

Obsah / Content

Články / Articles

Kodifikujme deskriptivně – a pěstujme tím jazykovědu! Reakce na článek Luboše a Vojtěcha Veselých

Let's codify language descriptively – and by doing that, let's practice linguistics!
A response to paper by Luboš and Vojtěch Veselý /3

Martin Beneš, Ondřej Dufek

Distributiva typu *pobít*, *pomřít*, *povyhazovat* v současné češtině

Distributive verbs such as *pobít* “to kill all”, *pomřít* “to die completely” (about a group of living beings), *povyhazovat* “throw everything away” in contemporary Czech /19

Ondřej Bláha

Variantní tvary sloves vzoru „sází“ a procesuální adjektiva od nich utvořená

Variant verb forms of the type “sází” and processual adjectives derived from them /27

Tereza Nádassy, Robert Adam

Indikativ a kondicionál po spojce *aniž*

The indicative and the conditional moods in sentences with the conjunction *aniž* /38

František Štícha

Několik poznámek ke kontradiktorické antonymii u sloves

Several remarks on the contradictory oppositions with verbs /50

Vojtěch Veselý

Recenze / Reviews

Jana Bílková, Ivana Kolářová, Miloslav Vondráček (eds.): *Lingvistika – korpus – empirie*

Praha: Ústav pro jazyk český, 2020 /66

Eva Höflerová

Zprávy / News

Co je v ČNK nového XI /74

Dominika Kovářková

Pokyny pro autory /77

Instructions for authors /78

Kodifikujme deskriptivně – a pěstujme tím jazykovědu! Reakce na článek Luboše a Vojtěcha Veselých

Martin Beneš, Ondřej Dufek

Ústav pro jazyk český AV ČR, v. v. i., Praha

benes@ujc.cas.cz; dufek@ujc.cas.cz

Let's codify language descriptively – and by doing that, let's practice linguistics! A response to paper by Luboš and Vojtěch Veselý

ABSTRACT: The paper responds critically to the article *Nekodifikujme, raději popisujme a vysvětlujme normu a úzus. Pěstujme jazykovědu, nepěstěme jazyk* (Let's not codify language, let's describe and explain its norms and usage instead: let's practice linguistics instead of refining language) by Luboš and Vojtěch Veselý (2019). We find the authors' statements problematic, hence our aim is to diagnose the argumentation malady they suffer from and propose complex treatment. In the initial text, we identify a "polemic succession syndrome" manifested by a few symptoms: *ignorance* of substantial part of language planning's field of interest, *lack of differentiatedness* or its confusing/misleading simplification, *resignation* from the challenges related to its research due to their difficulty, *rashness* in constructing the relation between corpora evidence (language usage) and language awareness/knowledge (norm) and *one-sided approach to language users* which the authors try to protect from accusation of lack of competence due to unclear reasons.

KEYWORDS: codification; language norms; language planning; language rules; language usage; competition of language units; polemic

KLÍČOVÁ SLOVA: kodifikace; normy; úzus; jazyková pravidla; jazykové plánování; konkurence jazykových prostředků; polemika

Distributiva typu *pobít*, *pomřít*, *povyhazovat* v současné češtině

Ondřej Bláha

Ústav pro jazyk český AV ČR, v. v. i., Praha

o.blaha@ujc.cas.cz

Distributive verbs such as *pobít* “to kill all”, *pomřít* “to die completely” (about a group of living beings), *povyhazovat* “throw everything away” in contemporary Czech

ABSTRACT: Distributive verbs are a semantically significant but infrequent category in contemporary Czech. The centre of this category are several dozens of verbs with the prefix *po-* and with the obligatory plural object (or subject). Within the SYNv8 corpus, the author found fifty-eight of these verbs occurring in journalistic and prosaic texts. The number of tokens related to these fifty-eight lemmas has been growing slightly in recent decades (1990–2014), as well as the number of verbs with two or more prefixes among these tokens. Compared to verbs as a whole, distributive verbs are characterized by certain specifics in grammatical behavior – they more often appear in the 3rd person, in plural and in passive voice.

KEYWORDS: Czech; grammar; verb; aktionsart; distributive verbs; prefixes

KLÍČOVÁ SLOVA: čeština; gramatika; sloveso; aktionsart; distributiva; prefixy

Variantní tvary sloves vzoru „sází“ a procesuální adjektiva od nich utvořená

Tereza Nádassy, Robert Adam

Katedra českého jazyka Pedagogické fakulty Univerzity Karlovy, Praha

tereza.nadassy@gmail.com

robert.adam@pedf.cuni.cz

Variant verb forms of the type “sází” and processual adjectives derived from them

ABSTRACT: Based on the corpus data of SYN2015, the paper explores the formal variety of processual adjectives derived from the verbs of the type “sází”. Considering the existence of the doublet *-ejí/-í* (3rd person, plural, indicative, present, active) – codified in stages during the 20th century in Standard Czech – it aims to verify whether this formal variety can be attested also in the case of adjectives (*-ející/-ící*). The results indicate major differences between the verbs with doublets codified in 1902, 1941 and 1958 on one hand and the rest of the type, where the doublets were codified uniformly in 1993, on the other.

KEYWORDS: deverbative processual adjectives; 4th verb class; written Czech; public printed texts; doublet

KLÍČOVÁ SLOVA: deverbativní adjektiva procesuální; 4. třída přítomní; psaná čeština; veřejné tištěné texty; dubleta

Indikativ a kondicionál po spojce *aniž*

František Štícha

Ústav pro jazyk český AV ČR, v. v. i., Praha
sticha@ujc.cas.cz

The indicative and the conditional moods in sentences with the conjunction *aniž*

ABSTRACT: The corpus statistics show that: (a) since 1990, the conditional mood with the conjunction *aniž* has been on the rise; (b) the conditional mood is more frequent in journalistic and technical texts than in fiction; (c) in fiction, the share of the conditional mood varies between different authors; this indicates that the semantics of the conditional mood in sentences with the conjunction *aniž* loses its validity with the progress of time; (d) the conditional mood in the clause with *aniž* is almost a norm in sentences containing the conditional mood in the head clause. The indicative mood in clauses with the conjunction *aniž* is used or prevails in contexts which suppress the hypothetical content of the clause due to its facticity.

KEYWORDS: indicative; conditional; mood; conjunction *aniž*; hypothetical

KLÍČOVÁ SLOVA: indikativ; kondicionál; slovesný způsob; spojka *aniž*; hypotetický

Několik poznámek ke kontradiktorické antonymii u sloves

Vojtěch Veselý

Pedagogická fakulta, Univerzita Hradec Králové

vojtech.vesely@uhk.cz

Several remarks on the contradictory oppositions with verbs

ABSTRACT: The paper deals with the question whether contradictory oppositions occur not only between state verbs (*mít* 'have' – *chybět* 'lack'), process verbs (*spát* 'sleep' – *bdít* 'be awake', *mluvit* 'speak' – *mlčet* 'be silent') and event verbs (*opustit* 'leave' – *zůstat* 'stay, remain', *vzít* 'take' – *nechat* 'leave') but also in such cases where one verb denotes a state/process and the other verb denotes an event which terminates the respective state/process. There could potentially be many such pairs. The paper sheds light on three pairs of verbs of this kind: *zazlívat* 'blame' – *odpustit* 'forgive', *pamatovat si/se* 'remember' – *zapomenout* 'forget', *žít* 'live' – *zemřít* 'die'. Of them only the pair *žít* 'live' – *zemřít* 'die' can be considered contradictory in the systemic sense. However, the categorial discrepancy between both verbs is alleviated by the fact that the respective state is conceptualized as an event. In the case of the other two pairs of verbs, the state/process verbs are conceived of as event verbs only in particular contexts, therefore there is no contradiction in the systemic sense. The author considers it probable that in Czech there are not many verbal pairs similar to the pair *žít* 'live' – *zemřít* 'die'.

KEYWORDS: contradictory opposition; contrary opposition; presupposition; entailment; state verb; process verb; event verb

KLÍČOVÁ SLOVA: kontradiktorický protiklad; kontrární protiklad; presupozice; sémantická implikace; stavové sloveso; procesuální sloveso; mutační sloveso

Recenze / Reviews

Jana Bílková, Ivana Kolářová, Miloslav Vondráček (eds.): *Lingvistika – korpus – empirie*

Praha: Ústav pro jazyk český, 2020, 244 s.

Eva Höflerová

Pedagogická fakulta, Ostravská univerzita

eva.hoeflerova@osu.cz

Koncem roku 2020 vydal ÚJČ AV ČR publikaci *Lingvistika – korpus – empirie*. Jedna z úvodních stran nese jednoduché sdělení: **K poctě Františka Štíchy**. Sedmadvacet českých lingvistů shromáždilo své aktuální studie do jednoho svazku a přihlásili se tak k obsahu naznačenému titulem a podtitulem společné publikace.

Editoři uspořádali texty do pěti kapitol charakterizovaných lingvistickou disciplínou. Autoři statí s korpusem v různé míře pracují, mnozí jsou s ním však spjati i jako tvůrci, aktivně se podílejí na jeho obsahu, podobě a uživatelské přístupnosti.

I. Korpus – Empirie – Morfologie

Oddíl morfologický otevírá **Markus Giger** studií **Několik poznámek k přítomnému činnému v ruštině a v češtině**. Autor podává informaci o srovnávání výskytu příčestí přítomného činného ve dvou slovanských jazycích. České přítomné příčestí činné má bohatou „minulost i současnost“ – odrazem toho je terminologická různost: aktivní verbální adjektivum pro současný děj / procesuální adjektivum / zpřídavnělý přechodník přítomný / -cí-ové participium verbální. Komparace jeho funkcí v češtině s funkcemi v ruštině (na základě korpusových dat) ukazuje především shodnou funkci syntaktickou – ze slovesného původu si tvary v obou jazycích ponechávají schopnost valenčně poutat předmět, z tvarového „posunu“ k adjektivům schopnost stát jako preponovaný či postponovaný přívlástek. Autor – podobně jako jiní bohemisté/rusisté – očekává od dalšího výzkumu objevení mnohých paralel a rozdílů (ty by pak mohly být součástí vysvětlení i obecnějších textových jevů), je si však vědom, že výzkum musí nastat „v trojmezí tvarosloví, slovotvorby a syntaxe“.

Petr Karlík přispívá statí **Přechodníky v kontextu nominalizací**. V úvodu jsou sděleny formální a základní syntaktické rysy přechodníku, ovšem výběrově – s cílem zdůraznit klíčová témata studie. Je to souvislost jistých (vybraných) přechodníkových forem s obvyklejším adjektivizovaným přechodníkem a dále spjatost/nespjatost subjektu v maticové a přechodníkové větě. Nemělo by uniknout pozornosti, že autor zvolil k ověření svých předpokladů mj. speciální dotazník. Ten předložil neškoleným mluvčím. Výsledky potvrdily některé autorem předpokládané trendy v chápání vlastností přechodníkové formy. Vazba na další otázky související s nominálními rysy výpovědi s přechodníkem je naznačena ještě v tématu závěrečném – souvislost subjektů věty maticové a přechodníkové v případě, že obě propozice vykazují vztah příčina – následek. Pak je ale třeba využít přechodník schopný časovou prodlevu mezi příčinou a následkem vyjádřit, tedy přechodník minulý.

Synonymní/variantní tvary infinitivu jako problém konzistentní lemmatizace je téma příspěvku **Klára Osolsobě**. Přestože byly z pracovního vzorku pečlivě odstraněny irelevantní případy, bylo nutno definovat pět slovesných skupin (typů), kde se vyskytuje variantnost v infinitivním tvarotvorném formantu. Autorka stanovuje zásady lemmatizace pro automatické zpracování sledovaného tvaru. V zásadě navrhuje buď koncept vícenásobného (variantního) lemmatu, nebo evidenci dvěma lemmaty. Pro zájemce o detailní morfologickou komparaci jsou zajímavé tabulky, v nichž jsou slovesné tvary porovnávány. Je tak přesvědčivě stanovena „oprávněnost“ použití jednoho či druhého způsobu značkování. Text je doplněn množstvím příkladů, u některých typů soupisem všech vyskytujících se jednotek.

Zjistit, jak je tvar, jenž bývá předmětem odborných statí, užíván a „prožíván“ neškolenými uživateli jazyka, je vždy zajímavé. V příspěvku **Lucie Jílkové K užívání přičestí v přísudku: 70 dotazníků** jde o konkurenci přičestí (zejména) na (e)n – typ poškozen-poškozený. Respondenti hodnotili subjektivně prožívaný významový rozdíl v předložených rodově obměněných variantách jevu. Materiál dotazníku i závěry autorky vycházejí z postulátu, že se jedná vždy o přičestí, odpovědi respondentů obsahují i pojetí „adjektivní“, tedy chápání tvaru jako vyjádření vlastnosti (jako výsledku činnosti či děje). Vedle tohoto rozdílu se však, podle mého názoru, jako podstatné jeví i to, že respondenti patrně postrádají prostředek k vyjádření „předpřítomného děje“. Hledají formu, která by vyjádřila okamžik uzavření minulého děje „v průniku“ s aktuálním okamžikem promluvy. Nacházejí ho (někteří) právě v krátkém tvaru přičestí.

II. Korpus – Empirie – Syntax

Eva Hajičová, Jiří Mírovský a Barbora Štěpánková jsou autory statí **Několik poznámek ke slovosledu a aktuálnímu členění ve světle anglicko-českého**

paralelního korpusu. Autoři sledují slovoslednou proměnlivost v obou jazycích a porovnávají ji. V češtině není slovosled svázán pevnými gramatickými pravidly, jako je tomu v angličtině. Je flexibilní, ale podléhá logice aktuálního členění větného (a to i v modifikacích pro projev mluvený). Autoři srovnávají dodržení pravidel pro užití příslovečného určení místa a času. Rozdíly, které nacházejí, jsou evidentní, odpovídají vlastnostem obou jazyků. Pozornost čtenářů jistě vzbudí i další část textu, jež referuje o umístění fokalizátorů – bodů, z nichž se nad výpovědí „rozprostírá“ úhel pohledu na sdělovanou informaci. Komparující studie českých a anglických textů se v české části materiálu opírá o velmi kvalitní překlad. Její význam pro poznání reprezentativní formy jazyka je tedy nesporný. Nabízí se otázka, zda do jaké míry se odlišují vlastnosti sledovaných jevů v materiálu mluvených textů.

Reciproka ve slovníku a syntaxi je název studie, pod níž jsou podepsány **Jarmila Panevová, Markéta Lopatková a Václava Kettnerová.** Formanty/slova *se a si* jsou polyfunkční. Bohatě uplatnění morfologické, lexikální a syntaktické způsobuje, že se při popisu těchto formantů neužívá přesná (jednotná) terminologie, konstatují jinými slovy autorky. V textu se zaměřují na jeden typ funkce těchto slov – reciprocitu sloves. Vymezují tři skupiny sloves s rysem reciprocitu, přičemž skupina třetí je předmětem dalších výkladů – výkladů o získání reciprocitu syntakticky, tzv. reciprokací (u sloves *fotografovat, ošetřovat, poranit*). Autorky přesně vymezují podmínky reciprokace. Oprávněnost svého postupu dokládají komparačně – zdůrazňují sémantiku morfologicky a lexikálně obměněných recipročních vyjádření. Formanty *se a si* tak zde získávají přesnou morfologickou charakteristiku.

Přesnost, systémovost, systematičnost, to vše s vytříbeným smyslem pro jazykovou realitu. Výše konstatované rysy stojí za textem **Akuzativní reke předložky „o“ a její identifikace při morfologické anotaci českých textů.** Autorem je **Vladimír Petkevič.** Dvojí reke předložky „o“ (zajímavé je kvantitativní porovnání výskytu předložky v akuzativu a v lokálu) je morfologický jev, který je nutno analyzovat „z vnějšku“, pomocí rozklíčování syntaktických a sémantických vlastností zdrojového materiálu. Autor stanovuje přesná disambiguační pravidla pro stanovení reke akuzativní či lokálové v případech, že tvar jména následujícího po předložce je v akuzativu i lokálu totožný.

Josef Štěpán je autorem stati **Souvětí složené z více než dvou vět hlavních.** V úvodu důkladně a důrazně (ve vztahu k badatelskému kontextu) vymezuje odborná a metodologická východiska. Charakterizuje souvětí souřadné odlišením od jiných (sou)větných celků, jež jen formálně souvětí souřadné připomínají; odlišuje souvětí souřadné od věty jednoduché s několikanásobným přísudkem, od celku s juxtapozicí oznamovacích vět a od celku, kdy je druhá věta sémanticky vztažena k větě první, i když se tak děje bez explicitního spojovacího výrazu. Podrobně cha-

rakterizuje koordinaci hlavních vět, zejména koordinaci slučovací. Stranou nezůstává kombinace koordinačních poměrů (tak častá v uměleckých textech – ty jsou ostatně materiálovým zdrojem studie). Determinace v souvětí souřadném je determinací výsostně sémantickou (při spoluúčinnosti spojovacích výrazů). Zde se nabízí srovnání uměleckých textů s aktuální mluvenou řečí (jež je často v textech zpodobňována), paralely by byly jistě zajímavé. Autor svou studii opírající se o vlastní manuální korpus textů uzavírá dvěma jevy, které charakterizuje pomocí termínů oslabená koordinace a apozice hlavních vět.

III. Korpus – Empirie – Sémantika

Stále bude dost těch, kteří dostanou v první informaci o korpusu coby elektronickém souboru textů, jež je vhodný především pro získávání „tvrdých dat“. Inspiraci pro vysvětlení sémantiky textu v něm zpočátku hledat nebudou. Brzy však pochopí, že právě (a jedině) ponor do množství realizovaných funkcí jazykového prostředku je jedním z nejspolehlivějších korektivů pro úvahy sémantické.

Slovnědruhově transpozice, příklady sémanticky neostrých (syntakticky polyfunkčních) slovních jednotek a jejich přesná ukotvující charakteristika v konkrétním kontextu – to vše je doménou **Miloslava Vondráčka**. Není tomu jinak ani v příspěvku **Výrazy typu *běda* z pohledu slovnědruhového**. Kromě slova uvedeného v titulu článku se autor zabývá slovnědruhovou platností slov kontextově obdobných: *škoda*, *hanba*, *díky*. Po pečlivé komparaci lze výrazy do slovnědruhové kategorie zařadit, byť s jistou tolerancí k neostrým hranicím. Tento stav proměnlivosti a polyfunkčnosti by se však – podle autora – měl projevit v precizním morfologickém popisu slova ve slovnících a ve značkování materiálu pro korpus.

„Tak hlavně to zdraví“ přeje **Alena M. Černá** v závěru svého příspěvku oslavenci, ve studii **Lexém zdraví ve staré češtině**. Neznalec nepředpokládá, že slovo *zdraví* mělo ve staré češtině významů více. Autorka nejdříve podrobně popisuje význam slova jako jednoslovného lexému. Je jím rozuměn nejen tělesný stav obecně, ale i stav, který lze nazvat slovem žití/bytí, je jím rozuměn dokonce i proces uzdravování, ať jmenujeme alespoň některé výrazy pro situaci, kdy člověk není „na smrtedlné posteli“. Přenesený/přenášený význam slova *zdraví* je popsán v části věnované slovním spojením se slovem *zdraví*. To se stává součástí přátelského pozdravu na uvítanou či rozloučenou, je klíčovým slovem v charakteristice míry nemoci či zranění. Jeho užívání opravňuje k tvrzení, že slovo *zdraví* je ve staré češtině synonymem slova život.

Autor **Andrey Izotov** má v článku **Čeština bez příkras a korpusová revoluce** nesporně pravdu v tom, že korpus byl „revolucí“ a publikace *Čeština bez příkras*

vítaným oživením. K obojímu je však třeba, domníváme se, přistupovat s nadhledem (a se znalostí aktuální situace). Způsob zpracování dat, jak je autor na úctyhodném materiálu předkládá, je odrazem nedokončené debaty především v české lingvistice (lexikologii): V čem jsou (aspoň přibližně!) zřejmě hranice obecné češtiny, co je hovorová čeština... atd. Pamětníci si vzpomenou. Zahraničním bohemistům s tak důkladným přístupem k třídění materiálu (ve stati např. body 5.1–5.4) zůstávají čeští teoretikové mnoho dlužni.

Úvod příspěvku **K modálním významům slovesa *muset* v konstrukcích s bezagentními slovesy** (tedy seznámení s relevantními, avšak neustálenými termíny) naznačuje, že největší problém je ve stanovení míry pozornosti, jakou se badatel bude věnovat kontextovým souvislostem spojeným s modalitou. **Vojtěch Veselý** nezkoumá jen samotné konstrukce se slovesem *muset*, věnuje pozornost rozsáhlejšímu úseku textu – zpravidla souvětí. Vztah dvou vět v souvětí není jen vztahem obsahů, ale zcela přirozeně sděluje jak časovou posloupnost dějů (ta je často sama o sobě vyjádřením kauzality), tak – sémantikou přítomných sloves v infinitivu – podmíněnost sdělovaných jevů. Autor však usiluje o poznání i jiného rozměru slovesa *muset*: chce jemně rozlišit nutnost, pravděpodobnost, hypotetičnost – v konstrukci se slovesem bezagentním. Nabízí se otázka: Je-li sloveso *muset* v uvedené vazbě tak flexibilní, nebude více nezpochybnitelných dokladů v mluvených textech?

IV. Korpus – Empirie – Slovtvorba

Výzkumy založené na korpusovém materiálu probíhají i napříč jazyky. **Božena Bednaříková** seznamuje v anglicky psaném textu **Conversion and Its Onomasiological Value Cross-linguistically** s dílčími výsledky probíhajícího projektu *Konverze jako produktivní proces slovtvorby*, který má mj. zjistit potenciál konverze ve slovanských jazycích. Výsledky zkoumání deseti vybraných sloves a jejich slovtvorných sítí/hnízd (obsahujících slova vzniklá derivací i konverzí) ukázaly převahu slov vzniklých konverzí. Autorka uvádí oblasti, které se v projektu ukazují jako nosné (potenciál vidových obměn, nových slov v rámci Aktionsartu, verbálních adjektiv a substantiv). Naznačuje také, že se rýsuje souvislost mezi mentálními kategoriemi uživatelů jazyka/ů a jazykovými kategoriemi ve zkoumaných jazycích. Projekt tak může mít i přínos interdisciplinární.

Fungující struktura jazykového systému je oporou, kolem níž „žije“ konkrétní jazyk. Jeho uživatelé pravidla (v komunikaci) podle potřeby překračují a posléze i pozměňují, jazyk se „očisťuje od forem již nepotřebných, nebo bývá o jiné formy rozšířen. Slovtvorba je oblast, kde se mohou projevit i krátkodobé tvořivé (po)činy

uživatelů jazyka. **Ivana Bozděchová** svůj příspěvek o aktuálních výsledcích slovo-
tvorných kreativců **Možnosti a meze slovo tvorby češtiny (K neologickým
a okazionálním názvům osob)** věnovala úzkému výseku momentálně existující
inovované části slovní zásoby, výrazům z *Hacknuté češtiny*. (Bližší informace o po-
pulárním zdroji uvádí sama autorka v textu.) Studie v zásadě ověřuje, zda soubor
novotvarů – ať již jsou klasifikovány jako neologismy či jako okazionalismy – vyka-
zuje systémové znaky. Na zjištění sumarizovaném v závěru je pozoruhodné, v jaké
šíři možností se autoři slovo tvorných „novinek“ pohybují – jak je do maximální
míry využit princip analogie a anomálie včetně kvazikompozice. Reálná konsoli-
dace *hacknutých slov* v mateřském jazyce mluvčích je samozřejmě otázkou daleké
kontextem příznivé budoucnosti.

Slovo tvorba je pole se značným průnikem sémantiky a gramatické formy –
stanovení slovo tvorných kategorií oprávněně přitahuje pozornost již několik de-
setiletí. Text **Patrika Mittera K pojetí mutačních a transpozicičních slovo tvor-
ných kategorií** systematicky, pomocí promyšlené argumentace spěje především
k návrhu úpravy klasifikace lexikálních jednotek chápaných jako výsledek mutace
a transpozice. Zaujmu zejména ty pasáže, které obhajují chápání některých sé-
mantických skupin jako proces mutačně-transpoziciční. Za novátorské lze považovat
vytvoření mutačně-transpoziciční sémanticko-slovo tvorné skupiny názvů nemocí.

Výsledky studia dvou konkrétních typů deverbativ podávají **Veronika Kolá-
řová a Magda Ševčíková** v článku **Substantiva tvořená od sloves se supletivní-
mi kořeny. Korpusová studie deverbativ s kořeny klád a lož**. Předmětem jejich
výzkumu je utvářenost deverbativ od sloves rozdílných vidových charakteristik.
Nevěnují pozornost jen evidenci jevů a jejich slovo tvorným prostředkům, ale i změn-
ám, jež slovo tvorba přináší v organizaci výpovědi – sledují jejich (systémové a ne-
systémové) valenční chování.

**Slovesa od substantiv s časovým významem a jejich prefixální deriváty
v současné češtině** je název rozsáhlé materiálově bohaté studie **Ivany Kolářové**.
Dokládá existenci skupiny sloves, která není v mnoha mluvnících do příkladů slo-
ves odvozených od substantiv zahrnuta ani jediným příkladem. Je nutno doplnit,
že řada z nich jsou okazionalismy, jejich místo ve slovní zásobě teprve prověří čas.
Přesto si studiem textu a příkladů v něm uvedených lze uvědomit, do jaké míry jsou
již součástí našeho života slovesa od substantiv *čas, datum*, slovesa od názvů časo-
vých jednotek (*minuta, hodina...rok*), od názvů časových úseků a hranic (*přestávka,
pauza..., fáze*), slovesa od názvů ročních období, svátků a významných dnů. Práce
Ivany Kolářové kategorizuje sledovanou skupinu sloves sémanticky, ale i z hledis-
ka formy – zřetelně se jako produktivní vydělují sufixy *-it/nit, -ovat*. (A nejedná se,
samozřejmě, jen o slovesa s kořenem cizího původu.)

IV. Korpus – Empirie – Text

Použit (situačně textem ukotvené) kvantitativní údaje k analýze sémantiky nadvětných úseků textů – tedy při studiu pragmatiky textu – je velmi přínosné.

Diskurz (v zásadě dialog) a jeho markery (zpravidla produkty partikulizace) jsou tématem, jež snad nikdy nebude zpracováno zcela, proměny zdrojového materiálu jsou přímo závislé na měnících se kontextových faktorech. Přesně to však činí diskurzní markery tématem aktuálním, pro komunikační lingvistiku je doslova tématem „povinným“. Český komunikační kontext sdružující (někdy směřující) sociálně-kulturní vlivy nejrůznějšího druhu způsobuje, že lexikální složení markeru, tedy ta lexikální slova, jež ho vytvářejí, naplňují v komunikaci nejen svůj význam původní, ale provázejí nebo dokonce umožňují realizaci i jiných komunikačních záměrů. **Jana Bílková a Jiří Zeman** se v příspěvku **Diskurzní marker *no a* v mluvené češtině** podrobně z pozice výše naznačené věnují jednomu z nich.

Mnozí autoři příspěvků v této publikaci znají F. Štíchu jako blízkého spolupracovníka. Jejich studie jsou holdem jeho kolegiálností. Takovým textem je např. práce **Mileny Hebal-Jeziarské** z Varšavské univerzity. Charakterizuje v ní některé typické charakteristiky česko-polské mezikulturní komunikace.

Ukázkou prostupnosti (průniku) zájmů badatelů v okruhu oslavence, jemuž je recenzovaná publikace věnována, je text **Jany Hoffmannové Mluvená čeština v mediálním žánru „partnerského rozhovoru“: ještě jednou k neformálnosti a „kolokvializaci“**. Příspěvek není jen odpovědí na dotaz lingvisty (Štíchy) stojícího zdánlivě „stranou“ teorie textu, je užitečným poučením o současném úzu interpersonálních kontaktů v mediální komunikaci, ale i o úzu jejího zpracovávání. Za inspirativní lze považovat zejména v závěru stručně naznačené téma aktuální intelektualizace syntaxe mluvené češtiny.

Petr Mareš se pohybuje svým tématem ve zcela odlišném časoprostoru – hluboko v předminulém století, a to v korespondenci a publicistických textech Karla Havlíčka. Svým způsobem však taktéž v mediální komunikaci, a to komunikaci břitké, přesné, tvrdé. P. Mareš však zkoumá metatext – příklady dokládající nikoli vlastnosti Havlíčkem hodnocených textů, ale Havlíčkovu kreativitu ve formulacích prosté informace (která není jen prostá), ve formulacích žádostí a výzev a zejména hodnocení, které není zdaleka jen konstatováním na očekávané hodnotící škále. Text **Informace, výzvy, hodnocení: Metatextová vyjádření v korespondenci a publicistice Karla Havlíčka** by jistě poučila nejednoho současného mediálního textotvůrce.

V části poslední, šesté, ponechávající si – se záměrem zřejmým – syntaktickou formu titulu i všech kapitol (**Korpus – Empirie – František Štícha**), předkládá

Miloslav Vondráček v chronologickém sledu odborný a pracovní život oslavence, respektovaného kolegy. Na závěr je zařazena výběrová bibliografie Františka Štíchy.

Publikace *Lingvistika – Korpus – Empirie* je inspirujícím obrazem současného bádání v oblasti korpusové lingvistiky. Je ukázkou odborné kolegiality, a to i v případě nesouhlasných názorů na dílčí lingvistické otázky či materiálový zdroj, jenž je pojmenován v titulu. Je gestem přátelství a úcty k jubilantovi.

Co je v ČNK nového XI

Dominika Kováříková

*Ústav českého národního korpusu, Filozofická fakulta UK, Praha
kovarikova@korpus.cz*

V dalším pokračování seriálu o novinkách v projektu Český národní korpus představíme nový korpus publicistiky Karla Havlíčka a upozorníme na výrazně rozšířené mluvené korpusy ORATOR a ORTOFON, především se ale podobně jako v předchozích dílech zaměříme na nové uživatelské nástroje. Jedná se o tři nástroje, které byly vytvořeny v uplynulém roce: QuitaUp, GramatiKat a Akalex.

Korpus KH-noviny, který vznikl ve spolupráci s R. Adamem a byl zveřejněn v letošním roce, obsahuje veškeré publicistické texty Karla Havlíčka otištěné v Pražských novinách (1846–1848) a v Národních novinách (1848–1850). Celkový rozsah korpusu je téměř 1 200 000 pozic. V roce 2021 byly také podstatně rozšířeny dva mluvené korpusy – korpus monologů ORATOR a reprezentativní korpus neformální mluvené češtiny ORTOFON. V obou případech šlo o navýšení dvojnásobné (ORATOR v současné době obsahuje 1,2 milionu slov a ORTOFON 2,1 milionu).

Už v roce 2020 byla zveřejněna aplikace **QuitaUp**, která koncepčně navazuje na podobné nástroje QUITA a QUITA Online. Lingvistům a dalším zájemcům má sloužit jako jednoduchý nástroj pro výpočet vybraných stylometrických indikátorů, které kvantitativně vyjadřují některé vlastnosti zadaného textu (jediným omezením je správný formát zadávaného textu).

Mezi poskytované možnosti patří například výpočet slovního bohatství nebo výpočet tematické koncentrace či aktivity textu. Pro kvantifikaci slovního bohatství zadaného textu nabízí QuitaUp mj. prostý poměr typů a tokenů (tzv. TTR – type token ratio), který ovšem trpí tím, že jeho hodnota je závislá na velikosti textu, a proto se zde nabízejí i další možnosti. Jednou z nich je index MATTR (klouzavý průměr TTR v rámci okna, které se po textu posouvá), další možností je zTTR vycházející z porovnání s referenčními hodnotami TTR pro text dané délky.

Údaj o tematické koncentraci (TC) vyjadřuje míru zaměřenosti textu na hlavní téma. Toto ústřední téma je detekováno pomocí tzv. tematických slov, tedy autosé-

mantických slov, která se vyskytují v pásmu nad h-bodem (to je bod, který je definován jako místo ve frekvenční distribuci textu, kde se rank daného slova rovná jeho frekvenci; je vnímán jako hranice mezi synsémantikou a autosémantikou). Podobně je zaměřen i údaj o sekundární tematické koncentraci (STC) – ten je vhodnější např. pro kratší texty, ve kterých se nad h-bodem nevyskytují téměř žádná autosémantika.

Kvantitativní údaj o aktivitě textu vyjadřuje míru dějovosti textu (v kontrastu s dalším poskytovaným údajem, s deskriptivitou textu). Jde o poměr sloves k součtu sloves a adjektiv, která se v textu vyskytují.

Na začátku roku 2021 byla zveřejněna aplikace **GramatiKat**, která má sloužit lingvistům zabývajícím se gramatickými kategoriemi, ale může být velmi užitečná i například pro výzkum lexikologický či kognitivnělingvistický nebo pro výuku češtiny. Aplikace poskytuje přehled rozložení hodnot gramatických kategorií v češtině, tedy například rozložení pádů u substantiv, stupňování u adjektiv nebo negace u sloves. V současné chvíli jsou k dispozici informace o českých substantivech, adjektiva a slovesa se teprve zpracovávají.

U zhruba 15 tisíc substantiv z korpusu SYN2015 se sleduje procentuální zastoupení 7 pádů ve dvou číslech. Z těchto údajů se vyvozuje, jaké je standardní chování českých substantiv, tedy jak je který z pádů zastoupen u celého slovního druhu. Na tomto pozadí je pak možno zjišťovat, která ze substantiv mají zastoupení některého pádu vyšší, než je obvyklé. Například při vyhledání slova *čekárna* na první pohled zjistíme, že se v lokálu singuláru vyskytuje daleko častěji, než je to u českých substantiv obvyklé. Další slovo, *vymoženost*, se neobvykle často vyskytuje hned ve dvou pádech, dativu plurálu a instrumentálu plurálu (a celkově dává přednost spíše plurálu). Zajímavé je i původní substantivum *důsledek*, o kterém se díky GramatiKatu dozvíme, že figuruje hned ve dvou sekundárních předložkách *v důsledku* a *důsledkem*; ty tvoří celkem 60 % výskytů lemmatu *důsledek* (v tomto případě by bylo jistě vhodné uvažovat o změně lemmatizace a slovnědruhového zařazení). Ač je nástroj původně určený zvláště k morfoloickému výzkumu, takováto rychle vyhledaná informace může velmi dobře sloužit například lexikografům nebo pedagogům, zvláště učitelům češtiny jako cizího jazyka.

V nástroji jsou dále k dispozici interaktivní tabulky, které umožňují mj. vyhledávání velkého množství lemmat, v jejichž paradigmatu se některý tvar vyskytuje neobvykle často, nebo se naopak nevyskytuje vůbec, je v korpusu nedoložený.

Nástroj **Akalex** obsahuje seznam zhruba tisíce akademických slov a víceslovných jednotek, jako jsou například *analýza*, *hierarchie*, *teoretický*, *lze*, *charakterizovat*, *dokládat*, *v rámci*, *dle*, *neboť*, *viz*, *např.* apod. Tento seznam vznikl na základě kvantitativních kritérií, z nichž nejdůležitější roli hraje to, kolikrát je určité slovo

častější v akademických než neakademických textech a v kolika oborech se vyskytuje. Seznam tvoří základ pro vznikající slovník, který bude určen pro různé skupiny uživatelů, zvláště ale pro studenty vysokých škol.

Akalex obsahuje i tzv. databázovou část, tedy rozsáhlý materiál, který má ostatním lingvistům sloužit pro vlastní výzkum akademické slovní zásoby. Jde o téměř šedesát tisíc n-gramů různého rozsahu, které byly extrahovány z akademických textů v korpusu SYN2015 (to jsou takové texty, které mají v atributu textový typ hodnotu SCI). Kritéria pro zahrnutí n-gramů do této výzkumné části Akalexu jsou jednoduchá – n-gram musí být alespoň dvakrát častější v akademických než neakademických textech a musí se vyskytovat alespoň ve třetině oborů (8 ze 24 oborů), které jsou rozlišovány v SYN2015. Takto vzniklé tabulky n-gramů je možné prohledávat podle různých kritérií, například podle slovního druhu jednotlivých složek, podle poměru frekvencí v akademických a neakademických textech, podle rovnoměrnosti rozložení v textech nebo podle asociační míry PMI, pomocí které se vyhledávají kandidáti na kolokace. Uživatel může vyhledávat slova nebo smysluplné více-slovné jednotky typické pro akademické texty prostřednictvím kombinace hodnot jednotlivých atributů. Takto lze například vyhledat takové kombinace adjektiva se substantivem, které se vyskytují v minimálně 23 oborech z 24 a mají vysokou hodnotu asociační míry PMI (alespoň 7); výsledkem je pak zhruba tisíc záznamů, z nichž lze vybrat takové kombinace jako *charakteristické rysy, vzájemné vztahy, nedílnou součástí, odborná literatura* nebo *široké spektrum*.

Za pozornost stojí i rozsáhlejší n-gramy, například mezi několika málo 6-gramy, které splňují základní kritéria, patří i tyto dva: 1. *v tomto případě se jedná o* a 2. *je třeba vzít v úvahu že*.