

**SLOVKO 2017**  
**POČÍTAČOVÉ SPRACOVANIE PRIRODZENÉHO JAZYKA,**  
**KORPUSOVÁ LINGVISTIKA, TERMINOLÓGIA**  
**A E-TERMINOLÓGIA**

**Agáta Karčová – Miroslav Zumrík**

*Jazykovedný ústav Ľudovíta Štúra SAV*

*Panská 26, 811 01 Bratislava, e-mail: agata.karcova@korpus.juls.savba.sk,  
miroslav.zumrik@korpus.sk*

V dňoch 25. – 27. októbra 2017 sa v Bratislave uskutočnil 9. ročník medzinárodnej bienálnej konferencie SLOVKO 2017 s podtitulom *Počítačové spracovanie prirodzeného jazyka, korpusová lingvistika, terminológia a e-terminológia*, ktorú zorganizovalo oddelenie Slovenského národného korpusu Jazykovedného ústavu Ľ. Štúra Slovenskej akadémie vied v Bratislave. Konferenciu slávnostne otvorila riaditeľka Jazykovedného ústavu Ľ. Štúra Slovenskej akadémie vied Gabriela Múcsková a zástupkyňa vedúcej oddelenia Slovenského národného korpusu JÚĽŠ SAV Jana Levická.

Na konferencii odzneli 2 plenárne prednášky a 30 prezentácií v autorstve viac ako 60 vedcov zo Slovenska, z Českej republiky, Ruska, Chorvátska, Gruzínska, Nemecka, Rakúska, Poľska, Fínska a zo Švédska. Väčšina prezentácií a diskusií odznela v angličtine, niektoré príspevky boli prednesené aj v češtine alebo slovenčine. Jednotlivé vystúpenia boli rozdelené do sekcií podľa tematického zamerania, priestor na diskusiu po každom príspevku účastníci konferencie aktívne využívali nielen na ozrejenie si faktov, ale aj na vzájomnú inšpiráciu. V priebehu konferencie mali účastníci k dispozícii zborník príspevkov, v ktorom sa nachádza 31 štúdií v anglickom jazyku vybraných a odporúčaných na publikovanie členmi vedeckého výboru konferencie. Zborník vyšiel ako špeciálne tematické číslo Jazykovedného časopisu (SLOVKO 2017. Natural Language Processing, Corpus Linguistics, Terminology, e-Terminology. 2017, roč. 68, č. 2, 431 s., ISSN 0021-5597; prizvané redaktorky: K. Gajdošová, A. Žáková), ktorý je evidovaný vo viacerých medzinárodných databázach, napr. EBSCO a SCOPUS. Zborník je dostupný aj v elektronickej verzii na adrese <http://www.juls.savba.sk/ediela/jc/2017/2/jc17-02.pdf>.

Príspevky konferencie sa zameriavali predovšetkým na oblasť počítačového spracovania prirodzeného jazyka (Natural Language Processing; ďalej NLP), odprezentované boli aj teoretické a praktické projekty a výsledky výskumu v oblasti korpusovej lingvistiky či gramatiky. Viacerí autori sa venovali predstaveniu špecializovaných databáz a syntakticky anotovaných, hovorených a nárečových korpusov. Samostatnú oblasť tvorili príspevky z oblasti terminológie a e-terminológie, ktorú organizátori SLOVKA 2017 doplnili k základným okruhom konferencie v súvislosti s aktuálnym riešením

domáceho projektu VEGA s názvom *Analýza terminologickej práce Jána Horeckého ako inšpirácia pre terminologický manažment 21. storočia na Slovensku*.

V správe zachytávame všetky príspevky publikované v zborníku vrátane dvoch príspevkov, ktoré neboli prezentované v priebehu konferencie (autori O. Zuban a M. Langenbakh z Ukrajiny), obe plenárne prednášky a prezentáciu R. Garabíka z oddelenia SNK JÚLEŠ SAV, ktoré odzneli v priebehu podujatia, pričom sme ich zoradili podľa tematických okruhov.

Prvý tematický okruh tvoria príspevky venované tvorbe korpusov a databáz písaných aj hovorených komunikátov, štandardnej slovenčiny i nárečí, a to s anotáciami na úrovni morfematiky, morfológie či závislostnej syntaxe.

Sekciu otvoril Jan Hajič z Ústavu formálnej a aplikovanej lingvistiky Karlovej univerzity v Prahe, ktorý sa v plenárnej prednáške *Universal Dependencies: Common Morphology and Syntax for Multiple Languages* zaoberal témou závislostnej syntaxe a jej spracovaním vo viacerých jazykoch. Hlavným cieľom predstaveného projektu Universal Dependencies (Univerzálne závislosti) je vytvorenie konzistentnej gramatickej anotácie textov vo všetkých jazykoch, podpora multilingválneho výskumu a vývoj v oblasti NLP založený na už existujúcich štandardoch. Výsledkom trojročnej práce na projekte, do ktorého je zapojených viac než 200 spolupracovníkov, je 6 verzií súhrnných databáz so syntakticky anotovanými textami (vo forme závislostných stromov), 95 jednotlivých syntakticky anotovaných databáz v 57 jazykoch s viac ako 13 miliónmi tokenov. Podrobné informácie o tomto medzinárodnom projekte sú prístupné na stránke <http://universaldependencies.org/>.

Daniel Zeman z Matematicko-fyzikálnej fakulty Karlovej univerzity v Prahe v príspevku *Slovak Dependency Treebank in Universal Dependencies* opisuje konverziu syntakticky anotovaných textov Slovenského národného korpusu do anotačnej schémy univerzálnych závislostí. Doposiaľ bola konvertovaná iba malá časť dát, napriek tomu ide o prvý slovenský závislostný korpus, ktorý je verejne dostupný verejnosti. Autor príspevku uvádza viaceré výskumné projekty, pri ktorých bol tento korpus využitý.

Štúdia *Microsyntactic Annotation of Corpora and Its Use in Computational Linguistics Tasks* v autorstve Leonida Iomdina z Ruskej akadémie vied v Moskve sa zameriava na anotáciu idiomatických výrazov, ktoré majú v ruštine významnú syntaktickú funkciu, ale nie sú dostatočne spracované v lingvistických modeloch používaných napríklad pri strojovom preklade či sémantickej anotácii. Možným riešením tejto situácie je vytváranie korpusov anotovaných na úrovni mikrosyntaxe. Autor príspevku ukazuje, ako sa táto úloha rieši v hĺbkovo syntakticky anotovanom ruskom podkorpuse SynTagRus.

Jana Šindlerová z Matematicko-fyzikálnej fakulty Karlovej univerzity v Prahe v príspevku *Annotation of the Evaluative Language in a Dependency Treebank* analyzuje anotovanie tzv. evaluatívneho jazyka (*evaluative language*) v Pražskom zá-

vislostnom korpuse 2.0. Potenciálne evaluatívne uzly v závislostných stromoch sú identifikované automaticky, následne sú manuálne skontrolované anotátorom, ktorý ich v prípade chybnjej identifikácie zruší. Po korektnjej identifikácii k nim anotátor pridá informáciu o polarite, zdroji a celi evaluácie.

Skúsenosti z prípravy nového historického korpusu češtiny 19. storočia opisuje Anna Řehořková z Ústavu Českého národního korpusu Karlovej univerzity v Prahe v článku *Issues of POS Tagging of the (Diachronic) Corpus of Czech: Preparing a Morphological Dictionary*. Pri koncipovaní slovníka museli autori korpusu prehodnotiť pravidlá morfolologickej anotácie hlavne v prípade neplnovýznamových slov, ktorých slovnodruhové určenie je často problematické. Autori z viacerých pracovísk navrhujú zaviesť značku morfolologickej nejednoznačnosti (*ambiguity tag*), a to nielen pre historický korpus, ale aj pre korpusy súčasného jazyka.

Pracovníci viacerých prestížnych výskumných inštitúcií v Moskve Olga Lyaševskaya, Victor Bocharov, Alexey Sorokin, Tatiana Shavrina, Dmitry Granovsky, Svetlana Alexeeva v štúdií *Text Collections for Evaluation of Russian Morphological Taggers* opisujú spracovanie textových zdrojov v rámci iniciatívy MorphoRuEval-2017, ktorá smeruje k stimulovaniu vývoja technológií na automatické morfologické spracovanie ruštiny. Realizátori projektu sa sústredili na štandardizovanie všetkých dostupných korpusov ruského jazyka s ručnou morfologickou anotáciou do jedného spoločného formátu CONLL. Zdroje textových dát tvoria ručne značkovaný podkorpus Ruského národního korpusu, syntakticky anotovaný podkorpus SynTag-Rus, dáta z OpenCorpora.org a korpus GICR.

Pracovníci Ústavu Českého národního korpusu Karlovej univerzity v Prahe Zuzana Komrsková, Marie Kopřivová, David Lukeš, Petra Poukarová, Hana Goláňová v príspevku *New Spoken Corpora of Czech: ORTOFON and DIALEKT* predstavujú dva špecializované hovorené korpusy: ORTOFON – korpus spontánnej hovorenej češtiny, a DIALEKT – korpus českých dialektov. Nové hovorené korpusy češtiny obsahujú v porovnaní s existujúcimi podrobnejšiu anotáciu rozdelenú do niekoľkých paralelných vrstiev, v ktorých sú metadáta o každom hovoriacom uvedené samostatne. Výhodou je tiež viacúrovňový prepis poskytujúci všeobecnú ortografickú aj fonetickú transkripciu podrobne zachytávajúcu výslovnosť všetkých hovoriacich.

Róbert Sabo a Jakub Rajčáni z Ústavu informatiky Slovenskej akadémie vied v Bratislave v príspevku *Designing the Database of Speech Under Stress* opisujú prípravnú fázu budovania databázy hovorených prehovorov pod vplyvom reálneho stresu. Výbrani hovoriaci riešili komunikačnú úlohu, pričom sa u nich zaznamenávali viaceré psychofyzologické ukazovatele, napríklad zvýšenie srdcovej frekvencie a elektrodermálnej aktivity, čo doplnilo subjektívne sebahodnotenie. Zvolený prístup sa ukazuje ako vhodné východisko pre budovanie databázy, pričom je potrebné doplniť ho o ďalšie parametre a relevantné faktory.

Ján Staš, Daniel Hládek, Peter Vizslay a Tomáš Koctúr z Fakulty elektrotechniky a informatiky Technickej univerzity v Košiciach sa v štúdiu *TEDxSK and JumpSK: A New Slovak Speech Recognition Dedicated Corpus* venujú opisu špecializovanej databázy hovorených komunikátov v slovenčine určenej na vývoj automatického rozpoznávania reči. Anotovaná rečová databáza bola generovaná automaticky pomocou akustickej segmentácie reči založenej na analýze hlavných komponentov a automatickej transkripcii reči pomocou dvoch komplementárnych systémov na rozpoznávanie reči. Automatickou anotáciou získali 21,26 % nových rečových segmentov vhodných na pretrénovanie alebo adaptáciu už existujúcich akustických modelov.

Korpus gruzínskych dialektov predstavujú Marina Beridze, David Nadaraia a Lia Bakuradze (Jazykovedný inštitút A. Chikobavu Tbiliskej národnej univerzity v Tbilisi; *Georgian Dialect Corpus: Linguistic and Encyclopedic Information in Online Dictionaries*). Budovanie korpusu zahŕňalo vytvorenie technického rámca, zhromažďovanie textových dát vrátane diel lexikografického charakteru, ich jazykové spracovanie a digitalizáciu, navrhnutie anotačnej štruktúry a spôsobu morfosyntaktickej anotácie. V súčasnosti je Korpus gruzínskych dialektov platformou pozostávajúcou z korpusu dialektov, textovej knižnice a lexikografickej databázy.

Oksana Zuban z Národnej univerzity T. Ševčenko v Kyjeve v príspevku *Automatic Morphemic Analysis in the Corpus of the Ukrainian Language: Results and Prospects* vysvetľuje teoretické východiská a fungovanie automatizovaného systému morfematickej a derivačnej analýzy (ASMDA) pre ukrajinský jazyk, ktorý má viacero samostatných funkcií: slúži ako informačný systém, automatizovane anotuje text na morfematickej úrovni a využíva sa ako lingvistická báza pre frekvenčné slovníky. Načrtáva zároveň možnosti využitia tohto systému na vytváranie morfematických modelov v novom korpusovom slovníku a na štylistickú analýzu textov.

Ďalší tematický okruh tvoria štúdie, v ktorých sa na báze písaných a hovorených korpusov skúmajú jazykové javy z oblasti ortografie, gramatiky, sémantiky, štylistiky, frazeológie, etymológie, ale aj dynamiky jazyka, pragmatiky, didaktiky, komparatistiky či literárnej vedy.

Jaroslava Hlaváčová z Matematicko-fyzikálnej fakulty Karlovej univerzity v Prahe v článku *Golden Rule of Morphology and Variants of Wordforms* skúma pravopisné varianty, ich lematizáciu a morfológické značkovanie. Keďže pravopisné varianty majú spravidla identickú morfológickú značku (tag), rovnakej leme môžu zodpovedať dva rôzne tvary s rovnakým morfológickým opisom. Táto nejednoznačnosť sťažuje mnohé úlohy v oblasti počítačového spracovania prirodzeného jazyka. Autorka formuluje tzv. zlaté pravidlo morfológie, ktoré poskytuje riešenie, ako jednoznačne označovať všetky slovné tvary vrátane ich variantov.

Marie Mikulová, Eduard Bejček, Veronika Kolářová a Jarmila Panevová z Matematicko-fyzikálnej fakulty Karlovej univerzity v Prahe v článku *Subcategoriza-*

*tion of Adverbial Meanings Based on Corpus Data* demonštrujú zvolenú metodológiu na získanie podrobného, komplexného a prehľadného opisu významov priestorových adverbií vrátane zoznamu ich formálnych realizácií s príkladmi. Pražský závislostný korpus s ručnou anotáciou na úrovni hlbkovej syntaxe poskytuje bádateľom množstvo príkladov, na základe ktorých môžu spoľahlivejšie definovať ich významy a presnejšie ich kategorizovať.

Hana Žižková z Filozofickej fakulty Masarykovej univerzity v Brne sa v príspevku *Compound Adverbs as an Issue in Machine Analysis of Czech Language* zaoberá zloženými príslovkami v češtine. Autorka predstavuje výsledky korpusového výskumu zameraného na prípady jednoznačne zložených prísloviok, ktoré však systém automatickej morfologickej analýzy (AMA) ako také neidentifikoval, resp. ich chybné zaradil k iného slovnému druhu. Neidentifikované, resp. nesprávne zaradené prípady klasifikovala do niekoľkých skupín, vďaka čomu sa môže anotačný nástroj v budúcnosti zlepšiť.

Trojica autorov Edyta Jurkiewicz-Rohrbacher z Univerzity v Regensburgu a Univerzity v Helsinkách a Björn Hansen a Zrinka Kolaković z Univerzity v Regensburgu sa v štúdiu *Clitic Climbing, Finiteness and the Raising-Control Distinction. A Corpus-Based Study* zameriava na fenomén tzv. klitik (*clitic climbing*) v súvislosti s doplnkami typu  $da_2$  v súčasnej srbčine. Názory srbských vedcov na to, či je konštrukcia zložená z doplnku  $da_2$  a plnovýznamového slovesa prijateľná, sa rozchádzajú. Vďaka korpusom mohli lingvisti zistiť relevantnejšie údaje o používaní tejto zriedkavej konštrukcie a zaujať objektívnejšie stanovisko k otázke jej používania.

Valenciu substantív analyzujú Veronika Kolářová, Anna Vernerová a Jana Klímová z Matematicko-fyzikálnej fakulty Karlovej univerzity v Prahe a Jan Kolář z Matematického ústavu Českej akadémie vied v Prahe v štúdiu *Possible but not Probable: A Quantitative Analysis of Valency Behaviour of Czech Nouns in the Prague Dependency Treebank*. Skúmajú valenciu českých deverbatívnych substantív v Pražskom závislostnom korpusu (*Prague Dependency Treebank*), pričom odkazujú na valenčný slovník českých substantív NomVallex. Autori príspevku ukazujú, že relatívne zastúpenie kombinácií participantov je takmer totožné pre všetky pozorované typy podstatných mien, zároveň špecifikujú najčastejšie a najmenej časté kombinácie daných substantív.

Katarína Chovancová, Lucia Ráčková, Dagmar Veselá a Monika Zázrivcová z Filozofickej fakulty Univerzity Mateja Bela v Banskej Bystrici v štúdiu *Valency Potential of Slovak and french Verbs in Contrast* porovnávajú valenciu 15 najfrekvencovanejších slovenských plnovýznamových slovies a ich francúzskych ekvivalentov, pričom konštatujú, že valencie slovies v oboch jazykoch nevykazujú iba vzájomné odlišnosti, ale tiež sémanticko-syntaktickú podobnosť. Hlavný prínos štúdie spočíva podľa autoriek v jej kontrastívnom prístupe na úrovni gramatiky, sémantiky, syntaxe, kognitívnej a korpusovej lingvistiky.

Zdeňka Urešová, Eva Fučíková a Eva Hajičová z Ústavu formálnej a aplikovanej lingvistiky Karlovej univerzity v Prahe v príspevku *CzEngClass – Towards a Lexicon of Verb Synonyms with Valency Linked to Semantic Roles* predstavujú spoločný projekt, ktorého cieľom je vytvorenie česko-anglického slovníka slovesných synonym. Jednotlivé skupiny slovík v slovníku budú charakterizované pomocou tzv. sémantických rolí s priradenými argumentmi podľa valenčných slovníkov. Nevyhnutným zdrojom pre potreby projektu je česko-anglický paralelný korpus, ktorý umožňuje nové pohľady na kontextuálnu synonymiu českých a anglických slovík.

Margaryta Langenbakh z Národnej univerzity T. Ševčenka v Kyjeve sa v príspevku *Corpus-Based Semantic Models of the Noun Phrases Containing Words with 'Person' Marker* zaoberá mechanizmom, ktorý riadi konštrukciu lexikálne správnych slovných sekvencií. Osvetľuje pritom niektoré aspekty modelovania vzorcov substantívnej valencie v prípade štruktúr typu substantívum – substantívum, kde jedna alebo obidve zložky obsahujú sémantickú značku typu „osoba“. Príspevok sa opiera o jazykové zdroje Ukrajinského korpusu.

V článku *Modeling Semantic Distance in the Pattern Dictionary of English Verbs* Silvie Cinkovia a Zdeněk Hlávka z Matematicko-fyzikálnej fakulty Karlovej univerzity v Prahe porovnávajú, do akej miery sa tzv. vzorce použitia (*usage pattern*) 29 slovík zo Slovníka vzorcov anglických slovík (*Pattern Dictionary of English Verbs*) zhodujú s náhodnými konkordanciami týchto slovík. Autori sa sústreďujú na prípady, kde pre danú slovesnú konkordanciu prichádza do úvahy viac než jeden vzorec použitia.

Zuzana Komrsková (Ústav Českého národného korpusu Karlovej univerzity v Prahe; *What Does že jo (and že ne) Mean in Spoken Dialogue*) sa zaoberá významom kolokácií *že jo* a *že ne* v hovorených dialógoch. Obe kolokácie sú typické pre spontánnu konverzáciu a vyjadrujú viacero pragmatických funkcií, napr. neistotu rečníka alebo žiadosť o spätnú väzbu. Autorka identifikovala dané kolokácie v dátach hovorených korpusov ORAL2013 a DIALOG, pričom ich selektovala na základe ich umiestnenia v rámci výpovede a významového vzťahu k ich blízkeho kontextu, čím ich odlišila od ostatných náhodných spoluvýskytov spojky *že* a častice *jo/ne*.

V štúdiu *Correlative Conjunctions in Spoken Texts* sa Petra Poukarová (Ústav pre jazyk český Českej akadémie vied v Prahe) zaoberá korelačnými spojeniami *bud' – anebo, jednak – jednak* a ďalšími, ktoré predstavujú jeden z prostriedkov kohezie textu. Výskyt jednej zložky páru predpokladá použitie druhej v nasledujúcom texte. S využitím českého hovoreného korpusu ORAL2013 autorka skúmala používanie týchto spojení v prototypických spontánnych hovorených komunikátoch, ktoré sa všeobecne považujú za menej koherentné a fragmentárnejšie ako písané texty, a zistila, do akej miery sa v nich realizuje druhá zložka viacnásobných výrazov.

Vladimír Petkevič a Milena Hnátková z Filozofickej fakulty Karlovej univerzity v Prahe v príspevku *Morphological Disambiguation of Multiword Expressions and Its Impact on the Disambiguation of Their Environment in a Sentence* opisujú, ako komplexný systém pravidiel morfolologickej dezambiguácie Českého národného korpusu rozpoznáva jednotlivé komponenty kolokácií a frazém, t. j. viacslavných pomenovaní (*multiword expressions*). Demonštrujú, že aj malá časť správne anotovaných ustálených výrazov výrazne prispieva k spoľahlivejšiemu rozlišovaniu zvyšných komponentov vety, ktoré nie sú súčasťou týchto ustálených slovných spojení.

Victor Zakharov zo Štátnej univerzity v Petrohrade v príspevku *Slavic Phraseology: A View Through Corpora* skúma a porovnáva vybrané frekventované frazémy v ruštine, češtine a slovenčine. Ako bázy na skúmanie mu slúžia písané korpusy daných textov, pričom využíva aj možnosti automatickej extrakcie idiómov a kolokácií. Na základe vlastných zistení opisuje, ako sa dajú textové korpusy a korpusové metódy a nástroje účinne využiť na obohatenie existujúcich frazeologických slovníkov o nové jednotky aj ako môžeme určovať frekvenciu používania konkrétnych frazém.

V prezentácii *Náčrt problematiky cudzích slov v slovenčine* Radovan Garabík (v spoluautorstve s Agátou Karčovou) z Jazykovedného ústavu Ľ. Štúra Slovenskej akadémie vied v Bratislave predstavil výsledky skúmania frekvenčného zastúpenia prevzatých slov v slovenčine. Autori sa sústredili na najpočetnejšie slová ako reprezentatívnu vzorku slov: z korpusu prim-7.0-public-vyv vyseletovali 1999 najfrekventovanejších lexém a na základe dostupných slovníkových a iných diel doplnili ich pôvod. Následne analyzovali podiel cudzích slov v slovenčine z hľadiska ich prvotného pôvodu a samostatne pre zastúpenie jazykov, z ktorého boli slová do slovenčiny bezprostredne prevzaté.

Michal Křen z Ústavu Českého národného korpusu Karlovej univerzity v Prahe v príspevku *Grammatical Change Trends in Contemporary Czech Newspapers* predstavuje *corpus-driven* metódu (t. j. založenú na samotnom jazykovom materiáli) na zisťovanie nedávnych gramatických zmien v súčasnej českej tlači. Skúmal texty denníka *Mladá fronta* z rokov 1992 až 2014, čo predstavuje 825 miliónov tokenov. Výskumnou úlohou bolo objaviť najrelevantnejšie gramatické zmeny v češtine tak, ako sa prejavujú v meniacej sa frekvencii vybraných skupín morfológických značiek (tagov). Výsledky naznačujú, že jazyk českých novín sa v priebehu posledných dvoch desaťročí mení na „slovesnejší“ a neformálnejší.

Metódu automatického merania parametrov hlasného čítania rozpracoval Marek Nagy (Fakulta matematiky, fyziky a informatiky Univerzity Komenského v Bratislave; *Measuring and Improving Children's Reading Aloud Attributes by Computers*). Vo výskume využíva zarovnanie, pri ktorom sa spracúva textová transkripcia zvukového segmentu reči a určuje sa konkrétny čas, v ktorom sa vyskytnú jednotlivé slová (forced alignment). Monitoruje tempo a plynulosť čítania detí,

pričom výsledky sa využívajú pri ich individuálnej motivácii. Nástroj sa testuje a vylepšuje v spolupráci so základnými školami už 5 rokov a podľa autora prináša dobré výsledky.

Richard Změlík z Filozofickej fakulty Univerzity Palackého v Olomouci sa v príspevku *The Use of Authorial Corpora Beyond Linguistics* zameriava na otázku využitia kvantitatívnych a kvalitatívnych metód v literárnej vede. Predstavuje vlastnú koncepciu korpusu českej prózy a stanovuje hlavné komponenty takéhoto korpusu. Okrem toho sa autor venuje súčasným stochastickým metódam, ktoré zahraniční vedci používajú pri analýze literárnych textov, sústreďuje sa na metódu tzv. tematickej analýzy a uvažuje o využití metadát pripravovaného korpusu pre potreby literárnej vedy. Týmto spôsobom sa pripravovaný korpus pokúša včleniť do kontextu digitálnych literárnych štúdií (*Digital Literary Studies*), jednej z oblastí skúmania v rámci tzv. digitálnych humanitných vied.

Tretí tematický okruh podujatia predstavujú príspevky z oblasti terminológie, e-terminológie, terminologickej analýzy, terminologického manažmentu a teoretickej terminológie.

V plenárnej prednáške *Tendencies in Terminology* riaditeľ Európskej terminologickej asociácie EAFT Henrik Nilsson zo Švédskeho centra pre terminológiu v Štokholme predstavil výskumné úlohy svojich výskumných inštitúcií. Venoval sa nejasnej hranici medzi trendmi a tendenciami v terminológii. Opísal vzťah medzi pojmom, termínom, definíciou termínu a objektom reality, uviedol definície terminologickej analýzy, práce a terminologického manažmentu. Zaoberal sa tiež otázkou vzťahu terminológie a korpusovej lingvistiky. Za hlavné trendy v terminológii označil väčšiu orientáciu terminológov na prácu s korpusmi odborných textov, zvýšené využívanie metód kolektívnej práce, tzv. crowdsourcingu, a dôraz na všestranné využitie terminologických nástrojov. Upriamil pozornosť aj na používanie angličtiny v odbornej komunikácii vo Švédsku či na sociologické aspekty terminológie.

Inovatívny koncept slovníka ekvivalentov vytvoreného špeciálne pre prekladateľov právnych textov predstavuje Weronika Szemińska z Ústavu odbornej a medzikultúrnej komunikácie Varšavskej univerzity v príspevku *Helping the Translator Choose: The Concept of a Dictionary of Equivalents*. Opisuje charakteristické vlastnosti právnej terminológie a poukazuje na problematiku polysémie, synonymie či nekongruencie medzi národnými právnymi systémami, v dôsledku ktorých je preklad právnych textov mimoriadne náročný. Autorka navrhuje systém klasifikácie ekvivalentov, ktorý by mal obsahovať každý terminologický slovník slúžiaci prekladateľom právnych textov.

Jana Nová a Hana Mžourková (Ústav českého jazyka Českej akadémie vied v Prahe; *Terminology and Labelling Words by Subject in Monolingual Dictionaries – What Do Domain Labels Say to Dictionary Users?*) sa zameriavajú na problema-



tiku označovania odborných oblastí (*domain labels*) vo výkladových slovníkoch. Porovnávajú pritom spôsob, akým sa v súčasných výkladových slovníkoch označujú termíny z oblasti medicíny a príbuzných vied (farmácie, biológie, biochémie, psychológie). Okrem rozdielov medzi jednotlivými slovníkmi poukazujú na nekonzistencie v rámci jedného slovníka. Zvažujú výhody a nevýhody označovania oblastí, ako aj využiteľnosť týchto označení z hľadiska používateľov slovníkov.

Bruno Nahod a Perina Vukša Nahod z Ústavu chorvátskeho jazyka a jazykovedy a Mirjana Bjeloš z Nemocnice „Sveti Duh“ v príspevku *Three Aspects of Processing Ophthalmological Terminology in a “Small Language”: A Case of Croatian Term Bank Struna* predstavujú metodologické problémy, s ktorými sa stretli pri spracúvaní oftalmologickej terminológie v chorvátskej terminologickej databáze Struna (<http://struna.ihjj.hr/>). Úlohou databázy, ktorá bola spustená v roku 2012, je postupne normalizovať chorvátsku terminológiu všetkých odborných oblastí a poskytovať platformu na spoluprácu odborníkov v daných oblastiach: terminológov aj jazykovedcov.

Miroslav Zumrík z Jazykovedného ústavu Ľ. Štúra Slovenskej akadémie vied v Bratislave v príspevku *Ján Horecký's Approach to Language and Thinking* uvažuje o teoretických základoch, ktoré J. Horeckému slúžili ako východisko pri modelovaní vzťahu jazyka (termínov) a myslenia (pojmov). Autor sa venuje Horeckého definíciám termínov, pojmov a ich vzájomného vzťahu, ako aj vzťahu týchto zložiek k okolitej realite. Definície termínu a pojmu zverejnil Horecký v niektorých svojich významných štúdiách. Autor príspevku tiež poukazuje na zistené podobnosti a rozdiely v prepracovaných teóriách lingvistu J. Horeckého a českého logika P. Tichého.

Z uvedených troch tematických okruhov sa vyčleňuje širšie koncipovaný príspevok Agnes Kim z Katedry slavistiky Viedenskej univerzity a Ludwiga Breuera z Katedry germanistiky Viedenskej univerzity, ktorí v štúdiu *On the Development of an Interdisciplinary Annotation and Classification System for Language Varieties – Challenges and Solutions* predstavujú špeciálny výskumný projekt Nemčina v Rakúsku: Variácia – kontakt – percepcia. Náplňou projektu je skúmať používanie a subjektívne vnímanie nemeckého jazyka v Rakúsku, jeho variácie a zmeny v danom priestore, ako aj jeho kontakt s inými jazykmi. Autori poskytujú informácie o vytvorenej koncepcii anotačnej štruktúry aj klasifikácii jazykových variet.

V závere medzinárodnej konferencie SLOVKO 2017 sa zástupkyňa oddelenia SNK JÚLŠ SAV Jana Levická poďakovala všetkým zúčastneným za aktívnu účasť, vyjadrila spokojnosť z plynulého priebehu konferencie a organizátorom, recenzentom i posudzovateľom príspevkov poďakovala za ich čas a profesionálny prístup. Zároveň všetkých pozvala na plánovanú jubilejnú 10. konferenciu SLOVKO 2019 zameranú nielen na oblasť počítačového spracovania prirodzeného jazyka, korpusovú lingvistiku, ale aj na dynamiku jazyka a gramatické zmeny.