

# RECENZE A ZPRÁVY

---

## Ty jo!!! Fakt dost dobrá knížka o bezva tématu aneb Jak hodnotíme v češtině ????????????????????

Radek ČECH | Katedra českého jazyka FF OU

VESELOVSKÁ, Kateřina (2017): *Sentiment Analysis in Czech*. Praha: Ústav formální a aplikované lingvistiky, 159 s.

Většina našich jazykových projevů je výsledkem společenské konverzace zaměřené na udržování „sítě“ mezilidských vztahů – zkuste si třeba vybavit, o jakých tématech jste mluvili za poslední dva dny a kolik času jednotlivá témata zabrala. S jistou nadsázkou pak lze společně se Zrzavým et al. (2017, s. 120) tvrdit, že „[r]eálný jazyk je především nástroj k pomlouvání bližních“, tedy k hodnocení, vyjadřování postojů atp., nikoliv k neutrálnímu předávání informací a instrukcí. Hodnotící postoje jsou neodmyslitelně spojeny s emocemi. Ostatně vyjadřování emocí (a všeho, co s nimi souvisí) je nejčastějším projevem našeho jazykového chování, jak ukázali např. Dunbar et al. (1997, s. 240) v experimentálním výzkumu zaměřeném na to, o čem lidé mluví, srov.: „[M]ost social conversation time is devoted to statements about the speaker’s own emotional experiences and/or relationships or those of third parties not present.“ Popis výrazových prostředků sloužících k vyjadřování emocí, postojů a hodnocení není v lingvistice ničím novým. Je však třeba zdůraznit, že významný impuls pro hlubší pochopení tohoto fenoménu vzešel z oblasti počítačnické lingvistiky a souvisí s dostupností jazykových dat obsahujících evaluativní postoje na internetu po r. 2000 (blogy, zákaznické recenze, Twitter atp.), a především s praktickými potřebami marketingu, obchodu a společenské sféry. Od počátku tisíciletí se tzv. postojová analýza (*sentiment analysis*, *opinion mining*) stává jedním z nejživějších odvětví počítačového zpracování přirozeného jazyka (srov. Liu, 2012).<sup>1</sup>

Recenzovaná kniha Kateřiny Veselovské *Sentiment Analysis in Czech* je jedním z výsledků badatelského úsilí realizujícího se právě v tomto odvětví lingvistiky. Jak sama autorka uvádí v úvodní kapitole, má její kniha dva základní cíle: „The first one is to give a compact description of basic means of emotional language in Czech. The second one is to employ the findings concerning emotional language in computational applications, namely sentiment analysis, i.e. automatic extraction of emotions from text“ (s. 1). Tyto dva cíle pak určují základní strukturu knihy. V první části nazvané *Linguistic Structure* (kapitoly 3–7) je prezentována poměrně důkladná analýza prostředků sloužících k vyjadřování emocí a postojů v češtině, v části druhé s názvem *Computational Application* (kapitoly 8–9) jsou pak uvedeny příklady užití tzv. postojové analýzy při identifikaci emocionálních vyjádření v textu. Těmto dvěma hlavním částem předchází kapitola 2, v níž je představen teoretický rámec zasazující knihu do širšího kontextu jak české, tak světové lingvistiky.

---

<sup>1</sup> Pro seznámení se základními charakteristikami tohoto badatelského směru viz heslo „postojová analýza“ v Novém encyklopedickém slovníku češtiny zpracované Kateřinou Veselovskou (2017).

Součástí knihy jsou také přílohy obsahující popisy týkající se charakteru dat, způsobu jejich anotace a použité statistiky. Nejen samotná struktura knihy, ale také velmi srozumitelný způsob, jakým je napsána, dávají čtenáři velmi dobrou možnost proniknout do „tajů“ postojové analýzy. Navíc, autorka prezentuje celou problematiku jako proces, ve kterém předkládá nejen hotové výsledky svého bádání, ale popisuje i badatelské problémy, nástrahy a slabá místa. To činí knihu velmi čtivou, aniž by to bylo jakkoliv na úkor odborné kvality, ba právě naopak – čtenář má možnost sledovat, jak autorka o daných fenoménech přemýšlí a jaká řešení nabízí, což je nepochybně inspirativní.

V části zaměřené na analýzu prostředků sloužících k vyjadřování emocí postupuje v souladu s jednotlivými jazykovými plány. Zkoumá prostředky lexikální, morfosyntaktické, sémantické a pragmatické. Z lingvistického hlediska se mi jeví jako nejzajímavější kapitoly zaměřené na morfosyntax a sémantiku, v nichž se autorce daří ukázat, jak gramatické a sémantické struktury ovlivňují vyjadřování hodnotících postojů. Formuluje obecná pravidla popisující význam těchto struktur ve sledovaných projevech (např. „if you find an evaluative adjective which is a part of a verbominal predicate, its subject should be a target of evaluation“ (s. 37)), což jí umožňuje efektivně využít získané teoretické poznatky v praktických aplikacích. Velmi přesvědčivě působí začlenění poznatků valenční syntaxe do popisu fungování sloves vyjadřujících emoce. Při analýze sémantických a pragmatických faktorů se dostává až k samotným hraničním možnostem výzkumu a zejména jeho aplikace – jde například o rozdílné interpretace hodnotících vyjádření z hlediska mluvčího a příjemce textu či o vliv situačního kontextu a ironie. V závěru teoretické části je také představen způsob, jakým je postojová analýza implementována do Pražského závislostního korpusu, a dále navržen způsob, jak zachytit hodnotící struktury prostřednictvím formalismu konstrukční gramatiky. Zejména návrh anotace a její začlenění do Pražského závislostního korpusu je cenným příspěvkem celého výzkumu, protože umožňuje nejen pozorovat syntaktické charakteristiky hodnotících prostředků a jejich vztah k jiným jazykovým charakteristikám, ale dovoluje zejména testovat hypotézy související s postojovou analýzou.

Druhá část knihy je zaměřena na prezentaci praktické aplikace postojové analýzy. Nejdříve je představena tvorba slovníku hodnotících výrazů češtiny – Czech Subjectivity Lexicon (Veselovská – Bojar, 2013) –, dále je detailně popsán způsob anotace tří různých jazykových databází (Aktualne.cz, Česko-Slovenská filmová databáze, Mall.cz) a její statistické vyhodnocení. Autorka otevřeně odhaluje problémy, které se při manuálním anotování objevily, a popisuje způsoby jejich řešení. V této části ukazuje, že anotace sémantických charakteristik jazyka je opravdu netriviální problém. Jádrem druhé části knihy je pak samotná postojová analýza, v níž jsou představeny jednotlivé metody a kriticky zhodnoceny výsledky. Konkrétně jde o automatickou detekci polarit (automatic polarity detection) a o identifikaci cílů hodnocení (opinion target identification) v češtině. Kromě češtiny, jíž je věnována rozhodující část kapitoly, je analyzována také angličtina a v případě použití neuronových sítí i dalších šest jazyků. Také zde je vše podáno velmi přehledně, srozumitelně a s kritickým nadhledem.

V závěrečné části knihy jsou nejen shrnuty výsledky (kapitola 10), ale také nastíněny perspektivy dalšího výzkumu (kapitola 11). Autorka zde potvrzuje to, co je patrné při pročitání celé knihy – jde o záznam určité části badatelského procesu, který nekončí a před kterým stojí důležité otázky, jež jsou pro ni výzvou. I proto se jedná o knihu, která stojí za přečtení.

## LITERATURA

- DUNBAR, Robin I. M. – MARRIOTT, Atlanta – DUNCAN, Noel D. (1997): Human conversational behavior. *Human Nature*, 8(3), s. 231–246.
- LIU, Bing (2012): *Sentiment Analysis and Opinion Mining: Synthesis Lectures on Human Language Technologies*. Morgan and Claypool Publishers.
- VESELOVSKÁ, Kateřina – BOJAR, Ondřej (2013): *Czech SubLex 1.0* [online; LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University]. Cit. 18. 4. 2019. <<http://hdl.handle.net/11858/00-097C-0000-0022-FF60-B>>.
- VESELOVSKÁ, Kateřina (2017): Postojová analýza. In: Petr Karlík – Marek Nekula – Jana Pleskalová (eds.), *CzechEncy – Nový encyklopedický slovník češtiny* [online]. Brno: Masarykova univerzita. Cit. 18. 4. 2019. <[https://www.czechency.org/slovník/POSTOJOVÁ\\_ANALÝZA](https://www.czechency.org/slovník/POSTOJOVÁ_ANALÝZA)>.
- ZRZAVÝ, Jan – BURDA, Hynek – STORCH, David – BEGALLOVÁ, Sabine – MIHULKA, Stanislav (2017): *Jak se dělá evoluce: Labyrintem evoluční biologie*. Praha: Argo – Dokořán.

Katedra českého jazyka FF OU  
Reální 5, 701 03 Ostrava  
cechradek@gmail.com

## Bohemistické disertace roku 2018

Dissertations on Czech language of/from the year 2018

Martin BENEŠ, Ondřej DUFEK, Martina RYBOVÁ | Ústav pro jazyk český AV ČR

Michal HOŘEJŠÍ | Ústav českého jazyka a teorie komunikace FF UK

Naše řeč si v rubrice Recenze a zprávy tradičně klade za cíl informovat o nejnovějším dění v oboru jazykovědné bohemistiky. Vedle jednotlivých zpráv o konferencích, sériích přednášek apod. jsme se rozhodli začít zpravovat také o nejnovějších obhájených disertačních pracích spadajících do našeho oboru a jeho blízkého okolí. Disertace totiž podle našeho mínění do jisté míry mohou odrážet, či dokonce předjímat to, co se zanedlouho může stát nosným tématem či metodologickou perspektivou.

Vyzvali jsme proto všechna česká pracoviště, která realizují doktorské studijní obory vlastní či potenciálně blízké lingvistické bohemistice, aby nám poskytla soupis disertací obhájených v roce 2018. Tuto výzvu budeme každoročně opakovat. Z obdrženého souhrnu údajů vždy vybereme a shrneme ty práce, které jsou pro obor relevantní. Vybrané disertace řadíme podle pracovišť a v závěru doplňujeme abecední přehled obsahující školitele a oponenty.