **evaluativnosti** je tedy 29,2%;

- souhrnná frekvence všech partikulí s příznakem **intenzifikace** 32,6 %;

- souhrnná frekvence všech partikulí s příznakem **expresivity** je 17,8 %;

- nejméně četnou je třída emocionální (třída 7 v kombinaci s evaluativností a intenzifikací) – tvoří pouhých 1,2 % materiálu;
- faktuální partikule výrazně charakterizují dvě vlastnosti: vysoká frekvence tvarů a relativně malý počet slov, která se hodně opakují, což je právě nejvíce charakteristickým příznakem pro jádro partikulí – průměrný poměr všech tvarů partikulí vydělených počtem všech lemmat pro faktuální partikule (238:1) je vyšší než pro všechny partikule v PMK (185:1);
- evaluativní a voluntativní partikule při poměrně nižších frekvencích výskytů mají vysoký počet různých lemmat.

6. SYNTAGMATIKA, KOLOKABILITA, VALENCE A BEZPROSTŘEDNÍ OKOLÍ PARTIKULÍ

Tato kapitola je věnována obecně syntagmatice, kolokabilitě a valenci partikulí, tj. jejich schopnosti vázat se k dalším slovům ve výpovědi, anebo celé výpovědi, což těsně souvisí s místem jejich výskytu. Dále je zpracována kumulace partikulí, která je z hlediska syntagmatiky především textová a zachycuje jen náhodně se opakující textové kombinace dvou a více partikulí vedle sebe, na rozdíl od systémové, jejímž výsledkem jsou víceslovné partikule. V závěrečné podkapitole je představeno bezprostřední okolí všech partikulí, které se vyskytly v PMK.

Pokud jde o **místo výskytu** partikulí, z analýzy materiálu PMK vyplynulo, že jsou téměř všechny partikule polyfunkční, resp. mohou se vyskytnout na jakémkoli místě v promluvě, přičemž se pojí buď s větou, anebo se slovem. Pouze několik partikulí, všechny ovšem s velmi nízkou frekvencí v PMK, se vyskytlo jen na začátku promluvy: *a potom, mimochodem, né aby, eště že* a *už aby*.

- partikule se třikrát častěji kombinují s větou než se slovem. Adpropoziční počáteční pozice je dvakrát častější než adlexémová počáteční, adpropoziční koncová je až třikrát častější než adlexémová koncová;
- nejsilnější pozicí, 57,2 %, je **pozice počáteční** (nejfrekventovanější: *no*, *tak*, *a*, *to*, *jako*, *teda*, *prostě*, *no tak*, *ale*, *třeba*);
- na **koncovou pozici** připadá zhruba 13 % výskytů (nejfrekventovanější: *že jo, jo, no, viď, teda, nó, víš, ne, taky, jó*);
- zhruba třetina partikulí se vyskytla **uprostřed promluvy** (nejfrekventovanější: *teda, jako, prostě, asi, taky, třeba, že jo, zase, tak, vlastně*);

- nejvýraznější **adlexémové partikule** jsou *dokonce i, až, i, zvlášť, přímo, zejména, doslova* a *třeba i*, všechny další adlexémové partikule se v mnohem větší míře kombinují též i s větou.

Právě vzhledem k výrazné polyvalenčnosti drtivé většiny partikulí co do jejich místa výskytu, nás zajímaly procentuální proporce jejích výskytů na určitých pozicích:

- procentuálně nejvíce zastoupené partikule na pozici **adpropoziční počáteční** (cca 70 % výskytů) jsou *dyť, a, von/a/i/o, no tak, nicméně, víceméně, takže, navíc, teď/ka, to, jinak, tudle, aby, dyž už, ale, tak, přitom, ovšem, pak, víte*;

- na pozici **adpropoziční koncová** (cca 60 % výskytů) najdeme *že jó, viď, že jo, nebo tak, jo, že* (expresivní);

- na pozici **adlexémová počáteční** (cca 60 % výskytů) – *dokonce i, až, i, zvlášť, kór, dejme tomu, doslova, zejména, třeba i*;

- na pozici **adpropoziční jiná** (cca 50 % výskytů) – *absolutně, že* (faktuálně-evaluativní), *nevim, tak jako, přece, pravděpodobně, stejně, holt, zkrátka, totiž, klidně, normálně*.

- drtivá většina partikulí má tedy poměrně vyrovnaný poměr výskytů na různých pozicích, přičemž většinou jde o různé proporce adpropoziční počáteční a jiná v kombinací s adlexémovou počáteční (srov. Tab. 6.18 a 6.20-24).

- poslední pozice, s níž *Kódovník PMK* počítal, je pozice **blíže neurčená**: partikule na této pozici (nejfrekventovanější *jako, teda, tak, prostě, třeba, asi*) 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, nezapadá do žádného kontextu, mluvčí o ni spíše zavadí, přičemž se tok jeho myšlenek ubírá jiným směrem. Tento seznam by mohl být zajímavý pro neurolingvisty, kteří zkoumají, jak se slova vzájemně kombinují a zapojují do větších celků, frází a vět, respektive jak mozek produktora přitom zpracovává strukturální a sémantické celky, na jejichž základě jsou tyto informace adresátem posléze dekódovány.

Zde platí pouze jediné pravidlo: pokud je silná počáteční pozice, není koncová a obráceně. Adlexémové partikule vznikají ad hoc v kontextu, jsou tedy nepředvídatelné, jejich kolokace nejsou pevné, a tudíž nemají vysokou frekvenci. Všechny kolokáty adlexémové počáteční partikule *teda* mají frekvenci 1! Ustálené víceslovné kombinace můžeme tedy čekat jen u adpropozičních partikulí.

Textové kumulace partikulí

V celém korpusu se poměrně často vyskytly dvě a více partikulí těsně vedle sebe tvořící různě dlouhé řetězce partikulí. Pokud tyto výskyty nebyly označeny za víceslovnou jednotku (ta je

zpravidla kombinací dvou partikulí), která je ustálená a již považujeme za systémový jev, jedná se o textové kumulace. Zhruba 1/8 výskytů partikulí připadá na různě dlouhé řetězce partikulí – od dvou do pěti slov za sebou:

- nejpočetnější skupinou je řetězec dvou partikulí – celkem se vyskytlo 6948 takovýchto kontextů, což je 9,4 % všech výskytů partikulí. Řetězec tří partikulí je několikanásobně nižší – jedná se o 1836 kontextů, což činí 2,4 %. Ještě méně často se vyskytuje čtyřnásobná kumulace partikulí – 255krát, což je jen 0,3 %. A konečně pětinásobná kumulace partikulí se vyskytuje jen 35krát, což je pouhých 0,04 % všech výskytů partikulí.

- nejfrekventovanější kombinace dvojnásobné kumulace*: no a, no ale, tak jako, teda jako, to jako, teda vopravdu, a teď, a tak* apod.

Partikule se seskupují do řetězců, nepravidelných a nahodilých, na jejichž výskyt, stejně jako na jejich složení, má vliv celá řada faktorů. U čtyřnásobných kumulací jsme se pokusili blíže je specifikovat, i když tím zdaleka toto téma nebylo vyčerpáno, nýbrž spíše nastíněno. Pro monolog každopádně platí, že v případě jakéhokoli narušení plynulého toku vyprávění, na němž se mohou podílet jak procesy tvorby a strukturace myšlenek obecně, včetně různých digresí, až po schopnost adekvátního verbálně kognitivního vyjádření, ovšem také i různé vnější faktory ovlivňující komunikační situaci, jsou řetězce partikulí to první, co se mluvčímu nabízí jako jakýsi záchytný bod před tím, než si myšlenky znovu uspořádá. U dialogů je situace složitější tím, že jsou ve hře dva mluvčí a jejich vzájemná interakce. Řetězcem partikulí recipient často reagoval souhlasně či odmítavě na stanovisko produktora, anebo usměrňoval dialog na nové téma.

Bezprostřední okolí partikulí

Jednoznačně nejfrekventovanějším **levým doplněním** partikulí je **interpunkční znaménko** které tvoří až 39 % všech levých doplnění partikulí (vysoký výskyt čárky, tečky a otazníku před partikulemi jednoznačně potvrzuje, že se partikule zdaleka nejčastěji vyskytují na začátku celé promluvy anebo dalšího promluvového úseku). Dalšími nejčastějšími slovními druhy vyskytujícími se před partikulemi jsou další **partikule** (11 %), **verba** (11%) a spojky (9 %):

- z partikulí se nejvíce vyskytují *tak* (nejčastější kombinace: *tak jako /asi, /prostě /už*), *no (no a /ale /asi /prostě /dyť /možná), ale (to /jako /zase /jinak), jako* (*jako jo / třeba /taky /teda* atd.), *teda (teda jako /samozřejmě /vlastně /vopravdu / že jo* atd.), *asi (asi tak / taky* atd.), *prostě (prostě jako /tak /teda* atd.), *třeba (třeba jako /taky /že jo* atd.), *jo (jo a /jako /prostě* atd.*), taky (taky jako /vlastně*) atd.;

- z 11 % verb vyskytujících se těsně před partikulemi téměř polovinu tvoří různé tvary verba *být* s nejčastějším tvarem *je*. Druhým nejfrekventovanějším verbem je *mít* – jeho celková frekvence je však mnohem nižší – 1,13 %;
- nejčastějším slovním tvarem vyskytujícím se před partikulemi je *to* – z celkového počtu výskytů 8 % jich připadá na partikuli, jsou to příklady kumulací dvou partikulí (nejčastější *to jako, to už*).

Za partikulemi se více vyskytují prázdná pozice (reflexivní částice), verba, zájmena, adverbia, předložky, adjektiva, frazémy, číslovky, citoslovce a kategorie jiné (propria např.). Naproti tomu v porovnání s levým doplněním, za partikulemi je méně interpunkcí (3x méně) a spojek. Porovnáme-li procentuální podíly slovních druhů a dalších kategorií **před** a **za** partikulemi, stejný podíl mají pouze další **partikule** – 11 %.

Deset nejfrekventovanějších levých a pravých doplnění:

Levé doplnění		**Pravé doplnění**	
,	18298	,	7601
.	6773	to	2970
to	3004	**jo**	2622
a	2527	.	2354
že	2016	tak	2157
je	1748	**a**	1628
..	1624	je	1329
tak	1407	**se**	1304
no	1406	**v**	1275
ale	1387	**..**	1158

Další tabulka názorně představuje největší rozdíly v distribuci slov či interpunkcí vyskytujících se jako levé či pravé doplnění partikulí. Doplnění nejsou seřazena podle frekvence, nýbrž podle míry rozdílu od nejvyšší k nejnižší mezi levým a pravým doplněním:

Výrazně vyšší levé doplnění			**Výrazně vyšší pravé doplnění**		
zleva	**lemma**	**zprava**	**zleva**	**lemma**	**zprava**
86	tom	3	10	v	1275
159	který	12	19	na	1098
63	což	5	218	jo	2622
1406	no	198	103	ty	740
118	bejt	20	97	jenom	396
396	nebo	86	107	myslim	289
372	takže	75	88	ne	254
283	protože	65	170	i	426
112	kde	31	556	jako	1113
6773	.	2354	424	já	1083

7. HOMONYMIE V KONTEXTU – NOVÁ FUNKCE SLOVNÍHO DRUHU

Pro kontextovou analýzu jsou bezesporu nejzajímavější partikule homonymní s adverbii, kterých je ostatně také nejvíc. Vzhledem k velkému objemu materiálu jsme se rozhodli, že v této kapitole výběrově blíže probereme jen jeho část, zhruba polovinu (viz 7.2 – 7.5). Při výběru jsme se řídili třemi kritérii: frekvencí slov, z našeho hlediska zajímavými posuny významu vedoucími ke vzniku nové funkce slova v kontextu, čili homonymní dvojice slov, a v neposlední řadě jsme se snažili do bližšího průzkumu zařadit též slova, která při automatickém značkování stále ještě činí potíže spjaté s automatickým zjednoznačněním jejich slovnědruhové platnosti. První a třetí kritérium se do značné míry překrývají, neboť slova působící potíže při automatické disambiguaci mají velmi často vysokou frekvenci.

Jedním z dalších výsledků, k nimž jsme dospěli při kontextové analýze výrazů homonymních s adverbii, je poměrný **nárůst míry partikulizace** řady původních adverbií. Týká se to především výrazů *absolutně, akorát, konečně, maximálně, minimálně, naprosto, navíc, ňák, specielně, zdaleka* ap. (srov. Tab. 8.3). Některé výrazy naopak posílily ve funkci adverbia, například *přímo* nebo *už*, u něhož vnímáme poměrně silně vyhraněný časový význam, na rozdíl od autorů *Statistiky češtiny*, kteří na *už* nahlíželi pouze jako na partikuli. Po započtení přehodnocených tvarů adverbií se celková frekvence partikulí v PMK zvýšila zhruba o 1000 dalších tvarů.

Po analýze kombinatorického profilu několika desítek partikulí, jež jsou v PMK homonymní s adverbii, a kontextu, v němž se realizuje jedna ze dvou funkcí, lze vyvodit několik závěrů:

- u většiny výrazů je pro jejich zařazení do jedné z těchto funkcí naprosto klíčovým faktorem jejich **sémantické okolí**, často bezprostřední sémantická kombinace slov. V případě některých výrazů – např. *absolutně, čistě, klidně, pomalu, stejně, úplně* a *vážně* lze vytvořit seznamy nejběžnějších kolokací pro obě funkce slova. Jsou to především výrazy s vyhraněným významem, u nichž nedochází k jeho oslabení, nýbrž k nové sémantické kombinaci, v níž je patrný zcela jednoznačný významový posun (*absolutně* → *naprosto, vůbec ne; čistě* → *jen; klidně* → *bez problémů; pomalu* → *skoro, už; stejně* → *tak jako tak; úplně* → *naprosto*);

- v dalších případech signálem partikulizace slova může být **vybočení z pravidelné kombinatoriky slova** v jeho adverbiální funkci (např. *docela síla, maximálně do konce roku, přímo božský*). Dalším signálem partikulizace mohou být často se opakující **neobvyklé kombinace**, v nichž dochází buď k přehodnocení původního významu, anebo k významové redundanci (např. *dost dobře možné, absolutně nejlepší*);

- kromě sémantiky v některých případech roli hraje i **intonace,** případně **slovosled**, zvláště u *klidně*. Seznamy nejběžnějších kolokací, stejně jako nově vznikající kombinace, které jsou často důsledkem nepravidelné kombinatoriky slova, jsou významnými záchytnými body při automatickém rozpoznání partikulí v kontextu;

- pokud jde o **bezprostřední okolí slova,** rozdíly mezi kombinatorikou v případě adverbií či partikulí se vyskytují zhruba u poloviny výrazů. Dost často jde o podobné kombinatorické možnosti, ovšem s odlišnou frekvencí jejich realizace v obou případech: týká se to např. partikulí *ňák, naopak, normálně, pak, potom, zas/e*, v jejichž okolí se mnohem častěji vyskytují partikule než v případě, jsou-li se ve své původní slovnědruhové platnosti. Je to doprovázeno i výskytem pro novou funkci běžných kombinací či přímo kolokací, jak už jsme na to upozornili výše (např. *ňák moc, ňák* + PREP; *normálně* + S / PREP; *asi naopak, ba naopak, jako naopak, zase naopak, eště naopak,* evaluativní fráze *naopak mám pocit / připadá mi / zdá se mi / myslím si*; *pak/potom další, druhý, eště, zase; potom je, potom si myslím; a / ale / i když / třeba + zase*-PART+ *je / na druhou stranu / naopak; není zas tak* +ADJ apod.);

- v případě partikulí *teď* a *nakonec* je pro jejich jednoznačné určení někdy zapotřebí přihlédnout i k širšímu kontextu. Zde navíc mohou hrát roli i **gramatické kategorie verba**, hl. čas a způsob): v případě adverbia, kdy jde o lineární vyprávění se u *nakonec* nejčastěji vyskytuje préteritum, u *teď* prézens a futurum; ve funkci partikule ve výčtu či hodnocení se u *nakonec* často vyskytuje kondicionál (typický i pro řadu dalších partikulí, např. *pomalu*), zatímco se u partikule *teď* ve vyprávění často objevuje préteritum;

- někdy je pro funkci partikule charakteristický i výrazně **negativní kontext**, v němž se partikule, na rozdíl od stejně znějících adverbií, vyskytují mnohem častěji: jde především o výrazy *absolutně*, *ňák*, *naprosto* a *stejně*, méně často *pomalu;*

- v případě modálního výrazu *možná* při určování jeho slovnědruhové platnosti sehrála roli grafická stránka, čili **způsob zápisu**: *možná, že* mnohem častěji bylo anotováno jako adverbium, na rozdíl od *možná* těsně zakomponovaného do proudu řeči;

- na funkci partikule se v kontextu mluveného projevu může podílet i **narušená větná stavba** (nedokončená výpověď, elipsa, různá vybočení ze souvislého toku vyprávění, digrese, tápání při hledání lepší formulace a adekvátnějšího uchopení myšlenky): např. *to je taková ta vtloukaná furt jasně, dobře, počkej, počkej* ap.

- pouze čtyři výrazy – *třeba, akorát, tak* a *jinak* – mají zcela **jiný kombinatorický profil**, a tudíž je jejich rychlá identifikace v kontextu poměrně snadná: je + ***třeba*-ADV** + INF se vyskytuje až v 75 % případů; po partikuli *třeba* se infinitiv sice může vyskytnout, nemůže ji však v tomto případě předcházet tvar verba *být*: např. *každý z nich něco umí, třeba učit děti ve škole.* Nejčastější kombinatorický profil pro další tři partikule je IZ/KONJ/PART + PART a časté kombinace jsou např. *a/ale/no/asi + tak*-PART, *akorát/jinak si myslím* apod. *Třeba* ve funkci partikule má dva významy: voluntativní nebo modální. *Akorát, tak* a *jinak* jako partikule mají hlavně spojovací funkci a vyskytují se převážně na začátku promluvy.

Partikule se obecně mnohem častěji vyskytují **na začátku výpovědi**, což v plynulém proudu vyprávění při zápisu mluveného jazyka není na první pohled patrné. Iniciálovou pozici však často signalizují interpunkční znaménka, jež se v případě partikulí jednoznačně vyskytují mnohem častěji, než jde-li o adverbium.

8. PŘEHLED VÝSLEDKŮ – PARTIKULE V SYN2005 A V SLOVNÍCÍCH ČESKÉHO JAZYKA V POROVNÁNÍ S PMK

V této kapitole porovnáváme výskyt partikulí v PMK s jejich výskytem v korpusu SYN2005 a zároveň uvádíme, do jaké míry jsou partikule zohledněny v slovnících českého jazyka.

PMK-SYN2005

Mimo jiné jsme porovnali výskyt prvních 100 nejfrekventovanějších slov v obou korpusech:

- pouze **pětina slov** se vyskytuje v obou korpusech v podobném pořadí (slova jsou uvedena abecedně): *ale, asi, aspoň, ať, hlavně, eště **(ještě)**, jenom, klidně, naopak, naprosto, ne, pak, pochopitelně, radši, rozhodně, skutečně, stejně, tak, to, vůbec, zkrátka, zvlášť;*

- další pětina slov se vyskytuje v srovnatelném pořadí (slova jsou uvedena podle pořadí v PMK, číslo za pomlčkou je pořadím v SYN2005): *třeba* 6-25, *taky* 11-23, *vlastně* 12-26, *i* 14-1, *už* 20-2, *samozřejmě* 22-33, *vopravdu* 23-31 ***(opravdu)**, ani* 24-6, *určitě* 25-40, *právě* 27-12, *snad* 30-19, *možná* 33-24, *dyť* 37-54 ***(vždyť)**, skutečně* 41-37, *že* 43-57, *bohužel* 56-70, *prej* 59-32 ***(prý)**, co* 64-49, *celkem* 68-50, *přímo* 76-63, *prakticky* 84-68, *jakoby* 96-78

Partikule typické pro **psaný korpus**:

- **slova nevyskytující se v první stovce v PMK (35)**

  alespoň, copak, doslova, jen, jistě, již, když, konečně, naštěstí, nejméně, nejspíš, nepochybně, nicméně, nikoliv, mimochodem, ono, ostatně, patrně, pravděpodobně, prosím, prý, především, přinejmenším, případně, rovněž, spíše, také, takřka, též, víte, však, vždyť, zejména, zdaleka, že

- **slova nejvíce vzdálená v PMK (12)**

  tedy 35-11, *až* 38-8, *ovšem* 58-22, *totiž* 61-13, *přece* 62-29, *například* 67-16, *dokonce* 75-20, *teprve* 91-34, *zřejmě* 78-30, *také* 94-5, *přitom* 99-45, *navíc* 100-73

Partikule typické pro **mluvený korpus**:

- **slova nevyskytující se v první stovce v SYN2005 (36)**

a tak, absolutně, akorát, dejme tomu, dost, dyť, fakt, holt, jako, jó, každopádně, myslim, nakonec, naprosto, nebo tak, né, no tak, normálně, nó, nó tak, ňák, potom, prej, přece jen, řekněme, tak ňák, takle, třeba i, viď, víceméně, víš, vono, vyloženě, zrovna, že jo, že jó

- **slova nejvíce vzdálená v SYN2005 (14)**

  no 1-43, *teda* 4-84, *prostě* 5-36, *a* 7-39, *jo* 10-67, *zase* 16-44+ *zas* 90, *spíš* 26-47 ***(spíše)***-*51, vono* 28-92 ***(ono)***, *takže* 31-66, *úplně* 39-69, *teď* 44-100, *jinak* 46-75, *jaksi* 50-98, *docela* 57-81

Kromě obecněčeských tvarů partikulí pro mluvený jazyk jsou charakteristické i víceslovné partikule, jež se v první stovce v psaném korpusu vůbec nevyskytují (souvisí to mj. i s obtížnou identifikací těchto výrazů při automatickém značkování, srov. 8.2). V mluveném korpusu se jich vyskytlo 10: *že jo, no tak, a tak, přece jen, třeba i, dejme tomu, nebo tak, že jó, tak ňák, nó tak*

Distribuce variantních tvarů v obou korpusech:

Variantní tvary	**PMK – v %**		**SYN2005 – v %**	
aspoň : alespoň	**96,6**	3,4	43,9	**56,1**
jenom : jen	**83,1**	16,9	6,6	**93,4**
nikoliv : nikoli	**68,9**	31,1	-	-
přece : přeci	**78,1**	21,9	**97,2**	2,8
spíš : spíše	**97,7**	2,3	46,1	**53,9**
teda : tedy	**91,2**	8,8	6,2	**93,8**
teprve : teprv	**76,8**	23,2	**97,7**	2,3
zase : zas	**76,1**	23,9	**78,5**	21,5

Značně vyrovnané rozložení *aspoň : alespoň* a *spíš : spíše* v SYN2005 svědčí o již oslabeném prioritním postavení tvarů *alespoň* a *spíše* v psaném projevu. Úplně stejnou četnost v mluveném i v psaném korpusu mají jen *zase* a *zas*.

PMK-SSJČ-SSČ

Oba slovníky nejméně reflektují vlastní partikule a partikule homonymní s adverbii, což zřejmě souvisí i s ne vždy jednoduše podchytitelným sémantickým posunem, k němuž při partikulizaci dochází, obzvlášť v případě nedostatku rozsáhlých dat. Je jistě zajímavé, že, jak už jsme uvedli, SSJČ např. u hesla *nakonec* jako slovní druh uvádí jen částici, u SSČ se u *nakonec* zase objevila i slovnědruhová platnost příslovce. Posuny ve vnímání slovnědruhové platnosti mnoha výrazů lze sledovat z tabulek v 8. kapitole.

Na základě dat z PMK je vidět, že se v SSČ nevyskytuje více než polovina partikulí homonymních s adverbii. Můžeme sice připustit, že u poloviny lemmat komunikační význam nemusí být tak výrazný (např. *dohromady, dost, furt, ňák, pěkně, pořád, teď, úplně*), ovšem více než polovina lemmat dle zjištění má v kontextu naprosto jednoznačně částicovou povahu (např. *absolutně* 91,9 %, *čistě* 89,6 %, *docela* 44,8 %, *doslova* 86,6 %, *fakticky* 75 %, *jinak* 58,3 %, *konečně* 62 %, *konkrétně* 80,5 %, *maximálně* 63,8 %, *minimálně* 69,2 %, *naprosto* 88,5 %, *navíc* 75,8 %, *nejmíň* 70 %, *normálně* 76,5 %, *pomalu* 69,2 %, *prakticky* 83 %, *přitom* 60,3 %, *přímo* 68,2 %, *radši* 94,9 %, *relativně* 47,6 %, *zdaleka* 96,5 %), která musela už být patrná i v době vzniku SSČ.

Poměr zastoupení jednoslovných vlastních partikulí vyskytujících se v PMK jen ve funkci partikule (nehomonymních) v obou slovnících procentuálně vypadá následovně:

Vlastní jednoslovné partikule v PMK	**SSJČ**		**SSČ**	
	nehomonymní	homonymní	nehomoymní	homonymní
	7 %	19,7 %	36 %	24 %

Závěr

Rozbor materiálu vyskytujícího se v PMK nám poprvé poskytl cenné kvantitativní údaje. Pouze zevrubná kvantitativní analýza nabízí hlubší a nezkreslený vhled do užití partikulí v mluveném jazyce, ať jde o jejich frekvenci, nejčastěji užívané tvary, poměry mezi variantními tvary, kumulační řetězce více partikulí anebo ustálené víceslovné jednotky, či místo výskytu partikulí v promluvě a jejich kombinatorický potenciál. Při popisu různých aspektů spjatých s výskytem partikulí v promluvě jsme se především snažili zachytit jejich funkci – od modální přes celou škálu funkcí faktuálních, intenzifikačních, expresivních, emocionálních až po funkci čistě kontaktní. Většina funkcí je těsně spjatá s vyjádřením postoje mluvčího vůči obsahu sdělení či partnerovi v promluvě. Neméně důležitou je funkce referenční odkazující k různým úsekům promluvy. Poukázali jsme však i na nadužívání některých výrazů (populárně značených jako „berličky“ či slova výplňková) odrážejících jazykovou nevytříbenost mluvčích, často ovšem souvisejících i s určitou mírou nejistoty, váhavosti či rozpačitosti, ne nutně jen neobratnosti produktora. Materiál PMK nám umožnil studovat výskyt partikulí v rozsáhlých kontextech, v nichž bylo možné lépe osvětlit kontextovou podmíněnost vzniku nové funkce slovního druhu a ilustrovat procesy partikulizace četnými příklady z kontextu.

Na základě detailního rozboru mnoha aspektů výskytu partikulí v PMK a jejich kontextového užití nám šlo o to nabídnout ucelený pohled na tento slovní druh v mluveném

jazyce. Takto široký záběr poukázal jednak na mnohovrstvenost problematiky, zároveň však naznačil možné další směry bádání v této oblasti. Po detailním popisu druhů a funkcí partikulí, jejich místa v promluvě, jejich typického okolí, kombinatoriky, míry partikulizace některých výrazů a hlavně po zevrubné kvantitativní analýze obsažené v této práci by se výzkum měl ubírat hlavně k širším kontextům – v některých částech práce jsme na tento směr již poukázali (srov. např. 6. kapitolu). Takto nastavený směr bádání – analýza dlouhých úseků mluveného projevu – by mohl přinést i nové výsledky týkající se role partikulí v mluveném diskursu. V rozboru míry homonymie mezi partikulemi a dalšími slovními druhy v kontextu, hlavně spojkami, je bezesporu nutné pokračovat. Dluhem zůstává podrobnější zpracování modu věty, v nichž se vyskytují partikule, ovšem též lepší zachycení partikulí ve slovnících českého jazyka potažmo učebnicích češtiny pro cizince, v nichž by při osvojování partikulí důraz měl být kladen na kvalitní učebnicové texty a autentické ukázky živého mluveného jazyka. Ve všech případech pouze rozsáhlá kontextová analýza přinese další kýžené plody.

BIBLIOGRAFIE

ADAMOVIČOVÁ, A. – PEŠIČKA, J., 2006, Běžně mluvený jazyk ve výuce češtiny pro cizince, In: *Sborník Asociace učitelů češtiny jako cizího jazyka*, (eds. B. Štindlová, J. Čemusová), Praha: Akropolis, s. 81-86

ADAMOVIČOVÁ, A. – HASIL, J., 2006, Čeština jako cizí jazyk – zpracování úrovně B2, *Bohemystika*, č. 3, s. 248-254

ADAMOVIČOVÁ, A., 2007, Tudjice i turcizmi u romanu Omerpaša Latas Ive Andrića i njihov prevod na češki (Cizí slova a turcizmy v novele Ivo Andriće Omerpaša Latas a jejich překlad do češtiny), *In Balto-Slavicum Pragense*, (ed. J. Marvan), Praha: Univerzita Karlova v Praze – Filozofická fakulta; s. 129-137

ČERMÁK F. a kol. (ADAMOVIČOVÁ, A. – PEŠIČKA, J. – SAVICKÝ, P. – ŠIMANDL, J. – ŠONKOVÁ, J., 2007, *Frekvenční slovník mluvené češtiny,* Praha: Karolinum, 510 s. [FSMČ]

ADAMOVIČOVÁ, A. – ČERMÁK, F. – PEŠIČKA, J., 2008, *Čítanka mluvené češtiny*. Praha: Karolinum, 298 s.

ADAMOVIČOVÁ, A., 2010, Překlad - od interpretace k deformaci originálu a jeho dopad na korpusová data, In *Mnohojazyčný korpus InterCorp: Možnosti studia*. 1. vyd. Praha: Nakladatelství Lidové noviny, s. 250-260

ADAMOVIČOVÁ, A., 2011: Zpracování gramatiky českého jazyka v učebních materiálech češtiny pro cizince, In *NŘ* 5, roč. 94, s. 256-258.

ADAMOVIČOVÁ, A., 2011, Běžně mluvená čeština a její frekvenční slovník, aneb Jak lidi v Praze dovopravdy mluvěj, *Přednášky z 54. běhu LŠSS,* (ed. M. Hrdlička), Praha: FF UK, s. 45-62

ADAMOVIČOVÁ, A., 2011: Prepozice *za* v češtině a v srbštině: jejich funkce, význam a vzájemný poměr. In *Korpusová lingvistika 1, InterCorp*, (ed. F. Čermák), Praha: NLN, s. 144-168

ADAMOVIČOVÁ, A., 2011: Překlad poezie – poezie v překladu. In *Psáno do oblak*, Sborník k nedožitým sedmdesátinám prof. Jana Kuklíka. (eds. J. Hasil, J, M. Hrdlička), Praha: Karolinum, s. 307-312

ADAMOVIČOVÁ, A., 2012: K prezentaci prepozic ve výuce češtiny pro cizince, In *Usta ad Albim BOHEMICA* XII, č. 2, Ústi nad Labem: Pedagogická fakulta UJEP, s. 8-19. Dostupno z < http://www.pf.ujep.cz/files/Usta_2-2012.pdf>

ADAMOVIČOVÁ, A, 2013: *Nebojte se češtiny*, (3. rozšířené a upravené vydání), Praha: Karolinum, 318 s.

ADAMOVIČOVÁ, A. – IVANOVOVÁ, D., 2013: *Basic Czech I*, (3. rozšířené vydání), Praha: Karolinum, 193 s.

ADAMOVIČOVÁ, A. – IVANOVOVÁ, D., 2013*: Metodické pokyny k učebnici Basic Czech I*, Praha: Karolinum, s. 19.; Dostupno z <http://www.cupress.cuni.cz/ink2_stat/index.jsp?include=podrobnosti&id=10588&zalozka=6>

ADAMOVIČOVÁ, A. – IVANOVOVÁ, D., 2013: *Doplňková cvičení k Basic Czech I, lekce 1-6*, Praha: Karolinum, 87 s.; Dostupno z <http://www.cupress.cuni.cz/ink2_stat/index.jsp?include=podrobnosti&id=10588&zalozka=6

ADAMOVIČOVÁ, A. - HRDLIČKA, M. – IVANOVOVÁ, D., 2014: *Basic Czech II*, (3. upravené vydání), Praha: Karolinum, 270 s.

ADAMOVIČOVÁ, A. 2014: *Doplňková cvičení k Basic Czech II, lekce 7-13*, Praha: Karolinum, 72 s.; Dostupno z <http://www.cupress.cuni.cz/ink2_stat/index.jsp?include=podrobnosti&id=11969

ADAMOVIČOVÁ, A., 2014: Partikule vs. adverbia – rozpoznání slovního druhu. In *Radost z jazyků – Sborník příspěvků k 75. narozeninám prof. Františka Čermáka*. (eds. V. Petkevič, A. Adamovičová, V. Cvrček), Praha: Nakladatelství Lidové noviny, s. 188-200

ADAMOVIČOVÁ, A., 2014.: Mezijazyková interference a vymezení negativních pravidel v počátečních fázích osvojování češtiny. In *Čeština jako cizí jazyk VII. Materiály ze VII. mezinárodního sympozia o češtině jako cizím jazyku*, (eds. J. Hasil, M. Hrdlička), Praha: Univerzita Karlova, nakladatelství Euroslavica, s. 114-122.

ADAMOVIČOVÁ, A. – HAJÍČKOVÁ, Z., 2016: Rozvoj dovednosti psaní v češtině jako cizím / druhém jazyku v akademickém prostředí, *CASALC Review,* s. 94-101; Dostupné z http://casajc.cz/article/show/25

ADAMOVIČOVÁ, A – HRDLIČKA, M., 2016: *Basic Czech III* (druhé vydání), Praha: Karolinum, 324 s.

ADAMOVIČOVÁ, A. – HAJÍČKOVÁ, Z., 2016: Česká verze projektu WriLab, In *Wrilab2 A Didactical Approach to Develop Text Competences in L2,* Zürich: LIT Verlag GmbH, s. 75-101

ADAMOVIČOVÁ, A., 2017: Čeština v korpusech, *Přednášky z 60. běhu LŠSS* (ed. S. Škodová), Praha: FF UK, s. 44-60

ADAMOVIČOVÁ, A., 2017: Partikule v Pražském mluveném korpusu, *SALi,* č. 8, Praha: FFUK, s. 88-108