



Akademický slovník současné češtiny

a software pro jeho tvorbu

aneb Slovníky a jejich uživatelé v 21. století

Sborník abstraktů z workshopu

Praha 29.–30. listopadu 2016

Sborník byl vydán s podporou grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny (DF13P01OVV011).

Editorky: MICHAELA LIŠKOVÁ, VERONIKA VODRÁŽKOVÁ, ZUZANA DĚNGEOVÁ

Korektura českých textů: HANA MŽOURKOVÁ, BARBORA PROCHÁZKOVÁ

Korektura slovenských textů: MAGDALENA KROUPOVÁ, BARBORA PROCHÁZKOVÁ

Překlad předmluvy do angličtiny a korektura anglických textů: TAMAH SHERMAN

Předtisková úprava: MARCELA HLADÍKOVÁ

Grafické zpracování obálky a logo workshopu: ALŽBĚTA ŠIRŮČKOVÁ

Vydal: Ústav pro jazyk český AV ČR, v. v. i., Praha 2016

ISBN 978-80-88211-02-0

Neoprávněné užití tohoto díla je porušením autorských práv a může zakládat občanskoprávní, správněprávní, popř. trestněprávní odpovědnost.

OBSAH

| | |
|--|----|
| Předmluva | 5 |
| Foreword | 10 |
| BARBORA ALBRECHTOVÁ Jak média zacházejí s variantními prostředky | 17 |
| KAMIL BARBIERIK Tvorb a, struktura a redakce hesla v DWS <i>Alexis</i> | 19 |
| MARTIN BENEŠ Zdroje a kritéria stylových charakteristik jazykových prostředků a jak je zjistit | 22 |
| VLADIMÍR BENKO Informačné technológie v projekte SSSJ | 24 |
| VÁCLAV CVRČEK Složení korpusu SYN2015 a jeho zpracování | 26 |
| ZUZANA DĚNGEOVÁ Revize hesel v prostředí webové aplikace <i>Alexis</i> | 28 |
| BRONISLAVA CHOCHOLOVÁ Opis prác na jednom zväzku <i>Slovníka súčasného slovenského jazyka</i> | 30 |
| ALEXANDRA JAROŠOVÁ Norma v kontexte spisovnosti a kodifikácie | 33 |
| VLADIMÍR JARÝ Specifika automatického generování slovníkového výstupu | 35 |
| STANISLAVA KLOFEROVÁ <i>Slovník nářečí českého jazyka</i> – úskalí lexikografické a dialektologické abstrakce | 38 |
| PAVLA KOCHOVÁ – ZDEŇKA OPAVSKÁ <i>Akademický slovník současné češtiny</i> | 40 |
| DOMEN KRVINA – MIJA MICHELIZZA Definitions of classifying adjectives in the new dictionary of the Slovenian standard language | 43 |
| MICHAL KŘEN Frekvenční slovníky češtiny | 47 |
| TOMÁŠ LIŠKA Představení DWS <i>Alexis</i> | 49 |
| MICHAELA LIŠKOVÁ Exemplifikace termínů v <i>Akademickém slovníku současné češtiny</i> | 51 |

| | |
|--|----|
| LUCIE MACHÁLKOVÁ – VERONIKA VODRÁŽKOVÁ Využití <i>Českého národního korpusu</i> při tvorbě ASSČ | 53 |
| MATEJ METERC – NATAŠA JAKOP Lexikografické spracovanie frazeologických variantov v novom slovníku slovenského spisovného jazyka | 55 |
| JANA NOVÁ – HANA GOLÁŇOVÁ Zpracování regionálně příznakových lexémů v ASSČ | 57 |
| ANDREJ PERDIH – JANOŠ JEŽOVNIK Na cestě k novému slovníku slovenského jazyka | 59 |
| HANA PROKŠOVÁ Konkurenční předložkové vazby v současné češtině | 61 |
| MARTIN PROŠEK Slovník češtiny pojatý jako moderní „learner’s dictionary“? | 63 |
| KAMILA SMEJKALOVÁ Morfologické varianty ve vztahu k normě a kodifikaci | 64 |
| MARTIN ŠEMELÍK Lesk a bída německo-českých terminologických slovníků | 66 |
| MILENA ŠÍPKOVÁ <i>Slovník nářečí českého jazyka</i> : charakteristika, vymezení, materiálová základna, výběr heslových slov | 68 |
| BARBORA ŠTĚPÁNKOVÁ Pro koho jsou synsémantika ve slovníku | 70 |
| VERONIKA ŠTĚPÁNOVÁ Výslovnost v českém výkladovém slovníku: co a jak | 72 |
| MARIE VACHKOVÁ Nad zpracováním hudební terminologie v německo-české databázi | 75 |
| VOJTĚCH VESELÝ K významům a funkcím reflexivního prvku <i>si</i> | 77 |

PŘEDMLUVA

Slovník je celý vesmír v abecedním pořádku.
Anatole France

Lexikografové se navzájem častují různými okřídlenými slovy, např. chtěl-li Bůh někoho potrestat, udělal ho slovníkářem, slovník je sérií kompromisů nebo slovník se nedá dokončit, lze jen přestat jej dělat. Ventilují tak svou frustraci nad nutností relativně jednoduchého lexikografického popisu něčeho tak složitého a nepravidelného, jako je živý jazyk. Přestože teoretické problémy neberou konce, dobrý slovník je především vydaný slovník, jak to lapidárně formuloval dánský lexikograf Henning Bergenholtz v rozhovoru pro *Jazykovědné aktuality*.

Na nový všeobecný výkladový slovník akademického typu reflektující současnou slovní zásobu Češi čekají už dlouho. Ovšem není očekáván jen tak ledajaký slovník, nýbrž slovník, který bude dostatečně rozsáhlý, spolehlivý a vstřícný ke čtenáři; který přispěje ke kultivaci naší mateřštiny, pomůže při nejrůznějších jazykových pochybách, s nímž bude radost pracovat při školní výuce a který možná potěší nečekanými jazykovými objevy při namátkové četbě. Solidní slovník, který bude stejně dobře jako laikům sloužit i odborníkům a který bude významnou oporou při vzniku dalších slovníkových děl. Zkrátka, na nový lexikon jsou kladeny vysoké požadavky.

Dostát jim se snaží pracovníci oddělení současné lexikologie a lexikografie ÚJČ AV ČR při přípravě *Akademického slovníku současné češtiny* (od roku 2012). Cesta k novému slovníku nebyla a není snadná z mnoha – jednoduše i jen velmi těžko představitelných – důvodů. Krátce před internetovým publikováním prvních hesel (počínajících písmenem A) se v prostorách Akademie věd České republiky na pražské Národní koná ve dnech 29.–30. 11. 2016 workshop *Akademický slovník současné češtiny a software pro jeho tvorbu aneb Slovníky a jejich uživatelé v 21. století*. Cílem workshopu je seznámit lingvistickou obec se zásadami zpracování nového lexikonu a umožnit odbornou diskusi, dále představit software pro psaní slovníku *Alexis*, který byl za uplynulých pět let vyvinut, a v neposlední řadě pak poskytnout prostor dalším lexikografům (tvořícím všeobecné slovníky jiných jazyků či slovníky speciální) i jiným jazykovědcům, aby sdíleli své zkušenosti a názory na to, co je dobrý slovník.

Workshop je podpořen grantovým projektem Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

V rukou držíte sborník rozšířených abstraktů. Nyní stručně shrneme, čemu se účastníci ve svých příspěvcích věnují. Zatímco abstrakty jsou ve sborníku řazeny pro přehlednost abecedně podle příjmení (prvního) autora, v následujícím přehledu upřednostníme tematické seskupení stanovené programem workshopu.

V ústředním příspěvku **Pavla Kochová** a **Zdeňka Opavská** syntetizujícím způsobem představují vznikající *Akademický slovník současné češtiny* (dále ASSČ) i přínosy lexikologicko-lexikografického výzkumu, který tvorbu nového slovníku nutně provází. Cílem ASSČ je komplexně popsat slovní zásobu současné češtiny v rozsahu 120 000–150 000 hesel. Primární materiálovou základnu tvoří synchronní korpus psaných textů SYN Ústavu Českého národního korpusu. Předpokládaným uživatelem tohoto všeobecného výkladového slovníku je středoškolsky vzdělaný rodilý mluvčí. Koncepční zásady jsou shrnuty v publikaci *Kapitoly z koncepce Akademického slovníku současné češtiny* (2016). Jednotlivé části slovníku budou postupně publikovány na internetu a počítá se i s knižní podobou dokončeného slovníku.

Systém pro psaní slovníku (DWS) *Alexis* prezentují tři zástupci sedmičlenného programátorského týmu. **Tomáš Liška** nastiňuje architekturu lexikografického softwaru, který vznikl na míru podle požadavků lexikografů. *Alexis* umožňuje nejen základní práci s hesly, ale také řadu nadstavbových funkcí usnadňujících revizní a redakční práce. Od počátku roku 2017 bude tento systém pro psaní slovníku volně dostupný. **Kamil Barbierik** se zaměřuje na popis tvorby, struktury a redakce hesla v DWS *Alexis*. Pro zpracovatele je důležité snadné ovládání softwaru a jeho přehledná struktura. V redakční fázi zpracování je vítána mj. funkce historie změn, uspořádaně zaznamenávající všechny přijaté i odmítnuté návrhy na úpravy různých prvků mikrostruktury hesla, zadávané interními i externími pracovníky. **Vladimír Jarý** přibližuje specifika automatického generování slovníkových výstupů v DWS *Alexis*. Prostřednictvím výstupního modulu je stanoven způsob, jakým se slovníková data zobrazují koncovému uživateli. Pozornost je věnována výstupu elektronickému a tištěnému, webovému a mobilnímu a konečně výstupu XML.

Softwarová problematika nemůže být při současných lexikografických diskusích opomíjena. Jak trefně poznamenává **Vladimír Benko**, pojem „počítačová lexikografie“ se v poslední době vyprazdňuje, protože „nepočítačová lexikografie“ už vlastně ani neexistuje. Ve svém příspěvku pojednává o dvou subtématech, a to o reprezentaci lexikálních dat v *Slovníku současného slovenského jazyka* (dále SSSJ) a o optimalizaci kolokačních gramatik pro korpusový manažer *Sketch Engine*.

Netriviálním tématem praktické organizace slovníkových prací se zabývají dvě autorky. **Bronislava Chocholová** popisuje proces tvorby jednoho dílu SSSJ: začíná přípravou hesláře, pokračuje koncipováním hesel, redakčními pracemi a odbornými konzultacemi. Po zapracování připomínek následuje recenzentské čtení, rozhodnutí kodifikační komise, ediční práce, zalomení textu a příprava doprovodných textů. Souběžně se připravuje další díl slovníku – přibližně tři roky před vydáním aktuálního dílu slovníku se začíná pracovat na dílu následujícím. **Zuzana Děngeová** se soustředí na spolupráci lexikografů s oborovými experty. Tato kooperace je nezbytnou součástí prací na všeobecném slovníku; předchází české výkladové slovníky obvykle uvádějí i několikastránkové seznamy spolupracujících odborníků. Autorka z týmu ASSČ ozřejmuje postup odborných revizí v redakčním modulu databáze *Alexis*.

Značná pozornost je věnována otázkám kodifikace a variantnosti ve slovníku. Jedna z vedoucích redaktorek SSSJ **Alexandra Jarošová** se zabývá normou v kontextu spisovnosti a kodifikace. Norma je v koncepci SSSJ vymezena jako soubor prostředků, jejichž noremnost je potvrzena statistickou mírou jejich výskytu. Kodifikace se realitě spisovných situací nejlépe přiblíží tehdy, když nebude fixovat prvky tzv. ideální normy (nepodloženého konstrukt) ani ignorovat výrazy, které jsou součástí spisovného úzu. Příspěvek **Kamily Smejkalové** zpracovává téma variantnosti na morfologické rovině jazyka ze dvou hledisek: z hlediska autora slovníku coby tvůrce kodifikace a z hlediska běžného mluvčího, který se o údaje ve slovníku opírá při tvorbě textu. Oba aktéři totiž vzájemně ovlivňují svou činnost, a tím i normu spisovného jazyka. Bude diskutována otázka, jak relevantní informace o normě vůbec získávat – *Český národní korpus* nemůže podat celkový obraz současné situace. Toto téma zajímá i **Barboru Albrechtovou**. Analyzuje, jakým způsobem mohou redakční pravidla velkých redakcí tištěných deníků ovlivnit obraz jazykové variantnosti v korpusech řady SYN. Korektorské zásahy mají bezpochyby vliv na jazykový materiál, z něhož lexikografové při přípravě slovníku vycházejí; autorka promýšlí důsledky tohoto stavu pro slovníkovou kodifikaci.

Tématem vystoupení **Hany Prokšové** jsou možnosti lexikografického zpracování konkurenčních předložkových vazeb v současné češtině. Podle její zkušenosti z jazykové poradny činí právě tato dynamická oblast mluvčím značné obtíže, a ve všeobecném výkladovém slovníku by proto uvítala co nejvíc informací o valenci a vazebnosti včetně stylového zhodnocení variant. **Martin Prošek** informuje o lingvistické strukturaci nové databáze dotazů v jazykové poradně ÚJČ. Databáze by měla vymezit, které oblasti jazykového systému jsou uživateli považovány za problematické a které naopak nikoli. Autor se dále zamýšlí nad tím, jak lze v moderním elektronickém slovníku řešit rozpor mezi požadavkem na stabilitu kodifikace a možnostmi elektronických médií, u kterých lze publikovaný obsah průběžně měnit. Řešení shledává ve vytvoření propracovaného popisného aparátu, který by v jednotlivých heslech umožňoval vyznačit jevy dlouhodobě stabilní, jevy příznačné jen pro určité období, jevy nové apod.

Lexikografické zpracování údajů o výslovnosti a stylu patří mezi diskutovaná témata. Hlavním zdrojem výzkumu spisovné výslovnosti by se měly stávat zejména veřejné oficiální projevy, které jsou považovány za kultivované, prestižní a adekvátní dané komunikační situaci. **Veronika Štěpánová** přehledně analyzuje možnosti a meze zdrojů vhodných k získávání aktuálních výslovnostních dat pro potřeby nového českého výkladového slovníku (zejména korpusey DIALOG, Monolog a fonetická databáze českých anglicismů). **Martin Beneš** se zabývá zdroji a kritérii stylové charakteristiky jazykových prostředků. Přibližuje svou představu, jak o ne/správnosti a stylové charakteristice jazykových prostředků rozhodovat pomocí experimentálních metod na základě recepce, nikoli na základě úzu (produkce).

Průnik korpusové lingvistiky a lexikografie je spojnicí následujících tří příspěvků. **Lucie Machálková** a **Veronika Vodrážková** přibližují tvorbu ASSČ se zřete-

lem k využívání dat z Českého národního korpusu i nástrojům na získávání těchto dat (webové rozhraní *KonText*, nástroj *Word Sketch* v aplikaci *Sketch Engine*). Je osvětleno, jak se využívá korpusové opory při několikafázové tvorbě hesláře a při zpracování konkrétních hesel. **Michal Křen** informuje o dvou frekvenčních slovnících češtiny vzniklých na datech Českého národního korpusu (*Frekvenční slovník češtiny*, 2004; *Frequency Dictionary of Czech: Core Vocabulary for Learners*, 2011). Nejprve uvádí faktory, které je třeba při vytváření frekvenčního slovníku zohlednit, aby výsledek nebyl pouhým seznamem slovních tvarů nebo lemmat. Dále popisuje oba frekvenční slovníky, jejich styčné body a odlišnosti v rozsahu, v doprovodných informacích, ve zdrojových datech a ve způsobu zpracování. **Václav Cvrček** se zaměřuje na korpus SYN2015 ve vztahu k jeho využití v lexikologickém a lexikografickém výzkumu. Tento stomilionový korpus je posledním přírůstkem v řadě korpusů ČNK mapujících současnou psanou češtinu. Koncepte jeho složení a zpracování doznala změn, které se týkají jak jeho poziční a strukturální anotace, repertoáru atributů, proporcí textových typů, tak celkového pojetí reprezentativnosti.

Další oddíl je věnován terminologii ve slovníku. **Marie Vachková** představuje zpracování hudební terminologie ve *Velké německo-české lexikální databázi* zpracovávané v Ústavu germánských studií FF UK. Zaměřuje se mj. na tyto body: výběr termínů do hesláře; terminologické varianty; rozsah informační nabídky heslové statě; problémy ekvivalence, polysémie a homonymie; synonyma na ose časové a regionální; metajazyk a encyklopedická složka; kolokace; podíl psaného a mluveného odborného jazyka. Součástí příspěvku je demonstrace vybraných muzikologických hesel. **Martin Šemelík** podrobuje kritice německo-české terminologické slovníky. Dospívá k závěru, že většina terminologických slovníků na našem trhu nesplňuje standardy moderní terminografie. Ve svém příspěvku však jen nepopisuje status quo, ale též navrhuje konkrétní kroky, které by překladovou terminografií zkvalitnily. **Michaela Lišková** se zabývá slovníkovou exemplifikací terminologických hesel. Přestože tvorba terminologické exemplifikace není pro lexikografa vždy snadná (některé požadavky na příkladovou část hesla jsou až protichůdné, je tedy třeba přistupovat na smysluplné kompromisy), pro uživatele slovníku je její existence jednoznačným přínosem. Problematika je ilustrována na typově různých heslech z rukopisu ASSČ.

Dialektologií v lexikografii a lexikografií v dialektologii se zabývají čtyři autorky. **Jana Nová** a **Hana Goláňová** představují přístup ke zpracování regionálně příznakových lexémů v ASSČ. K jejich označování se používají dva typy stylistických kvalifikátorů. Kvalifikátor *region.* s upřesněním oblasti (čes., mor., slez.) je určen pro lexémy s velkým areálem, pokrývajícím území určité nářeční skupiny nebo zasahujícím do více nářečních skupin. Kvalifikátor *dialekt.* je pak určen pro lexémy s menším nebo nesouvislým areálem. **Milena Šipková** a **Stanislava Kloferová** ve dvou příspěvcích zpravují o *Slovníku nářečí českého jazyka* (SNČJ). Tento unikátní lexikon vzniká od roku 2011 v dialektologickém oddělení ÚJČ v Brně. Je koncipován jako celonárodní, nediferenční, tj. relativně úplný; zahrnuje výrazy související s materiální a duchovní lidovou kulturou z celého území českého

národního jazyka. Zpracovává slovní zásobu z období přibližně od poloviny 19. do konce 20. století. Reflektována jsou úskalí lexikografické a dialektologické abstrakce; princip lemmatizace v nářečním slovníku umožňuje uspořádat bohatou hláskovou variabilitu nářeční lexikální zásoby tak, aby nebyl jazykový materiál podán roztržštěně, bez systémových souvislostí. Vyzdvížena je kulturní i badatelská hodnota SNČJ.

Samostatný blok je věnován přípravám nového slovníku slovinského jazyka. **Andrej Perdih** a **Janoš Ježovnik** představují tvorbu tohoto lexikonu s názvem *Slovar slovenskega knjižnega jezika*, třetí vydání (dále SSKJ³). SSKJ³ zachycuje všeobecně rozšířenou slovní zásobu současné slovinštiny, předpokládaný rozsah je 100 000 hesel. Koncepce slovníku byla publikována roku 2015 a nyní se pracovníci oddělení lexikologie Ústavu slovinského jazyka Frana Ramovše soustředí na vlastní psaní hesel. **Domen Krvina** a **Mija Michelizza** se zaměřují na možnosti výkladu významu relačních přídavných jmen v SSKJ³. **Matej Meterc** a **Nataša Jakop** tříbí lexikografické zpracování frazeologických variant v SSKJ³. Práci jim usnadňuje celoslovný sociolingvistický výzkum, na jehož základě bylo vytvořeno slovinské paremiologické minimum, obsahující necelých tisíc jednotek.

Poslední oddíl přesahuje do gramatiky. **Vojtěch Veselý** si za téma svého příspěvku určil významy a funkce reflexivního prvku *si*. Pro lexikografické zpracování sloves je klíčová otázka, zda je reflexivní prvek součástí lexému jako slovo-tvorný nebo významově prázdný morfém, nebo zda se ke slovesu připojuje jen v některých jeho užitích (jako reflexivní zájmeno, příp. jako tvarotvorný morfém). V prvním případě je reflexivní prvek součástí slovesného lemmatu, v druhém případě se reflexivní tvar či reflexivní konstrukce uvádí nanejvýš v příkladové části hesla. **Barbora Štěpánková** připomíná, že chápání synsémantik co do dělení či šíře není v české lingvistice jednotné či ustálené (zejména u částic, ale i u spojek). Slovníkář musí pečlivě zvážit, kterou ze stávajících klasifikací aplikovat, či zda raději nevytvořit klasifikaci novou. Dále je třeba brát v potaz, že jednojazyčný výkladový slovník vždy do určité míry supluje roli mluvnice či pravopisných pravidel, a podrobnost zpracování synsémantických hesel přizpůsobit uživatelským potřebám.

Autorům i uživatelům slovníků přejeme uspokojení ze zacházení s lexikony; autorům hodně sil, vzhledu a týmového ducha, uživatelům pak, aby snadno a rychle nacházeli právě ty informace, které hledají.

Michaela Lišková

FOREWORD

A dictionary is the entire universe in alphabetical order.
Anatole France

Lexicographers shower each other with various winged words, e.g. that when God wanted to punish someone, he made him a dictionary-maker, a dictionary is a series of compromises, or that you can't finish a dictionary, you can only stop working on it. In this way, they ventilate their frustration over the necessity of a relatively simple lexicographic description of something as complicated and irregular as a living language. Though the theoretical problems are never-ending, as the Danish lexicographer succinctly put it in an interview for the Czech journal *Jazykovědné aktuality* [Linguistics News], a good dictionary, above all, is a published dictionary.

The Czechs have been waiting a long time for a new general explanatory dictionary reflecting the contemporary lexicon. Of course, they are not expecting just any old dictionary, but rather, a dictionary that will be sufficiently extensive, dependable and reader-friendly; that will contribute to the cultivation of their mother tongue, help in the most various moments of linguistic uncertainty, that will be a pleasure to use in teaching, and that might even delight them with unexpected linguistic discoveries when just leafing through it at random. A respectable dictionary that will serve experts just as well as laymen and will be a significant pillar in the creation of further dictionaries. In short, great demands have been placed on the new lexicon.

The employees of the Department of Contemporary Lexicology and Lexicography of the Czech Language Institute, Czech Academy of Sciences, have been trying to fulfill them in the preparation of the *Akademický slovník současné češtiny* [Academic Dictionary of Contemporary Czech, hereafter ADCC] since 2012. The journey toward the new dictionary has not been and still is not an easy one for many reasons, some easily imaginable, others very difficult to imagine. Shortly before the publication of the first entries on the internet (starting with the letter A), the workshop *Akademický slovník současné češtiny a software pro jeho tvorbu aneb Slovníky a jejich uživatelé v 21. století* [The Academic Dictionary of Contemporary Czech and the software for its creation or Dictionaries and their users in the 21st century] is taking place in the building of the Czech Academy of Sciences on Národní street in Prague on November 29–30, 2016. The aim of the workshop is to familiarize the linguistics community with the principles of the development of the new lexicon and to enable expert discussion, present the dictionary-writing software *Alexis* developed over the past five years, and last but not least, to provide space for other lexicographers (who are creating general dictionaries of other languages or special dictionaries) and linguists to share their experiences and opinions on what makes a good dictionary.

The workshop is supported by the grant project of the Program of Applied Research and Development of National and Cultural Identity (abbreviated in

Czech as NAKI) of the Czech Ministry of Culture *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* [A new journey toward a modern monolingual explanatory dictionary of contemporary Czech] (DF13P01OVV011).

You are holding in your hands a book of extended abstracts. We will now briefly summarize the topics dealt with by the authors of the contributions. While the abstracts in this book are ordered alphabetically by the (first) author's surname for the sake of clarity, in the overview below we give preference to the thematic grouping determined by the workshop program.

In the main text, **Pavla Kochová** and **Zdeňka Opavská** offer a synthesizing presentation of the emerging *Akademický slovník současné češtiny* [Academic Dictionary of Contemporary Czech] as well as the contributions to lexicological and lexicographic research that necessarily come along with the creation of this new dictionary. The aim of the ADCC is to describe the contemporary Czech lexicon of 120,000–150,000 entries. The primary material basis consists of the synchronic corpus of written texts (SYN) from the Institute of the Czech National Corpus. The anticipated user of this general explanatory dictionary is a native speaker with a secondary-school education. The conceptual principles are summarized in the publication *Kapitoly z koncepce Akademického slovníku současné češtiny* [Chapters from the Conception of the Academic Dictionary of Contemporary Czech] (2016). The individual parts of the dictionary will be gradually published on the internet, and the publication of the finished dictionary in book form is also planned.

Three representatives of the seven-member programming team present the dictionary-writing system (DWS) *Alexis*. **Tomáš Liška** outlines the architecture of the lexicographic software that was custom-created based on the requirements of the lexicographers. *Alexis* enables not only the basic work with the entries, but also a series of superstructural functions facilitating the revision and editing work. This DWS will be freely accessible from the beginning of 2017. **Kamil Barbierik** focuses on the description of the creation, structure and editing of entries in the DWS *Alexis*. The easy operation of the software and its transparent structure are important for dictionary authors. In the editing phase, the function showing the history of changes, recording all accepted as well as rejected editing suggestions for various elements of the microstructure of the entry, entered by internal and external employees, in an orderly manner. **Vladimír Jarý** explains the specifics of the automatic generation of dictionary outputs in the DWS *Alexis*. Through the output module, the way in which the dictionary data are displayed for the end user is established. Attention is devoted to the output in electronic, printed, web, mobile and finally XML form.

In contemporary lexicographic discussions, the issue of software cannot be neglected. As **Vladimír Benko** aptly notes, the concept of “computer lexicography” has become empty in recent times, because “non-computer lexicography” does not actually exist any longer. In his contribution, he discusses two sub-themes, the representation of lexical data in the *Slovník súčasného slovenského jazyka* [Dictionary of the Contemporary Slovak Language] (hereafter DCSL) and the optimization of collocation grammars for the corpus manager *Sketch Engine*.

Two authors devote attention to the non-trivial topic of the practical organization of work on dictionaries. **Bronislava Chocholová** describes the process of the creation of one volume of the DCSL: it begins with the preparation of the list of dictionary entries, continues with the conception of the entries, the editorial work and specialized consultations. The integration of the comments is followed by the review procedure, the decision of the codification committee, editing work, pagination of the text and the preparation of the accompanying texts. The next volume of the dictionary is prepared concurrently – approximately three years prior to the publication of the current volume of the dictionary, work on the next volume begins. **Zuzana Děngeová** concentrates on the collaboration between lexicographers and specialized experts. This cooperation is a necessary part of the work on the general dictionary; the previous Czech explanatory dictionaries usually offer a list of collaborating experts which is sometimes up to several pages long. The author from the ADCC team elucidates the process of specialized revisions in the editing module of the *Alexis* database.

Significant attention is devoted to questions of codification and variation in the dictionary. One of the head editors of the DCSL, **Alexandra Jarošová**, deals with the norm in the context of standard language and codification. The norm is delimited in the conception of the DCSL as a set of means, the normative character of which has been confirmed by the statistical degree of their occurrence. Codification will approach the reality of standard-language situations the most when it does not stabilize elements of the so-called ideal norm (an unsupported construct) or ignore expressions which are a part of standard usage. The contribution by **Kamila Smejkalová** examines the topic of variation on the morphological level of language from two perspectives: from the perspective of the author of a dictionary as a creator of codification and from the perspective of the ordinary speaker, who relies on the information in the dictionary when creating a text. That is, both actors influence the other's activity, and in doing so, also the standard language norm. The question of how to gain relevant information about the norm will be discussed – the Czech National Corpus cannot provide an overall picture of the contemporary situation. **Barbora Albrechtová** is also interested in this topic. She analyses how the editing rules of the large editorial teams of printed daily newspapers can influence the depiction of linguistic variety in the SYN corpus series. Copy editing undoubtedly has an influence on the linguistic material the lexicographers use as the basis for the preparation of the dictionary; the author thinks over the consequences of this state of affairs for dictionary codification.

The topic of the presentation by **Hana Prokšová** concerns the possibilities for the lexicographic treatment of competing prepositional phrases in contemporary Czech. According to her experience from the language consulting center, it is this dynamic area that creates significant problems for speakers, which is why she would welcome the maximum of information on valency and binding tendencies in the general explanatory dictionary, including the stylistic evaluation of variants. **Martin Prošek** provides information on the linguistic structuration of the new

database at the language consulting center of the Czech Language Institute. The database should delimit which areas of the language system are considered problematic and which, conversely, are not. The author further considers how to address the contradiction between the demand for codification stability and the affordances of the electronic media, in which published content can be gradually changed. He sees a solution in the creation of a highly-developed apparatus for description which would enable the marking of long-term stable phenomena, phenomena characteristic only for a certain period, new phenomena, and the like.

The lexicographic treatment of information on pronunciation and style is a frequently discussed topic. The main source of research on standard pronunciation should become, above all, public official speech, which is considered to be cultivated, prestigious and adequate for the given communication situation. **Veronika Štěpánová** provides an analytical overview of the possibilities and limits of the sources appropriate for gaining up-to-date pronunciation data for the needs of the new Czech explanatory dictionary (above all the corpora DIALOG, Monolog and the phonetic database of Czech Anglicisms). **Martin Beneš** deals with the sources and criteria of the style characteristics of linguistic means. He explains his notion of how to decide about the in/correctness and stylistic characterization of linguistic means through the use of experimental methods on the basis of reception, not on the basis of use (production).

The intersection of corpus linguistics and lexicography is what connects the next three contributions. **Lucie Machálková** and **Veronika Vodrážková** familiarize readers with the creation of the ADCC with consideration given to the use of data from the Czech National Corpus as well as to the tools used to gain this data (the web interface *KonText*, the *Word Sketch* tool in the *Sketch Engine* application). They explain how the corpus support materials are used in the multi-phase creation of the list of entries and in the processing of specific entries. **Michal Křen** provides information on two frequency dictionaries of Czech based upon data from the Czech National Corpus (*Frekvenční slovník češtiny*, 2004; *Frequency Dictionary of Czech: Core Vocabulary for Learners*, 2011). He first presents the factor which must be considered in the creation of a frequency dictionary if the result is not to be a mere list of word forms or lemmas. He further describes both frequency dictionaries, their points of contact and differences in extent of supplemental information provided, in source data and the manner of treatment. **Václav Cvrček** focuses on the corpus SYN2015 in relation to its use in lexicological and lexicographic research. This 100-million-word corpus is the latest edition in the series of corpora in the CNC that map out contemporary written Czech. The conception of its composition and processing has undergone changes concerning both its positional and structural annotations, repertoire of attributes, text type proportions, as well as its overall concept of representativeness.

A further section is devoted to terminology in the dictionary. **Marie Vachková** presents the treatment of musical terminology in the *Velká německo-česká lexikální databáze* [Large German-Czech Lexical Database] being created at the Institute of Germanic Studies, Faculty of Arts, Charles University. She focuses, among

others, on these points: the selection of terms for the list of entries; terminological variants; the extent of the information offered in the entry text; problems of equivalence, polysemy and homonymy; synonyms on the temporal and regional axis; metalanguage and the encyclopedic component; collocations; the proportion of written and spoken specialized language. The contribution includes selected musicological entries. **Martin Šemelík** provides a critique of German-Czech terminological dictionaries. He reaches the conclusion that most terminological dictionaries on the Czech market do not fulfill the standards of modern terminography. In his contribution, however, he does not only describe the status quo, but also proposes concrete steps which would increase the quality of bilingual/translation terminography. **Michaela Lišková** deals with the dictionary exemplifications of terminological entries. In spite of the fact that the creation of terminological exemplification is not always easy for the lexicographer (some requirements for the example section of the entry are even contradictory and it is thus necessary to reach sensible compromises), for the dictionary user, its existence is a definite benefit. The issue is illustrated using entries of various types from the ADCC manuscript.

Four authors devote their texts to dialectology in lexicography and lexicography in dialectology. **Jana Nová** and **Hana Goláňová** present an approach to the treatment of regionally marked lexemes in ADCC. There are two types of stylistic qualifiers used to mark them. The qualifier *region.* with the specification of the region (*čes., mor., sléz.*) [Bohemia, Moravia, Silesia] is used for lexemes with a large area, covering the territory of a certain dialect group or belonging to more than one dialect group. The qualifier *dialekt.* is then used for lexemes with a smaller or a disjointed area. **Milena Šipková** and **Stanislava Kloferová** report on the *Slovník nářečí českého jazyka* [Dialect Dictionary of the Czech Language] (DDCL) in two contributions. Work on this unique lexicon has been going on since 2011 at the Department of Dialectology of the Czech Language Institute in Brno. It is conceived of as covering the whole country, non-differentiating, i.e. relatively complete; it contains expressions related to the material and religious folk culture from the entire territory of the Czech national language. It covers the lexicon from the period approximately from the mid-19th century to the end of the 20th century. The pitfalls of lexicographic and dialectological abstraction are reflected; the principle of lemmatization in the dialect dictionary enables the organization of the rich phonetic variability of the dialectal lexicon in a way that the language material is not presented in a fragmented manner, without systemic contexts. The cultural and research value of the DDCL is emphasized.

A separate section is devoted to the preparations of the new dictionary of Slovenian. **Andrej Perdih** and **Janoš Ježovnik** present the creation of this lexicon with the title *Slovar slovenskega knjižnega jezika*, tretja izdaja [the 3rd edition of the Dictionary of the Standard Slovenian Language] (hereafter DSSL³). The DSSL³ captures the generally expanded lexicon of contemporary Slovenian, with an anticipated volume of 100 000 entries. The dictionary conception was published in 2015 and now the employees of the Lexicology Department of the Fran Ramovš

Institute of the Slovenian Language are concentrating on their own entry-writing. **Domen Krvina** and **Mija Michelizza** focus on the possibilities for explaining the meaning of relational adjectives in the DSSL. **Matej Meterc** and **Nataša Jakop** are refining the lexicographic treatment of phraseological variants in the DSSL. Their work has been made easier by a nationwide Slovenian sociolinguistic research project, on the basis of which the Slovenian paremiological minimum, containing just under one thousand units, was created.

The final section extends into the area of grammar. **Vojtěch Veselý** chose as his topic the meanings and functions of the reflexive element *si*. A key question for lexicography is whether the reflexive element is a part of the lexeme as a derivational or a semantically empty morpheme, or whether it joins to the verb only in some cases of its use (as a reflexive pronoun or in some cases as an inflectional morpheme). In the first case, the reflexive element is part of the verb lemma, in the second case, the reflexive form or construction is presented, at most, in the example section of the dictionary. **Barbora Štěpánková** reminds us that the understanding of synsemantic words is not unified or stabilized (above all in the case of particles, but also in the case of conjunctions), in terms of neither their classification nor their scope. A dictionary-maker must carefully consider which of the existing classifications to apply or whether it would not be better to create a new classification. In addition, it is necessary to consider the fact that monolingual explanatory dictionary always serves as a substitute for a grammar or orthography manual, and to tailor the treatment of the degree of detail of the synsemantic entries to the needs of users.

We hope the authors and dictionary users are satisfied with their use of the lexicons. We wish the authors energy, insight, and team spirit. As for the users, may they be able to quickly and easily find the information they are looking for.

Michaela Lišková

JAK MÉDIA ZACHÁZEJÍ S VARIANTNÍMI PROSTŘEDKY

BARBORA ALBRECHTOVÁ

*oddělení jazykové kultury, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
albrechtova@ujc.cas.cz*

Key words: corpus, mass media, language variants

Autoři každého slovníkového díla jsou postaveni před otázku, jak se vyrovnat s variantními prostředky odrážejícími dynamiku jazyka. Při rozhodování, které prostředky uvést a které ne, lexikografům pomáhají různé zdroje, jako nejvýznamnější zdroj při přípravě nového *Akademického slovníku současné češtiny* lze označit databázi *Českého národního korpusu*. Heslář připravovaného slovníku se totiž zakládá na třech referenčních korpusech řady SYN, a to konkrétně SYN2000, SYN2005 a SYN2010 (Michalec, 2014).

Jedním ze tří hlavních makrotypů textů zachycených v korpusech jsou publicistické texty. Masmédia pracují s určitými žánry, jež jsou institucionalizovány, a výsledná podoba textů tak podléhá „žánrové politice“ daného média (Kaderka, 2013). Tato institucionalizace se promítá i do jazykové podoby textů, v rámci redakce bývají sdílána pravidla, jak zacházet s určitými jazykovými prostředky. Tato pravidla obvykle vytvářejí a na jejich dodržování pak dohlížejí korektoři, kteří bývají běžnou součástí velkých redakcí.¹ Již v minulosti se poukazyvalo na to, že korektorské zásahy mají významný vliv na výslednou podobu textů (např. Jedlička, 1950; Hlavsa, 1990).²

Příspěvek se zaměří na to, jakým způsobem mohou redakční pravidla velkých redakcí tištěných deníků (data byla získána od seriózních deníků jako MF DNES, Hospodářské noviny apod.) ovlivnit obraz jazykové variantnosti v korpusech řady SYN. Bude sledovat, jak jednotlivé redakce přistupují k otázce variantnosti, z jakých zdrojů čerpají a do jaké míry se interní pravidla redakcí promítají do konečné podoby komunikátů. Předběžné výsledky ukazují, že ačkoliv finální podoba textů neodpovídá pravidlům redakce vždy, jsou některá z těchto pravidel v textech jednotlivých redakcí i napříč redakcemi preferovaná (varianty bez kvantit, např. *leukemie*, v textech MF DNES, varianta *bestseller* s dvěma -ll- v textech MF DNES a Hospodářských novin apod.). Cílem tohoto příspěvku je upozornit na možný posun obrazu variantnosti určitých lexikálních jednotek v korpusových datech a zamyslet se nad možnými důsledky pro kodifikaci, jako je zejména uvádění příznaků u variant, v krajním případě samotné zachycení či nezachycení varianty ve slovníku.

¹ Pravidla mohou mít materiální podobu (slovníčky, nástěnky apod.) nebo mohou fungovat jako sdílené povědomí o tom, jak se s určitými jazykovými prostředky běžně zachází.

² Jedlička explicitně upozorňuje na to, že korektorské zásahy mají vliv na materiál, z něhož lexikografové čerpají při přípravě slovníku.

SEZNAM LITERATURY A ZDROJŮ

- HLAVSA, Z. (1990): K jazykové stránce komunikace v hromadných sdělovacích prostředcích. *Slovo a slovesnost*, 51, s. 124–130.
- JEDLIČKA, A. (1950): Zjišťování mluvnické normy soudobé spisovné češtiny. *Naše řeč*, 34, s. 121–132.
- KADERKA, P. (2013): Komunikační vzorce a komunikační žánry. In: S. Čmejrková a kol. (eds.), *Studie k moderní mluvnici češtiny 1. Pragmatické aspekty češtiny*. Olomouc: Univerzita Palackého v Olomouci, s. 83–101.
- KŘEN, M. – ČERMÁK, F. – Hlaváčová, J. – HNÁTKOVÁ, M. – JELÍNEK, T. – KOCEK, J. – KOPŘIVOVÁ, M. – NOVOTNÁ, R. – PETKEVIČ, V. – PROCHÁZKA, P. – SCHMIEDTOVÁ, V. – SKOUMALOVÁ, H. – ŠULC, M. (2014) [online]: *Korpus SYN, verze 3 z 27. 1. 2014*. Praha: Ústav Českého národního korpusu FF UK. Cit. 2016-09-01. <<http://www.korpus.cz>>.
- MICHALEC, V. (2014): Poznámky k výstavbě hesláře připravovaného výkladového slovníku současné češtiny. In: K. Dudová (ed.), *Varia 22*. Nitra: Katedra slovenského jazyka Filozofickej fakulty UKF a Slovenská jazykovedná spoločnosť pri JÚLŠ SAV, s. 274–282.
- RULFOVÁ, M. (1986): Variantnost v tvarosloví. In: J. Petr (ed.), *Mluvnice češtiny 2. Tvarosloví*. Praha: Academia, s. 266–267.

TVORBA, STRUKTURA A REDAKCE HESLA V DWS ALEXIS¹

KAMIL BARBIERIK

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze
Břehová 7, 115 19 Praha 1
kamil.barbierik@foxcom.eu*

Key words: dictionary database, dictionary writing system, dictionary entry editing

Příspěvek se věnuje popisu nástrojů, které DWS *Alexis* nabízí pro tvorbu a další zpracování hesla.

Základním nástrojem pro práci s heslem je zpracovatelský formulář. Jedná se o sadu vstupních polí reprezentujících elementy mikrostruktury hesla, zorganizovaných tak, aby formulář usnadňoval orientaci slovníkáře. Přehlednost je docílena zejména postupným hierarchickým shlukováním vstupních polí na několika úrovních. Na první úrovni je formulář rozdělen do čtyř částí, a to:

1. **Záhlaví formuláře**, které obsahuje informace o hesle jako celku: datum vytvoření, datum editace, jméno zpracovatele, zařazení hesla do příslušného hesláře, forma slovníkového výstupu, lexikografická poznámka atd. Z hlavičky je také přístupná většina funkcí pro heslo jako např. historie korektur či archiv.

2. **Varianty hesla**. Lze pracovat s jednou nebo více variantami hesla. První varianta je hlavní a podle ní se řídí řazení ve slovníku. Při přidávání varianty lze využít funkci kopírování poslední varianty, která vytvoří novou variantu s předvyplněnými údaji. Pořadí variant lze při zpracování měnit.

3. **Významy hesla**. Lze vytvořit a editovat libovolný počet významů. Významy se v první úrovni zobrazují ve formě minimalizované karty, kde lze editovat jen samotný výklad významu; pro editaci dalších detailních informací o významu je třeba kartu rozbalit. Lze přidávat i podvýznamy vázané na jednotlivé významy.

4. **Sekce připojení**. Na rozdíl od editace mikrostruktury pomáhá tvořit makrostrukturu slovníku, a to připojením různého druhu: propojení, přihnízdování nebo linkování. Připojení v systému umožňují přiřadit i heslo, které zatím v databázi neexistuje; bude propojeno po svém vzniku.

Zastavme se krátce u nástrojů pro připojování hesel.

a) Nástroj pro editaci umožňuje k významu připojit synonyma nebo antonyma spolu s označením, zda se jedná o neutrální nebo příznakové slovo.

b) Dalším připojením je propojení slova nebo části textu výkladu významu k jeho základnímu tvaru. Tento proces, tzv. lemmatizace, je automatizován využí-

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

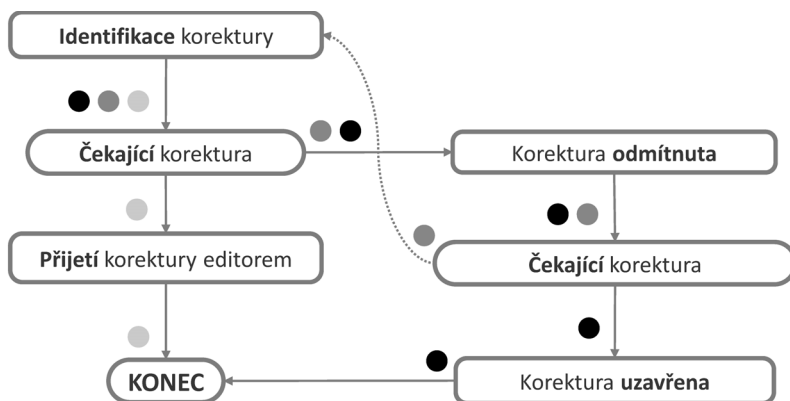
tím externího nástroje *Morphodita*, avšak neřeší případy, kdy je základové slovo homonymní. Existuje proto i možnost ručního připojení, které má před automatickým připojením přednost.

c) Zajímavým způsobem připojení je tzv. skupinkování významů. Tato funkce umožňuje definovat propojení významů nebo jejich částí, které jsou pak mezi sebou sdílené a tvoří skupinku s naprosto stejnými výklady významů. Editace sdíleného významu kteréhokoli člena skupinky způsobí stejnou změnu v ostatních členských významech.

Tato propojení výrazně napomáhají udržovat v datech konzistenci.

Nezbytnou součástí tvorby slovníku je redakce hesel. Systém *Alexis* implementuje redakční modul, který umožňuje zadat korekturu na konkrétní místo mikrostruktury hesla a také poskytuje rozhraní pro další práci s korekturou.

Cyklus redakce začíná zadáním korektury. Korektura pak čeká na reakci zpracovatele, který má možnost ji buď přijmout, nebo zamítnout. V případě přijetí je korektura aplikována a technický cyklus redakce končí. Jestliže zpracovatel s korekturou nesouhlasí, může ji odmítnout. Korektor se pak na základě argumentů uvedených do příslušného pole rozhodne, zda korekturu s případným poupravením znovu zadá, nebo ji vyhodnotí jako nevhodnou a heslo uzavře bez jejího aplikování do hesla, čímž cyklus končí (viz obrázek níže: diagram technického cyklu redakce neznázorňuje možnost přímé aplikace korektury, která se využívá zejména při opravách překlepů). Každá akce a komunikace v redakčním modulu se zaznamenává do historie, čímž modul umožňuje nahlížet do historie vývoje konkrétní části mikrostruktury hesla. Redakční modul podporuje i redakci externími odborníky, pro něž může být vytvořen seznam relevantních hesel; přístup do databáze pak umožňuje pouze práci s daným seznamem. Interní lexikografové tak mají korektury i od externích odborníků přímo v systému a mohou s nimi běžným způsobem pracovat.



SEZNAM LITERATURY A ZDROJŮ

- BARBIERIK, K. – DĚNGEOVÁ, Z. – HOLCOVÁ HABROVÁ, M. – JARÝ, V. – LIŠKA, T. – LIŠKOVÁ, M. – VIRIUS, M. (2014): Simple and Effective User Interface for the Dictionary Writing System. In: A. Abel – C. Vettori – N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus*. 15–19 July 2014, Bolzano/Bozen, s. 125–135.
- BARBIERIK, K. – DĚNGEOVÁ, Z. – JARÝ, V. – LIŠKA, T. – LIŠKOVÁ, M. – VIRIUS, M. (2014): Dictionary Writing System: Output and Editorial Modules. In: J. Ministr – M. Tvrdíková (eds.), *Proceedings of IT for Practice 2014, 9–10 October 2014*. Ostrava: VŠB – Technical University of Ostrava, s. 17–23.
- BARBIERIK, K. – BODLÁK, M. – DĚNGEOVÁ, Z. – JARÝ, V. – LIŠKA, T. – LIŠKOVÁ, M. – NOVÝ, J. – VIRIUS, M. (2015): The Current Status of the Development of the ALEXIS Dictionary Writing System. In: K. Gajdošová – A. Žáková (eds.), *Natural Language Processing, Corpus Linguistics, Lexicography*. Lüdenscheid: RAM-Verlag, s. 17–25.

ZDROJE A KRITÉRIA STYLOVÝCH CHARAKTERISTIK JAZYKOVÝCH PROSTŘEDKŮ A JAK JE ZJISTIT

MARTIN BENES

*oddělení jazykové kultury, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
benes@ujc.cas.cz*

Key words: common knowledge, intersubjectivity, language rules (norms), linguistic correctness, ontological “socialism”

Príspevek zaprvé predstavuje tzv. „socialistickou“ konceptualizaci předmětné oblasti lingvistiky v aplikaci na problematiku jazykové správnosti (resp. stylových charakteristik jazykových prostředků), zadruhé popisuje zdroje/příčiny ne/správnosti (resp. stylové charakteristiky) jednotlivých jazykových prostředků a identifikuje její kritéria a zatřetí načrtává empirický program, který by bylo v návaznosti na tato zjištění žádoucí realizovat.

1. Podle „socialistické“ konceptualizace „jazyka“ (Itkonen, 1978; 2003; 2008) existují v předmětné oblasti lingvistiky dva typy entit. Zaprvé jazyková **pravidla (normy)**, která nejsou přístupná pomocí smyslů, existují tzv. v **intersubjektivitě** myslí jednotlivých uživatelů jazyka, a to v podobě **sdílených, tříúrovňových znalostí**. Rozlišují se pravidla fonologická, sémantická, gramatická a distribuční (stylová). Pravidla jsou implicitní, podvědomá a liší se od svého explicitního, vědomého zachycení v tzv. větách o pravidlech. Zadruhé tu jsou **jazykové prostředky**, které pravidla předepisují/zapovídají užívat při odpovídajícím jazykovém jednání, jehož prostřednictvím se tyto prostředky dostávají do času a prostoru, kde jsou pak vnímatelné smysly. Soubor *všech* těchto prostředků tvoří **úzus**.

2. Je-li jednání pomocí určitého prostředku v souladu s příslušným sdíleným pravidlem, je takový prostředek správný, pokud ne, je nesprávný. Stylová charakteristika prostředku se přitom konceptualizuje jako jeho správnost (patřičnost) v těch a těch typech komunikačních situací (a nesprávnost v typech jiných). Co do jazykové ne/správnosti existují tzv. jasné a tzv. nejasné případy. Jde-li o **jasný případ**, znají všichni členové řečového společenství příslušné pravidlo, a tedy i způsob správného jazykového jednání s jistotou. Jde-li však o **nejasný případ**, nemáme obvykle o ne/správnosti (stylové charakteristice) daného prostředku žádné spolehlivé poznatky, neboť o ní nelze rozhodnout na základě intuice, ale je k tomu třeba výzkum. Tyto prostředky je žádoucí zmapovat a charakterizovat. V nejasném případě může být stav např. takovýto: V úzu existuje $\pm 30\%$ výskytů varianty X a $\pm 70\%$ výskytů varianty Y. To může být způsobeno dvěma modelovými situacemi. Jednak tím, že $\pm 100\%$ členů řečového společenství se řídí pravidlem „použij prostředek Y“, ale v $\pm 30\%$ případech ho $\pm 100\%$ z nich porušuje; jednak tím, že existuje $\pm 30\%$ členů řečového společenství řídicích se (novým) pravidlem „použij prostředek X“ a $\pm 70\%$ členů řečového společenství

řídících se (starým) pravidlem „použij prostředek Y“, přičemž obě skupiny tato pravidla dodržují v ± 100 % případech. V první situaci je oněch ± 30 % výskytů prostředku X pro všechny členy řečového společenství (stále) nesprávných; za správný všichni považují jen prostředek Y, i když relativně často chybní. Ve druhé situaci je prostředek X pro ± 30 % členů řečového společenství (už) správný a pro ± 70 % členů řečového společenství (stále, ale možná jen zatím) nesprávný, zatímco prostředek Y je pro stejných ± 30 % členů řečového společenství (už) nesprávný a pro oněch ± 70 % členů řečového společenství (stále ještě) správný. Ve druhé situaci **nelze říct, zda je prostředek X, resp. Y ne/správný absolutně**, ale jen relativně, tj. jak pro kterého člena řečového společenství.

3. O ne/správnosti (stylové charakteristice) nejasných prostředků nelze rozhodovat přímočaře jen na základě úzu, tj. produkce, ale primárně na základě recepce (srov. Bermel – Knittl, 2012). Je nutno experimentálními metodami ověřovat, zda prostředek, jehož status (stylová charakteristika) nás zajímá, je podle členů řečového společenství v souladu, či nesouladu s pravidlem, které v daném případě sdílí. Je-li při určitém jednání možné v řečovém společenství využít víc pravidel (viz druhou situaci výše), je dobré vědět něco o tom, kolik kterých členů (= s jakými sociolingvistickými charakteristikami) sdílí to či ono pravidlo. Kvůli relevanci výsledků je třeba testovat minimálně desítky osob. Z toho důvodu je nutné upřednostnit metody kvantitativní (zejm. psycholingvistické) před kvalitativními. Testování se nabízí založit na tom, že prostředek, který není v souladu s příslušným pravidlem, by měl v testovaných osob vyvolat tzv. **nenulovou reakci** (Homoláč, 1998, s. 141). Experimenty by pak měly hledat nějaký její experimentálně detekovatelný korelát.

SEZNAM LITERATURY A ZDROJŮ

- BERMEL, N. – KNITTL, L. (2012): Morphosyntactic variation and syntactic constructions in Czech nominal declension: corpus frequency and native-speaker judgements. *Russian Linguistics*, 36, s. 91–119.
- HOMOLÁČ, J. (1998): K pojetí normy u Ivy Nebeské. *Naše řeč*, 81, s. 138–147.
- ITKONEN, E. (1978): *Grammatical Theory and Metascience: A Critical Investigation into the Methodological and Philosophical Foundations of 'Autonomous' Linguistics*. Amsterdam: John Benjamins.
- ITKONEN, E. (2003): *What is Language? A Study in the Philosophy of Linguistics*. Åbo: Åbo Akademis tryckeri.
- ITKONEN, E. (2008): The central role of normativity in language and linguistics. In: J. Zlatev – T. P. Racine – C. Sinha – E. Itkonen (eds.), *The Shared Mind: Perspectives on Intersubjectivity*. Amsterdam – Philadelphia, PA: John Benjamins, s. 279–305.

INFORMAČNÉ TECHNOLOGIE V PROJEKTE SSSJ¹

VLADIMÍR BENKO

*oddelenie súčasnej lexikológie a lexikografie, Jazykovedný ústav Ľudovíta Štúra SAV
Panská 26, 811 01 Bratislava
vladob@juls.savba.sk*

Key words: collocation profiles, corpus, DWS, monolingual dictionary, Sketch Engine

Úvod

Pojem „počítačová lexikografia“ sa v poslednom období začal vyprázdňovať, pretože „nepočítačová lexikografia“ už vlastne ani neexistuje. Informačné (a komunikačné) technológie sa obyčajne zúčastňujú na všetkých fázach lexikografických projektov a informatická gramotnosť patrí medzi základné kompetencie každého člena lexikografického tímu. Vo veľkých projektoch je navyše potrebné, aby členom tímu bol aj informatik s príslušným zázemím v relevantných oblastiach (počítačová a korpusová lingvistika, reprezentácia štruktúrovaných dát, nástroje na riadenie projektu a pod.).

Vzhľadom na šírku problematiky sa v našom príspevku zameriame na dve oblasti, v ktorých sa náš prístup líši od podobných projektov akademických slovníkov, a to na reprezentáciu lexikálnych dát a na optimalizáciu kolokačných gramatík pre korpusový manažér *Sketch Engine*.

Príprava heslových statí

Jednou z charakteristík slovníkových projektov je ich dlhodobosť – doba tvorby jednozväzkového slovníka nezriedka dosahuje 10 rokov, pri mnohozväzkových slovníkoch sa čas obyčajne počíta v desaťročiach. Na druhej strane informačné a komunikačné technológie sa rozvíjajú veľmi rýchlo, preto riešenia uskutočnené na začiatku projektu sa časom môžu ukázať ako neadekvátne. Práce na *Slovníku súčasného slovenského jazyka* – projektu Jazykovedného ústavu Ľudovíta Štúra SAV – sa začali v 90. rokoch v inštitúcii prakticky „nezasiahnutej“ informatizáciou, do ktorej práve dorazili prvé počítače triedy PC XT. Ak to porovnáme s dneškom, keď každý lexikograf má na stole pracovnú stanicu s viacerými monitormi, ktorá je napojená cez lokálnu sieť na jazykové zdroje (textové korpusy, elektronické slovníky, lingvistické publikácie a pod.) uložené na ústavných serveroch a ktorá je cez rýchle internetové linky pripojená na externé zdroje v internete, nemáme vôbec pocit, že prešlo len 20 rokov.

V súčasnosti väčšina veľkých lexikografických projektov používa špecializovaný softvér – DWS (*Dictionary-Writing System*), ktorý obyčajne kombinuje editor textu kódovaného pomocou jazyka XML s databázou umožňujúcou efektívne

¹ Príspevok vznikol v rámci grantového projektu VEGA č. 2/0015/14 *Slovník súčasného slovenského jazyka – 5. etapa (Koncipovanie a redigovanie slovníkových hesiel a s tým spojený lexikologicko-lexikografický výskum)*.

uchovávanie dát a ich vyhľadávanie podľa jednotlivých informačných polí príslušných databázových záznamov.

V našom projekte uplatňujeme prístup založený na jednoduchších dátových štruktúrach zakódovaných pomocou proprietárneho jazyka triedy LML (*Light-weight Markup Language*) a štandardného textového editora rozšíreného o funkcie farebného zvýraznenia syntaxe, čo nám umožňuje výraznejšie odlíšiť jednotlivé elementy heslovej state (Benko, 2001). V príspevku uvedieme charakteristiku nášho riešenia a porovnáme výhody a nevýhody oboch prístupov.

Optimalizácia kolokačných profilov

Jednou zo základných činností lexikografov pri tvorbe heslových statí výkladového slovníka je analýza výskytov príslušného heslového slova v reálnych kontextoch.

Vyhodnotenie všetkých výskytov lexikálnej jednotky v neustále sa zväčšujúcich textových korpusoch je však reálne len pri heslových slovách zo spodnej časti frekvenčného zoznamu. Napríklad vo frekvenčnom zozname získanom z najnovšej verzie najväčšieho korpusu SNK *prim-7.0-public-all* nachádzame takmer 39 000 jednotiek frekvenciou výskytu ≥ 1000 , čo obyčajne predstavuje dve tretiny heslára bežného jednozväzkového výkladového slovníka. Nie je preto prekvapujúce, že tu začíname pociťovať ako najväčší problém nadbytok dát, ktoré nie je možné spracúvať tradičnými postupmi a na ktorých analýzu potrebujeme nové nástroje. Jedným z takýchto nástrojov je *Sketch Engine* (SkE) umožňujúci sumarizovať kolokačné správanie lexikálnych jednotiek vo forme tzv. slovných skíc (*word sketches*), ktoré sa v našej lexikografickej tradícii označujú ako *kolokačné profily* (Kilgarriff et al., 2004).

Predpokladom použitia nástroja SkE je okrem morfosyntakticky anotovaného korpusu dostatočnej veľkosti existencia súboru pravidiel (tzv. kolokačnej gramatiky), pomocou ktorých sa v kolokačnom profile pre zvolené slovo zobrazujú gramatické a kolokačné vzťahy. V rámci projektu SSSJ používame nástroj SkE od roku 2007, keď kolokačná gramatika pre slovenčinu ešte neexistovala a gramatika pre češtinu nebola momentálne k dispozícii. Vyvinuli sme vlastnú kolokačnú gramatiku, ktorá postupom času prekonala na základe spätnej väzby od lexikografov niekoľko modifikácií, bola optimalizovaná na lexikografické použitie a stala sa základom pre kompatibilné gramatiky aj pre iné jazyky (Benko, 2014). V príspevku uvedieme hlavné charakteristiky nášho prístupu k tvorbe kolokačných gramatík.

ZOZNAM LITERATÚRY A ZDROJOV

- BENKO, V. (2001): Počítačová podpora slovenských lexikografických projektov – retrospektívny pohľad. In: A. Jarošová (ed.), *Slovenčina a čeština v počítačovom spracovaní*. Bratislava: Veda, s. 181–194.
- BENKO, V. (2014): Compatible Sketch Grammars for Comparable Corpora. In: A. Abel – C. Vettori – N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus*. 15–19 July 2014, Bolzano/Bozen, s. 417–430.
- KILGARRIFF, A. – RYCHLÝ, P. – SMRŽ, P. – TUGWELL, D. (2004): The Sketch Engine. In: G. Williams – S. Vessier (eds.), *Proceedings of the eleventh EURALEX International Congress EURALEX 2004 Lorient, France, July 6–10, 2004*. Lorient: Université de Bretagne-Sud, s. 105–116.
- Slovenský národný korpus – prim-7.0-public-all*. Bratislava: Jazykovedný ústav L. Štúra SAV (2015) [online]. Cit. 2016-09-01. <<http://korpus.juls.savba.sk>>.

SLOŽENÍ KORPUSU SYN2015 A JEHO ZPRACOVÁNÍ

VÁCLAV CVRČEK

*Ústav Českého národního korpusu, Filozofická fakulta Univerzity Karlovy
Panská 7, 110 00 Praha 1
vaclav.cvrcek@ff.cuni.cz*

Key words: corpus design, Czech, representativeness, written language

Stomilionový korpus SYN2015 (Křen et al., 2015) je nejnovějším přírůstkem v řadě korpusů *Českého národního korpusu* (dále ČNK) mapujících současnou psanou češtinu. Vzhledem k tomu, že reprezentativní stomilionové korpusy této řady jsou v rámci ČNK vůbec nejpoužívanější, lze očekávat, že SYN2015 bude v následujících pěti letech sloužit jako hlavní zdroj informací pro korpusově orientovaný výzkum češtiny. Oproti svým předchůdcům (zejm. SYN2010) doznala koncepce jeho složení a zpracování řady změn, které se týkají jak jeho poziční a strukturální anotace, repertoáru atributů, proporcí textových typů, ale také celkového pojetí reprezentativnosti. Příspěvek shrne nejdůležitější odlišnosti korpusu SYN2015 ve vztahu k jeho využití v lexikologickém a lexikografickém výzkumu.

Korpus se svým složením znatelně odlišuje od svých předchůdců, stomilionových reprezentativních korpusů SYN2000 až SYN2010. Nejde přitom ani tak o to, že texty ve všech korpusech jsou vzájemně disjunktní (žádný text není obsažen ve dvou korpusech), změnila se především koncepce metaanotace textů i metodika jeho vyvažování. V oblasti žánrové a text-typové klasifikace bylo třeba přistoupit k řadě změn, které byly vynuceny postupnou proměnou spektra publikovaných textů. SYN2015 jako korpus, který si klade za cíl reprezentovat současnou tištěnou (tj. veřejnou) psanou češtinu, tak na nejobecnější úrovni klasifikace textů obsahuje vedle krásné literatury (fiction) a publicistiky také třetinu literatury oborové. Ta se do značné míry překrývá s dřívější odbornou literaturou, její záběr je ovšem trochu širší (obsahuje např. texty zaměřené na profesionály v různých oborech, které ovšem nemají striktně vědecký ráz). Podstatné změny zaznamenal i mediální trh (nárůst počtu tzv. lifestyleových časopisů), na něž musí reagovat i reprezentativní korpus; vedle klasických periodik (celostátních i regionálních) se tak v SYN2015 objevuje i rozšířené spektrum těchto nově vzniklých časopisů. Složení korpusu se nově neodvíjí od výsledků sociologických průzkumů čtenosti, je na místo toho založeno na rovnoměrném zastoupení všech hlavních textových typů a podtypů (podrobněji viz Cvrček et al., 2016).

Z hlediska anotace jednotlivých tokenů byl korpus SYN2015 obohacen vedle již tradiční lemmatizace a morfologické anotace (obojí ve vylepšené verzi) také o značkování syntaktické. To je založeno na schématu vypracovaném pro PDT (*Prague Dependency Treebank*), dále upravené a zjednodušené pro potřeby uživatelů ČNK. Novinkou je také přidání atributu *proc*, který nese informaci o tom,

který z nástrojů provedl finální morfologickou disambiguaci daného tokenu, což může napomoci při odhadování spolehlivosti tagování (pravidlová disambiguace je tradičně považována za spolehlivější než stochastická). Podrobněji o anotaci viz Křen et al., 2016.

SEZNAM LITERATURY A ZDROJŮ

- CVRČEK, V. – ČERMÁKOVÁ, A. – KŘEN, M. (2016): Nová koncepce synchronních korpusů psané češtiny. *Slovo a slovesnost*, 77, s. 83–101.
- KŘEN, M. – CVRČEK, V. – ČAPKA, T. – ČERMÁKOVÁ, A. – HNÁTKOVÁ, M. – CHLUMSKÁ, L. – JELÍNEK, T. – KOVÁŘÍKOVÁ, D. – PETKEVIČ, V. – PROCHÁZKA, P. – SKOUMALOVÁ, H. – ŠKRABAL, M. – TRUNEČEK, P. – VONDŘIČKA, P. – ZASINA, A. (2015) [online]: *SYN2015: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Cit. 2016-09-01. <<http://www.korpus.cz>>.
- KŘEN, M. – CVRČEK, V. – ČAPKA, T. – ČERMÁKOVÁ, A. – HNÁTKOVÁ, M. – CHLUMSKÁ, L. – JELÍNEK, T. – KOVÁŘÍKOVÁ, D. – PETKEVIČ, V. – PROCHÁZKA, P. – SKOUMALOVÁ, H. – ŠKRABAL, M. – TRUNEČEK, P. – VONDŘIČKA, P. – ZASINA, A. (2016) [online]: Representative Corpus of Contemporary Written Czech. In: N. Calzolari et al. (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, s. 2522–2528. Portorož: ELRA. Cit. 2016-09-01. <http://www.lrec-conf.org/proceedings/lrec2016/pdf/186_Paper.pdf>.

REVIZE HESEL V PROSTŘEDÍ WEBOVÉ APLIKACE ALEXIS¹

ZUZANA DĚNGEOVÁ

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
dengeova@ujc.cas.cz*

Key words: cooperation among lexicographers and experts, dictionary writing system (DWS), expert review

Akademický slovník současné češtiny si klade za cíl podat uživatelům podrobný a komplexní popis lexikálních jednotek současné češtiny. Shromáždění a správné uvedení všech informací o hesle klade na lexikografy značné lingvistické i mimo-lingvistické nároky, zároveň se nelze obejít bez cenné spolupráce s experty z oboru jazykovědy a oborů dalších.

Spolupráce s odbornými konzultanty je nezbytnou součástí prací na slovníku, výkladové slovníky obvykle uvádějí i několikastránkové seznamy spolupracujících odborníků². Forma, jakou konzultace s odborníky probíhají, nebývá ve slovnících a slovníkové literatuře popsána, případně jen značně útržkovitě, např. v úvodních textech *Slovníku spisovného jazyka českého* se na s. XIX uvádí: „...[odborní poradci] se této práci věnovali soustavně, při četných poradách (...) s nevšední ochotou odpovídali na mnohé dotazy písemné i telefonické...“; také J. Filípec (1958, s. 224) stručně poznamenává: „Svou péčí o odborné názvosloví zainteresoval [PSJČ] mnoho odborníků k systematické práci na úseku odborné terminologie (uvedených spolupracovníků je 65 kromě vědeckých ústavů, institucí, úřadů a závodů) (...)“. Více k podobě komunikace mezi lexikografy a experty z jiných oborů bychom se mohli dozvědět snad jen ze vzpomínek a vzpomínkových textů lexikografů, nicméně k tomuto tématu nebývá upírána téměř žádná pozornost.

Předpokládáme, že konzultace probíhaly ústně i písemně, písemné materiály pak byly ve formě psané, tištěné, předávány ve formátu textových dokumentů. Následovaly konzultace, zanesení úprav do rukopisu slovníku.

Při rozvíjení specializovaného lexikografického softwaru *Alexis*, v jehož prostředí ASSČ vzniká, jsme se zaměřili také na otázku usnadnění komunikace a spolupráce mezi lexikografy a odbornými konzultanty. Cílem bylo minimalizovat práci lexikografů spojenou se zanášením změn z textových dokumentů – generovaných v odpovídajícím formátu z databáze *Alexis* a posílaných konzultantům – do zdrojové databáze. Chtěli jsme také eliminovat chyby způsobené přenosem textových dokumentů mezi počítači s různými operačními systémy.

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

² Viz např. Seznam spolupracovníků SSJČ na s. XX–XXI, předmluva ASCS na s. 5, Předmluva PSJČ na s. X–XI.

Za tímto účelem byl v prostředí softwaru *Alexis* upravován redakční modul, který v první řadě slouží k redakčním úpravám hesel (a pak také k vzájemným zpracovatelským revizím). V redakčním modulu probíhá revidování hesel – hesla lze prohlížet v tiskovém náhledu a ke konkrétním částem hesla (např. ke všem jednotlivým tvarům tvaroslovné charakteristiky, k jednotlivým výkladům významu polysémních hesel, k jednotlivým exemplifikačním blokům) lze zadat návrh na úpravu, který je adresován autorovi hesla, případně může být cílen na jiné zpracovatele. Upravený redakční modul je pak využíván při revizích hesel odbornými konzultanty, ať už z řad spolupracovníků z Ústavu pro jazyk český AV ČR, v. v. i., nebo externími spolupracovníky vně Ústavu. Odborným konzultantům je umožněn přístup do webové aplikace *Alexis* prostřednictvím uživatelského účtu pro externisty – jedná se o přístup s omezenou možností prohlížení dat a bez zapisovacích a editačních práv. Externí uživatelé mohou získat přístup s možností prohlížet přidělená hesla v textových dokumentech (role *Nahlížení*), nebo navíc s možností zasílat lexikografům korektury a poznámky k heslům (role *Odborník*).

Díky redakčnímu modulu mohou externisté zapisovat své návrhy na úpravy a připomínky přímo v prostředí webové aplikace *Alexis* a zpracovatelé slovníku mohou na jejich základě jednoduše upravovat hesla přímo v databázi. Předávání textových materiálů a následné zanášení úprav tímto odpadá. K dalším přínosům redakčního modulu patří možnost okamžité editace hesel; možnost ke konkrétním uživatelským účtům nastavit cílového adresáta / cílové adresáty korektur; možnost k jednotlivým heslům připsat konkrétní otázky pro konzultanty; dále je uchovávána historie všech úprav, které byly na daném hesle (nebo na konkrétním poli hesla) zadány a jak (a kým) byly zpracovány. Více o funkcích redakčního modulu a práci s ním, včetně manuálu pro externí uživatele softwaru *Alexis*, bude představeno v příspěvku.

SEZNAM LITERATURY A ZDROJŮ

- Akademický slovník cizích slov* (1995). Praha: Academia. (ASCS)
- BARBIERIK, K. – BODLÁK, M. – DĚNGEOVÁ, Z. – JARÝ, V. – LIŠKA, T. – LIŠKOVÁ, M. – NOVÝ, J. – VIRIUS, M. (2015): The Current Status of the Development of the ALEXIS Dictionary Writing System. In: K. Gajdošová – A. Žáková (eds.), *Natural Language Processing, Corpus Linguistics, Lexicography*. Lüdenscheid: RAM-Verlag, s. 17–25.
- FILIPEČ, J. (1958): Akademický Příruční slovník jazyka českého dokončen. *Slovo a slovesnost*, 19, s. 211–224.
- KOCHOVÁ, P. – OPAVSKÁ, Z. (eds.) (2016): *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i.
- Příruční slovník jazyka českého* (1935–1957). Praha: Státní nakladatelství, Školní nakladatelství, Státní pedagogické nakladatelství. (PSJČ)
- Slovník spisovné češtiny pro školu a veřejnost* (1978). Praha: Academia. (SSČ)
- Slovník spisovného jazyka českého* (1960–1971). Praha: Nakladatelství ČSAV. (SSJČ)
- ZÍSKAL, A. (1938): Závěrečný příspěvek k lexikografické teorii a praxi (Podrobný slovník současného spisovného jazyka a uspořádání hesel v něm). *Slovo a slovesnost*, 4, s. 212–222.

OPIS PRÁC NA JEDNOM ZVÄZKU SLOVNÍKA SÚČASNÉHO SLOVENSKÉHO JAZYKA¹

BRONISLAVA CHOCHOLOVÁ

oddelenie súčasnej lexikológie a lexikografie, Jazykovedný ústav Ľudovíta Štúra SAV
Panská 26, 811 01 Bratislava
bronah@juls.savba.sk

Key words: monolingual dictionary, time schedule, workflow

Štandardný zväzok *Slovníka súčasného slovenského jazyka* (ďalej SSSJ) má rozsah približne 1 000 tlačенých strán v stanovenej typografickej úprave, čo sa rovná objemu asi 4 000 normostrán a 7 mil. bajtov textu skoncipovaného v elektronickej podobe. Počet heslových slov s komplexnou informáciou o jednotkách lexikálnej zásoby slovenčiny sa v jednotlivých zväzkoch líši: prvý zväzok a–g obsahuje 30 293 heslových slov, druhý zväzok h–l ich má 25 418 a zatiaľ posledný, tretí zväzok obsahuje 16 815 heslových slov.

Na slovníku sa súbežne vykonávajú práce na dvoch zväzkoch. Na skoncipovaných častiach aktuálneho zväzku SSSJ prebiehajú redakčné práce v troch cykloch. Autorky, ktoré nerobia internú redakciu aktuálneho zväzku, začínajú popri vnášaní redaktorských zásahov do autorského konceptu koncipovať aj ďalší zväzok. To znamená, že približne tri roky pred vydaním aktuálneho zväzku sa začína pracovať na ďalšom zväzku. Redakčné práce sa robia priebežne s koncipovaním a dobiehajú ešte rok a pol po skoncipovaní. Prededitačné a editačné práce a kompletná príprava tlačových podkladov zväzku trvajú približne rok.

Celý proces tvorby jedného zväzku zahŕňa tieto etapy:

1. Príprava heslára – ide o sumár heslárov z existujúcich slovníkov v elektronickej podobe (*Slovník slovenského jazyka, Krátky slovník slovenského jazyka, Pravidlá slovenského pravopisu*).

2. Koncipovanie hesiel – vypracovanie heslových statí v objeme približne 4 000 normostrán, pričom každá heslová stať obsahuje tieto štandardné komponenty: 1. záhlavie ukazujúce pravopisnú podobu heslového slova s informáciou (v prípade potreby) o jeho výslovnosti a pôvode a s podrobnou gramatickou charakteristikou, 2. údaje o subkategoriatných gramatických príznakoch (napr. počítateľnosť, hromadnosť, stupňovateľnosť, prísudková pozícia heslového slova a iné), 3. informáciu o sémantickej spájateľnosti (napr. výlučná spájateľnosť slova s pomenovaniami označujúcimi ľudí alebo predmety), 4. informáciu o povahe slova z hľadiska jazykových, komunikačných (funkčných) a štýlových noriem,

¹ Príspevok vznikol v rámci grantového projektu VEGA č. 2/0015/14 *Slovník súčasného slovenského jazyka – 5. etapa (Koncipovanie a redigovanie slovníkových hesiel a s tým spojený lexikologicko-lexikografický výskum)*.

5. sémantizačné poznámky encyklopedického a pragmatického charakteru, 6. výklad významu, 7. usúvzťažnenie so synonymami a opozitami, 8. typické spojenia a citátové doklady, 9. ustálené spojenia rozdelené do niekoľkých funkčných typov, 10. prihniezdované slová a relevantné odkazy; norma, ktorú by mala každá autorka pri koncipovaní dodržať (ak nevykonáva na pracovisku iné vedeckovýskumné alebo vedecko-organizačné činnosti), je 800–1 000 heslových slov za rok.

3. Redakčné práce, ktoré zahŕňajú:

a) priebežnú počítačovú redakciu textu sledujúcu dodržiavanie stanovenej lexikografickej techniky; vykonáva sa pomocou validačných programov špeciálne vyvinutých pre potreby SSSJ;

b) internú redakciu textu spočívajúcu v opravách zabezpečujúcich štandardnú a porovnateľnú kvalitu každého hesla; každá interná redaktorka číta len pridelenú časť skoncipovaného textu; po zapracovaní pripomienok internej redakcie pripraví autorky text na vedeckú redakciu;

c) vedecká redakcia textu sledujúca správnosť a primeranosť spracovania jazykového materiálu z hľadiska aplikácie jazykovednej teórie a metodológie, pričom každá redaktorka číta celý rukopis.

4. Odborná konzultácia – terminologické heslá pripomienkujú externí odborní konzultanti (približne 70 konzultantov zo 100 odborov), pričom táto fáza trvá približne rok. Pripomienky jednotlivých odborníkov zapracúvajú buď autorky do svojho konceptu, alebo editorky v rámci záverečných prác.

5. Recenzenské čítanie – celý skoncipovaný text čítajú traja recenzenti – lingvisti, a to približne tri mesiace. Ich pripomienky zapracúvajú editorky v rámci záverečných prác.

6. Zasadnutia kodifikačnej komisie – na základe vopred pripravených podkladových materiálov riešia členovia komisie problémové javy; výsledky zapracúvajú editorky v rámci záverečných prác.

7. Editačné práce – zjednocujúce zásahy, ktoré vykonáva šesť editoriek a hlavná redaktorka pomocou objektivej analýzy a počítačových nástrojov v celom skoncipovanom texte a ktorých cieľom je dodržanie jednotného spracovania desiatok štruktúrnych a obsahových parametrov každého hesla. Záverečné zjednocovacie práce celého textu prebiehajú v dvoch fázach:

- slovnodruhovú editáciu: zjednocovanie slovnodruhovej a sémanticko-formálnej príslušnosti hesiel; sledovanie konzistentnosti a proporčnosti spracovania jednotlivých informačných zón hesiel: grafická podoba slova, varianty, gramatický aparát v záhlaví, gramaticko-sémantický aparát medzi záhlavím a výkladom, výslovnosť;
- abecednú editáciu: zjednocovanie vybraných lexikálno-gramatických javov, ako sú napr. vlastné mená, ustálené slovné spojenia, frazeológia, synonymá, antonymá, homonymá, funkčno-štylistické javy; okrem toho sa do textu vnášajú aj pripomienky odborných konzultantov a recenzentov a výsledky zo zasadnutí kodifikačnej komisie.

8. Zalomenie textu – najskôr sa urobia tzv. stĺpcové korektúry, pričom opravy vnášajú do textu v elektronickej podobe ešte editorky, následne sa text odovzdá grafikovi, ktorý pripraví stránkové korektúry a vnáša do textu zásahy korektoriek.

9. Príprava príloh, tirážnych a úvodných strán, grafickej inštrukcie pre používateľa uverejnenej na predsádkach slovníka a obálky vrátane anotácie na zadnej strane slovníka.

ZOZNAM LITERATÚRY A ZDROJOV

Interné materiály oddelenia súčasnej lexikológie a lexikografie. Bratislava: Jazykovedný ústav Ľudovíta Štúra SAV.

NORMA V KONTEXTE SPISOVNOSTI A KODIFIKÁCIE¹

ALEXANDRA JAROŠOVÁ

*oddelenie súčasnej lexikológie a lexikografie, Jazykovedný ústav Ľudovíta Štúra SAV
Panská 26, 811 01 Bratislava
sasaj@juls.savba.sk*

Key words: codification, monolingual dictionary, literariness, norm, Slovak language

1. Chápanie normy v podstate osciluje medzi: **1. útvarom** s určitými kladne chápanými vlastnosťami a **2. hodnotiacim meradlom**, vzorom, ktorý hovoriaci prirodzene pociťujú ako záväzný. Táto oscilácia vystihuje aj skutočnú ambivalentnú povahu normy, ktorú J. Dolník výstižne nazýva „hodnotiace a hodnotové vnímanie jazyka“ (Dolník, 2010, s. 101).

1.1. Ak jazykový prostriedok použitý v úze vnímame ako taký, ktorý sa v danej situácii bežne používa, teda ako dobrý (so zreteľom na dosiahnutie určitého cieľa), potom ho „priradujeme“ k norme ako k súboru pravidelne používaných, „normálnych“ jazykových prostriedkov². Súbor jazykových prostriedkov, ktoré sa konvencionalizovali na základe **pravidelného používania** (= **hodnotiace kritérium**) tvorí **normu ako útvar**. Takto, teda z hľadiska úzu a „útvarového“, na normu nazerá pražský funkcionalizmus.

1.2. Jazykový útvar vyhovujúci určitej (kladne vnímanej) charakteristike nadobúda v momente jeho reflexie **hodnotový status**. Je vzorom jazykového správania, etalónom s regulačnou silou. Je nielen realizáciou časti možnosti, ktoré systém ponúka, ale aj modelom pre ďalšiu materializáciu týchto možností (Coseriu, 1979).

2. Niektorým jazykovedcom je dopredu jasné, ako by mala (spisovná) norma vyzeráť, a podľa toho „explicitne vyberajú“ spisovný úzus, ktorý by mal byť „prameňom poznávania normy“ (Ružička, 1967, s. 15). V takom prípade sa norma ako hodnotovo vnímaný súbor prostriedkov, spontánne tvorený (niekedy v súčinnosti s poradenskou činnosťou jazykovedcov) v samotnom jazyku, mení na „**ideálny útvar**“ autoritatívne regulovaný jazykovedcami neberúcimi dostatočný ohľad na väčšinový úzus. Tento **ideálny hodnotiaci štandard** sa hyperbolizuje do podoby **pravidla určujúceho správnosť/nesprávnosť prostriedkov**. Normovosť prostriedkov sa externalizuje, teda narába sa s ňou tak, akoby nebola prítomná v samotnom jazyku, ale bola výtvorom jazykovedcov.

¹ Príspevok vznikol v rámci grantového projektu VEGA č. 2/0015/14 *Slovník súčasného slovenského jazyka – 5. etapa (Koncipovanie a redigovanie slovníkových hesiel a s tým spojený lexikologicko-lexikografický výskum)*.

² Ak hovoríme, že „vnímame jazykový prostriedok ako normálny“, teda hovoríme o **pocite** normálnosti, naznačujeme tým, že pred svojím zexplicitnením v úze je norma súčasťou jazykového povedomia človeka, napr. v podobe jazykového citu.

3. Koncepcie, s ktorými sa na Slovensku pristupovalo k tvorbe výkladových slovníkov, pokryli postupne všetky vymenované chápania normy. *Slovník súčasného slovenského jazyka* (ďalej SSSJ) vzniká teda v istom kontinuálnom spisovno-jazykovom prostredí charakterizovanom striedavým uplatňovaním protichodných kodifikačných postojov. SSSJ si musel vymedziť pojmovú náplň tradičných konceptov z oblasti teórie spisovného jazyka – *norma*, *funkčnosť*, *zaužívanosť* a pod. Treba mať na zreteli, že preskriptívna kodifikácia (delimitácia jazykových prostriedkov na spisovné a nespisovné) je súčasťou kultúry používateľov slovenčiny. *Normu* vymedzujeme ako súbor prostriedkov, ktorých normovosť je potvrdená štatistickou mierou ich výskytu. Vychádzame teda z *frekvencie* výskytu jazykového prostriedku v úze, z jeho *zaužívanosti*. *Funkčnosť* výrazu chápeme v prvom rade ako jeho adekvátnosť v istej komunikačnej situácii (formálnej alebo neformálnej). Kombinácia týchto dvoch kritérií nám umožňuje konštatovať, že určitý výraz (napríklad „nesystémový bohemizmus“ *zdielať* < *sdíleť*) je frekventovaný a používa sa aj v tzv. spisovných situáciách (navrhované slovenské ekvivalenty sú komunikačne menej efektívne). Slovo *zdielať* je súčasťou **reálnej, ale nekodifikovanej spisovnej normy (spisovného úzu)**. Slovo *pretekovisko*, uvedené v *Krátkom slovníku slovenského jazyka*, sa v úze nevyskytuje. Kodifikovala sa **systémová potencia**, nie reálne slovo. Je súčasťou **ideálnej normy**. V úze sa používa spojenie *pretekárska/dostihová dráha* alebo zaužívaný bohemizmus *závodisko*. Kritérium *systémovosti* môže poslúžiť pri hodnotení variantných terminologických neologizmov (*legislatívkyňa*?/*legislatívka*).

V rámci lingvistických opisov sa ako určité nezlyhávajúce **kritérium spisovnosti** javí **kodifikácia** chápaná technicky ako evidencia daného stavu normy. Realite spisovných situácií sa kodifikácia najväčšmi priblíži vtedy, ak nebude fixovať prvky **ideálnej normy** ani ignorovať výrazy, ktoré sú súčasťou **spisovného úzu**.

ZOZNAM LITERATÚRY A ZDROJOV

- COSERIU, E. (1979): System, Norm und ‚Rede‘. In: U. Petersen (ed.), *Sprache – Strukturen und Funktionen*. 3. vyd. Tübingen: Narr, s. 45–59.
- DOLNÍK, J. (2010): *Teória spisovného jazyka so zreteľom na spisovnú slovenčinu*. Bratislava: Veda.
- RUŽIČKA, J. (1967): Problémy jazykovej kultúry. In: J. Ružička (ed.), *Kultúra spisovnej slovenčiny*. Bratislava: Vydavateľstvo Slovenskej akadémie vied, s. 9–22.

SPECIFIKA AUTOMATICKÉHO GENEROVÁNÍ SLOVNÍKOVÉHO VÝSTUPU¹

VLADIMÍR JARÝ

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze
Břehová 7, 115 19 Praha 1
vladimir.jary@foxcom.eu*

Key words: cache, dictionary writing system, printed output, web interface

Výstupní modul představuje jednu z nejdůležitějších součástí systému pro tvorbu slovníku *Alexis*. Tento modul je založen na sadě pravidel, která definují způsob, jakým se slovníková data prezentují koncovému uživateli. V systému je implementováno několik typů výstupů; v příspěvku se věnujeme především 1. elektronickému a tištěnému, dále pak 2. webovému a mobilnímu a 3. XML.

amarylis (2. j. -su, 6. j. -su) m. než. <lat.>, **amarylis** (neskl.) s. <lat.>

pokojevá rostlina s velkou cibulí, dlouhým bezlistým stonkem a nápadnými velkými zvonkovitými květy zprav. červené nebo bílé barvy: *koupit si cibule amarylis, amarylisu; amarylis kvete od ledna do dubna*

amarylis -su (6. j. -su) m. než. i neskl. s. <lat.> pokojevá rostlina s velkou cibulí, dlouhým bezlistým stonkem a nápadnými velkými zvonkovitými květy zprav. červené nebo bílé barvy: *koupit si cibule amarylis, amarylisu; amarylis kvete od ledna do dubna*

Obrázek 1. Porovnání elektronického (vlevo) a tištěného (vpravo) výstupu

1. Elektronický a tištěný výstup jsou založeny na formátu HTML, který definuje sémantiku dokumentu, a značkovacím jazyce CSS, který definuje vzhled dokumentu. Elektronický výstup funguje jako vstupní bod do korektorského modulu (Barbierik et al., 2013). Tento výstup jsme implementovali jako interaktivní webovou stránku; jeho součástí jsou všechny informace zadané na kartě hesla včetně metainformací zahrnujících autora hesla, datum poslední změny nebo poznámek pro odborné korektory.

Naproti tomu tištěný výstup je implementován v podobě statického PDF dokumentu (i v tomto případě používáme kombinaci HTML + CSS, výsledný dokument je do PDF souboru převeden nástrojem *wkhtmltopdf*). Tento výstup zobrazuje hesla tak, jak budou prezentována v tištěné podobě slovníku. Na rozdíl od elektronického výstupu nejdou do tištěné podoby hesla všechny prvky

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

2. **Formát PDF výstupu** slouží jako základ pro webový výstup a výstup pro mobilní zařízení. Tyto výstupy budou zpřístupněny veřejnosti. Oba výstupy jsou propojeny s externím lemmatizačním nástrojem *Morphodita* (Straková – Straka – Hajič, 2014), který automaticky vytvoří v textech výkladů významu a exemplifikací proklíky na základní tvary slov.

3. Konečně **XML výstup** transformuje data do stromové struktury a umožní propojit slovník s jinými jazykovědnými nástroji.

Elektronický a tištěný výstup *Alexisu* jsou již od roku 2013 každodenně používány lexikografy. Webový výstup, výstup pro mobilní zařízení a XML výstup jsou připraveny k nasazení do veřejného provozu. Od počátku roku 2017 bude software *Alexis* volně dostupný.

SEZNAM LITERATURY A ZDROJŮ

- BARBIERIK, K. et al. (2013): A new path to a modern monolingual dictionary of contemporary Czech: the structure of data in the new dictionary writing system. In: K. Gajdošová – A. Žáková (eds.), *Natural Language Processing, Corpus Linguistics, E-learning, Proceedings of the conference Slovo 2013*, Bratislava: Slovenská akadémia vied, Jazykovedný ústav Ľudovíta Štúra, s. 9–26.
- BARBIERIK, K. et al. (v tisku): Cache Module for the Dictionary Writing System. In: J. Janech – J. Kostolny – T. Gratkowski (eds.), *Proceedings of the 2015 Federated Conference on Software Development and Object Technologies*.
- STRAKOVÁ, J. – STRAKA, M. – HAJIČ, J. (2014): Open-Source Tools for Morphology, Lemmatization, POS Tagging and Named Entity Recognition. In: K. Bontcheva – Z. Jingbo (eds.), *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, Baltimore, Maryland: Association for Computational Linguistics, s. 13–18.

SLOVNÍK NÁŘEČÍ ČESKÉHO JAZYKA – ÚSKALÍ LEXIKOGRAFICKÉ A DIALEKTOLOGICKÉ ABSTRAKCE

STANISLAVA KLOFEROVÁ

*dialektologické oddělení, Ústav pro jazyk český AV ČR, v. v. i.
Veveří 97, 602 00 Brno
kloferova@iach.cz*

Key words: dialectology, Dictionary of Czech Dialects, lemma, lemmatization

Slovník nářečí českého jazyka (dále SNČJ) je nediferenční slovník, který zaznamenává lexikální zásobu nářečí na celém území národního jazyka, tj. v Čechách, na Moravě a ve Slezsku. Je prvním slovníkem svého druhu v češtině. Jeho jedinečnost se odvíjí od charakteru lexikální zásoby, kterou zachycuje. Musí obsáhnout nejen nářeční lexikální zásobu pokud možno v co největší úplnosti, ale rovněž v její územní variabilitě: musí hierarchizovat v odstupu roviny hláskoslovné, tvaroslovné a lexikální včetně frazeologických a idiomatických jevů. Uvedené zásady jsou představeny v koncepci SNČJ, zveřejňovaném v jeho elektronické verzi od února 2016 na webových stránkách Ústavu pro jazyk český AV ČR, v. v. i., (dostupný z: <<http://snjcj.ujc.cas.cz>>).

Slovník je tvořen heslovými odstavci, které sestávají ze záhlaví a z těla hesla. Heslové záhlaví obsahuje lemma (v případě morfologické variability i více lemmat), jeho gramatickou charakteristiku, tj. informaci o gramatickém rodu u substantiv, o vidu u sloves; u ostatních slovních druhů je uvedena jen zkratka slovního druhu. Následuje výčet registrovaných nářečních variant. Tělo hesla tvoří výčet významů s exemplifikací, výčet dokladů z frazeologie a idiomatiky a přihnízdovaná hesla (výrazivo určitého slovtvorného charakteru řadící se pod základové lemma). Heslový odstavec je uzavřen odkazy na důležitá lexikografická díla obsahující totožné lemma.

Lemma je tvořeno konstruktem, jehož podoba je výsledkem dialektologické abstrakce. Zastřešuje všechny zaznamenané hláskové realizace v nářečích. Nemusí věrně odrážet podobu zjištěného dokladu. Je zapsáno spisovně, tj. ve shodě s pravopisem. Za lemmatem jsou v závorce uvedeny nářeční realizace slov, jež lze vřadit pod zastřešující lemma; jako první varianta se uvádí podoba shodná s lemmatem (je-li zachycena). Nářeční varianty jsou uvedeny v českém dialektologickém přepisu (viz *Český jazykový atlas. Dodatky*, 2011, s. 411).

Podoba lemmatu u domácích slov se řídí etymologií příslušného výrazu. Doklad *bapčina* ‚ztrouchnivělé částečky dřeva v dutých stromech, trouch‘, odvozený od substantiva *babka*, je vřazen pod lemma **babčina**, doklad *bábrdle* ‚babička‘ (s vloženou hláskou *d* ve skupině *rl*) pod lemma **bábrle**. Východomoravské doklady *břuchál*, *břucháu* stejně jako středomoravský *břuchál*, *břochal* a varianta *bruchál* z kopaničářské nářeční oblasti jsou vřazeny pod lemma **břichál**: všechny obměny jsou výsledkem nářečního vývoje hlásek *r*, *u* a *l* v příslušných regionech.

Varianty v záhlaví jsou uváděny v nářeční realizaci s výjimkou asimilačních změn na konci slova: doklady *guláš*, *guláž* jsou variantami lemmatu **guláš**: koncové písmeno zde znázorňuje finální konsonant v nepřímých pádech, tj. gen. sg. -še/-ša, popř. -že/-ža.

Dialektologická abstrakce se týká také morfologické pozice hlásky, a to v koncovce feminin, která jsou v nářečích v Čechách zakončena po tzv. měkkých souhláskách na *-e* (*babice*, *chvíle*), kdežto na Moravě a ve Slezsku na *-a* (*babica*, *chvíla*). Jde o reflex uskutečněné (*-e*) nebo neuskutečněné (*-a*) přehlásky 'a (> ě) > e. Formy původní, nepřehlasované, jsou vřazeny pod lemmata s přehláskou: **babice** a **chvíle**. Lemma je takto utvořeno i v případě, že existuje pouze původní, nepřehlasovaná podoba evidovaná z Moravy, popř. Slezska (*bakaňica* 'ušpiněné dítě', lemma **bakanice**).

Důležitým pramenem při vytváření lemmatu je evidence slova v *Slovníku spisovného jazyka českého* (1989; dále SSJČ). Tato skutečnost však může vést k rozbití etymologického hnízda nebo povědomí o něm, třebaže jsme si vědomi, že SNČJ není dílem etymologickým. Ne vždy je totiž slovo v SSJČ zapsáno v podobě blízké etymologickému východisku. Jako příklad lze uvést doklady s náslovným *b-*, *p-* přejaté z němčiny: Nářeční doklady *biglovat* a *piglovat* 'žehlit' jsou vřazeny pod lemma **piglovat**, neboť varianta *piglovat* je (jako jediná z obou) evidována jako spisovná na rozdíl od (etymologicky) výchozí varianty *biglovat* s náslovným *b-* (něm. *bügeln*). Doklad *biglarka* 'žehlírka' je však vřazen pod lemma **biglářka**, neboť v SSJČ není registrována *piglárka* (resp. *piglarka*), kterou bychom mohli očekávat, protože existuje sloveso *piglovat*. Lemma zde tedy vychází z etymologicky náležitého německého východiska s počátečním *b-*, lemmatizován (v návaznosti na vývoj v dialektech) je též slovotvorný sufix tvořící činitelské jméno (*-arka* → *-árka*).

Každou ze zaznamenaných nářečních variant včetně jejich příslušnosti k určitému lemmatu a heslu lze vyhledat v online SNČJ ve vyhledávacím souboru Rejstřík.

Je zřejmé, že lemmatizace je v nářečním slovníku principiálním postupem, který umožňuje uspořádat bohatou hláskovou variabilitu nářeční lexikální zásoby. Jinými postupy, než je metoda dialektologické lemmatizace, by byl jazykový materiál představen roztržštěně, v jednotlivinách, aniž by byl vsazen do systémových souvislostí.

SEZNAM LITERATURY A ZDROJŮ

- Český jazykový atlas. Dodatky (2011). Praha: Academia.
Slovník nářečí českého jazyka (2016–) [online]. Cit. 2016-09-19. <<http://sncj.ujc.cas.cz/>>.
Slovník spisovného jazyka českého 1–8 (1989). Praha: Academia.

AKADEMICKÝ SLOVNÍK SOUČASNÉ ČEŠTINY¹

PAVLA KOCHOVÁ – ZDEŇKA OPAVSKÁ

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
kochova@ujc.cas.cz, opavska@ujc.cas.cz*

Key words: Czech language, dictionary-making principles, lexicography, monolingual dictionary

V úvodní přednášce je syntetizujícím způsobem představen *Akademický slovník současné češtiny* (dále ASSČ), který vzniká v oddělení současné lexikologie a lexikografie Ústavu pro jazyk český AV ČR, v. v. i., od r. 2012; od roku 2013 do konce roku 2016 pak s podporou grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

ASSČ patří k všeobecným výkladovým slovníkům středního rozsahu (s plánovaným objemem 120 000–150 000 hesel). Řadí se k slovníkům akademického typu, v nichž se podává podrobný a komplexní popis lexikálních jednotek. Cílem ASSČ je popsat slovní zásobu současné češtiny, užívanou ve veřejné i běžné komunikaci, se zaměřením na lexikální jednotky, které jsou v těchto komunikačních sférách běžně rozšířené. Předpokládaným uživatelem slovníku je středoškolsky vzdělaný rodilý mluvčí. Primární materiálovou základnu představuje synchronní korpus psaných textů SYN Ústavu Českého národního korpusu (ÚČNK) o rozsahu 2,2 miliardy slov, dále se čerpá z mediálního archivu společnosti Newton Media, a. s., (archiv celostátních i regionálních periodik a přepisů televizních a rozhlasových pořadů), z korpusů řady ORAL ÚČNK, z webových korpusů *Aranea* a z internetových textů. Přihlíží se k materiálovým databázím a sbírkám oddělení současné lexikologie a lexikografie. Slovník bude postupně a průběžně publikován v elektronické podobě na internetu. Zároveň se, v návaznosti na lexikografickou tradici, na závěr lexikografické práce počítá s knižní podobou slovníku.

ASSČ je v pořadí čtvrtým všeobecným výkladovým slovníkem vznikajícím v Ústavu pro jazyk český. Od vydání 1. sešitu typově nejbližšího *Slovníku spisovného jazyka českého* ho dělí téměř šedesát let a od 1. vydání zatím posledního všeobecného výkladového slovníku, jednosvazkového *Slovníku spisovné češtiny pro školu a veřejnost*, ho odděluje téměř čtyřicet let.² Stejně jako předcházející všeobecné výkladové slovníky akademického typu je i ASSČ svébytným lexiko-

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

² Podrobně k historii českého slovníkářství viz Hladká (2007).

grafickým dílem. V koncepci nového slovníku se odrážejí jednak změny, k nimž v lingvistických disciplínách (zejména lexikologii a lexikografii³, ale též morfologii, syntaxi, stylistice a dalších) v tomto časovém období došlo, jednak nové možnosti, které současný technický pokrok otvírá pro tvorbu a publikaci výkladových slovníků.⁴ Při přípravě ASSČ lze tedy v mnohém navázat na koncepci a zpracování hesel v předcházejících výkladových slovnících, ale mnohé aspekty je třeba řešit nově nebo odlišně. K základním koncepčním principům slovníku, v nichž došlo ke změně, patří výstavba hesláře, který je nově založen primárně na korpusech, akcentace samostatnosti lexémů – ústup od hnízdování a uvádění významu u jednotek dříve hnízdovaných, zdůrazňování autonomnosti více-slovných lexikálních jednotek, dále detailnější podávání gramatických údajů – zejména morfologických údajů a údajů o valenci, důraz na exemplifikační část heslové stati.⁵

Významnou proměnou prošla také organizace lexikografické práce v souvislosti s novými technickými možnostmi, které umožnily přechod od slovníku doby předpočítačové ke slovníku doby počítačové, od papírového média k elektronické formě. Pro ASSČ to znamená, že je od počátku vytvářen a zpracováván prostřednictvím programu *Alexis*, specializovaného softwaru pro tvorbu slovníku (DWS).⁶

Závěr přednášky je zaměřen na přínosy souběžného lexikologicko-lexikografického výzkumu, který tvorbu nového slovníku nutně provází a díky němuž se prohlubuje popis některých nominačních a strukturních typů, vyplňují se mezery dosavadního popisu a reflektuje se dynamika slovní zásoby.

SEZNAM LITERATURY A ZDROJŮ

- ABEL, A. – KLOSA, A. (2012): The Lexicographic Working Environment in Theory and Practice. In: R. V. Fjeld – J. M. Torjusen (eds.), *Proceedings of the 15th EURALEX International Congress: Oslo, 7–11 August 2012*. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo, s. 1–23.
- ATKINS, B. T. S. – RUNDELL, M. (2008): *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- BENEŠOVÁ, L. – KŘEN, M. – WACLAWIČOVÁ, M. (2013): *ORAL2013* [online]. Cit. 2016-09-01. <<http://www.korpus.cz>>.
- BENKO, V. (2015): *Araneum Bohemicum Maximum, Araneum Bohemicum Maius, Araneum Bohemicum Minus* [online]. Cit. 2016-09-01. <<http://www.korpus.cz>>.
- ČERMÁK, F. (2013): Poznámky k tvorbě slovníku na základě korpusu. *Jazykovedný časopis*, 64, s. 133–150.

³ Viz např. Čermák (2013), Hanks (2010).

⁴ Srov. Abel – Klosa (2012), Atkins – Rundell (2008).

⁵ Detailně o koncepci slovníku pojednává kolektivní monografie *Kapitoly z koncepce Akademického slovníku současné češtiny* (2016). Prolegomena k připravovanému slovníku byla publikována v časopise *Naše řeč* (viz Kochová – Opavská, 2016).

⁶ Tento software je vyvíjen (rovněž s podporou uvedeného grantového projektu) v oddělení současné lexikologie a lexikografie ÚJČ AV ČR, v. v. i. Jeho prezentaci jsou věnovány speciální příspěvky (srov. Barbierik, Jarý, Liška v tomto sborníku).

- HANKS, P. (2010): Compiling a Monolingual Dictionary for Native Speakers. *Lexikos*, 20 (AFRILEX-reeks/series, 20), s. 580–598.
- HLADKÁ, Z. (2007): Lexikografie. In: J. Pleskalová a kol. (eds.), *Kapitoly z dějin české jazykovědné bohemistiky*. Praha: Academia, s. 164–198.
- KOCHOVÁ, P. – OPAVSKÁ, Z. (2016): Akademický slovník současné češtiny. *Naše řeč*, 99, s. 57–83.
- KOCHOVÁ, P. – OPAVSKÁ, Z. (eds.) (2016): *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i.
- KOPŘIVOVÁ, M. – WACLAWIČOVÁ, M. (2006): *ORAL2006* [online]. Cit. 2016-09-01. <<http://www.korpus.cz>>.
- KŘEN, M. – ČERMÁK, F. – HLAVÁČOVÁ, J. – HNÁTKOVÁ, M. – JELÍNEK, T. – KOCEK, J. – KOPŘIVOVÁ, M. – NOVOTNÁ, R. – PETKEVIČ, V. – PROCHÁZKA, P. – SCHMIEDTOVÁ, V. – SKOUMALOVÁ, H. – ŠULC, M. (2014): *Korpus SYN, verze 3 z 27. 1. 2014* [online]. Cit. 2016-09-01. <<http://www.korpus.cz>>.
- Mediální archiv NEWTON Media, a. s. Cit. 2016-09-01. <<http://www.newtonmedia.cz>>.
- Slovník spisovné češtiny pro školu a veřejnost* (1978). Praha: Academia.
- Slovník spisovného jazyka českého* (1960–1971). Praha: Nakladatelství ČSAV.
- WACLAWIČOVÁ, M. – KOPŘIVOVÁ, M. – KŘEN, M. – VÁLKOVÁ, L. (2008): *ORAL2008* [online]. Cit. 2016-09-01. <<http://www.korpus.cz>>.

DEFINITIONS OF CLASSIFYING ADJECTIVES IN THE NEW DICTIONARY OF THE SLOVENIAN STANDARD LANGUAGE

DOMEN KRVINA – MIJA MICHELIZZA

leksikološka sekcija, Inštitut za slovenski jezik Frana Ramovša ZRC SAZU
Novi trg 4, SI-1000 Ljubljana
domen.krvina@zrc-sazu.si, mmija@zrc-sazu.si

Key words: classifying adjectives, definitions, dictionary, Slovene

Introduction

The paper presents a specific subtype of genus proximum-differentia specifica dictionary definition, used in the 3rd edition of the Dictionary of the Slovenian Standard Language (SSKJ³) in the definitions of classifying adjectives. This definition subtype can be characterized as a developed form of relational definition as it represents the semantic relationship between the adjectival motivational basis (usually a noun or a verb from which the adjective is formed) and the noun as the core of a multi-word expression (cf. Vidovič Muha, 2000, pp. 65–66, 310–327).

Definitions of classifying adjectives

The basic – and relatively stable – semantic relationships are as follows:

- a) spatiality,
- a1) temporality,
- b) being made of something,
- c) (indirect) intention.

Examples:¹

- (a) **mědvretěncni**: *ki je med vretenci*: medvretenčni prostor, medvretenčna odprtina
- (a1) **prědpremiěrni**: *ki je pred premiero*: predpremierna predstava, uprizoritev
- (b) **pehōtni**: /.../ *ki je sestavljen zlasti iz pehote*: pehotni bataljon, oddelek, polk
- (c) **jūšni**: /.../ 2. *ki je namenjen za juho*: jušni dodatek, jušni rezanci, jušna govedina
2.1 *ki je namenjen za strežbo, kuhanje juhe*: jušni krožnik, jušni lonec, jušna posoda

The meaning of a classifying adjective can include several basic semantic relationships, as in *kožni*:²

kožni:

- 1. *ki je na koži, v koži*: kožni pigment; kožne celice, plasti, pore 1.1 *ki se pojavlja na koži, prizadene (najprej) kožo*: kožni melanom, rak; kožne bolezni, reakcije, spremembe 1.2 *ki se izvaja na koži*: alergijsko kožno testiranje
- 2. *ki je namenjen za kožo*: kožni pripravek
- 3. *ki je tak kot pri koži*: hlačke, nogavice kožne barve

¹ The examples are taken from Gliha Komac et al. (2016).

² The examples are taken from Gliha Komac et al. (2015).

When there are no basic semantic relationships to be found, the relational definition can still be developed with certain subtypes of semantic relationships if supported by corpus material (e.g. *monsunski*: ‘ki spremlja, je značilen za’ ‘accompanied by, typical of’). *Monsun* also serves as a definition example of a classifying adjective formed from a lexeme that might not be widely known (such as *koža*, *premiera*, *vitamin*, cf. also Landau, 2001, s. 160) – therefore an extension of the definition with a (full) explanation of this lexeme is helpful in saving time otherwise required to look it up elsewhere in the dictionary (Krvina, 2015, pp. 414).

monsunski monsúnska monsúnsko *prévnik* [monsúnski]
 POMEN
 iz geografije: ki je v zvezi z monsunom, (tj.) deževnim obdobjem, zlasti v južni in jugovzhodni Aziji, ko piha veter s hladnega morja na ogreto kopno
 ▶ **prid. beseda** + sam. beseda
 monsúnsko obdobje
 • Kadar v Indiji nastopi *monsúnsko* obdobje, več dni lije kot iz šikafa.
 a) iz geografije: ki spremlja monsun, je značilen za monsun
 ▶ **prid. beseda** + sam. beseda
 monsúnski dež, nalliv | monsúnski oblak | monsúnski vetrovi | monsúnске padavine | monsúnске poplave | monsúnsko deževje | monsúnsko vreme
 • Podnebje Indijske podceline je povsem odvisno od *monsúnskih* vetrov.
 • Pakistan so prizadele najhujše *monsúnске* poplave v zgodovini, ki so zahtevale najmanj 800 smrtnih žrtev, prizadele pa so še milijon ljudi na severozahodu države
 • Poplave je povzročilo močno *monsúnsko* deževje, ki traja že šesti dan.
 b) ki je tak kot pri monsunu
 ▶ **prid. beseda** + **sam. beseda**
 monsúnski nalliv
 • Cesta se je zaradi skorajda *monsúnskega* nalliva v 1. etapi spremenila v olimpijski bazen.

Figure 1: Classifying adjective *monsunski*³ in the SSKJ³

Sometimes one of the basic semantic relationships (such as being made of something), which in such cases often have a somewhat broader and more concise definition (‘ki vsebuje zelo veliko, načrtno zgoščeno količino’ ‘contains very large, intentionally concentrated amount’), can be followed by a certain semantic relationship subtype, as in *vitaminski*: ‘ki temelji na’ ‘based on’.

vitaminski vitaminska vitaminsko *prévnik* [vitaminski]
 POMEN
 ki je v zvezi z vitamini
 ▶ **prid. beseda** + sam. beseda
 vitaminska sestava | vitaminska vrednost
 • Smiselno je pregledati *vitaminsko* sestavo na deklaraciji ter izbrati preparate, ki so dopolnjeni tudi z minerali in mikroelementi.
 • Če želite ohraniti *vitaminsko* vrednost živil, jih za dlje ne izpostavljajte vplivu zraka in svetlobe, ne kuhajte jih predolgo in ne pogrevajte.
 a) ki vsebuje zelo veliko ali načrtno zgoščeno količino vitaminov
 ▶ **prid. beseda** + sam. beseda
 vitaminski dodatek, nadomestek | vitaminski koktajl, napitek, pripravek, preparat | vitaminski kompleks | vitaminski izdelki | vitaminska bomba | vitaminska hrana, prehrana | vitaminska injekcija | vitaminska kombinacija | vitaminska krema, maska, obloga | vitaminska pijača | vitaminska ploščica | vitaminske kapljice, pilule, tablete, tabletkе | vitaminske snovi | vitaminsko dopolnilo
 • Jeseni in še posebej pozimi večkrat poseže po vitaminih in *vitaminskih* dodatkih.
 • V *vitaminskih* tabletah je učinkovina močno zgoščena, zato je treba skrbno paziti, koliko posameznega vitamina zaužijemo.
 ▶ **prid. beseda** + sam. beseda
 vitaminski in mineralni, vitaminski in rudninski, beljakovinski in vitaminski
 • Nekateri ljudje menijo, da bodo hranila, ki jih ne dobijo iz osiromašenega prehranjevanja, nadomestili z *vitaminskimi* in mineralnimi prehranskimi dodatki.
 • Danes je v prodaji veliko *vitaminskih* in rudninskih dodatkov, vendar je le nekaj odstotkov naravnih.
 b) ki temelji na povečani uporabi, uživanju vitaminov
 ▶ **prid. beseda** + sam. beseda
 vitaminska kura, terapija | vitaminska nega
 • Cilj *vitaminske* terapije je normalizirati in okreptiti delovanje imunskega sistema.
 • Koži lahko vitalnost povrne z *vitaminsko* nego, ki vpliva na njeno notranje mikro kroženje in ji omogoča več sijaja.

Figure 2: Classifying adjective *vitaminski* in the SSKJ³

³ These figures represent the current, not necessarily final state of the entries in the SSKJ³.

Quite common are cases where an adjective is not solely classifying or non-classifying; very often a relational meaning⁴ follows immediately after the primary non-relational meaning. Such a relational meaning is based on a number of stable multi-word expressions (functioning very much like separate lexemes); for each one a separate definition can be made as well. However, these possible separate definitions would have many points in common and are thus – if possible – brought together into a single and usually more general definition. This does not apply to cases where the meaning of a multi-word expression cannot be deduced from the words forming it (e.g. *zeleno zlato* ekspr. 'hmelj').

Such an example is *polzaprt* with its primary non-relational meaning (1. 'ki ni povsem zaprt' 'not entirely closed') which is followed by a number of relational meanings: 2. 'ki ni povsem obdan s čim' 'not entirely surrounded by something', 3. 'ki pri svojem delovanju, uporabi predvideva, dopušča določena odstopanja, olajšave, ugodnosti' 'envisaging, allowing for exemptions, privileges in its functioning, use', 4. 'ki je ločen od okolice, vendar dopušča, omogoča vstopanje ali izstopanje posameznih elementov' 'isolated from surroundings but allowing, enabling entering or leaving of certain elements'.

pólzaprť pólzapřta pólzapřto prsevná [pólzapřãt]
 POMEN
 1. ki ni povsem zaprt
 ▶ prid. beseda + sam. beseda
 polzaprtne oči
 • Pisala je lahko tudi s *polzaprtimi* očmi in s premori vmes.
 2. v obliki **polzaprti ki ni povsem obdan s čim**
 ▶ prid. beseda + sam. beseda
 polzaprtno morje
 • Skrb zbujajoče so razmere na območju Sredozemlja, ki je kot *polzaprtno* morje že v precejšnji meri degradirano.
 3. v obliki **polzaprti ki pri svojem delovanju, uporabi predvideva, dopušča določena odstopanja, olajšave, ugodnosti**
 ▶ prid. beseda + sam. beseda
 • Ker *polzaprti* ureditev pobiranja cestnine ni najbolj učinkovita, bodo dolgovi še prevladovali.
 • Iz *polzaprtga* oddelka so jo zaradi napake prestavili nazaj v zaprtga.
 ▶ vezni glag. + prid. beseda v **rodilniku** + sam. beseda v **rodilniku**
 biti polzaprtga tipa
 • Novo zdravilišče bo *polzaprtga* tipa, namenjeno ljudem, ki imajo težave z luskavico, pa tudi drugim, ki cenijo zdravilnost termalne vode.
 4. v obliki **polzaprti ki je ločen od okolice, vendar dopušča, omogoča vstopanje ali izstopanje posameznih elementov**
 ▶ prid. beseda + sam. beseda
 polzaprti krog
 • Sistem na *polzaprti* krog pošilja izdihani zrak skozi filter, ki absorbira ogljikov dioksid, nato pa mu doda svež zrak.
 • Voda kroži v *polzaprti* sistemu.
 • Sporazum je *polzaprti* narave.

Figure 3: Adjective *polzaprt* in the SSKJ³

Conclusion

This article has presented a number of solutions for definitions of classifying adjectives which are used in the 3rd edition of the Dictionary of the Slovenian Standard Language (SSKJ³). These solutions represent a certain development in relational definitions beyond those in the 1st and 2nd edition of Dictionary of the Slovenian Standard Language (SSKJ), a hint to which is given in the previous edi-

⁴ In Slovene a relational meaning is reflected in the form of the adjective: it takes the ending *-(n)i* as opposed to *-Ø*, e.g. *nemi, zeleni, polzaprti* : *nem, zelen, polzaprt*, in singular masculine nominative and with non-animate nouns also in accusative.

tions by the separation of multi-word expressions of classifying adjectives with nouns using slashes. Apart from basic semantic relationships (spatiality, temporality, being made of something, (indirect) intention) forming the meaning of classifying adjectives, the corpus material may point to other subtypes of relational definitions (such as being based on, being accompanied by, typical of) worthy of recognition.

REFERENCES

- GLIHA KOMAC, N. – JAKOP, N. – JEŽOVNIK, J. – KLEMENČIČ, S. – KRVINA, D. – LEDINEK, N. – MIRTIČ, T. – PERDIH, A. – PETRIC, Š. – SNOJ, M. – ŽELE, A. (2015): *Koncept novega razlagalnega slovarja slovenskega knjižnega jezika. Različica 1.1*. [online]. Ljubljana: Inštitut za slovenski jezik Frana Ramovša ZRC SAZU. Cit. 2016-08-18. <<http://www.fran.si/novi-sskj>>.
- GLIHA KOMAC, N. – JAKOP, N. – JEŽOVNIK, J. – KLEMENČIČ, S. – KRVINA, D. – LEDINEK, N. – METERC, M. – MICHELIZZA, M. – MIRTIČ, T. – PERDIH, A. – PETRIC, Š. – SNOJ, M. – ŽELE, A. (2016): Novi slovar slovenskega knjižnega jezika – predstavitev temeljnih konceptualnih izhodišč. In: F. Marušič – P. Mišmaš – R. Žaucer (eds.), *Škrabčevi dnevi 9*. Nova Gorica: Založba Univerze v Novi Gorici, pp. 16–33. Cit. 2016-10-26. <http://www.ung.si/media/storage/cms/attachments/2016/10/21/13/45/01/Zbornik-%C5%A0D9_akt.16_splet.pdf>.
- KRVINA, D. (2015): Slovarske razlage: med treznim recikliranjem starega in preudarnim uvajanjem novega. In: M. Smolej (ed.), *Slovnica in slovar – aktualni jezikovni opis. Obdobja 34*. Ljubljana: Znanstvena založba FF Univerze v Ljubljani, pp. 409–415.
- LANDAU, I. S. (2001): *Dictionaries. The Art and Craft of Lexicography*. Second Edition. Cambridge et al.: Cambridge University Press.
- VIDOVIČ MUHA, A. (2000): *Slovensko leksikalno pomenoslovje: govornica slovarja*. Ljubljana: Znanstveni inštitut Filozofske fakultete Univerze v Ljubljani.

FREKVENČNÍ SLOVNÍKY ČEŠTINY

MICHAL KŘEN

Ústav Českého národního korpusu, Filozofická fakulta Univerzity Karlovy
Panská 7, 110 00 Praha 1
michal.kren@ff.cuni.cz

Key words: Czech language, dispersion, frequency dictionary, lemmatization, representativeness

Cílem příspěvku je charakterizovat dva frekvenční slovníky češtiny vzniklé na datech *Českého národního korpusu* (Čermák – Křen et al., 2004; Čermák – Křen et al., 2011). V první části se zaměříme na základní faktory, k nimž je potřeba při vytváření frekvenčního slovníku přihlídnout, aby výsledek nebyl pouhým seznamem slovních tvarů nebo lemmat, který dnes lze snadno vygenerovat pomocí nástrojů, jako je např. *KonText*. Těmito faktory jsou především vygenerovativnost korpusových dat a jejich vztah k jazykové realitě, dále spolehlivost nástrojů použitých při anotaci korpusu (zejména lemmatizace a morfologického značkování; nutnost rozsáhlých manuálních revizí popisuje Křen, 2006) a konečně disperze, tj. rovnoměrnost rozložení lexikálních jednotek v korpusu (Gries, 2008).

Druhá část příspěvku bude věnována popisu obou frekvenčních slovníků, jejich styčných bodů a odlišností v rozsahu, v doprovodných informacích, ve zdrojových datech a ve způsobu zpracování, zvláště s ohledem na výslednou podobu hesláře. Oba slovníky spojuje několik společných rysů, zejména použití průměrné redukované frekvence (ARF; Savický – Hlaváčová, 2002) jako prostředku pro stanovení běžnosti slov, který bere v úvahu disperzi. Na druhou stranu je mezi nimi také řada rozdílů způsobených předpokládaným publikem (rodilí mluvčí, resp. ti, kteří se česky teprve učí), požadavky vydavatelů (počet hesel, doprovodné informace) i dostupnými daty: oba slovníky jsou založeny především na korpusech psané češtiny (SYN2000, resp. SYN2005), novější z nich ale pokrývá i neformální mluvený jazyk (korpuse řady ORAL). S tím souvisí také odlišný způsob zpracování lexikálních jednotek, které jsou typické pouze pro určitou varietu češtiny (příznakovost), a jejich vliv na uspořádání hesláře. Při srovnání heslářů pak ukážeme důsledky zmíněných rozdílů mezi oběma slovníky pro (ne)zařazení konkrétních lexikálních jednotek.

SEZNAM LITERATURY A ZDROJŮ

- ČERMÁK, F. – KŘEN, M. et al. (2004): *Frekvenční slovník češtiny*. Praha: Nakladatelství Lidové noviny.
- ČERMÁK, F. – KŘEN, M. et al. (2011): *Frequency Dictionary of Czech: Core Vocabulary for Learners*. London: Routledge.
- GRIES, S. (2008): Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics*, 13, s. 403–437.

- KŘEN, M. (2006): Frequency Dictionary of Czech: A Detailed Processing Description. In: M. Šimková (ed.), *Insight into Slovak and Czech Corpus Linguistics*. Bratislava: Veda, s. 16–25.
- SAVICKÝ, P. – HLAVÁČOVÁ, J. (2002): Measures of word commonness. *Journal of Quantitative Linguistics*, 9, s. 215–231.

PŘEDSTAVENÍ DWS *ALEXIS*¹

TOMÁŠ LIŠKA

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze
Břehová 7, 115 19 Praha 1
tomas.liska@foxcom.cz*

Key words: dictionary database, dictionary writing system, lexicography

Lexikografické dílo *Akademický slovník současné češtiny* vzniká v prostředí specializovaného softwaru *Alexis* (mj. Barbierik et al., 2014a; Barbierik et al., 2014b; Barbierik et al., 2015). V příspěvku bude představena architektura softwaru včetně jeho jednotlivých modulů.

Ústředními objekty slovníku jsou hesla definovaná svou mikrostrukturou. Na heslo jsou navázány v hierarchické (nikoli planární) struktuře jednotlivé varianty hesla a také významy korespondující s daným heslem. Celý soubor hesel tvoří hlavní heslář a hesláře další: seznamy proprií (rodných jmen, zeměpisných jmen ad.) a zkratk a značek.

Software *Alexis* obsahuje řadu sofistikovaných nástrojů. Zmíňme se o některých z nich. Pro filtrování dat poskytuje software *Alexis* sadu nástrojů souhrnně pojmenovanou filtry. Základní hledání je umožněno pomocí rychlého filtru, který bleskově vyhledá požadovaný výraz ve zvoleném poli mikrostruktury hesla. V situacích, kdy je třeba vyfiltrovat podmnožinu hesel dle negenerických kritérií, lze využít ruční filtr, kde podmnožinu hesel definuje explicitně uživatel výčtem. Nejkomplexnějším filtrovacím nástrojem je modul xFiltr, který nabízí uživateli tvorbu logických výrazů přes celou mikrostrukturu hesla s využitím logických spojek, negace a zástupných znaků. Každá vyfiltrovaná podmnožina záznamů může být pojmenována a uložena pro pozdější použití. Uživatelé mohou filtry mezi sebou i sdílet.

Základní operací s hesly je jejich editace; ta přímo souvisí s dalšími nástroji, které jsme v softwaru *Alexis* implementovali. Každá změna hesla je systémem zaznamenána a opatřena metadaty. O automatické zpracování změn se stará modul verzování. Pro uživatele je jeho činnost zcela transparentní, neboť běžný uživatel pracuje vždy s aktuální podobou údajů hesla. V pozadí jsou všechny minulé verze každého hesla archivovány a slouží ke komparačním (v rámci hesel) nebo statistickým (v rámci týmu) účelům.

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

Dalším klíčovým nástrojem pro práci s hesly je redakční modul, který pokrývá kompletní proces redakční práce, všechny navrhované korektury uchovává, umožňuje editorovi a redaktorovi nad navrhovanými změnami diskutovat a změny lze na heslo aplikovat či je odmítnout. Modul eviduje kompletní životní cyklus korektury od jejího vzniku až po její úplné vypořádání. Aplikované redakční změny jsou promítnuty do hesla. Zde dochází k integraci funkcí editace hesla, redakce a verzování.

Systém *Alexis* je multiuživatelský, proto disponuje tzv. inteligentní správou uživatelů. Rozlišujeme uživatele interní (lexikografové pracující na slovníku) a externí (odborníci na terminologii, korektoři). Každé skupině uživatelů nebo konkrétnímu uživateli je možno přiřadit práva přístupu k datům a k operacím v systému.

Systém *Alexis* je provozovaný jako webová aplikace běžící na serveru. Uživatelé k aplikaci přistupují prostřednictvím standardního webového prohlížeče. Je zajištěna centralizovaná správa dat, centralizované zálohování, řízené aktualizace aplikace a struktur dat, řízení přístupu uživatelů k aplikaci apod. *Alexis* neklade na uživatele nároky na technické a programátorské dovednosti (uživatelé nejsou nuceni si instalovat software na svůj počítač, nemusí provádět samostatné aktualizace ani zálohy).

Lexikografové potřebují hesla vidět v náhledu pro slovník: k tomu slouží generátor slovníkového výstupu. Generátor implementuje nejen rozsáhlou sadu pravidel a výjimek pro správné generování výstupu samotného, ale i řešení zobecněného výstupu, na který jsou aplikována pravidla pro konkrétní typ výstupu pro jeho finální zformátování. Jinak řečeno: výstup je vygenerován dle jedné sady pravidel pro strukturální podobu hesla a jeho vizuální podoba je v druhém kroku generování přizpůsobena tomu, zda se generuje výstup pro webový prohlížeč, náhled k tisku, výstup pro redakční práci, elektronický výstup pro strojové napojení na jiné systémy či jednu z mobilních aplikací pro přenosná zařízení.

Systém *Alexis* je vyvíjen jako software na míru pro uspokojení náročných požadavků lexikografů. Výsledkem týmového úsilí programátorů a lexikografů je unikátní softwarové dílo, které slovníkářům v jejich práci významně pomáhá, a mohou se tak plně soustředit na předmět svého výzkumu.

SEZNAM LITERATURY A ZDROJŮ

- BARBIERIK, K. et al. (2014a): Simple and Effective User Interface for the Dictionary Writing System. In: A. Abel – C. Vettori – N. Ralli (eds.), *Proceedings of the XVI EURALEX International Congress: The User in Focus*. 15–19 July 2014, Bolzano/Bozen, s. 125–135.
- BARBIERIK, K. et al. (2014b): Dictionary Writing System: Output and Editorial Modules. In: J. Ministr – M. Tvrdíková (eds.), *Proceedings of IT for Practice 2014, 9–10 October 2014*. Ostrava: VŠB – Technical University of Ostrava, s. 17–23.
- BARBIERIK, K. et al. (2015): The Current Status of the Development of the ALEXIS Dictionary Writing System. In: K. Gajdošová – A. Žáková (eds.), *Natural Language Processing, Corpus Linguistics, Lexicography*. Lüdenscheid: RAM-Verlag, s. 17–25.

EXEMPLIFIKACE TERMÍNŮ V AKADEMICKÉM SLOVNÍKU SOUČASNÉ ČEŠTINY¹

MICHAELA LIŠKOVÁ

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
liskova@ujc.cas.cz*

Key words: exemplification, lexicography, monolingual dictionary, terminology

Slovníková exemplifikace se stává předmětem soustředěnějšího odborného zájmu (srov. Vachková, 2014); řada dílčích problémů však dosud zůstává bez důkladné metalexikografické reflexe (jde mj. o specifika exemplifikace kolokviálních, vulgárních nebo terminologických lexikálních jednotek). V tomto příspěvku se budeme věnovat právě exemplifikaci termínů, a to na materiálu rukopisu *Akademického slovníku současné češtiny* (dále ASSČ).

Zatímco v předchozích akademických výkladových lexikonech češtiny ani v prvních dvou dílech *Slovníku současného slovenského jazyka* (2006, 2011) nejsou terminologická hesla vždy příklady opatřena, při tvorbě koncepčních zásad ASSČ bylo stanoveno, že exemplifikace termínů bude obligatorní (Lišková – Procházková, 2016). Slovníková exemplifikace totiž potvrzuje existenci slova, konkretizuje výklad významu a pomáhá rozlišit významy u polysémních slov, postihuje gramatické vlastnosti lexému a jeho kolokabilitu, ukazuje stylové a komunikačně-pragmatické charakteristiky dané lexikální jednotky. To vše je jistě žádoucí nabídnout i u terminologických hesel a významů. Proč se tedy některé slovníky exemplifikace některých termínů zříkají? Ústředními požadavky na slovníkový příklad jsou typičnost, srozumitelnost, ilustrativnost a přiměřená délka kontextu, k čemuž se příp. přidávají zásady jako exemplifikovat všechny významové odstíny či dokládat valenci. Z toho vyplývá, že sestavení exemplifikačního bloku není vždy jednoduché, neboť některé z požadavků vykazují až protichůdné tendence; je tak třeba přistupovat i k různým kompromisům.

Věcná správnost lexikografického zpracování terminologických hesel a významů by měla být zajištěna spoluprací s oborovými experty (Lišková – Nová, 2016), jejichž kontrolní činnost se netýká jen výkladu významu, ale i exemplifikační části heslového odstavce. Tato revize se jeví jako nezbytná, neboť (méně nápadné) věcné chyby nemusejí být lexikografy v autentickém jazykovém materiálu rozpoznány; v ASSČ jsou zpracovány termíny z cca sta oborů, z nichž některé procházejí bouřlivým vývojem (chemie, informatika, medicína ap.).

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

Na omezeném prostoru alespoň v náznavu ilustrujeme typově různou exemplifikaci termínů třemi příklady konkrétních hesel z rukopisu ASSČ: jde o 1. stručnou, podle zvolené koncepce však dostatečnou exemplifikaci pomocí pouze jednoho příkladu (diskutována bude míra jeho encyklopedičnosti); 2. exemplifikaci dokládající variantní heslo; 3. exemplifikaci obsahující přenesené užití (dosud nelexikalizované, ale nikoli singulární užití dané lexikální jednotky).

aorist -tu (6. j. -tu) m. než. <řec.> lingv.

jednoduchý slovesný tvar vyjadřující minulý děj bez ohledu na jeho ukončenost nebo neukončenost: *aorist v češtině zanikl v patnáctém století*

bulimie -ie ž. <lat.> <řec.> med.

(mentální) bulimie psychicky podmíněná porucha příjmu potravy spočívající v záchvatovitém přejídání a následném vynuceném zvracení: *příznaky bulimie; onemocnět mentální bulimií; některé metody hubnutí mohou vést až k mentální anorexii nebo bulimii; lidé trpící bulimií mají obvykle narušené vnímání vlastního těla*

drúza -zy ž. <něm.> miner.

nepravidelně srostlá skupina krystalů na společném podkladu: *drúzy křemene; na výstavě bude možné zhlédnout velké drúzy křišťálů a achátů; přen. proti nebi se rýsovala drúza třpytivých mrakodrapů*

Závěrem konstatujeme, že přestože tvorba terminologické exemplifikace není pro lexikografa vždy snadná, pro uživatele slovníku je její existence jednoznačným přínosem.

SEZNAM LITERATURY A ZDROJŮ

Akademický slovník současné češtiny. Rukopis. Praha: Ústav pro jazyk český AV ČR, v. v. i. (ASSČ).

LIŠKOVÁ, M. – NOVÁ, J. (2016): Obecné zásady zpracování terminologie v ASSČ. In: P. Kochová – Z. Opavská (eds.), *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i., s. 176–179.

LIŠKOVÁ, M. – PROCHÁZKOVÁ, B. (2016): Exemplifikace. In: P. Kochová – Z. Opavská (eds.), *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i., s. 154–160.

Slovník súčasného slovenského jazyka, A–G, H–L (2006, 2011). Bratislava: Veda.

VACHKOVÁ, M. (2014): Z lexikografické dílny *Velké německo-české lexikální databáze*: Slovníkový příklad. *Jazykovědné aktuality*, roč. LI (3–4), s. 58–70. Cit. 2016-09-01. <<http://jazykovednesdruzeni.cz/JA3414.pdf>>.

VYUŽITÍ ČESKÉHO NÁRODNÍHO KORPUSU PŘI TVORBĚ ASSČ¹

LUCIE MACHÁLKOVÁ – VERONIKA VODRÁŽKOVÁ

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
machalkova@ujc.cas.cz, vodrazkova@ujc.cas.cz*

Key words: Czech National Corpus, dictionary sources, dictionary entry, exemplification, list of entries, monolingual dictionary

Cílem příspěvku je přiblížit tvorbu *Akademického slovníku současné češtiny* (dále ASSČ, více o slovníku Kochová – Opavská, 2016a; Kochová – Opavská, 2016b; Lišková, 2015). V ohnisku pozornosti budou ty postupy, které využívají dat z *Českého národního korpusu* i nástrojů na získávání těchto dat (zejména webové rozhraní *KonText*).

Dosavadní akademické všeobecné slovníky (tj. PSJČ, SSJČ, SSČ) využívaly jako základní materiálový zdroj cílenou excerpci. Jazykový korpus jako primární základna slovníku je v oblasti akademických výkladových slovníků češtiny novum. Kolektiv zpracovatelů ASSČ využívá korpusové opory zejména ve dvou etapách: (1) při několikafázové tvorbě hesláře a (2) při zpracování konkrétních hesel.

(1) Heslář je vytvářen po jednotlivých písmenech. Prvotní heslář je generován na základě tří vyvážených korpusů: SYN2000, SYN2005, SYN2010. Podmínkou generování je frekvence výskytu 5 a více ve všech korpusech dohromady. Automaticky získaný výsledek je dále zpracováván a vyhodnocován. Nejprve je očištěn od chybně lemmatizovaných položek. Poté je heslář daného písmene postoupen zpracovatelům. Zpracovatelé dále pracují s korpusem SYN (spojení všech synchronních psaných korpusů) i jinými zdroji dat. Pro korpus SYN využívají zejména vyhledávací nástroj *KonText* a nástroj *Word Sketch* v aplikaci *Sketch Engine*. V *KonTextu* je možné vyhodnocovat frekvenci lexikální jednotky nebo její zastoupení ve zdrojích. Do hesláře ASSČ nejsou zařazovány např. úzce specializované termíny, efemérní výrazy, výrazy sice frekventované, ale pouze v rámci minimálního počtu zdrojů, apod. Zpracovatel může také heslář rozšířit o jednotky, které sice patří k málo frekventovaným, ale dotváří slovotvorné hnízdo, zpravidla se jedná o adverbia. Jako doplněk při rozhodování o zařazení či nezařazení jednotky je využíván také např. korpus mluveného jazyka ORAL, mediální archiv společnosti Newton Media, a. s., webový korpus *Araneum Bohemicum*, webový korpus *WebCorp* nebo neologická databáze *Neomat*.

(2) Jednotka, která se stává heslářovou položkou, je dále zpracovávána jako samostatné heslo. Rozhraní *KonText* i nástroj *Word Sketch* v aplikaci *Sketch Engine* (oba pro ČNK) slouží v této fázi ke stanovení podoby lemmatu (včetně jeho variant), ke zjištění tvaroslovných charakteristik, valence a kolokability a k určení

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

stylistické příznakovosti (pomocí cílené analýzy jazykového materiálu). Tyto nástroje také slouží sémantické analýze, napomáhají při rozpoznávání významové struktury lexému, při určení polysémie, homonymie, synonymie a antonymie. U frekventovaných jednotek je výzkum těchto jevů prováděn na náhodně vybraném vzorku o minimálním počtu 300 konkordancí.

Korpusová data jsou zohledňována také při identifikaci víceslovných lexikálních jednotek a jejich zpracování. V neposlední řadě je ČNK stěžejním zdrojem pro výběr dokladů/příkladů do exemplifikační části hesla.

SEZNAM LITERATURY A ZDROJŮ

- Akademický slovník současné češtiny*. Rukopis. Praha: Ústav pro jazyk český AV ČR, v. v. i. (ASSČ). BENKO, V. (2015): *Araneum Bohemicum Maximum* [online]. Cit. 2016-09-01. <<http://www.korpus.cz>>.
- CVRČEK, V. – VONDŘIČKA, P. (2011): *SyD – Korpusový průzkum variant* [online]. Praha: Ústav Českého národního korpusu FF UK. Cit. 2016-09-01. <<http://syd.korpus.cz>>.
- Korpusy řady SYN* (syn2000, syn2005, syn2010, syn) (2000, 2005, 2010, 2014) [online]. Praha: Ústav Českého národního korpusu FF UK. <<http://www.korpus.cz>>.
- Databáze excerpčního materiálu Neomat* (2015) [online]. Verze 3.0. Praha: oddělení současné lexikologie a lexikografie Ústavu pro jazyk český AV ČR, v. v. i. Cit. 2016-09-01. <<http://neologismy.cz>>.
- Google [online]. Palo Alto (CA, USA): Google Inc. Cit. 2016-09-01. <<http://www.google.cz>>.
- KOCHOVÁ, P. – OPAVSKÁ, Z. (2016a): *Akademický slovník současné češtiny. Naše řeč*, 99, s. 57–83.
- KOCHOVÁ, P. – OPAVSKÁ, Z. (eds.) (2016b): *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i., s. 14–26.
- LIŠKOVÁ, M. (2015): *Akademický slovník současné češtiny a softwarový nástroj pro jeho tvorbu*. In: Z. Děngeová – P. Vališová (eds.), *Proměna jazyka a jeho výzkumu v době nových médií a technologií* [online]. Cit. 2016-09-01. Praha: Ústav pro jazyk český AV ČR, v. v. i., s. 42–48. <<http://www.ujc.cas.cz/miranda2/export/sitesavcr/ujc/sys/galerie-download/sbornik-brling.pdf>>.
- MACHÁLKOVÁ, L. – PROCHÁZKOVÁ, B. (2016): *Základní technika práce s korpusovým materiálem*. In: P. Kochová – Z. Opavská (eds.), *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i., s. 17–19.
- Mediální archiv NEWTON Media*, a. s. Cit. 2016-09-01. <<http://www.newtonmedia.cz>>.
- MICHALEC, V. (2014): *Poznámky k výstavbě hesláře připravovaného výkladového slovníku současné češtiny*. In: K. Dudová (ed.), *Varia XXII: zborník plných príspevkov z XXII. kolokvia mladých jazykovedcov: Nitra, 5.–7. 12. 2012*. Nitra: Katedra slovenského jazyka Filozofickej fakulty UKF, Slovenská jazykovedná spoločnosť pri JÚLŠ SAV, s. 273–281. Cit. 2016-09-01. <http://www.juls.savba.sk/ediela/varia/22/Varia22_abstrakty.pdf>.
- Oborová databáze ODE* (2009). Interní pracovní materiál. Praha: oddělení současné lexikologie a lexikografie Ústavu pro jazyk český AV ČR, v. v. i. Cit. 2016-09-01.
- Příruční slovník jazyka českého* (1935–1957). Praha: Státní nakladatelství, Školní nakladatelství, Státní pedagogické nakladatelství. (PSJČ)
- Slovník spisovné češtiny pro školu a veřejnost* (1978; 2., opr. a dopl. vyd. 1994; 3., opr. vyd. 2003). Praha: Academia (1997, 2004, 2005) [CD-ROM]. Voznice: LEDA. (SSČ)
- Slovník spisovného jazyka českého* (1960–1971). Praha: Nakladatelství ČSAV. (SSJČ)
- WebCorp* (1999–2013) [online]. Birmingham City University: Research and Development Unit for English Studies. Cit. 2016-09-01. <<http://www.webcorp.org.uk/live>>.

LEXIKOGRAFICKÉ SPRACOVANIE FRAZELOGICKÝCH VARIANTOV V NOVOM SLOVNÍKU SLOVINSKÉHO SPISOVNÉHO JAZYKA

MATEJ METERC – NATAŠA JAKOP

leksikološka sekcija, Inštitut za slovenski jezik Frana Ramovša ZRC SAZU
Novi trg 4, SI-1001 Ljubljana
matej.meterc@zrc-sazu.si, natasa.jakop@zrc-sazu.si

Key words: lexicography, phraseme, phraseology, phraseography, paremiology, Slovene, general standard dictionary

Nový slovník slovinského spisovného jazyka venuje osobitnú pozornosť frazeologickým jednotkám. Náčrt konceptu nového výkladového slovníka slovinského spisovného jazyka (Gliha Komac et al., 2015, s. 56) predstavuje komplexný opis každej frazémy. Frazému chápe ako ustálenú a reprodukovateľnú viacslovnú pomenovaciú konštrukciu nevetnej alebo vetnej povahy s frazeologickým významom, výraznou konotatívnou zložkou a expresívnou hodnotou. Fungovanie všetkých v slovníku prezentovaných frazém a ich variantov v súčasnom úze je overené ich frekvenciou v slovinskom jazykovom korpuse *Gigafida*. Hlavná podoba frazémy (tzv. „nulový variant“ alebo „frazelogická lema“) ako aj všetky varianty sú uvedené v reprezentatívnej podobe.

Slovník sa snaží zachytiť všetky typy frazeologických variantov podľa slovinskej typológie (Kržišnik, 1987), ktorá je vo veľkej miere ovplyvnená Mlacekovou (1984) typológiou. Predstavuje ortografické, morfológické, lexikálne, syntaktické a tiež pragmatické varianty (napr. vykanie a tykanie u pragmatických frazém). V rámci lexikálnej variantnosti prezentuje aj varianty spôsobené rozdielom medzi dokonavým a nedokonavým vidom slovíes, ak takéto varianty jazykový korpus potvrdzuje. Redaktori posudzujú frazeologické varianty aj z hľadiska jazykovej normy, avšak s rešpektovaním faktu, že frazémy sú z formálnej ako aj sémantickej stránky jednotky anomálne (Čermák, 1993, s. 46–51; Kržišnik, 2010, s. 93) a že ich vzťah k jazykovej norme je špecifický (Kržišnik, 1996). Každý variant sa musí v jazykovom korpuse objavovať natoľko často, aby redaktor mohol posúdiť, že naozaj ide o ustálenú formu, a nie o aktualizáciu (Kržišnik, 1987). Okrem toho musí ísť o prototypové použitie frazémy (Čermák, 2007, s. 572–573).

Lexikografický proces zahŕňa overovanie frazém, ktoré boli zaradené do *Slovníka slovinského spisovného jazyka* (SSKJ) alebo do *Slovníka slovinských frazém* (Keber, 2011), ale aj hľadanie doteraz neregistrovaných jednotiek. Veľký dôraz sa kladie na to, aby bol na základe korpusových údajov lexikograficky posúdený čo najväčší počet potenciálnych variantov frazémy. Preto treba kombinovať rôzne vyhľadávacie postupy. Okrem toho, že si redaktori všimajú, v akých podobách bola frazéma doteraz zachytená – ak vôbec bola – používajú napr. vyhľadávanie

častých kolokátov v kontexte každého komponentu frazémy. Veľa frazém aj ich variantov počas svojej práce s korpusom nájdú aj iní lexikografi, nielen frazeológovia. Lexikografickú prácu s variantmi parémií zlahčuje sociolingvistický výskum 918 jednotiek, podľa ktorého bolo určené slovinské paremiologické minimum (Meterc, 2014, 2016). V prípade tejto časti frazeológie sa odpovede okolo 450 respondentov z celého Slovinska používajú ako vstupné údaje pre korpusový výskum paremiologických variantov, ktorý je vo vzťahu k výskumu iných typov frazém špecifický (Đurčo, 2014).

ZOZNAM LITERATÚRY A ZDROJOV

- ČERMÁK, F. (1993): Povaha univerzálného ve frazeologii a idiomatice. In: E. Krošláková (ed.), *Frazeológia vo vzdelávaní, vede a kultúre. Phraseology in Education, Science and Culture: Zborník materiálov z II. medzinárodnej frazeologickej konferencie, Nitra 14.–16. september 1992*. Nitra: Fakulta humanitných vied, s. 45–52.
- ČERMÁK, F. (2007): *Frazeologie a idiomatika česká a obecná*. Praha: Karolinum.
- ĐURČO, P. (2014): Paremiologija i korpusnaja lingvistika. *Vestnik novgorodskogo gosudarstvennogo universiteta: Filologičeskije nauki*, 77, s. 13–17.
- Gigafida (2016) [online]. Cit. 2016-09-12. <<http://www.gigafida.net>>.
- GLIHA KOMAC, N. – JAKOP, N. – JEŽOVNIK, J. – KLEMENČIČ, S. – KRVINA, D. – LEDINEK, N. – MIRTIČ, T. – PERDIH, A. – PETRIC, Š. – SNOJ, M. – ŽELE, A. (2015): *Osnutek koncepta novega razlagalnega slovarja slovenskega knjižnega jezika*. Ljubljana: ZRC SAZU.
- KEBER, J. (2011): *Slovar slovenskih frazemov*. Ljubljana: ZRC SAZU.
- KRŽIŠNIK, E. (1987): Prenovitev kot inovacijski postopek. *Slava*, 1/1, s. 49–56.
- KRŽIŠNIK, E. (1996): Norma v frazeologiji in odstopi od nje v besedilih. *Slavistična revija*, 44/2, s. 133–154.
- KRŽIŠNIK, E. (2000): Idiomska beseda ali frazeološka enota. *Slavistična revija*, 58/1, s. 83–94.
- METERC, M. (2015): Online questionnaire providing information on most well-known and well-understood proverbs in Slovene language. In: V. Jesenšek – P. Grzybek (eds.), *Zora 97. Phraseologie im Wörterbuch und Korpus. Phraseology in Dictionaries and Corpora*. Maribor: Univerza v Mariboru, s. 205–216.
- METERC, M. (2016): The Slovene Paremiological Optimum – New Empirical Research Tools and the Augmentation of the Field of Minimum-oriented Research. *Proverbium*, 33, s. 319–338.
- MLACEK, J. (1984): *Slovenská frazeológia*. Bratislava: Slovenské pedagogické nakladateľstvo.
- Slovar slovenskega knjižnega jezika* (2016) [online]. Cit. 2016-09-12. <<http://www.fran.si>>.

ZPRACOVÁNÍ REGIONÁLNĚ PŘÍZNAKOVÝCH LEXÉMŮ V ASSČ¹

JANA NOVÁ – HANA GOLÁŇOVÁ

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
nova@ujc.cas.cz*

*dialektologické oddělení, Ústav pro jazyk český AV ČR, v. v. i.
Veveří 97, 602 00 Brno
golanova@ujc.cas.cz*

Key words: dialectism, monolingual dictionary, lexicography, regionalism, stylistics

Lexikální jednotky, které nemají celonárodní povahu, jsou do hesláře ASSČ vzhledem k zaměření slovníku a jeho omezenému rozsahu zařazovány limitovaně, pouze při dostatečně vysoké frekvenci v jazykovém materiálu. Zpravidla nezařazuje regionální varianty odlišné pouze hláskoslovně (moucha – **mucha**, vrabčák – **brabčák**) ani pravidelné tvaroslovné obměny (čepice – **čepica**). Při dostatečné frekvenci zařazujeme rodové varianty (**brambor** – **brambora**, **okurka** – **okurek**) a některé slovotvorně odlišné lexémy (čes. seřvat (koho) – mor. **dořvat** (koho)). Největší část regionálních jednotek v ASSČ tvoří lexikální dialektismy a regionalismy: jednak regionálně diferencní pojmenování pojmů z běžného života (šeřík – **bez**, brambor – **bandor** – **erteple** – **kobzole** – **zemák**), jednak názvy regionálně specifických reálií známé i mimo původní region (**čagan**, **frgál**; SSSJ používá pro lexémy tohoto typu označení *etnografizmy*).

Při výkladu významu názvů regionálních reálií se uplatňuje popisná výkladová perifráze (srov. **čagan** ‚dřevěná hůl se sekyrkou‘), v případě regionálně diferencních pojmenování užíváme často výklad neutrálním spisovným synonymem (srov. **bez 4**. ‚šeřík‘). Exemplifikace je zpracována podle stejných zásad jako u ostatních hesel (srov. Lišková – Procházková, 2016).

K označování regionálně příznakových lexémů používáme dva typy stylistických kvalifikátorů umístěných v záhlaví hesla nebo příslušného významu (srov. Lišková, 2016, s. 74–75). Kvalifikátor *region*. s upřesněním oblasti (čes., mor., slez.) je určen pro lexémy s velkým areálem, pokrývajícím území určité nářeční skupiny (Bělič, 1972) nebo zasahujícím do více nářečních skupin, např. celomoravské výrazy (**dědina**, **stolař**). Kvalifikátor *dialekt*. je určen pro lexémy s menším nebo nesouvislým areálem. Pro upřesnění územní platnosti těchto lexémů zvažujeme jednak stručný popis areálu nebo jeho nejvýznamnější části v záhlaví hesla (**bezinka 2**. dial. (stř. Čechy) ‚černý bez‘), jednak uvedení regionu ve výkladu významu (**frgál** dial. ‚velký koláč... tradičně pečený na Valašsku‘). U některých

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

lexémů s výrazně rozptýleným areálem je však popis výskytu problematický. Při uvádění stylistických kvalifikátorů je nutné hodnotit i různou stylovou platnost v jednotlivých regionech českého jazykového území, např. výraz **erteple** je z hlediska spisovného jazyka hodnocen jako zastaralý, ale v některých nářečích je neutrální (srov. ČJA 3, s. 268), moravismy **dědina**, **stolař** jsou tradičně přijímány jako regionální spisovné varianty, ale např. výraz **zavazet** nikoli (srov. Kloferová 1996; Jančák 2001 aj.) apod.

SEZNAM LITERATURY A ZDROJŮ

- BALHAR, J. a kol. (eds.) (1992–2011): *Český jazykový atlas 1–5, Dodatky*. Praha: Academia.
- BĚLIČ, J. (1972): *Nástin české dialektologie*. Praha: SPN.
- BUZÁSSYOVÁ, K. – JAROŠOVÁ, A. (eds.) (2006): *Slovník súčasného slovenského jazyka*. 1. diel A–G. Bratislava: VEDA.
- GOLÁŇOVÁ, H. (2011): Ke zpracování regionálních lexikálních jednotek v lexikální databázi Pralex. In: J. Světlá – A. Jarošová – A. Rangelova (eds.), *Česká a slovenská výkladová lexikografia na začiatku 21. storočia*. Brno: Jazykovedný ústav L. Štúra SAV a Tribun EU, s. 77–84.
- HLADKÁ, Z. (2002): Dialektismus. In: P. Karlík – M. Nekula – J. Pleskalová (eds.), *Encyklopedický slovník češtiny*. Praha: Nakladatelství Lidové noviny, s. 110.
- JANČÁK, P. (2001): K jazykově zeměpisné charakteristice česko-moravských a česko-slezských protikladů v slovní zásobě. *Naše řeč*, 84, s. 175–183.
- KLOFEROVÁ, S. (1996): Nářeční slovní zásoba a status spisovnosti. In: R. Šrámek (ed.), *Spisovnost a nespisovnost dnes*. Brno: Masarykova univerzita, s. 69–75.
- LIŠKOVÁ, M. (2016): Stylistická charakteristika. In: P. Kochová – Z. Opavská (eds.), *Kapitoly z koncepcie Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i., s. 154–161.
- LIŠKOVÁ, M. – PROCHÁZKOVÁ, B. (2016): Exemplifikace. In: P. Kochová – Z. Opavská (eds.), *Kapitoly z koncepcie Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i., s. 154–161.

NA CESTĚ K NOVÉMU SLOVNÍKU SLOVINSKÉHO JAZYKA

ANDREJ PERDIH – JANOŠ JEŽOVNIK

leksikološka sekcija, Inštitut za slovenski jezik Frana Ramovša ZRC SAZU
Novi trg 4, SI-1001 Ljubljana
andrej.perdih@zrc-sazu.si, janos.jezovnik@zrc-sazu.si

Key words: corpus linguistics, general dictionary, lexicography, Slovenian language

Príspevek predstavuje začátek tvorby výkladového slovníku slovinského spisovného jazyka s názvem *Slovar slovenskega knjižnega jezika*, tretja izdaja (dále SSKJ³), který od roku 2015, kdy byla zveřejněna jeho koncepce, vzniká v lexikologickém oddělení Institutu slovinského jazyka Frana Ramovše ZRC SAZU. SSKJ³ je slovník středního rozsahu s předpokládaným počtem 100 000 heslových slov. Slovník zachycuje všeobecně rozšířenou slovní zásobu současné slovinštiny. Částečně navazuje na předchozí vydání slovníku *Slovar slovenskega knjižnega jezika* (dále SSKJ¹, 1. vydání 1970–1991, 2., doplněné a částečně opravené vydání 2014), přitom však využívá nových metodologických východisek a lexikografických řešení.

Od roku 1991, kdy vyšel poslední svazek SSKJ¹, byl v lexikologickém oddělení Institutu slovinského jazyka Frana Ramovše ZRC SAZU publikován větší počet různých typů slovníků (retrográdní slovník (1996), pravopisná příručka se slovníkem (2001), slovník frazémů (2011), slovník neologismů (2012, 2013), právě vychází synonymický slovník (2016)). I když bylo v roce 2014 uveřejněno druhé vydání SSKJ, je potřeba znovu popsat slovní zásobu současné slovinštiny stále aktuální.

Základním materiálovým zdrojem při lexikografickém zpracování je miliardový korpus *Gigafida*, při práci jsou zčásti používány i korpusy *Kres* a *Nova beseda*. Pro analýzu korpusových zdrojů se používají funkce nástroje *Sketch Engine*. K získání údajů o výslovnosti, dynamickém a melodickém přízvuku a v určitých případech i ohýbání se využívá dotazníkové metody. Etymologická vysvětlení se většinou opírají o existující etymologické údaje.

Vzhledem ke změně typu a obsahu zdrojů (elektronické korpusy), změně primární podoby slovníku z tištěné na elektronickou a vzhledem k určitým koncepčním rozhodnutím došlo ve srovnání se SSKJ¹ k některým změnám mikrostruktury slovníku. Do hesláře jsou nově zařazeny i symboly a zkratky. Nově jsou zapracovány nahrávky spisovné výslovnosti a skloňovací–přízvuková paradigma včetně dynamického i melodického přízvuku a videonahrávky slovinského slovníku znakového jazyka. Místo relačního typu výkladů u odvozenin jsou použity plné výklady, dodatečně je nabídnuta informace o odvozeninách včetně odkazů na daná hesla. Exemplifikace v SSKJ¹ sice byla řazena podle syntaktických vztahů, ale nový slovník tyto vztahy navíc explicitně vyjmenovává a zároveň doplňuje i exemplifikací ve formě vět z korpusu (upravených pouze výjimečně). Slovník nově obsahuje i etymologická vysvětlení všech heslových slov.

V celkové mikrostruktuře slovníku jsou zahrnuty následující údaje: heslové slovo, slovnědruhové zařazení, tvaroslovné údaje, varianty, výslovnost ve dvou rozdílných transkripcích (s nahrávkou), odkaz na video znakového jazyka, informace o případné zkratce, stylistická charakteristika a údaje o oblasti užití slov, výklad významu, jména taxonu (botanické a zoologické nomenklatury), synonyma, kolokace (i s popisem gramatických vztahů), exemplifikace dokumentovaná příklady z korpusu, normativní poznámky (upozornění na odchylky od aktuální kodifikace), tabulky ohýbání heslového slova včetně dynamického a melodického přízvuku, víceslovné jednotky, slovtvorné údaje a etymologie.

Rychlejší zpracovávání hesel je umožněno extrakcí vybraných údajů z existujících slovníků a databází (slovtvorné vztahy, etymologické údaje, frazémy, výslovnost, varianty), přičemž bereme v úvahu, že tímto způsobem získané údaje nemusí být ani úplné, ani správné pro zastaralost údajů nebo mohou být do slovníků zařazeny podle východisek neshodných se SSKJ³.

Slovník bude postupně publikován na slovníkovém portálu *Fran* (www.fran.si), kde je momentálně volně dostupných 31 slovníků, které vznikly nebo byly digitalizovány v Institutu slovinského jazyka. Kromě základních funkcí, které portál nabízí, budou poskytnuty interní odkazy (na jiná hesla: přes synonyma, slovtvorné údaje, odkazy v etymologii), odkazy na bližší informace v etymologickém slovníku v rámci portálu *Fran*, odkazy na nahrávky slovinského slovníku znakového jazyka mimo portál *Fran* a zvukové nahrávky výslovnosti heslových slov.

SEZNAM LITERATURY A ZDROJŮ

- AHAČIČ, K. – LEDINEK, N. – PERDIH, A. (2015): *Fran: The Next Generation Slovenian Dictionary Portal*. In: K. Gajdošová – A. Žáková (eds.), *Natural Language Processing, Corpus Linguistics, Lexicography. Eighth International Conference. Bratislava, Slovakia, 21–22 October 2015. Proceedings*. Bratislava: RAM-Verlag, s. 9–16. <http://korpus.sk/~slovko/2015/Proceedings_Slovko_2015.pdf>.
- GLIHA KOMAC, N. – JAKOP, N. – JEŽOVNIK, J. – KLEMENČIČ, S. – KRIVINA, D. – LEDINEK, N. – MIRTIC, T. – PERDIH, A. – PETRIC, Š. – SNOJ, M. – ŽELE, A. (2015): *Koncept novega razlagalnega slovarja slovenskega knjižnega jezika*. Različica 1.1. Ljubljana: Inštitut za slovenski jezik Frana Ramovša ZRC SAZU. <http://www.fran.si/179/novi-slovar-slovenskega-knjiznega-jezika/datoteke/Potrjeni_koncept_NoviSSKJ.pdf>.
- GLIHA KOMAC, N. – JAKOP, N. – JEŽOVNIK, J. – KLEMENČIČ, S. – KRIVINA, D. – LEDINEK, N. – MICHELIZZA, M. – METERC, M. – MIRTIC, T. – PERDIH, A. – PETRIC, Š. – SNOJ, M. – ŽELE, A. (2016): *Novi slovar slovenskega knjižnega jezika – predstavitev temeljnih konceptualnih izhodišč*. In: F. Marušič – P. Mišmaš – R. Žaucer (eds.), *Škrabčevi dnevi 9*. Nova Gorica: Založba Univerze v Novi Gorici, s. 16–33. <http://www.ung.si/media/storage/cms/attachments/2016/10/21/13/45/01/Zbornik-%C5%A0D9_okt.16_splet.pdf>.
- PERDIH, A. (2016): *Splošni razlagalni slovarji slovanskih jezikov*. Ljubljana: Založba ZRC, ZRC SAZU. Zbirka Lingua Slovenica 8.
- Slovar slovenskega knjižnega jezika* (1970–1991). Ljubljana: SAZU, Državna založba Slovenije (SSKJ¹).
- Slovar slovenskega knjižnega jezika. Druga, dopolnjena in deloma prenovljena izdaja* (2014). Ljubljana: ZRC SAZU, Cankarjeva založba (SSKJ²).
- Slovar slovenskega knjižnega jezika* 2016. Elektronska objava tretje izdaje, rastoči slovar na www.fran.si (2016). Ljubljana: Inštitut za slovenski jezik Frana Ramovša ZRC SAZU (SSKJ³).

KONKURENČNÍ PŘEDLOŽKOVÉ VAZBY V SOUČASNÉ ČEŠTINĚ

HANA PROKŠOVÁ

oddělení jazykové kultury, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
proksova@ujc.cas.cz

Key words: language change, preposition, prepositional phrase

Můj příspěvek bude klást otázky po oprávněnosti a možnostech zachycení vývojových tendencí českého předložkového systému ve slovnících. Zachycovat ve výkladovém slovníku určitého jazyka aktuální dynamiku, nebo dokonce vývoj probíhajících změn v jazyce zčásti predikovat, je nesmírně obtížné. Lze tak citlivě činit u vývoje lexikálního významu, případně u určitých systematických změn v tvarosloví. Zaznamenat aktuální vývojové tendence vztahů významově-gramatických často není v možnostech autorů slovníku, přestože by jeho uživatelé mnohdy tento popis, případně i predikci uvítali. Ukázkovou výzvou je zachycení dynamiky předložkového systému. Prvotní otázkou zůstává, zda vůbec a jak detailně vztahy vyjadřované předložkami do základního výkladového slovníku zapracovávat i na jiných místech, než jsou hesla věnovaná samotným předložkám. Z hlediska pracovnice jazykové poradny mohu konstatovat, že uživatelé češtiny nemají vždy jasnou představu, které jazykové informace ve slovníku mohou nalézt, a právě vazby, tedy informace o širším okolí užitých slov, ve slovnících a základních příručkách, jako jsou *Pravidla českého pravopisu*, hledají marně. Slovník vazeb mnozí tazatelé neznají nebo ho nemají k dispozici.¹ Ve slovnících přehledových bývá valence řešena v rámci jednotlivých hesel v exemplifikaci. Oblast exemplifikace je tak stěžejní pro případné zachycení případných vývojových tendencí.

Pravděpodobně nejsilnější změnou novodobého předložkového systému češtiny je progresse předložky *na* zejména ve vazbách adverbálních, jak je vymezuje František Čermák (Čermák, 1996, s. 32). Mnohé z konkurencí, do nichž předložka *na* vstupuje, shrnuje už Marie Čechová ve svém článku z roku 1981, některým tendencím pronikání této předložky se věnuji i ve své (vznikající) disertační práci. Předložka *na* setrvale proniká do adverbálních spojení dynamických: *na prodejnu, na bar, na pobočku*, i statických: *na prodejně, na baru, na pobočce, na e-shopu*. V těchto spojeních je už vnímána jako přijatelná, byť stylově nižší, zatímco například ve spojení se substantivem *obchod*, které je sémanticky blízké substantivům *prodejna* či *e-shop*, případně se substantivem *banka* předložka *na* tradičně užívaným předložkám běžně nekonkuruje. Uživatelé jazyka a ostatně ani jazykovědci tak

¹ Významnou pomocí v této oblasti pro laické uživatele i pro odborníky může být elektronický slovník aktuálně užívaných vazeb *Valency Lexicon of Czech Verbs* VALLEX, který vzniká v Ústavu formální a aplikované lingvistiky MFF UK.

nemohou vývojové tendence chápat výhradně jako prediktabilní a jednotné posuny v systému. V příspěvku se pokusím na tyto drobné nejednotnosti poukázat.

Specifická je například situace u vazby se substantivem *sklad*. Jak SSJČ, tak SSČ uvádějí pouze vazbu *na skladě* (SSJČ: *mít /zboží/ na skladě*; SSČ: *zboží /je/ na skladě*), PSJČ uvádí *na skladě* i *ve skladě*, nikoli však explicitně jako varianty. Jazyková poradna ovšem v rámci systematizace a s ohledem na četné výskyty považuje i variantu *ve skladě/skladu* za přípustnou, přestože pro ni není ve dvou jmenovaných slovnících opora. Konkurence těchto dvou vazeb ukazuje ještě jemnou odlišnost doslovného užití a přeneseného užití, v němž je spojení *být/mít na skladě* užíváno pro vyjádření prosté přítomnosti a nemůže být spojením *ve skladě/skladu* nahrazeno (např. *pro každého má na skladě zcela osobní radu*²). Relativně častým dotazem v poradně je rovněž konkurence *jít do bazénu* × *jít na bazén*, která podle introspekce mnoha mluvčích odráží rozdíl v chápání *bazénu* jako komplexního areálu a *bazénu* jako samotné umělé nádrže s vodou. Ani v jednom ze slovníků není ani jeden z příkladů uveden.

Domnívám se, že dosavadní koncepce slovníků uživatele o tyto cenné a objektivně často vyhledávané informace může „ochuzovat“. Úskalí zpracování těchto jazykových informací však samozřejmě spočívá v možném rozsahu tištěného slovníku i v tom, že základním cílem výkladových slovníků doposud nebylo cílit na popis syntagmatických vztahů v jazyce. Pro uživatele slovníku by však právě reflexe vybraných vývojových tendencí v této oblasti mohla být užitečná. Vzhledem k této proměnlivosti a k základnímu charakteru konkurence v jazyce by i v případě začlenění „novějších“, variantních vazeb bylo vhodné citlivě zhodnotit jejich stylovou příslušnost, respektive hierarchii. Ta je u popisu konkurenčních jevů uživateli jazyka podle zkušeností jazykové poradny velmi žádaná, přestože z pohledu mnohých jazykovědců nikoli nezbytná.

Další zajímavou otázkou v oblasti zpracování předložkového systému je pak reflexe víceslovných sekundárních předložek ve slovníku. Jejich počet za poslední desetiletí kontinuálně narůstá³ a slovníky dosud zpracovávaly jako samostatná heslová slova pouze frekventované jednoslovné sekundární předložky.

SEZNAM LITERATURY A ZDROJŮ

- BLATNÁ, R. (2009): *Víceslovné předložky v současné češtině*. Praha: NLN.
- ČECHOVÁ, M. (1981): K expanzi předložky *na* v současné jazykové praxi. *Naše řeč*, 64, s. 33–45.
- ČERMÁK, F. (1996): Systém, funkce, forma a sémantika českých předložek. *Slovo a slovesnost*, 57, s. 30–46.
- KŘEN, M. – CVRČEK, V. – ČAPKA, T. – ČERMÁKOVÁ, A. – HNÁTKOVÁ, M. – CHLUMSKÁ, L. – JELÍNEK, T. – KOVÁŘIKOVÁ, D. – PETKEVIČ, V. – PROCHÁZKA, P. – SKOUMALOVÁ, H. – ŠKRABAL, M. – TRUNEČEK, P. – VONDŘIČKA, P. – ZASINA, A. (2015) [online]: *SYN2015: reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu FF UK. Cit. 2016-09-01. <<http://www.korpus.cz>>.

² Doklad z korpusu SYN2015.

³ Viz nedokonalý, přesto rozšiřující seznam Renaty Blatné *Víceslovné předložky v současné češtině*.

SLOVNÍK ČEŠTINY POJATÝ JAKO MODERNÍ „LEARNER’S DICTIONARY“?

MARTIN PROŠEK

*oddělení jazykové kultury, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
prosek@ujc.cas.cz*

Key words: dictionary, learner’s dictionary, dictionary making, language consulting, language database, electronic reference materials

V příspěvku vyjdeme z dosavadních výsledků práce na lingvistické strukturaci nové databáze dotazů v jazykové poradně Ústavu pro jazyk český AV ČR, v. v. i. Databáze by měla přesněji než kdykoli v minulosti vymezit, které oblasti jazykového systému jsou uživateli považovány za problematické a které naopak nikoli. Už v první fázi budování struktury databáze vzniká specifická mentální mapa jazykového systému, která zachycuje ty jeho oblasti, které uživatelé jazyka považují za problematické. Právě na takto vytipovaných problematických oblastech bývají založeny tzv. learner’s dictionaries, které jsou obohaceny o informační rámečky s podrobnějšími výklady problematických jevů, jež mají uživateli pomoci vyhnout se případným nedostatkům při jazykové produkci. Klademe si tedy otázku, zda a v jaké formě by podobné řešení bylo funkční a žádoucí i ve slovnících češtiny pro rodilé mluvčí. Odpovědi hledáme na konkrétních příkladech.

Dále hledáme odpověď na otázku, jak lze v moderním elektronickém slovníku řešit rozpor mezi tradičním požadavkem na stabilitu kodifikace a možnostmi elektronických médií, u kterých lze publikovaný obsah průběžně měnit podle aktuálních potřeb. Z hlediska uživatele je totiž svrchovaně žádoucí, aby měl jistotu, že při používání elektronických zdrojů poučení o jazyce nebude muset neustále sledovat aktualizace, protože informace v nich publikované se mohou každým okamžikem změnit. Zrovna tak silné jsou však mezi uživateli jazyka požadavky, aby v elektronických zdrojích našli poučení o jevech zcela nových nebo o aktuálních vývojových tendencích u jevů již dříve zachycených v kodifikaci. Oběma těmto oprávněným požadavkům lze podle našeho názoru vyjít vstříc vytvořením propracovaného popisného aparátu, který by přímo v jednotlivých heslech umožňoval vyznačit jevy dlouhodobě stabilní, jevy nově se objevivší (nově pronikající grafické podoby, nové významy apod.), jevy příznačné jen pro určité období (např. pojmenování přívrženců politických stran, které ve veřejném prostoru působily jen dočasně), jevy efemérní, okázalné atd.

SEZNAM LITERATURY A ZDROJŮ

Interní databáze dotazů jazykové poradny. Praha: Ústav pro jazyk český AV ČR, v. v. i.
PHILLIPS, J. – PROŠEK, M. (2011): The Oxford Studijní slovník – a New Semi-Bilingual Dictionary for Czech Learners of English. *Studie z aplikované lingvistiky*, 2, s. 29–42.

MORFOLOGICKÉ VARIANTY VE VZTAHU K NORMĚ A KODIFIKACI

KAMILA SMEJKALOVÁ

*oddělení jazykové kultury, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
smejkalova@ujc.cas.cz*

Key words: codification, language norm, morphological variants

Příspěvek se zabývá tématem variantnosti na morfoložické rovině jazyka ze dvou hledisek: z hlediska potenciálního autora slovníku coby tvůrce kodifikace a z hlediska běžného mluvčího, který se o údaje ve slovníku či jiné kodifikační příručce opírá při tvorbě textu. Je důležité pohlížet na téma variant z hlediska činnosti a rozhodování obou aktérů, protože vzájemně ovlivňují svou činnost, a tím i normu spisovného jazyka. Pokud se v normě spisovného jazyka (případně širěji v úzu) objevují varianty, je mluvčí při tvorbě konkrétního spisovného textu nucen vybrat jednu z nich. Mluvčí se při výběru z variant řídí různými kritérii, která závisí především na jeho pojetí jazykové správnosti a na postoji k ní: frekvence v úzu (co je častější, to je lepší), vlastní intersubjektívni norma (pojetí toho, co je spisovné a co není, případně obecněji, co je adekvátní či neadekvátní v daném typu textu na základě zkušeností a znalostí jazyka), soulad s platnou kodifikací (co je ve slovníku, to je považováno za správné), respektive s interní normou instituce či redakce apod. Pro účel tohoto příspěvku budu počítat s takovým typem uživatele jazyka, který má ambici ve svém textu používat spisovné jazykové prostředky a za spisovné považuje v největší míře právě prostředky kodifikované. Tak jako se musí mluvčí rozhodnout pro určité řešení v situaci, kdy jazykový prostředek může mít dubletní morfoložickou formu, je i při tvorbě kodifikace třeba se vyrovnat s existencí variant. Zatímco běžný mluvčí vybírá jednu variantu vhodnou pro konkrétní komunikát, tvůrce kodifikace se při tvorbě kodifikační příručky (slovník, *Pravidla českého pravopisu*, mluvnice) rozhoduje, zda doporučit pro spisovné jazykové projevy jednu variantu (a kterou) a ostatní ohodnotit jako nespisovné, nebo zda umožnit mluvčím výběr z více variant, přičemž tyto varianty mohou být hodnoceny jako stylově rovnocenné nebo může být jedna varianta hodnocena jako stylově neutrální, základní a ostatní mohou mít stylové příznaky: v tradičním pojetí jsou to např. varianty hovorové, knižní apod. V příspěvku se zamýšlíme nad tím, na základě kterých kritérií by tvůrci kodifikační příručky měli činit taková rozhodnutí, pokud se rozhodnou, že je činit chtějí. Krátce nastíním nejčastější typy situací, ve kterých je třeba existenci variant řešit.¹

¹ Nepočítáme mezi ně případy, kdy se v úzu setkáme s tvarem, který je gramaticky defektní, a jde tedy patrně o překlep či chybu nerodilého mluvčího (*vidím *pse – vidím psa*). V takovém případě o variantách ani nelze mluvit.

1. Na jedné straně je varianta považovaná za spisovnou, na druhé varianta považovaná za nespisovnou: *mladý muž – mladej muž*. Drtivá většina případů je jasně rozlišitelná. V této skupině může nastat problém v těch případech, kdy jedna z variant je stylově periferní a je na pomezí stylových vrstev, obvykle mezi vrstvou hovorovou a nespisovnou, respektive můžeme váhat, zda vůbec případy tohoto typu do této skupiny patří (spis. *táhl* – nespis. *táhnul*, ale spis. *bouchl* – spis. *bouchnul*).

2. Na jedné straně je varianta hodnocená jako spisovná, stylově neutrální, bezpříznaková, na straně druhé varianta spisovná, ale s určitým stylovým příznakem: *pláči – pláču, děkuji – děkuju, policisté – policisti*. I zde se můžeme setkat s hraničními případy, které by mluvčí mohli řadit spíše do první kategorie, kde jedna z variant je hodnocena jako nespisovná. Právě tvary *děkuju a policisti* jsou mluvčími často vnímány jako tvary pro spisovný text nevhodné. Z druhé strany naopak může některý mluvčí variantu považovanou za neutrální hodnotit jako knižní (*policisté, děkuji*) a poměr obou dublet převrátit.

3. Obě (všechny) varianty jsou považovány za stylově rovnocenné: *do Mělníka – do Mělníku, v průjezdě – v průjezdu, bez Wolkera – bez Wolkra*. Stylová rovnocennost však neznamená rovnocennost v oblasti frekvence či zaměnitelnost v určitých situacích či kolokacích. Např. u osobních jmen hraje při volbě koncovek roli i rodinná tradice.

Rozhodnou-li se tvůrci slovníku doporučovat určitou variantu (nebo určité varianty) jako preferovanou, např. pomocí jejího pořadí v hesle, je třeba si stanovit kritéria, na základě nichž tak budou činit. Za nejpodstatnější z nich považujeme noremnost daného jazykového prostředku. Toto kritérium úzce souvisí s frekvencí ve spisovných textech, ale i s intersubjektivním vnímáním daného prostředku uživateli jazyka. A zároveň je aplikace tohoto kritéria spojena s problémem, jak relevantní informace o normě vůbec získávat, protože současná situace, kdy se autoři příruček opírají převážně o *Český národní korpus*, nemůže podat celkový obraz.

SEZNAM LITERATURY A ZDROJŮ

- DANEŠ, F. (1979): Postoje a hodnotící kritéria při kodifikaci. In: J. Kuchař a kol. (eds.), *Aktuální otázky jazykové kultury v socialistické společnosti*. Praha: Academia, s. 79–91.
- HOMOLÁČ, J. – MRÁZKOVÁ, K. (2014): K stylistickému hodnocení jazykových prostředků, zvláště lexikálních. *Slovo a slovesnost*, 75, s. 3–38.
- JELÍNEK, M. (1979): Posuny ve stylistické charakteristice jazykových prostředků a jejich kodifikace. In: J. Kuchař a kol. (eds.), *Aktuální otázky jazykové kultury v socialistické společnosti*. Praha: Academia, s. 109.
- KNAPPOVÁ, M. (2008): *Naše a cizí příjmení v současné češtině*. 2. vydání. Liberec: AZ KORT.
- NEBESKÁ, I. (2003): *Jazyk – norma – spisovnost*. 2., doplněné vydání. Praha: UK.

LESK A BÍDA NĚMECKO-ČESKÝCH TERMINOLOGICKÝCH SLOVNÍKŮ

MARTIN ŠEMELÍK

*Ústav germánských studií, Filozofická fakulta Univerzity Karlovy
nám. Jana Palacha 2, 116 38 Praha 1
martin.semelik@ff.cuni.cz*

Key words: Czech, German, lexicography, terminology

Naléhavost kontrastivního zachycení terminologických subsystémů slovní zásoby různých párů jazyků je dána především praktickými potřebami překladu. Tento příspěvek v nutně komprimované podobě tematizuje problémy dnešní překladové terminografie¹, a to na příkladu jazykového páru němčina-čeština.² Ačkoli vedoucí postavení angličtiny v roli univerzálního jazyka dnešní vědy nelze zpochybňovat, je zároveň pravda, že v oborech jako např. filozofie, historie, literární věda apod. je němčina i nadále důležitým jazykem. Její význam v České republice navíc posiluje bezprostřední blízkost německého jazykového prostředí.

Německo-česká terminografie – totéž ale v podstatě platí i o terminografii dalších jazykových párů – je celkově mimořádně konzervativní, a to způsobem, který mnohdy až vyráží dech. V důsledku toho se potýká s celou řadou dnes již chronických problémů, a to především proto, že křečovitě lpí na stereotypních a zastaralých řešeních, která se mj. ve světle zkušeností překladatelů odborných textů již dávno ukázala jako nesprávná. Jedná se např. o otázky související s výběrem inventarizovaných položek, ekvivalentací, zachycením lexikálních vztahů (především terminologických synonym) apod.

Tento příspěvek je svým charakterem kontemplativní (popisující status quo) i konstruktivní (zahrnující návrh řešení) zároveň. Vychází z analýzy informační nabídky heslových statí vybraných německo-českých terminologických slovníků, které jsou dnes na trhu k dispozici, a kontrastuje ji s informační nabídkou heslových statí slovníků z produkce lexikografické sekce Ústavu germánských studií FF UK v Praze.³ Ukazuje se přitom mj. následující:

(a) Je třeba se vzdát mylné představy, že k psaní kvalitních překladových terminologických slovníků stačí být odborníkem daného oboru a dobře ovládat příslušné jazyky. Mnohá iritující faux pas současných terminologických slovníků jsou důsledkem absentujícího dohledu ze strany profesionálních lexikografů. Má-li terminologický slovník hladce sloužit svému účelu, musí vycházet ze spolupráce mezi experty daného oboru a lexikografy vnímavými vůči potřebám uživatelů slovníků.

¹ K základům terminografie viz Bergenholtz (1994) a Machová (1995) (inter alia).

² Viz též Šemelík (2011) a Vachková (2011), z nichž tento příspěvek vychází.

³ Srov. např. Kommová (2005) a Střelková – Suchánková et al. (2014).

(b) Terminologické slovníky musí začít svým uživatelům nabízet (relativně) kompletní jazykový servis, mají-li být opravdu funkčním nástrojem práce s jazykem. I odborný překladový slovník má usilovat o všestranný popis slovní zásoby. Neexistuje žádný důvod, aby nároky kladené na terminologický slovník byly v tomto ohledu nižší než na překladový slovník obecný.

(c) Je třeba se zamýšlet nad možnostmi prezentace terminografických dat ve formě online databank, které umožní propojení slovníku s dalšími zdroji, např. encyklopediemi, a které nebudou jen slepě kopírovat layout tištěných slovníků.⁴

Naprostá většina terminologických slovníků na našem trhu nespĺňuje standardy, jež by měly být u terminografických děl dnešní doby samozřejmostí. Výše zmíněné body (a–c) snad mohou přispět ke zlepšení tohoto suboptimálního stavu. Předpokladem je ale zcela zásadní změna paradigmatu v oblasti překladové terminografie. Čas k ní nazrál již před velmi dlouhou dobou.

SEZNAM LITERATURY A ZDROJŮ

- BERGENHOLTZ, H. (1994): Zehn Thesen zur Fachlexikographie. In: B. Schaefer – H. Bergenholtz (eds.), *Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern*. Tübingen: Gunter Narr Verlag, s. 43–56.
- BERGENHOLTZ, H. – NIELSEN, S. – TARP, S. (eds.) (2009): *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang.
- KOMMOVÁ, J. (2005): *Chemický slovník německo-český*. Praha: Karolinum.
- MACHOVÁ, S. (1995): Terminografie. In: F. Čermák – R. Blatná (eds.), *Manuál lexikografie*. Jinočany: H&H, s. 137–157.
- STŘELKOVÁ, V. – SUCHÁNKOVÁ, E. et al. (2014): *Německo-česká hudební terminologie on-line*. Praha: Filozofická fakulta Univerzity Karlovy. Cit. 2016-09-01. <<https://lexarchiv.ff.cuni.cz/slovník-hudebnich-hesel/>>.
- ŠEMELÍK, M. (2011): Krátké zamyšlení nad bolestmi současné německo-české terminografie. *Jazykovědné aktuality*, 50 (3–4), s. 87–95. Cit. 2016-09-01. <<http://jazykovedne-sdruzeni.cz/JA1134.pdf>>.
- VACHKOVÁ, M. (2011): *Das große akademische Wörterbuch Deutsch-Tschechisch. Ein erster Werkstattbericht*. Frankfurt am Main et al.: Peter Lang, s. 83–102.

⁴ K této problematice podrobně viz Bergenholtz – Nielsen – Tarp (2009).

SLOVNÍK NÁŘEČÍ ČESKÉHO JAZYKA: CHARAKTERISTIKA, VYMEZENÍ, MATERIÁLOVÁ ZÁKLADNA, VÝBĚR HESLOVÝCH SLOV

MILENA ŠIPKOVÁ

*dialektologické oddělení, Ústav pro jazyk český AV ČR, v. v. i.
Veveří 97, 602 00 Brno
sipkova@iach.cz*

Key words: nationwide dialect dictionary, dialectology, material basis

V roce 2011 začali pracovníci dialektologického oddělení Ústavu pro jazyk český AV ČR, v. v. i., v Brně zpracovávat *Slovník nářečí českého jazyka* (dále SNČJ). Po popisu gramatického systému od J. Běliče a geolingvistickém zachycení českých dialektů v šestisvazkovém *Českém jazykovém atlase* (dále ČJA) jde o třetí nejdůležitější úkol tradiční české dialektologie.

Nářeční slovník má mnoho společného se slovníky výkladovými, je to svým způsobem slovník výkladového typu, je ovšem obtížné stanovit jeho hranice. SNČJ je koncipován jako **celonárodní, nediferenční, tj. relativně úplný**. Zahrnuje výrazy související s materiální a duchovní lidovou kulturou z celého území českého národního jazyka, tj. jazyka domácích obyvatel historických zemí Čechy, Morava a české části Slezska; nezahrnuje lexikum českých enkláv v zahraničí. Do slovníku je v maximální možné míře včleněno lexikum tradičních profesí a řemesel. SNČJ zpracovává slovní zásobu z období přibližně **od poloviny 19. do konce 20. století; jádro tvoří materiál z 1. poloviny 20. století**. V širokém slova smyslu je slovníkem současným; můžeme říci, že zahrnuje nářeční slovní zásobu v průběhu zhruba **150 let**.

Celonárodní nářeční slovník má svá **specifika**: jeho **cílem** je (1) zachytit, vyložit a doložit **významy** a významové odstíny slov v nářečích českého národního jazyka, (2) zaznamenat **teritoriální rozšíření** nářečního lexika na území Čech, Moravy a Slezska (včetně jeho diferencovanosti formální); problematice spojené s tímto úkolem je věnován příspěvek Stanislavy Kloferové. Primárně jde o slovník elektronický, do budoucna se počítá s knižní podobou. Část, která byla začátkem letošního roku zveřejněna na internetu, obsahuje hesla písmen A–C (necelých osm tisíc hesel); internetová verze slovníku (sncj.ujc.cas.cz) umožňuje jeho případné doplňování: podněty veřejnosti včetně dalších, dosud nezaznamenaných dokladů s přesným významem a lokalizací jsou vítány na adrese dialekt@iach.cz.

Přestože až na výjimky (např. *bačovský*) SNČJ nezpracovává toponyma ani výrazy od nich odvozené, umožňují odkazy na slovníky pomístních jmen (*Slovník pomístních jmen v Čechách* a *Slovník pomístních jmen na Moravě a ve Slezsku*) nepřímo obsáhnout i oblast anoikonymie/mikrotoponymie, jejíž nedílnou součástí

nářeční lexikum je. Navíc počítáme s tím, že elektronická verze SNČJ umožní jeho vzájemné propojení s *Českým jazykovým atlasem* i s oběma výše zmíněnými slovníky pomístních jmen, čímž ještě více vynikne jeho komplexní charakter.

Celostní lexikografické zachycení a zpracování nářeční slovní zásoby patří k akutním a nejpotřebnějším akcím bohemistiky na poli uchování našeho kulturního dědictví, neboť v důsledku pronikavých poválečných společenských změn na vesnici a ve venkovském životě vůbec naše nářečí jakožto živé útvary národního jazyka rychle a nenávratně mizí. Jako živé výrazové prostředky se uchovávají jen ta slova, která jsou pro život lidí (zejména) na venkově důležitá, potřebná a užitečná. Z tohoto pohledu má tak SNČJ dvojí základní význam: kulturní a badatelský. **Kulturní hodnota** slovníku spočívá v tom, že dokumentuje a prokresluje historický obraz minulých dob: ukazuje, jak naši předkové vnímali, segmentovali a klasifikovali svět, co v něm poznávali, rozlišovali a co pro svůj život potřebovali jazykově fixovat (např. na 30 sloves pro význam ‚pršet‘). **Badatelská hodnota** slovníku tkví v tom, že v nemalé míře poskytuje podněty a poznatky k postihu vývojové dynamiky a zákonitostí lexikálního systému češtiny a jejich vnějších a vnitřních stimulů. Speciálně pak je tento slovník nepostradatelný pro bádání etymologická a slavistická, pro studia jazykové interference a jazykových kontaktů i pro zkoumání běžně mluveného jazyka. Slovník je však cenným přínosem i pro jiná, nelingvistická bádání, zvláště etnologická a historická. Materiál slovníku je dobře využitelný i ve výuce na všech stupních škol.

SNČJ vychází z dokladů obsažených v nářeční databázi. Nejrozsáhlejším materiálovým zdrojem je *Archiv lidového jazyka* (ALJ), systematicky budovaný od roku 1952. Obsahuje cenný materiál vyexcerpovaný především z dialektologických prací (rukopisných, tištěných i elektronických), z anketních dotazníků, ČJA, dále ze sběrů nářečního lexika prováděných odbornými pracovníky dialektologického oddělení (doplňkové terénní výzkumy) a dopisovateli. Součástí databáze jsou v menší míře i doklady z beletrie. Materiál se průběžně doplňuje. Uvedené zdroje obsahují cca 1,5 milionu záznamů. Důležitým pramenem je *Databáze souvislých nářečních textů* a *Archiv zvukových záznamů nářečních promluv*.

SNČJ jako slovník nediferenční zaznamenává jak lexikální zásobu, která je vlastní pouze dialektům, tzv. dialektismy (specifické, vlastní lexikální, sémantické), tak i výrazy shodné se spisovným jazykem. Druhá část příspěvku je věnována vymezení zpracovávaného lexika, problematice výběru heslových slov a zdůvodnění zvoleného přístupu.

SEZNAM LITERATURY A ZDROJŮ

Slovník nářečí českého jazyka (2016–) [online]. Cit. 2016-09-19. <<http://sncj.ujc.cas.cz/>>.

PRO KOHO JSOU SYNSÉMANTIKA VE SLOVNÍKU¹

BARBORA ŠTĚPÁNKOVÁ

*oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
stepankova@ujc.cas.cz*

Key words: conjunctions, dictionary, particles, prepositions, synsemantic words

Vymezení cílového uživatele patří mezi výchozí rozhodnutí při vzniku jednojazyčného výkladového slovníku. Odvíjí se od něj rozsah slovníku, způsob výkladu významu či volba používané terminologie. Hlavní pozornost je přitom zaměřena na způsob zpracování autosémantických slovních druhů, které jsou primární náplní slovníku.

Synsémantika (tj. předložky, spojky a částice), jakož i další pomocné slovní druhy (zájmena či číslovky) jsou ve slovnících spíše nezbytným doplňkem.

Synsémantika, která Filipec (1985, s. 77) označuje za slova gramaticko-lexikální, se vyznačují relačním typem slovnědruhového významu (tamtéž, s. 78). Tento jejich význam se pak ve slovníku zaznamenává prostřednictvím popisu jejich gramatické, pragmatické či vztahové funkce. Ve srovnání s autosémantikou (či citoslovci) jejich uchopení vyžaduje mnohem větší míru abstrakce.

Ve stávajících českých výkladových slovnících nalezneme u zpracovaných synsémantik různá pojetí – od výkladu synonymem (např. u některých částic v SSČ), přes zařazení do typu (u některých spojek v SSČ), po detailní výklad polysémního hesla (hlavně u předložek v SSJČ).

Při rozhodnutí, jak přistupovat k synsémantikům při tvorbě *Akademického slovníku současné češtiny*, je nutné brát v potaz např. následující faktory:

1. Chápání synsémantik co do dělení či šíře není v české lingvistice jednotné či ustálené (zejména u částic, ale i u spojek). S tím souvisí i terminologická nejednotnost. Slovníky však obvykle své výklady staví právě na jakési gramatické zkratce – přiřazení k typu. Je pak vhodné přiklonit se k některé ze stávajících klasifikací, či vytvořit klasifikaci novou?

2. Jednojazyčný výkladový slovník vždy do určité míry supluje roli mluvnice či pravopisných pravidel. Hledá uživatel spojky ve slovníku proto, aby věděl, kde psát čárky? Je otázkou volby, jak explicitně se tyto informace ve slovníku promítnou. S tím souvisí i způsob zprostředkování gramaticky komplikovaných oblastí jazyka (včetně případného zjednodušení).

3. Digitální (potažmo databázové) zpracování slovníku v kombinaci s množstvím korpusově dostupných textů umožňuje autorům přinést uživateli opravdu

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

velkou šíři doložených významů slov. Tyto možnosti však zároveň představují úskalí ve stanovení hranice při výběru relevantních položek či významů.

SEZNAM LITERATURY A ZDROJŮ

- ČERMÁK, F. – BLATNÁ, R. (eds.) (1995): *Manuál lexikografie*. Jinočany: H&H.
FILIPEC, J. – ČERMÁK, F. (1985): *Česká lexikologie*. Praha: Academia.
Slovník spisovné češtiny pro školu a veřejnost (1978). Praha: Academia. (SSČ)
Slovník spisovného jazyka českého (1960–1971). Praha: Nakladatelství ČSAV. (SSJČ)

VÝSLOVNOST V ČESKÉM VÝKLADOVÉM SLOVNÍKU: CO A JAK¹

VERONIKA ŠTĚPÁNOVÁ

*oddělení jazykové kultury, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
stepanova@ujc.cas.cz*

Key words: codification, Czech language, dictionary, language corpora, language databases, lexicography, model speakers, phonetics, pronunciation, standard

S ohledem na uvádění výslovnostních podob stojí současná česká lexikografie před dvěma hlavními výzvami:

- 1) volba způsobu zaznamenávání/transkribování výslovnosti (tj. JAK);
- 2) výběr konkrétních zvukových forem, respektive zdrojů pro slovníková data (tj. CO).

První z otázek již byla podrobně rozebrána zejména ve *Slově a slovesnosti* (Štěpánová, 2013) a v interních diskusích nad koncepcí *Akademického slovníku současné češtiny*, bylo by však jistě přínosné se k tomuto tématu ještě vrátit v rámci širší diskuse, které by se zúčastnili i odborníci působící jinde než v Ústavu pro jazyk český AV ČR, v. v. i.

Jako ještě závažnější a naléhavější se nyní jeví druhé téma, tj. reálné využití v dnešní době dostupných (ne však vždy ideálních) zdrojů pro získání aktuálních výslovnostních dat, která je vhodné ve slovníku uvádět. Spoléhat pouze na výslovnostní podoby zaznamenané staršími zdroji, případně na subjektivní povědomí zpracovatelů slovníku není udržitelné, zároveň však dosud scházejí komplexní korpusové zdroje mluveného jazyka, a zejména rozsáhlejší fonetické výzkumy v dané oblasti.

Hlavním zdrojem výzkumu spisovné výslovnosti by se v současnosti měly stávat zejména veřejné oficiální projevy (srov. např. *Encyklopedický slovník češtiny*, 2002, s. 91), v nichž se mluvčí snaží o co nejlepší srozumitelnost a které jsou považovány za kultivované (po všech stránkách), prestižní a adekvátní dané komunikační situaci. Jako vhodné aktéři normy se pro tyto účely jeví tzv. modelová mluvčí, kterými jsou především profesionální uživatelé jazyka (srov. např. Ammon, 1995; tuto koncepci formování spisovnosti v českém prostředí představil V. Dovalil, 2006 a 2012). V českém prostředí se lze spíše setkávat s termínem řečový/mluvní vzor, jehož vymezení je však širší (srov. zvláště Palková, 2008–2009, s. 21). Za modelového mluvčího je možné pokládat jen takový řečový vzor, který je profesionálním uživatelem jazyka usilujícím o kultivovaný projev, jenž je

¹ Příspěvek vznikl v rámci grantového projektu GAČR č. 16-06012S a s podporou RVO 68378092.

prezentován veřejně v oficiální komunikační situaci a bývá předem připravený; nejčastěji se jedná o rozhlasové a televizní hlasatele, moderátory atp.

V přístupu ke kodifikaci zvukové roviny jazyka lze stále vhodně uplatňovat tři známá Danešova kritéria (srov. Daneš, 1977): noremnost, adekvátnost (přiměřenost) k funkcím, systémovost. Noremnost výslovnostních jevů je důležité nejen zkoumat na zvukových materiálech získaných od tzv. modelových mluvčích, ale zároveň i ověřovat jejich přijatelnost v jazykovém společenství. Hledisko funkční adekvátnosti zde představuje především dobrá srozumitelnost dané formy s ohledem na konkrétní komunikační situaci. Kritérium systémovosti je pak důležité brát v úvahu při formulování obecných pravidel, stanovování variant atp. Jako zásadní se jeví zejména při popisu nových, dosud neustálených jevů, jakými mohou být např. výslovnostní podoby přejímek.

Nejen kvůli svému rozsahu a náročnosti nemohou být tyto práce svěřeny pouze jednotlivci a nemohou být kvalitně zvládnuty bez systematických dlouhodobých a nepřetržitých kolektivních výzkumů jazykové praxe. Těchto činností by se zřejmě měli opětovně zhostit pracovníci Ústavu pro jazyk český AV ČR, v. v. i., ve spolupráci s jinými lingvisty z celého Česka a s dalšími odborníky z praxe (čímž by došlo k znovuzřízení orgánu typu ortoepické komise). Nemělo by však jít pouze např. o krátkodobé projekty, ale nejlépe o soustředěné týmové práce v rámci znovuoobnoveného ústavního úkolu.

Přednostně vyřešena musí být zejména metodologie sběru a analýzy relevantních a dostatečných jazykových dat, bez nichž nejsou kvalitní výstupy představitelné. Z dostupných korpusových zdrojů se jako nejvhodnější jeví zejména korpusy DIALOG a *Monolog*. Zvláště pro případy, kdy je třeba ověřit výslovnostní podobu např. konkrétního, nepříliš frekventovaného slova cizího původu nebo vlastního jména, je však nezbytné v současnosti hledat i jiné než klasické lingvistické korpusové zdroje. Za jeden z nich lze považovat zvukový archiv Českého rozhlasu, v němž je možné sofistikovaně vyhledávat, a to i např. jen u vybraných (modelových) mluvčích. Příslušná prezentační webová aplikace sice není zaměřena primárně jazykovědně, avšak podobné postupy by mohly nabídnout velice zajímavý a cenný zdroj reálných dat, která jsou využitelná i lingvisticky. Další (paralelní) možností, jak získat výslovnostní údaje do vznikajícího slovníku, je využít výsledků jednorázových výzkumů věnujících se konkrétní užší oblasti slovní zásoby. Např. jedním z výstupů z právě probíhajícího projektu GA ČR *Fonologie českých anglicismů* bude fonetická databáze českých anglicismů, která by měla kromě dalších informací obsahovat i aktuální výslovnostní doporučení sestavená na základě kombinace výzkumu úzu a normy, excerptů dat z dosud publikovaných zdrojů, aplikací obecných pravidel fonologické adaptace anglicismů (a přejímek vůbec; srov. Duběda a kol., 2014) do češtiny atp.

Vzhledem k plánované elektronické publikaci budoucích lexikografických výstupů je rovněž namístě počítat s možností připojení zvukových nahrávek, jak je to již běžné v mnohých cizojazyčných slovnících. Jako důležitá a nezastupitelná se jeví zejména vzájemná intenzivní spolupráce lexikografů s fonetiky.

SEZNAM LITERATURY A ZDROJŮ

- AMMON, U. (1995): *Die deutsche Sprache in Deutschland, Österreich und der Schweiz. Das Problem der nationalen Varietäten*. Berlin, New York: Walter de Gruyter.
- DANEŠ, F. (1977): K dvěma základním otázkám kodifikace. *Naše řeč*, 60, s. 3–13.
- DOVALIL, V. (2006): K úvahám o spisovném/standardním jazyku: pohled z poněkud širší perspektivy. *Slovo a slovesnost*, 67, s. 96–103.
- DOVALIL, V. (2012): Nad Dolníkovou Teorií spisovného jazyka o konceptuálních a metodologických problémech ve výzkumu jazykových norem a spisovné variety. *Slovo a slovesnost*, 73, s. 135–146.
- DUBĚDA, T. – HAVLÍK, M. – JÍLKOVÁ, L. – ŠTĚPÁNOVÁ, V. (2014): Loanwords and Foreign Proper Names in Czech: a Phonologist's View. In: J. Emonds – M. Janebová (eds.), *Language Use and Linguistic Structure. Proceedings of the Olomouc Linguistics Colloquium 2013*. Olomouc: Palacký University, s. 313–321.
- Encyklopedický slovník češtiny* (2002). P. Karlík – M. Nekula – J. Pleskalová (eds.). Praha: Nakladatelství Lidové noviny.
- Korpus DIALOG 1.2* (2015) [online]. Praha: ÚJČ AV ČR. Cit. 2016-08-26. <<http://ujc.dialogy.cz>>.
- Korpus Monolog 1.1* (2016) [online]. Praha: ÚJČ AV ČR. Cit. 2016-08-26. <<http://monolog.dialology.org>>.
- PALKOVÁ, Z. (2008–2009): Výslovnost současné češtiny a kultura řeči. *Český jazyk a literatura*, 59, s. 18–26.
- Prezentační webová aplikace pro testování přístupu k archivu* (2014) [online]. Liberec: Technická univerzita v Liberci. Cit. 2016-08-26. <<https://naki.ite.tul.cz>>.
- ŠTĚPÁNOVÁ, V. (2013): Pojetí výslovnosti v českých výkladových slovnících (výzva k diskusi). *Slovo a slovesnost*, 74, s. 279–297.

NAD ZPRACOVÁNÍM HUDEBNÍ TERMINOLOGIE V NĚMECKO-ČESKÉ DATABÁZI

MARIE VACHKOVÁ

Ústav germánských studií, Filozofická fakulta Univerzity Karlovy
nám. Jana Palacha 2, 116 38 Praha 1
marie.vachkova@ff.cuni.cz

Key words: bilingual lexical database, German-Czech lexicography, LSP, musicology, terminography, written and spoken usage

Zpracování německo-české databáze hudební terminologie byl jeden z úkolů, které si vytkla lexikografická sekce při Ústavu germánských studií FF UK v Praze. Průběžná integrace terminologických celků do *Velké německo-české lexikální databáze* vede již od začátku prací mj. i k úvahám nad speciálními slovníky, které by přesahovaly nabídku pouhých glosářů; zpracování termínů, které je ovlivněno lingvistickým pohledem a poznatky moderní metalexikografie, se snaží vytvořit ke stávajícím tištěným i internetovým zdrojům alternativu, která je vhodná pro uživatelský okruh jevící zájem jak o muzikologii, tak i o němčinu, a nikoliv jen o letmou informaci bez solidního popisného aparátu.

Nároky na rozsah makrostruktury a výstavbu hesla (databázového vstupu) odrážejí otestované potřeby i výchozí situaci předpokládaného adresáta a jsou diskutovány s odborníky. Vzhledem k tomu, že slovník hudební terminologie, který by dostal nárokům kladeným na teoretická východiska a lingvistické parametry, chybí u nás jak v tištěné, tak elektronické podobě, bylo nutno zmapovat existující terminologické zdroje v obou jazycích a vystavět dílo na zelené louce. Za tímto účelem vznikla rešerše německých a českých muzikologických termínů ve formě informačně bohaté diplomové práce od muzikoložky a germanistky Veroniky Střelkové, která spolu se zpracovanými hesly a cenným dokladovým materiálem uvádí zájemce do základních zpracovatelských otázek.

Tento příspěvek se zaměřil jen na jejich reprezentativní výběr, a to na **a.** komplikovanou cestu při sestavování hesláře vzhledem k rozmanitosti muzikologických disciplín, kterou ovlivnila kritéria pro vznik lexikálního minima, **b.** úvahy k rozsahu informační nabídky heslové statě, **c.** jednotlivé body databázového vstupu se specifickými problémy (zvl. ekvivalence, polysémie, homonymie), **d.** úlohu metajazyka a encyklopedické složky. Na základě příkladů pak budou doloženy vybrané aspekty týkající se **e.** kolokací, zvláště pak dosud opomíjených verb a verbonominálních spojení i jejich forem v exemplifikaci, **f.** podílu psaného a mluveného odborného jazyka, **g.** problémů terminologických variant, **h.** synonym na ose časové a regionální. Příspěvek zakončí představení elektronické databáze hudebních termínů.

SEZNAM LITERATURY A ZDROJŮ

- BALTER, G. A. (1976): *Fachwörterbuch Musik / Deutsch-Russisch und Russisch-Deutsch*. Moskva, Leipzig: Deutscher Verlag für Musik.
- BATUŠEK, J. – HOROVÁ, E. (1988): *Česko-německý slovník hudební terminologie*. Praha: SPN.
- BATUŠEK, J. – HOROVÁ, E. (1989): *Německo-český slovník hudební terminologie*. Praha: SPN.
- KROHN, D. (1992): *Grundwortschätze und Auswahlkriterien: Metalexikographische und fremdsprachendidaktische Studien zur Struktur und Funktion deutscher Grundwortschätze*. Göteborg: Acta Universitatis Gothoburgensis.
- STŘELKOVÁ, V. (2012): *Zur Problematik des lexikalischen Minimums am Beispiel der Terminologie der Musikwissenschaft*. Diplomová práce. Praha: Filozofická fakulta Univerzity Karlovy.
- TĚŠILOVÁ, M. (1998): Die tschechischen Fachsprachen im 20. Jahrhundert und Ihre Erforschung: eine Übersicht. In: H. E. Wiegand – L. Hoffmann – H. Kalvenkämpfer (Hrsg.), *Fachsprachen: ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft*. Berlin: Walter de Gruyter, s. 1545–1551.
- SPIŠAR, J. (1996): *Anglicko-český hudební slovník s ilustrovaným prologem*. Ostrava: Montanex.
- VACHKOVÁ, M. a kol. *Velká německo-česká lexikální databáze*. Praha: Ústav germánských studií, Filozofická fakulta Univerzity Karlovy.
- VACHKOVÁ, M. – KOMMOVÁ, J. (2008): Fachwortschätze im allgemeinen Deutsch-Tschechischen Wörterbuch. In: M. Vachková (Hrsg.), *Beiträge zur bilingualen Lexikographie*. Praha: Univerzita Karlova, s. 205–220.
- WERNER, R. (1988): Das Problem der Äquivalenz im zwei- und im mehrsprachigen Fachwörterbuch. In: H. E. Wiegand – L. Hoffmann – H. Kalvenkämpfer (Hrsg.), *Fachsprachen: ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft*. Berlin: Walter de Gruyter, s. 1853–1884.

K VÝZNAMŮM A FUNKCÍM REFLEXIVNÍHO PRVKU *SI*¹

VOJTĚCH VESELÝ

oddělení současné lexikologie a lexikografie, Ústav pro jazyk český AV ČR, v. v. i.
Letenská 4/123, 118 51 Praha 1
vvesely@ujc.cas.cz

Key words: accusative case, dative case, reflexive morpheme, reflexive pronoun

Badatelé, kteří se věnují problematice zvrtných zájmen a zvrtných sloves z hlediska gramatického nebo lexikografického, se primárně zaměřují na významy a funkce reflexivního prvku *se*. Interpretace reflexivního prvku *si* je přitom neméně obtížná. V příspěvku se pokusím zodpovědět některé (teoretické i praktické) otázky, které vyvstaly v souvislosti se zpracováváním reflexivních sloves dativního typu pro vznikající *Akademický slovník současné češtiny*.

Pro lexikografické zpracování sloves je klíčová otázka, zda je reflexivní prvek součástí lexému jako slovotvorný nebo významově prázdný morfém, nebo zda se ke slovesu připojuje jen v některých jeho užitích (jako reflexivní zájmeno, příp. jako tvarotvorný morfém). V prvním případě je reflexivní prvek součástí slovesného lemmatu, v druhém případě se reflexivní tvar / reflexivní konstrukce uvádí nanejvýš v příkladové části hesla, tj. u slovesa nereflexivního.

Významy a funkce reflexivního prvku *si* budou v příspěvku popsány na pozadí významu a funkcí reflexivního prvku *se*. Položím si otázku, nakolik jsou funkce obou výrazů analogické. Pokud bychom například *se* v konstrukcích typu *umýt se* traktovali nikoli jako větný člen, ale jako tvarotvorný morfém, bylo by možné stejně hodnotit *si* v konstrukcích typu *umýt si (vlasy)*? Zdá se, že reflexivní prvek *si* má výraznější sémantiku než reflexivní prvek *se*: primárně označuje toho, kdo má z děje prospěch či neprospěch, srov. *vidět se (v zrcadle)* vs. *koupit si dům*; tento rozdíl vyplývá z obecné sémantiky akuzativu a dativu. S prospěchovým významem nejspíše geneticky souvisí význam prožívání děje, zájmu na ději, libosti apod., který prvek *si* vyjadřuje jako slovotvorný morfém / částice: *ležet si, plavat si, kouřit si*. Tento význam lze do jisté míry identifikovat také u sloves, u nichž prvek *si* vyjadřuje absenci externího adresáta: *číst si, zpívat si, hvízdát si* aj.

Další otázkou je obligatornost/fakultativnost reflexivního prvku *si*. Jako fakultativní se zpravidla hodnotí reflexivní prvek *se* v konstrukcích typu *venku (se) chumelí, žebrák (se) belhal ulicí* apod. Existují v češtině reflexivní slovesa dativního typu, jejichž význam by se po vypuštění reflexivního prvku nezměnil?

¹ Příspěvek vznikl v rámci grantového projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) Ministerstva kultury ČR *Nová cesta k modernímu jednojazyčnému výkladovému slovníku současné češtiny* (DF13P01OVV011).

SEZNAM LITERATURY A ZDROJŮ

- DOKULIL, M. a kol. (1986): *Mluvnice češtiny 1*. Praha: Academia.
- KARLÍK, M. – NEKULA, M. – PLESKALOVÁ, M. (eds.) (2002): *Encyklopedický slovník češtiny*. Praha: Nakladatelství Lidové noviny.
- ŠTÍCHA, F. (1981): Zvratná slovesa dativního typu. *Naše řeč*, 64, s. 186–191.
- VESELÝ, V. (2016): Zpracování reflexivních sloves a reflexivních konstrukcí. In: P. Kochová – Z. Opavská (eds.), *Kapitoly z koncepce Akademického slovníku současné češtiny*. Praha: Ústav pro jazyk český AV ČR, v. v. i., 2016, s. 119–123.
- WAGNER, R. (2012): *Reflexivní slovesa v češtině a prognóza jejich ekvivalentů v němčině*. Nепublikovaná disertační práce. Brno: Filozofická fakulta Masarykovy univerzity.

