

# TÝŽDEŇ VEDY A TECHNIKY V SLOVENSKOM NÁRODNOM KORPUSE

**Katarína Gajdošová**

*Slovenský národný korpus Jazykovedného ústavu Ľ. Štúra SAV, Panská 26  
811 11 Bratislava, e-mail: katarinag11@gmail.com*

Týždeň vedy a techniky na Slovensku je už niekoľko rokov známy ako týždeň, v ktorom sa slovenské vedecké pracoviská viac otvárajú verejnosti, aby popularizovali výsledky svojej práce. Počas tohto týždňa môže široká verejnosť vstúpiť na miesta, kam zvyčajne nechodí a dokonca aj tam, kam sa inokedy počas roka nedá dostať vôbec. Aj Slovenská akadémia vied SAV otvára brány svojich pracovísk v tomto týždni už niekoľko rokov, predovšetkým počas Dní otvorených dverí na jednotlivých akademických pracoviskách.

Do tohtoročného Týždňa vedy a techniky na Slovensku sa oddelenie Slovenského národného korpusu Jazykovedného ústavu Ľ. Štúra SAV v Bratislave (ďalej SNK) zapojilo zorganizovaním série prednášok, prezentácií a praktických seminárov pre odbornú, ale aj širšiu verejnosť, ktoré prebiehali v dňoch 10. – 13. novembra 2014. Podujatie bolo inšpirované podnetmi zo strany viacerých záujemcov dozvedieť sa viac informácií o jednotlivých novších aj starších projektoch SNK. Cieľom podujatia bolo nielen informovať, ale aj podnietiť diskusiu o zaujímavých témach a nadviazať spoluprácu.

V pondelok 10. 11. 2014 predniesla M. Šimková prednášku *AUTORI, TEXTY A JAZYKY V SLOVENSKOM NÁRODNOM KORPUSE*, ktorá bola v poradí už druhou informačnou prednáškou určenou pre poskytovateľov textov do celku SNK. Po prednáške nasledovala diskusia s prítomnými poskytovateľmi, z ktorých viacerí zmluvne obnovili spoluprácu so SNK a priniesli svoje nové diela v elektronickej podobe na zaradenie do korpusu.

Utorok 11. 11. 2014 sa začal všeobecnou prednáškou M. Šimkovej *SLOVENSKÝ NÁRODNÝ KORPUS A JEHO DATABÁZY*, v ktorej stručne predstavila jednotlivé korpusy celku SNK, ale aj špecifické databázy ako napríklad paradigmy podstatných mien, ktoré sú používateľom dostupné na stránke SNK aj prostredníctvom webovej stránky s lingvistickými zdrojmi.

Predpoludnie druhého dňa podujatí patrilo prezentácii J. Levickej *SLOVENSKÁ TERMINOLOGICKÁ DATABÁZA*, v ktorej informovala o tejto špecifickej databáze sústreďujúcej na jednom mieste termíny z rôznych vedných oblastí. Tie tvoria konkrétni odborníci v spolupráci s jazykovedcami. J. Levická hovorila nielen o štruktúre terminologického záznamu a technických možnostiach databázy, ale aj o problémoch, na ktoré databázové spracovanie terminológie jednotlivých vedných odborov môže poukázať a ktoré je možné v databáze tohto typu riešiť.

Popoludní bol pre vopred prihlásených siedmich záujemcov pripravený ďalší zo série praktických seminárov pre záujemcov, ktorí sa chceli naučiť pracovať s písanými korpusmi SNK, tento raz určený pre začiatočníkov.

V stredu 12. 11. 2014 dopoludnia pripravilo oddelenie SNK pre širšiu verejnosť 2 prezentácie: K. Gajdošová o projekte KORPUS NÁREČÍ – o zhromažďovaní textového nárečového materiálu, jeho spracovaní korpusovým spôsobom, a R. Garabík o projekte SLOVNÍK SÉMANTICKÝCH VZŤAHOV TYPU WORDNET, na ktorej sa mohli záujemcovia dozvedieť okrem iného napríklad to, že v slovníku sú spracované sémantické vzťahy najfrekvencovanejších slovenských podstatných a prídavných mien, sloviac a prísloviac, že každý slovenský synset je prepojený s anglickým ekvivalentom z anglického WordNetu, (Princetone), ale aj to, že tento slovenský sémantický slovník je prepojený s litovským WordNetom, ktorý vznikol v rámci projektu Slovak Online.

V popoludňajšom bloku predstavil R. Garabík v prezentácii PARALELNÉ SLOVENSKO-CUDZOJAZYČNÉ KORPUSY koncepciu tvorby paralelných korpusov v SNK, textovú náplň jednotlivých korpusov, možnosti vyhľadávania a špecifiká konkrétnych jazykových párov textov. Na prednášku nadväzoval praktický seminár VYHĽADÁVANIE V PARALELNÝCH KORPUSOCH pre vopred prihlásených siedmich záujemcov. K. Gajdošová pripravila pre účastníkov zaujímavé úlohy zamerané aj na vyhľadávanie pomocou regulárnych výrazov, prostredníctvom ktorých sa účastníci seminára zoznámili s osobitosťami vyhľadávania v paralelných korpusoch vo vyhľadávacom nástroji NoSketch Engine.

Štvrtok 13. 11. 2014 bol v rámci tradície Jazykovedného ústavu L. Štúra SAV nazvaný Dňom otvorených dverí, počas ktorého pripravilo oddelenie SNK pre návštevníkov viacero zaujímavých prezentácií. Prvú prednášku s názvom VIETE, ČO JE KOLOKÁCIA? (O KONCIPOVANÍ SLOVNÍKA SLOVNÝCH SPOJENÍ SLOVENČINY) predniesli kolektívne D. Majchráková, K. Chlpíková a K. Bobeková, ktoré sa venujú príprave SLOVNÍKA SLOVNÝCH SPOJENÍ SLOVENČINY. V prezentácii zazneli informácie o východiskách a koncepcii tvorby tohto slovníka, pôvodne vypracovanej P. Ďurčom, ale aj problémoch, ktoré v minulosti vznikali či ktoré sa priebežne objavujú ako nové. Široká diskusia priniesla viaceré zaujímavé otázky, ale aj hodnotenia dôležitosti práce na takomto type slovníka a očakávania tlačenej aj elektronickej podoby slovníka pre verejnosť.

V prezentácii HISTORICKÝ KORPUS SLOVENČINY hovoril R. Garabík o spôsobe tvorby a náplni tohto diachrónneho korpusu slovenčiny a stručne previedol prítomných aj rôznymi obdobiami predpisovného a spisovného obdobia slovenčiny. Následne zdôraznil, že Historický korpus slovenčiny bol v októbri 2014 rozšírený okrem iného aj o diela H. Gavloviča VALASKÁ ŠKOLA MRAVÚV STODOLA a J. I. Bajzu WESELÉ UČINKI A REČEŇ, KTERÉ K STRÁWEŇŮ TRÚCHLIVÍCH HODIN ZEBRAL A VIDAL J. I. BAJZA.

Prezentácia ZVUKOVÝ ARCHÍV SLOVENSKÝCH NÁREČÍ, ktorú predniesla A. Karčová, vyvolala medzi prítomnými návštevníkmi širokú diskusiu. Hoci je projekt tvorby tohto archívu jedným zo začínajúcich sa projektov oddelenia a archív sa tvorí len krátky čas, obsahuje už dnes niekoľko desiatok zvukových nárečových záznamov z rôznych lokalít Slovenska. Archivácia takéhoto materiálu si vyžaduje osobitnú koncepciu, navrhnutie metainformácií o zvukových záznamoch a systematickú prácu pri ich zhromažďovaní, digitalizovaní, triedení, anotácii, archivácii a prístupňovaní na vedeckovýskumné ciele v oddelení SNK. Diskusia podnietila niektorých z prítomných záujemcov, aby do Zvukového archívu slovenských nárečí SNK odporučili viaceré hodnotné nárečové zvukové záznamy.

Podvečerná PREZENTÁCIA PRODUKCIE ODDELENIA SNK sa niesla v príjemnej atmosfére. Okrem knižných starších titulov zborníkov z konferencie SLOVKO a ďalších podujatí a projektov mohli účastníci vidieť a niektorí si aj odniesť len nedávno vydané tituly JAZYKOVÉDNÉ ŠTÚDIE XXXI. ROZVOJ JAZYKOVÝCH TECHNOLOGIÍ A ZDROJOV NA SLOVENSKU A VO SVETE (10 ROKOV SLOVENSKÉHO NÁRODNÉHO KORPUSU) a zborník AKTUÁLNE OTÁZKY TERMINOLÓGIE MARKETINGU z rovnomených vedeckých konferencií. Okrem knižnej produkcie predstavili pracovníci SNK aj nástroje, ktoré boli vyvinuté v oddelení SNK. Najviac zaujal najnovší nástroj na vizualizáciu kolokácií KOLOKAT, ale záujem bol aj o ďalšie nástroje: nástroj na automatické dopĺňanie diakritiky do textu DIAKRITIK, prekladač do štúrovskej slovenčiny LUDEVÍT či nástroje na tvorbu a rekonštrukciu morfológických značiek používaných v SNK.

Počas Týždňa vedy a techniky v Slovenskom národnom korpuse pripravilo oddelenie SNK pre verejnosť dohromady 12 podujatí. Spolu sme na našom pracovisku privítali 117 hostí z viacerých bratislavských aj mimobratavských pracovísk (Banská Bystrica, Žilina, Varšava, Nitra, Srbsko). Rozsah poskytnutých informácií, podnetov v diskusiách, rozvinutých spoluprác a poďakovaní za prácu pri tvorbe korpusov a databáz sa nedá vyjadriť v číslach. Teší nás, že podporovatelia, poskytovatelia textov, používatelia a všetci priaznivci SNK vedia, že dvere do Slovenského národného korpusu nebývajú otvorené len raz v roku, ale sú otvorené každý deň pre tých, ktorí sa chcú o korpuse niečo dozvedieť, pomôcť mu rozširovať sa a vylepšovať sa. Zároveň sme radi, že počas tohto osobitného týždňa sme mohli byť koncentrovane k dispozícii počas prezentácií a diskusií všetkým, ktorí sa s nami chceli stretnúť práve v týždni, v ktorom sa intenzívnejšie predstavuje slovenská veda.