

CHYTŘE NA SLOVNÍKY ZPRÁVA Z 6. MEZINÁRODNÍ KONFERENCE ELEX¹



Loňský lexikografický zefýr zavál evropské slovníkáře na nejzápadnější cíp Evropy, do portugalské Sintry, pouhých sedm osm kilometrů vzdušnou čarou od Cabo da Roca (kde notně fouká a kam jsme ovšemže také zavítali rozšířivše početné zástupy turistů). Na přelomu září a října se tu konal v pořadí šestý ročník bienální lexikografické konference eLex. Motto *Smart lexicography* — tedy chytrá lexikografie — bylo zvoleno záměrně tak, aby bylo význačné; sami organizátoři ho na konferenčním webu vysvětlují takto: v poslední dekádě se zásadně rozšířilo užívání chytrých telefonů — a slovníkáře jako by tento trend zaskočil, slovníky rychlému nástupu nových digitálních formátů nestačí a zůstávají pozadu. To je bohužel smutná pravda: pro mnohého lexikografa splývá online slovník s tištěnou formou převedenou do elektronické verze a vloženou na internet, v horším případě s pdfkem dostupným na webu. Nedozírný potenciál online prostředí tak většinou zůstává ladem a ke skutečným *digital-born* slovníkům je ještě daleká cesta, podmíněná nezbytnou změnou v uvažování autorů slovníků. (Čest výjimkám!) Zaklínáme-li se ustavičně uživatelem jakožto alfou a omegou slovníkářské práce, pak musí být jeho potřeby skutečně zohledněny. A pokud si uživatel navykl — ó jak rychle a ochotně! — na smartphony a nedá bez nich už ani ránu, pak je potřeba brát tento formát za výchozí a přizpůsobit mu své dosavadní zvyklosti, především co se struktury nabízených informací a grafického vzhledu týče.² Jinak hrozí, že slovníky budou vytlačeny „vševědoucím“ dotazovým řádkem Google; co dnes uživatelé nenajdou rychle a pohodlně, to je nezájímá — a možná se ukáže, že se bez slovníků většina lidí dost dobře obejde. Berme to však nikoliv jako důvod ke skepsi a defetismu, ale spíš jako výzvu pro slovníkáře, aby šli s dobou a své návyky, leckdy dosti zbytnělé a zautomatizované, zrevidovali a osvěžili novými podněty.

Druhý možný výklad sloganu posledního eLexu je chytré užití, resp. znovuvyužití slovníkového obsahu. Kompilace slovníku bývá většinou dlouhodobý a vyčerpávací proces, přitom ona nesmírná práce přinese výsledek jen jednorázový a klopotně nasbíraná data už leží mnohdy dále nevyužita a nepočítá se s jejich dalším upotřebením (odhlédneme-li od eticky problematického opisování ze starších slovníků). Dnešní iniciativa některých lexikografů vytěžit z jedněch dat co nejvíce je proto logická a hodna podpory i následování. Nová slovníková díla mohou vznikat ze starších zdrojů, resp. tatáž data mohou posloužit jako pramen vícero typům slovníků, z kterých by si uživatel mohl vybírat přesně podle svých představ. V rámci současného

1 Publikace vznikla za podpory projektu „Progres Q10. Jazyk v proměnách času, místa a kultury“ a projektu „Kreativita a adaptabilita jako předpoklad úspěchu Evropy v propojeném světě“, reg. č. CZ.02.1.01/0.0/0.0/16_019/0000734, financovaného z Evropského fondu pro regionální rozvoj.

2 Viz též Kosem, I. — Lew, R. — Müller-Spitzer, M. — Ribeiro Silveira, M. — Wolfer, S. — Local Partners (2019): The image of the monolingual dictionary across Europe. Results of the European survey of dictionary use and culture. *International Journal of Lexicography*, 32, 1, s. 92–114.



trendu e-lexikografie se totiž počítá s tím, že slovník bude přizpůsobitelný každému uživateli přesně podle jeho potřeb, které se člověk od člověka liší, a to často podstatně: jinak bude slovník využívat středoškolsky vzdělaný rodilý mluvčí, jinak cizinec učící se místnímu jazyku, jinak zas vědecký pracovník atd.

Právě možnost přizpůsobit si nastavení svým potřebám (např. omezení na rychlou a přímočarou informaci o základních významech, nebo naopak rozšíření o etymologický údaj či o okrajové významy ze starších slovníků) je jedním z oněch úžasných výdobytků, jež tištěné médium neumožňovalo, a rozhodně by nebylo rozumné se takovýchto prostředků zříkat. Aby se však mohlo týčtž, ale i starších dat používat opakovaně a uzpůsobitelně, je zapotřebí zajistit především technickou stránku věci: vyvinout nástroje podporující kompatibilitu různých formátů, API nástroje apod., na obecnější rovině přiblížit lexikografii sémantickému webu.

A nejen to, nezbytná je i dostatečně robustní infrastruktura, která by poskytla potřebnou podporu lexikografům — a ti z ní následně těžili ve prospěch svých cílových uživatelů. I ta existuje: jmenuje se ELEXIS a má za cíl „zprístupňování slovníků, lingvistických dat a jazykových nástrojů evropským komunitám“. ³ Asi nepřekvapí, že za ní stojí agilní Slovinci, konkrétně Simon Krek s Iztokem Kosemem (ten je také už po léta hlavním organizátorem eLexu), neméně klíčovou roli hraje také český zástupce (firma Lexical Computing v čele s Milošem Jakubíčkem). Některé projekty na ELEXIS přímo navázané byly představeny jako samostatné příspěvky; za všechny doporučujeme především nástroj Elexifinder, ⁴ což je snadno prohledávatelná databáze výstupů nejen čistě lexikografických, ale také těch z oborů se slovníkářstvím úzce souvisejících. Elexifinder umožňuje uživatelům prohledávat odborné články a videa z konferencí za pomoci klíčových slov a dalších podmínek (konkrétní konference, autor, jazyk/region apod.), všechny zdroje jsou volně stažitelné, a to včetně pečlivě oamotovaných videozáznamů, mj. z posledních konferencí eLex a Euralex. ⁵ Takto bohatý a dostupný zdroj dosud nebyl k dispozici a jistě ho ocení každý (nejen) evropský lexikograf.

Samotné konferenci předcházel jednodenní workshop *Collocations in Lexicography: Existing solutions and future challenges*. Znova tu přišla na přetřes otázka vymezení a klasifikace kolokací a víceslovných jednotek vůbec spolu s metodami jejich automatického vyhledávání. Workshop nenabídl ucelené řešení otázky víceslovných jednotek; autoři víceméně jednomyslně konstatovali, že ani léta výzkumů nepřinesla uspokojivé odpovědi na to, jak víceslovné jednotky vymežit, klasifikovat — a především spolehlivě vyhledat.

Další tři dny to byl už dokonalý pelmel lexikografických témat s úctyhodným rozptylem: od diachronní lexikografie přes neografiu až po lexikografii překladovou či

³ Více viz <https://elex.is/objectives/>. Zde jsou cíle shrnuty pod tři slogany: už žádné objevování Ameriky (*No more reinventing the wheel*), propojování nepropojeného (*Connecting the disconnected*) a osvěta (*Educating*).

⁴ <http://er.elex.is/>

⁵ Aktuálně jsou na webu kromě videozáznamů volně dostupné kompletní sborníky ze všech (!) ročníků konferencí Euralex (od r. 1983) i eLex (od r. 2009) a celé ročníky tří lexikografických časopisů v digitalizované podobě.



specializovanou (terminografie, studentské slovníky aj.); představovaly se projekty jak zdárně dokončené (např. finský slovník *sanakirja.fi*), tak i ty ve fázi zrodu či teprve chystané; kromě velkých jazyků byly zmíněny i ty malé a ohrožené (lívština, sámština, udmurtština, zapotečtina aj.). Už vzhledem k tematickému zaměření konference se často objevovala problematika layoutu, ale též metodiky práce. Čím dál více se počítá s aktivní rolí uživatelů v rámci lexikografického procesu: pojmy jako *crowdsourcing* či *citizen science* jsou skloňovány rok od roku častěji. Opakovaně zmínovány byly rovněž aplikace pomáhající při produkci textu (*writing assistants*), i to je způsob, jak vycházet vstříc aktuálním potřebám uživatelů.

Krátce se zastavme u plenárních přednášek. O první se postarala reprezentantka „domácího týmu“ z univerzit v Coimbre a Lisabonu, **Margarita Correia** [ová], která představila jakási „Pravidla portugalského pravopisu“. A byť její příspěvek vycházel na náš vkus z příliš preskriptivních pozic, byl zajímavý sám o sobě díky skutečnosti, že portugálština má — na rozdíl třeba od češtiny — polycentrickou normu: kromě samotného Portugalska má status úředního jazyka ještě v bývalých koloniích, tj. v Brazílii, Angole, Mosambiku, Guineji-Bissau, na Svatém Tomáši a Princově ostrově, Kapverdách, v Macau a ve Východním Timoru. Sladit různé pravopisné normy v jednu jistě nebylo jednoduché, a přesto se lusofonním lingvistům — již před třiceti lety — podařilo dobrat konsensu (*O Acordo Ortográfico da Língua Portuguesa de 1990*). **Alexander Geyken**, vedoucí redaktor úctyhodného webového slovníku němčiny DWDS.de, přehledně shrnul stav současné německojazyčné lexikografie, s akcentem na online prostředí a ony výše zmíněné „chytré“ a moderní trendy. **David Baines**, s až misionářským vzezřením a zápalem, představil křesťanskou mezinárodní neziskovou organizaci SIL, jež iniciuje lingvistický výzkum méně známých a užívaných jazyků, překlad Bible do nich a obecně i růst gramotnosti v zaostalých oblastech. **Maciej Piasecki** z Vratislavské technické univerzity zase nabídl perspektivu WordNetu jakožto relačního sémantického slovníku, ve kterém jsou lexikální významy, fungující jako základní stavební bloky celé struktury, primárně definovány hustou sítí vzájemných lexikálně-sémantických vztahů. Jako posledního — a rozhodně nejneortodoxnějšího — jmenujme **Matta Kohla**, sympatického Američana a nositele Ceny Adama Kilgarriffa pro rok 2019,⁶ který představil svou kratochvilnou zábavu: online hiphopový slovník.⁷ Nestává se příliš často, že by přednášející mluvil za doprovodu beatu a rapu linoucího se z reproduktorů — a publikum osvěžující exkurz mimo akademický rámec do této svérázné subkultury vesměs ocenilo. Ba co víc, k improvizovanému rapu se nechal strhnout i hlavní organizátor Iztok Kosem!

Z dílčích příspěvků už jen stručně zmíníme těch pár, které nás zaujaly. **Anna Dziewianko** [vá] se věnovala problematice reklam a jejich cílenosti na uživatele v online slovnících. Je jasné, že bezplatné webové slovníky se bez těchto reklam neobejdou, a pozornost lexikografů tak této zdánlivě marginálii patří po zásluze.

Jaka Čibej a **Špela Arhar Holdt** [ová] představili projekt automaticky sestaveného tezauru slovinštiny, který by měl být do použitelné podoby postupně upravován

6 Bližší informace o této ceně, udělované od r. 2016 v rámci eLexu, najdete zde: <https://kilgarriff.co.uk/prize/>

7 <https://therightrhymes.com/>



za pomoci veřejnosti: uživatelé hlasují o adekvátnosti jednotlivých synonym, čímž mění jejich pořadí, a podléjí se tak přímo na tvorbě slovníku.

Za rozsáhlejší výzkumný tým předvedla **Ana Frankenberg-Garcia**[ová] práci s textovým editorem *ColloCaid*, který pisatelé nabízí typické akademické kolokace přímo při psaní odborného textu. Nástroj je prozatím k dispozici v angličtině, ale autoři vybídli odbornou veřejnost ke spolupráci a k přípravě verzí v dalších jazycích.

Nejeli jsme jen jako pasivní přihlížeči: česká (a slovenská) lexikografie byla zastoupena mj. společným příspěvkem **Vladimíra Benka** a **Michala Škrabala** *Make my (Czechoslovak Word of the) Day*. Cílem projektu je přístupnou formou — každodenním zveřejňováním českého a slovenského ekvivalentu vybraného slova na webu — navazovat na společnou česko-slovenskou jazykovou tradici.

Doznáváme upřímně: byly to příjemně strávené čtyři dny, vyšperkované pěšími výlety k oceánu i večerním koupáním. Útulné prostředí hotelu Vila Galé bylo snad až zbytečně honosné, navíc vzhledem k lokalitě — daleko mimo centrum města⁸ — i značně nepraktické. Ostatně i sami organizátoři měli jisté pochybnosti o správnosti takovéto volby. Příští eLex by se měl konat v Brně, a jistě bude rozdílný, zda bude situován v centru města, k odlehlejšímu Pryglu, či dokonce k Veverské Bítýšce, kde člověk možná pookřeje v lůně přírody, zato šalinou se neprojede...

Michal Škrabal | Ústav Českého národního korpusu, Filozofická fakulta Univerzity Karlovy |
Nám. J. Palacha 2, 116 38, Praha 1
ORCID ID: 0000-0001-9358-3361
michal.skrabal@ff.cuni.cz

Dominika Kovářiková | Ústav Českého národního korpusu, Filozofická fakulta Univerzity Karlovy |
Nám. J. Palacha 2, 116 38, Praha 1,
ORCID ID: 0000-0002-4419-6901
dominika.kovarikova@ff.cuni.cz

8 V tomto ohledu je těžko překonatelný 4. ročník eLexu, konaný v srpnu 2015 na jihoanglickém hradě Herstmonceux. Podle pamětníků měl tento ročník nezapomenutelnou, až harrypotterovskou atmosféru, ale dostat se do obchodu či do restaurace bylo velmi nešťastné a dobrodružné.