



**Jazykovedný ústav Ľudovíta Štúra
Slovenskej akadémie vied**



**Ústav pro jazyk český Akademie věd
České republiky, v. v. i.**

ČESKÁ A SLOVENSKÁ VÝKLADOVÁ LEXIKOGRAFIA NA ZAČIATKU 21. STOROČIA

Súbor príspevkov v rámci medzinárodného projektu
Princípy a metódy tvorby výkladového slovníka

Tribun EU
Jazykovedný ústav Ľudovíta Štúra
Bratislava 2011

Redakcia:

Mgr. Jindra Světlá, Ph.D.

Mgr. Alexandra Jarošová, CSc.

Albena Rangelova, CSc.

Recenzenti:

doc. PhDr. Tatjana Grigorjanová, CSc.

Mgr. Jana Benkovičová, CSc.

© Klára Buzássyová, Edith Čonosová, Hana Goláňová, Martina Habrová, Jana Hašanová, Nicol Janočková, Alexandra Jarošová, Eva Kolovecká, Marta Koutová, Jana Letafková, Zdeňka Opavská, Helena Pernicová, Eva Podruhová, Barbora Procházková, Albena Rangelova, Jindra Světlá, Zdeňka Tichá, Vojtěch Veselý, Radek Volejník, 2011
Obálka © Vladimír Benko, 2011

© Jazykovedný ústav Ľ. Štúra SAV, 2011

© Tribun EU, 2011

Príprava tlačových podkladov: Mgr. Vladimír Radik

Tlač: Tribun EU s. r. o., Cejl 32, 602 00 Brno, Česká republika

ISBN ...

OBSAH (CONTENT)

Na úvod.....	7
Jindra Světlá	9 – 18
Struktura a zpracování hesel v lexikální databázi Pralex Structure and Processing of Entries in the Pralex Lexical Database	
Jindra Světlá	19 – 27
Budování hesláře lexikální databáze Pralex Building the List of Entries of the Pralex Lexical Database	
Jindra Světlá, Barbora Procházková	29 – 38
K základnímu zpracování substantiv v lexikální databázi Pralex On Basic Processing of Nouns in the Pralex Lexical Database	
Marta Koutová	39 – 50
Exemplifikace substantiv v lexikální databázi Pralex (zpracování příkladové části hesla) Exemplification of Nouns in the Pralex Lexical Database (Processing of the Example Part of an Entry)	
Edith Čonosová	51 – 57
Zpracování zdvojnásobků v lexikální databázi Pralex Processing of Diminutives in the Pralex Lexical Database	
Martina Habrová	59 – 67
Zpracování antonym v lexikální databázi Pralex Processing of Antonyms in the Pralex Lexical Database	
Helena Pernicová	69 – 76
Zpracování názvů osob a jmen přechýlených v lexikální databázi Pralex Processing of Personal Names and Gender Inflected Names in the Pralex Lexical Database	
Hana Golářová	77 – 84
Ke zpracování regionálních lexikálních jednotek v lexikální databázi Pralex On Processing of Regional Lexical Units in the Pralex Lexical Database	
Eva Podruhá	85 – 93
Základní charakteristika kompozit v lexikální databázi Pralex Basic Characteristics of Compounds in the Pralex Lexical Database	
Vojtěch Veselý	95 – 103
Ke zpracování zájmen v lexikální databázi Pralex Processing of Pronouns in the Pralex Lexical Database	
Jana Letafková	105 – 114
Zpracování vlastních jmen v lexikální databázi Pralex Processing of Proper Names in the Pralex Lexical Database	
Radek Volejník	115 – 124
Zpracování hypokoristik v lexikální databázi Pralex Processing of Hypocorisms in the Pralex Lexical Database	
Eva Kolovecká	125 – 136
Zpracování přirovnání v lexikální databázi Pralex Processing of Similes in the Pralex Lexical Database	

Zdeňka Opavská	137 – 152
Zpracování víceslovných lexikálních jednotek v lexikální databázi Pralex (se zaměřením na zpracování nefrazeologických víceslovných lexémů)	
Processing of Multi-Word Lexical Units in the Pralex Lexical Database (with Focus on the Processing of Non-Phraseological Multi-Word Lexemes)	
Albena Rangelova	153 – 164
Typologie zkrácených útvarů v lexikální databázi Pralex	
Typology of Abbreviated Forms in the Pralex Lexical Database	
Zdeňka Tichá	165 – 173
Jednopísmenné zkratky a značky ve slovnících a v lexikální databázi Pralex	
One-Letter Abbreviations and Acronyms in Dictionaries and the Pralex Lexical Database	
Klára Buzássyová	175 – 191
K aspektom tvorenia antonymných lexém s prefixom <i>ne-</i> (nominácia, lexikalizácia, pragmatické komponenty)	
Some aspects of Formation of Antonymous Lexemes with Prefix <i>ne-</i> (Nomination, Lexicalization, Pragmatic Components)	
Jana Hašanová, Nicol Janočková	193 – 205
Sloveso ako lexikografická výzva	
Verb as a Lexicographic Challenge	
Alexandra Jarošová	207 – 218
Ustálené prirôvnania: problémy ich chápania a lexikografického zachytenia	
Similes: Problems of Their Understanding and Lexicographic Processing	

Na úvod

Predkladaný zborník príspevkov je výsledkom spolupráce českého a slovenského lexikografického kolektívu v rámci spoločného projektu Ústavu pro jazyk český Akademie věd České republiky, v. v. i. a Jazykovedného ústavu Ľ. Štúra Slovenskej akadémie vied *Princípy a metódy tvorby výkladového slovníka* (Principles and Methods of Compiling a Monolingual Dictionary)¹.

Obsah príspevkov predstavuje problémové okruhy, ktoré boli identifikované v priebehu práce s bohatým lexikálnom materiálom a prediskutované na spoločných pracovných stretnutiach uskutočnených striedavo na českom a slovenskom akademickom pracovisku (2009 – 2011). Autori sa pri riešení nastolených otázok pokúsili uplatniť nový, netradičný pohľad na jazykové javy, na princípy výstavby ich lexikologicko-lexikografického opisu a na používané počítačové nástroje.

Väčšiu časť rukopisu tvoria práce týkajúce sa lexikálnej databázy *Pralex*, na ktorej pracuje lexikografický kolektív v Ústave pro jazyk český AV ČR v Prahe v rámci výskumného zámeru *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521). V šestnástich príspevkoch sa podáva systematický výklad problematiky spracovania lexikálnej databázy ako fundamentálnej opory budúceho výkladového slovníka súčasnej češtiny, takže čitateľ dostáva komplexný obraz databázy, získava prehľad o jej štruktúre, nadobúda hlbšie znalosti o jej funkciách, o možnostiach jej využitia, ale dôkladne sa môže oboznámiť aj s problémami.

Tri slovenské príspevky vychádzajú zo skúseností z koncipovania a redigovania prvých troch zväzkov *Slovníka súčasného slovenského jazyka* (tlačou zatiaľ vyšli dva) a ponúkajú nové riešenia niektorých koncepčných otázok (reinterpretácia typov antonymie, opis vybraných slovesných kategórií s ohľadom na ich fungovanie v komunikácii, prehodnotenie spracovania ustálených prirovnaní).

Redakcia

1 Ide o druhú etapu pokračujúceho projektu. Publikačným výstupom prvej etapy (2006 – 2008) je zborník *Lexikografie v kontextu informační společnosti* (2008).

Struktura a zpracování hesel v lexikální databázi Pralex¹

Jindra Světlá

Ústav pro jazyk český AV ČR, v. v. i.

The Structure and Treatment of the Entries in the Pralex Lexical Database

The article deals with the program as well as content part of the theoretical conception of capturing Czech lexis in the form of a lexical database. It briefly outlines the development of the Pralex software and describes the overall structure of the Pralex LDB in terms of its macrostructure as well as microstructure and the main principles of the treatment of database items (the data obligatorily provided with every entry). The ending is devoted to the potential of the application of this lexical database not only for a future monolingual dictionary and specialised dictionaries but also for various linguistic research projects.

1. Úvod

Cíl této databáze byl stanoven v zadání výzkumného záměru ÚJČ AV ČR, v. v. i., nazvaného *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století: jde o zpracování a výzkum současné české slovní zásoby v rovině teoretické a materiálové*². Z hlediska rozsahu se v tomto projektu uvažovalo o celkovém počtu 100 000 až 120 000 databázových položek. Tento počet však vycházel z předpokladu, že zakoupíme hotový lexikografický software i podpůrné programy a rovnou začneme databázi naplňovat. Z metodologického hlediska měla být databáze od začátku koncipována tak, aby se její věcný obsah, tj. informace obsažená v každé její položce, mohl stát **východiskem pro vypracování výkladového slovníku současné češtiny**, který svým obsahem bude sloužit jak pro běžné uživatele, tak pro účely zpracování přirozeného jazyka v informačních technologiích; jednotlivé položky databáze tedy mají sloužit jako **materiálový základ pro vypracování příslušných slovníkových hesel**. Z tohoto zadání také vyplývá především popisná funkce této databáze – zaměřujeme se na shromáždění a roztrídění jazykového materiálu dokládajícího především dnešní úzus (zatím výhradně na základě korpusu SYN³), nikoli na vyhodnocení současné lexikální normy; to bude cílem až následujícího úkolu, k němuž v této databázi připravujeme rozsáhlé podklady⁴.

Ve skutečnosti jsme se první dva roky projektu zabývali především vývojem a zprovozněním naší vlastní databáze, protože pro potřeby slovanských jazyků nic

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 K dalším úkolům v rámci tohoto výzkumného záměru viz Rangelova – Králík 2007, Rangelova 2008.

3 Od roku 2011 byl původní korpus SYN přejmenován na orig_syn.

4 Přihláška grantového projektu *Vytvoření koncepce nového akademického výkladového slovníku češtiny* byla podána v roce 2011. Koncepce má vzniknout v letech 2012–2014 a na ni by pak mělo navázat zpracování nového slovníku středního až velkého rozsahu (v elektronické i tištěné podobě).

vhodného neexistovalo. Nejprve tedy byla v letech 2005–2006 vyvinuta bohatě strukturovaná lexikální databáze, a to podle našeho vlastního návrhu ve spolupráci s Centrem zpracování přirozeného jazyka FI MU v Brně. Společně vyvinutý lexikografický software specializovaný pro potřeby češtiny byl nazván **PRALED** (název vznikl v Brně z označení “Pražská lexikální databáze”); tento **program** pro databázové zpracování slovní zásoby spolu s navrženým jádrem databáze a uživatelským rozhraním pro zápis dat, jejich zobrazení a vyhledávání tvoří **softwarovou část projektu**.⁵

Zároveň bylo pochopitelně třeba pro naši práci vytvořit i **teoretickou koncepci zachycení českého lexika formou lexikální databáze**. Ta je sice součástí zadání celého projektu, ale proto, že na její přípravu nebyl vymezen dostatek času na počátku úkolu (v souvislosti s potřebou vývoje vlastní databáze), tato koncepce vzniká po jednotlivých částech průběžně po celou dobu projektu v závislosti na postupu prací (podle toho, jaké úkoly aktuálně řešíme). Má dvě části – programovou (vývoj Praledu) a obsahovou – vymezení našeho pojetí a stanovení zásad pro zpracování jednotlivých slovních druhů a dílčích typů hesel.

V rámci **programové části databázové koncepce** jsme nejprve ve fázi vývoje programu navrhli celkovou strukturu databáze, formuláře pro zápis dat pro jednotlivé slovní druhy a v jejich rámci potřebné nástroje (samostatná okna) a v nich různé položky – pole pro zápis dat, v relevantních případech i kompletní dílčí nabídky (roletová menu), ze kterých se dají vybírat předem definované údaje. Při tvorbě databázové koncepce jsme se zpočátku věnovali souběžně i přípravě koncepce budoucího slovníku, který jsme pracovníčně nazvali Lexikon 21, zkráceně L21.⁶ Z tohoto důvodu (pro budoucí slovník) jsme rovnou navrhli všechny položky a nabídky potřebné pro komplexní lexikografický popis, i když nyní se všemi nepracujeme. Z časových i jiných důvodů jsme se totiž museli postupně omezit pouze na koncepci databázovou a vymezit jen zásady pro tzv. základní zpracování databázových položek, které není se slovníkovým zpracováním srovnatelné – má jiné cíle i parametry. Proto byly v roce 2007 vymezeny tzv. **povinné položky**, které v rámci výzkumného záměru vyplňujeme vždy – ty jsou součástí **tzv. základního zpracování** hesel v lexikální databázi (dále LDB) Pralex.

Druhou částí projektu (podle zadání výzkumného záměru jedinou, protože s vývojem naší vlastní databáze se v projektu původně vůbec nepočítalo) je **obsahová náplň** zprovozněné databáze, kterou jsme pro odlišení od části programové pracovníčně nazvali **PRALEX** (tato zkratka vznikla z označení “Pražský lexikon” – jde o *data-bázi slov, slovních tvarů a slovních spojení*).

Lexikální databázi Pralex jsme začali naplňovat lexikograficky relevantními daty až po zmíněné fázi vývoje a testování softwaru, tedy v průběhu roku 2007. K tomu bylo třeba se nejprve zaměřit na **stanovení zásad pro základní zpracování jednotlivých slovních druhů, resp. různých typů hesel v databázové podobě**. Průběžně zapracová-

5 Podrobně k etapám vývoje, struktuře heslového formuláře a funkcím LDB Praled viz Světlá 2008; viz též Pala et al. 2007.

6 Tento název jsme používali v letech 2005–2007, viz Světlá 2007 a další stati našeho kolektivu publikované ve sborníku *Computer Treatment of Slavic and East European Languages*, Bratislava 2007 (SLOVKO 2007), které byly zaměřeny spíše k uvažované koncepci slovníkové.

vávání koncepce lexikální databáze je nezbytnou součástí naší práce a její postuláty jsou ihned ověřovány na zkušebním vzorku a poté vtělovány do zpracovaných hesel. Je však třeba zdůraznit, že tato **obsahová koncepce** lexikální databáze je zaměřena výhradně na základní zpracování databázových položek, tedy jen v tom rozsahu, jak bylo stanoveno výzkumným záměrem. V žádném případě nemůže nahradit koncepci slovníkovou, která je obsahem přihlášky grantového projektu na léta 2012–2014.

Od začátku byl největší důraz kladen na základní zpracování substantiv⁷, protože jde o nejpočetnější slovní druh; některým dílčím typům či sémantickým skupinám substantiv jsme věnovali zvýšenou pozornost⁸ a naše zásady i jejich zpracování je podrobnější. Z dalších slovních druhů byla postupně – na základě předem připravených zásad – zpracována citoslovce, číslovky a zájmena⁹. Práce probíhaly v uvedeném pořadí; na každý z těchto méně obsáhlých souborů hesel se specializoval jeden člen zpracovatelského kolektivu. Vedle lexémů jednoslovných souběžně zpracováváme i vybrané lexikální jednotky víceslovné (frazémy¹⁰ a sousloví¹¹), a dále zkratky a značky¹², vlastní jména¹³, kromě toho byly do hesláře vloženy ze SSJČ i všechny podslovní části¹⁴; v roce 2011 se zaměříme na základní zpracování adjektiv a adverbíí, na závěr budou do databáze doplněna slovesa.

2. Celková struktura LDB Pralex

V rámci **makrostruktury** v LDB Pralex rozlišujeme **hesla jednoslovná a víceslovná, zkratky a značky**; jako samostatné položky uvádíme též **podslovní části** – ty však zatím pouze přebíráme podle SSJČ a nově je nezpracováváme ani nevyhodnocujeme. Celkové uspořádání heslových slov je abecední; hesla s počátečním velkým a malým písmenem se řadí v hesláři zvlášť, takže zde vlastně máme dvě abecedy (dva hesláře). I když bude možnost vyhledat vše najednou, primární nastavení je ale oddělené: buď pracujeme se souborem hesel začínajících malým písmenem (pak jde hlavně o soubor apelativ, ale patří sem i některé zkratky a značky), nebo velkým písmenem (tento soubor představují propria a iniciálové zkratky).

Databázové pojetí vyžaduje (zejména kvůli třídění dat v příkladové části hesla), aby každé lemma představovalo samostatnou položku databáze, což v praxi zna-

7 Ke zpracování substantiv viz zde společnou stať J. Světlé a B. Procházkové, k exemplifikaci stať M. Koutové.

8 K tomu viz zde stať E. Čonosové o zpracování zdobnělin, H. Pernicové o zpracování apelativních názvů osob a jmen přechýlených, M. Habrové o zpracování antonym a H. Goláňové o zpracování regionálních lexikálních jednotek.

9 Ke zpracování zájmen viz zde stať V. Veselého.

10 Ke zpracování přirovnání viz zde stať E. Kolovecké.

11 Ke zpracování sousloví viz zde stať Z. Opavské.

12 Ke zpracování zkratk a značek viz zde stať A. Rangelové a Z. Tiché.

13 Ke zpracování vlastních jmen viz zde stať J. Letafkové, ke zpracování hypokoristik stať R. Volejníka.

14 Viz zde stať E. Podruhové.

mená, že každá formální podoba (jednoslovnost vs. víceslovnost) i grafická podoba hesla (u variant) je v hesláři uvedena zvlášť. Tím je také vyloučen princip hnízdování. Za **samostatnou položku databáze** tedy považujeme každý jednoslovný lexém i všechny jeho varianty¹⁵, dále homonyma, zkratky, značky, podslovní části a vybrané víceslovné jednotky.

V rámci **mikrostruktury** u každé položky databáze podáváme konkrétní informace o heslovém slově a jeho manifestacích. Položky hesláře pracovním názvem "hesla", i když heslo v LDB Pralex není totožné se slovníkovým heslem; celkové zpracování našich hesel jako jednotek lexikální databáze je v mnohých směrech odlišné od zpracování hesel v tradičním výkladovém slovníku, přestože některé části hesla se slovníkovému zpracování přibližují. Struktura databázového hesla je – v souladu se slovníkovým pojetím struktury lexémů – hierarchická. Pochopitelně rozlišujeme hesla monosémní a polysémní, ale zvolený databázový přístup vyžaduje, aby i monosémní heslo mělo hierarchickou strukturu: nejprve se tedy u každé databázové položky uvádějí údaje platné pro celé heslo (na tzv. **kartu Heslo**), poté údaje vztahující se k jednotlivým významům (na tzv. **kartu Význam**). Každý význam je označen písmenem V a příslušnou číslicí, tzn. i jediný význam u monosémního hesla označujeme V1.

Od začátku jsme usilovali o to, abychom kromě ukládání údajů do **formuláře pro zápis dat** mohli pracovat i s uživatelsky přátelským **zobrazením** celého hesla, které by co nejvíce připomínalo strukturu heslové stati v tradičním výkladovém slovníku, resp. ve slovníku elektronickém¹⁶. K tomu nám slouží náhled, kde se zapsaná data okamžitě po uložení zobrazují. Vedle zobrazení jednoho či několika hesel v samostatných oknech pracujeme i s tzv. **Heslářem** – to je vstupní obrazovka, na které se formou tabulky zobrazují vyhledaná hesla pod sebou s nejdůležitějšími údaji uvedenými ve sloupcích (včetně administrace prací).

Kromě toho ovšem využíváme nejrůznější možnosti výpočetní techniky, které máme díky elektronickému zpracování dat k dispozici, včetně hypertextového propojení mezi hesly pro vyjádření jejich systémových vztahů, i různé funkce programu – např. funkci vyhledávání a třídění dat podle různých atributů.

3. Zpracování hesel v LDB Pralex

Tzv. základní zpracování každé databázové položky (pracovním tedy nazývané heslo) představuje složitý proces počínaje výběrem hesla a jeho zařazením do hesláře (s uvedením zdroje pro tento výběr), přes lemmatizaci (určení formy uváděné v záhlaví), vymezení typu hesla a označení slovního druhu, případně homonym a variant, až po práci s výkladovou definicí, poznámkovým aparátem a příkladovou částí hesla.

15 Jsme si vědomi toho, že samostatné uvádění variant představuje pouze pracovní, pomocné řešení v rámci této databáze, abychom při třídění materiálu mohli ke každé formální podobě lemmatu přiřadit příslušné doklady; při následném slovníkovém zpracování budou pochopitelně všechny varianty sloučeny do jediného hesla.

16 Srov. *Slovník spisovné češtiny pro školu a veřejnost* a *Velký slovník cizích slov* na CD-ROM z nakl. LEDA.

Poznámkový aparát na různých úrovních hesla slouží především pro zpracovatele-lexikografy: v jednotlivých poznámkách evidujeme nejrůznější informace, které později využijeme při slovníkovém zpracování. Některé poznámky mohou být ale určeny i pro uživatelskou veřejnost (např. encyklopedická poznámka u rodných jmen obsahuje údaj o jmeninách).

V relevantních případech propojujeme hesla mezi sebou pomocí tzv. **Seznamů**, které umožňují již v průběhu prací realizovat hypertextové odkazy (prokliky), a tudíž i snadné přechody mezi hesly, otevírání více hesel najednou na ploše obrazovky a jejich paralelní zpracování, vyhodnocování a současné revize či prohlížení. Tímto způsobem – pomocí **hypertextových odkazů** umístěných v jednotlivých seznamech – lze zachytit nejrůznější **systémové vztahy**, zejména homonymii, variantnost, synonymii či antonymii, ale i vztah mezi jednoslovnými a víceslovnými hesly (frazémy či souslovími a jejich komponenty), stejně tak jako vztahy mezi apelativy a proprii, nebo mezi zkratkami a jejich plnými názvy zpracovanými jako samostatná hesla apod.

Povinně vyplňované údaje (položky heslového formuláře)

V rámci tohoto základního zpracování, vymezeného výzkumným záměrem a přijatými zásadami, u každého zpracovávaného databázového hesla uvádíme:

a) na kartě Heslo:

Lemma: lemma odpovídá reprezentativnímu tvaru u slovníkových hesel, ale v databázi se v poli pro lemma uvádí zásadně jediný tvar, čili bez variant (ty jsou v LDB Pralex uvedeny samostatně a pak hypertextově a pomocí souhrnného odkazového hesla propojeny mezi sebou) a bez tzv. podhesel, která bývají ve výkladových slovnících někdy připojována k výchozí heslové stati – např. homonyma ve sdružené heslové stati¹⁷;

Homonymii: homonyma značíme římskou číslicí a pracovně (pro potřeby LDB Pralex) je čísujeme v pořadí podle slovních druhů, počínaje substantivem (všechna homonyma v rámci dané řady rovněž pomocí seznamu hypertextově propojujeme mezi sebou, aby se dala prohlížet zároveň);

Typ hesla: toto členění (heslo jednoslovné – víceslovné – podsl. část – zkratka/značka – odkazové) se zároveň promítá i do makrostruktury (viz výše); po označení typu hesla se otevře příslušný heslový formulář, který se u jednotlivých slovních druhů a typů hesel pochopitelně liší uváděním různých kategoriálních informací, ale zejména strukturou exemplifikace.

Slovní druh/typ: zařazení hesla ke **slovnímu druhu** až na některé výjimky zatím zachováváme v souladu se SSJČ, případně jen komentujeme difference či novější pojetí ve *Slovníku spisovné češtiny pro školu a veřejnost* (SSČ). K určení **typu lexikální jednotky** (dále LJ) – mimo rámec slovních druhů – patří rozhodnutí, zda jde v našem

17 K hesláři viz zde podrobněji naši stat' *Budování hesláře LDB Pralex*.

pojetí o sousloví, frazém, citátový výraz, vlastní jméno, zkratku nebo značku; technické řešení vyžadovalo, abychom tyto dílčí typy hesel zařadili do stejné nabídky (roletového menu) jako slovní druhy, protože v databázi potřebujeme kvůli možnosti vyhledávání údaje z různých nabídek vzájemně kombinovat (hlavně z hlediska jednoslovnosti a víceslovnosti);

Variantnost: v LDB Pralex pracovní registrujeme jako samostatné databázové položky (a propojujeme mezi sebou) téměř všechny variantní podoby uvedené v SSJČ a dalších zdrojích (při doplňování hesláře podle SSJČ vynecháváme pouze varianty označené už v tomto slovníku jako zastaralé). U všech variant (zatím bez ohledu na to, zda jde o varianty hláskové, grafické, pravopisné či slovtvorné) se nejprve zaměřujeme na dnešní stav úzu (z pohledu jejich doloženosti v korpusu SYN); celý soubor variant tak bude po stránce materiálové připraven k dalšímu vyhodnocování a okomentování z nejrůznějších aspektů (jak z hlediska časové osy, stylové příznakovosti, dnešní kodifikace či lexikální normy, tak i z hlediska dalšího výběru pro zařazení do slovníků různého typu, rozsahu a zaměření).

Zdroj pro heslář: tento údaj je pro nás pracovní informací o tom, odkud jsme dané lemma (z hlediska formy) převzali do našeho hesláře. Uvedený údaj má dvě části, protože vycházíme z konfrontace korpusových a vybraných referenčních zdrojů. U referenčních zdrojů se pro evidenci výskytu ve slovnících řídíme **zásadou prvního výskytu dané formy** ve výkladovém slovníku, počínaje SSJČ; vedle toho je též pro nás relevantním údajem i výskyt v používaných materiálových zdrojích, tedy to, zda je daná forma LJ uvedena ve *Frekvenčním slovníku češtiny* (dále FSČ) či nikoli, resp. zda se dané lemma vyskytuje v korpusu SYN, případně i v jiném zdroji.

b) na kartě Význam:

V **Poznámce k výkladu** nejprve všechny položky hesláře (databázová hesla) podrobně komentujeme z hlediska **výskytu** daného lemmatu a jeho variant v dostupných referenčních zdrojích, zejména výkladových slovnících; kromě toho zde registrujeme i celkové zpracování příslušných hesel, resp. jejich jednotlivých významů v předchozích slovnících. Jde o zpracovatelské poznámky, které nám velice usnadní další práci s heslem v rámci následného slovníkového zpracování, proto zde shromažďujeme v explicitní podobě¹⁸ nebo systémem odkazů a komentářů zejména informace o 1. registraci ve výkladovém slovníku (počínaje SSJČ)¹⁹, zda byla LJ v jednotlivých slovnících zpracována samostatně či přihrnázována, zda byla v daném slovníku zpracována s vlastním výkladem či bez výkladu, k jakému základovému slovu byla přihrnázována, resp. jaké varianty slovníky uvádějí, o co se dále můžeme opřít (např. o výklad u synonyma nebo u jiného hesla, o výklad v jiném slovníku) apod.

18 Některé z těchto informací nacházíme prostřednictvím slovníkového vyhledávače DEBDict, který je do Praledu integrován, ale pro náš rozhodovací proces potřebujeme zpracovatelské komentáře mnohem podrobnější, proto se nespolehneme pouze na tento nástroj, i když nám práci velice usnadňuje.

19 Tento údaj by měl korespondovat s tím, co uvedeme v poli Zdroj pro heslář.

Výklad významu: Jak již bylo řečeno, podle zadání výzkumného záměru bylo naším úkolem ke každé databázové položce (lemmatu) především shromáždit z korpusu relevantní doklady, které budou sloužit jako materiálový základ pro budoucí slovníkové zpracování. Nalezené korpusové doklady (tedy příklady na užití daného slova v konkrétních kontextech) rovnou vyhodnocujeme a třídíme k jednotlivým významům – právě proto jsou všechna hesla v databázi podrobně strukturována. K tomu jsme ale potřebovali vycházet z nějaké existující významové struktury, protože nové významové členění a definování jednotlivých významů slov se bude realizovat až při následném novém slovníkovém zpracování.

Většinou se v tomto směru opíráme o výklady i členění hesel podle SSJČ. Pokud SSJČ z nějakého důvodu nevyhovuje dnešnímu stavu, pracujeme i s definicemi z různých dalších slovníků; u nových významů výklad navrhujeme sami (alespoň ve formě pracovní definice, s níž ještě bude třeba dále pracovat). To znamená, že pro tvorbu výkladů nemáme jednotný metodologický základ (protože každý referenční zdroj vychází z odlišné koncepce), a proto rozhodně nelze tyto pracovní výkladové definice považovat za definitivní. Všechny výklady bude třeba v budoucím výkladovém slovníku zrevidovat, mnohé přeformulovat a sjednotit v souladu s budoucí koncepcí nového lexikografického popisu, která zatím neexistuje, a proto bude třeba ji nejprve vypracovat na začátku prací na následujícím úkolu (viz pozn. 3).

V případech potřeby výkladové definice převzaté z různých referenčních zdrojů již nyní částečně upravujeme, popřípadě na základě provedené analýzy korpusových dokladů přidáváme i další významy a navrhujeme nová dílčí řešení (vlastní návrh výkladové definice). Zároveň je ale třeba připomenout, že tato práce s významy a výkladovými definicemi vyplývá primárně z požadavku na roztřídění korpusových dokladů k dnešnímu úzu, které máme k dispozici z korpusu SYN a je pouze v kompetenci jednotlivých zpracovatelů, není podložena koncepčně a v žádném případě nemůže nahradit kompletní analýzu významové struktury každé lexikální jednotky, kterou bude třeba tak jako tak znovu provést při kompletním zpracování hesel v novém výkladovém slovníku. Zároveň je evidentní, že popis oblasti významového užití LJ má jiná východiska a pracovní postupy, než popis celé významové struktury.

V rámci stávajícího výzkumného záměru se nezaměřujeme na všechny kategorie, které bývají tradičně popisovány v obecném výkladovém slovníku, i když jsou pro ně příslušné položky v heslovém formuláři předem navrženy. Nezaměřujeme se např. na pragmatické aspekty významu, nevyhodnocujeme zatím stylové, oborové či komunikační charakteristiky, neřešíme ani otázku jazykové normy, variantnosti, spisovnosti a nespisovnosti, protože naplňovaná databáze je výlučně deskriptivní, nikoli preskriptivní, a naše současná práce je zaměřena hlavně na oblast významového užití a spojitelnost (kolokabilitu) heslových slov. Nelze proto očekávat, že LDB Pralex splní všechny funkce připisované výkladovému slovníku – kdyby na ni následně nenavázalo nové slovníkové zpracování, tato databáze zůstane především materiálovým podkladem pro další výzkum slovní zásoby jako celku. **V žádném případě LDB Pralex nelze považovat za hotový slovník**, i když některé parametry komplexního lexikografického popisu splňuje již nyní. Pro nás je cen-

né zejména to, že vedle utříděného lexikálního materiálu obsahuje též obrovské množství zpracovatelských poznámek a komentářů, které lze pokládat za přípravou fázi následujícího úkolu.

Příkladová část hesla – Exemplifikace: Z hlediska materiálových zdrojů spočívá jádro práce našeho lexikografického týmu ve vytěžení obsáhlých materiálových souborů – tří jazykových korpusů vytvořených v ÚČNK (SYN2000, SYN2005 a SYN-2006PUB) a shromážděných v elektronické podobě v korpusu SYN (který všechny tři uvedené korpusy zahrnuje), a to z různých aspektů: jak pro potřeby budoucího výkladového slovníku, tak pro další výzkum současného stavu české slovní zásoby a její dynamiky, ale také pro aplikace češtiny v informačních technologiích. Díky poměrně časové homogenosti materiálu můžeme na základě textů z posledních dvaceti až třiceti let ukázat, jak se slova v dnešní češtině užívají.

Protože máme jedinečnou příležitost pracovat s obrovským objemem lemmatizovaných a morfologicky značkových jazykových dat – s využitím klasických lexikografických metod v kombinaci s moderními softwarovými nástroji pro jejich vyhledávání a třídění (Bonito a Word Sketch Engine, dále WSE), pro budoucí slovník výkladový (i pro jiné využití, např. pro kolokační slovník) připravujeme velice podrobnou exemplifikaci z publicistických, uměleckých i odborných textů posledních desetiletí, které poskytují ucelený obraz dnešního užití lexikálních jednotek a jejich fungování v jazyce, a zachycují tak dynamicky pojatý současný jazyk. To znamená, že při našem databázovém zpracování nejde o pouhou excerpci větných kontextů k danému lemmatu tak, jak se ručně ke každému slovníku připravovala dříve, ani o úzký výběr typických příkladů sloužících v hotové slovníkové heslové stati k ilustraci, jak bylo obvyklé v dosavadních výkladových slovnících, ale o jakýsi mezistupeň, kdy na základě vyhodnocení nalezených korpusových příkladů rovnou přiřazujeme lexikograficky relevantní doklady k příslušným významům. Tyto doklady ukazují celou šíři spojitelnosti heslových slov a bude se s nimi podle potřeby pracovat dál.

V rámci LDB Pralex tedy jako výchozí, tzv. totální excerpci používáme celé texty shromážděné v korpusu v elektronické podobě a pak pracujeme s těmito kontexty dále. V příkladové části databázového hesla nazvané **Exemplifikace** evidujeme jak tzv. **citátové doklady**, čili autentické větné kontexty s uvedením datace, zdroje a typu textu, tak především (zejména u substantiv, adjektiv aj. autosémantik) tzv. **upravené doklady**, které představují minimální syntagmata až vybrané části vět, dokládající celou oblast významového užití heslových slov a jejich jednotlivých významů. Pro každý slovní druh nebo typ hesla jsme na základě podrobné analýzy materiálu navrhli specifickou strukturu exemplifikace (členění do samostatných bloků, číslovanych E1, E2 atd.)²⁰.

V rámci nového slovníkového zpracování se pak při další práci s korpusovým i jiným materiálem zaměříme hlavně na tzv. definiční kontexty, tedy ty, které nám pomohou při popisu významu naformulovat definici a daný význam dokreslit; ilustrovat dnešní užití pak budeme na základě výběru reprezentativních příkladů právě z těch dokladů, které byly shromážděny v LDB Pralex.

20 Částečně s oporou o publikaci Čermák – Holub 2005 aj.

4. Využití LDB Pralex pro budoucí výkladový slovník, případně pro dílčí slovníky speciální

Na základě rozsáhlého výchozího hesláře v LDB Pralex vzniká soubor částečně připravených hesel pro budoucí výkladový slovník, přičemž jednotlivé slovní druhy jsou v různém stadiu zpracovanosti. Souhrnně lze říci, že v rámci daného výzkumného záměru jsme se zaměřili především na budování hesláře a shromáždění materiálových podkladů pro další zpracování, ale zároveň postupně získáváme zkušenosti při práci s významovou strukturou hesla a s tvorbou výkladových definic.

Naše současná práce velmi usnadní začátek dalšího úkolu, protože budoucí slovník bude moci na naše analýzy a výsledky plynule navázat z mnoha aspektů: všechny zpracované položky databáze jsou okomentovány z hlediska doloženosti v korpusu, u každého hesla jsou naznačeny nové významy, pokud jsme byli schopni je podchytit (zatím jen na základě korpusu SYN), příkladová část ukazuje celou šíři spojitelnosti heslových slov, u mnohých hesel je bohatý poznámkový aparát ve vztahu k referenčním zdrojům atd.

Nové slovníkové zpracování tak bude moci pokračovat – po vymezení cíle, rozsahu a koncepce tohoto díla – především směrem k významové části hesla, tj. k vyhodnocení významového členění polysémních hesel, zatím v LDB Pralex strukturovaných hlavně na základě SSJČ a nových dokladů z korpusu, a bude se moci zaměřit na další změny a úpravy jejich významové struktury na základě dnes doložených významů, a také na potřebné úpravy výkladových definic všech hesel v souladu se současnou lexikální normou a aby jejich metajazyk nebyl ideologicky zatížený. Dále se naše budoucí práce bude muset pochopitelně zaměřit i na ostatní aspekty lexikografického popisu, kterými se v rámci stávajícího úkolu zatím nezabýváme.

Hesla zpracovaná v LDB Pralex budou využita též jako podklad pro dílčí výstupy – různé slovníky speciální. Rovněž budou sloužit jako utříděný materiál pro další lingvistické výzkumy. Ty položky databáze, které nebude třeba dále přehodnocovat a doplňovat, by po závěrečné revizi mohly být také postupně zpřístupňovány odborné veřejnosti. Nejprve bude na podzim roku 2011 veřejnosti zpřístupněna demoverze s vybranými ukázkovými hesly.

Dobu, kterou máme k dispozici pro práci na LDB Pralex, chápeme jako přípravnou fázi budoucího slovníku také z toho důvodu, že v ÚJČ AV ČR došlo v uplynulém desetiletí ke kompletní generační obměně lexikografického pracoviště a náš řešitelský kolektiv je dnes složen převážně z mladých, perspektivních pracovníků, kteří tímto způsobem získávají potřebné zkušenosti a znalosti pro další lexikografickou práci.

Literatura

- ČERMÁK, František – HOLUB, Jan: Syntagmatika a paradigmatika českého slova I. Valence a kolokabilita. 3. vyd. Praha: Karolinum 2005. 191 s.
- ČERMÁK, František: Kolokace v lingvistice. In: Studie z korpusové lingvistiky 2. Kolokace. Eds. F. Čermák – M. Šulc. Praha: Nakladatelství Lidové noviny 2006, s. 9–16.

- FILIPEC, Josef – ČERMÁK, František: Česká lexikologie. Praha: Academia 1985. 281 s. Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSC)
- JAROŠOVÁ, Alexandra: Problém vymedzenia kolokácií. In: Jazykovedný časopis 2007, roč. 58, č. 2, s. 81–102.
- Manuál lexikografie. Eds. F. Čermák – R. Blatná. Jinočany: H&H 1995.
- PALA, Karel – HORÁK, Aleš – RAMBOUSEK, Adam – RANGELOVA, Albena: Nové nástroje pro českou lexikografii – DEB2. In: Gramatika a korpus, F. Štícha, J. Šimandl (ed.), Praha: ÚJČ AV ČR 2007, s. 197–204.
- RANGELOVA, Albena: K úkolům výzkumného záměru *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století*. In: Lexikografie v kontextu informační společnosti. Sborník příspěvků z pracovního setkání Praha, 4.–6. 2007. Eds. A. Rangelova – J. Světlá – A. Jarošová. Praha: Ústav pro jazyk český AV ČR, v. v. i., 2008, s. 11–18.
- RANGELOVA, Albena – KRÁLÍK, Jan: Wider Framework of the Research Plan Creation of a Lexical Database of the Czech Language of the Beginning of the 21st Century. In: Computer Treatment of Slavic and East European Languages. Slovanské a východoeurópske jazyky v počítačovom spracovaní. Fourth International Seminar (Bratislava 25–27 October 2007). Proceedings. Eds. J. Levická – R. Garabík. Bratislava: Tribun 2007, s. 209–217.
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- Slovník spisovné češtiny pro školu a veřejnost. Red. V. Mejstřík a kol. Voznice: Leda spol. s r. o. 2005. (CD-ROM). (SSČ)
- SVĚTLÁ, Jindra: K návrhu, vývoji a funkcím lexikální databáze češtiny. In: Lexikografie v kontextu informační společnosti. Sborník příspěvků z pracovního setkání Praha, 4.–6. 2007. Eds. A. Rangelova – J. Světlá – A. Jarošová. Praha: Ústav pro jazyk český AV ČR, v. v. i., 2008, s. 19–32.
- SVĚTLÁ, Jindra: The possibilities and limits of lexicographical description of the Czech lexicon in database form. In: Computer Treatment of Slavic and East European Languages. Slovanské a východoeurópske jazyky v počítačovom spracovaní. Fourth International Seminar (Bratislava 25–27 October 2007). Eds. J. Levická – R. Garabík. Bratislava: Tribun 2007, s. 244–248.
- Velký slovník cizích slov. Verze slovníku 2.0. Red. J. Kraus – V. Petráčková. Voznice: Leda spol. s r. o. 2005. (CD-ROM). (VSCS)

Budování hesláře lexikální databáze Pralex¹

Jindra Světlá

Ústav pro jazyk český AV ČR, v. v. i.

Building the List of Entries of the Pralex Lexical Database

The list of entries of the Pralex lexical database (LDB) was at first built according to the *Frequency Dictionary of Czech*, later enriched by other entries from the *Dictionary of the Standard Czech Language* and continually enriched from the corpus SYN, too (so far optionally). To make the database an inventory of lexical units of contemporary Czech, as well as a base for a new middle-sized monolingual dictionary or other special dictionaries, we would like to expand it by choosing relevant lexical items from both newer reference resources (dictionaries) and material resources, e.g. other corpuses of written texts made by the Institute of the Czech National Corpus, various other databases (e.g. text archive Newton) or our own databases (excerpt database of neologisms and special excerpt database of various branches).

1. Úvod

V lexikografické praxi se pro vznik hesláře každého nového díla primárně nabízejí tři základní možnosti – buď se jeho heslář vybuduje na základě seznamu hesel z vlastní excerptce, anebo vznikne komparací vhodných referenčních zdrojů, tj. heslářů existujících výkladových a jiných slovníků, anebo se přistoupí ke kombinaci obou přístupů.

V české výkladové lexikografii byl vlastní sběr lexikálního materiálu realizován především pro *Příruční slovník jazyka českého* (dále PSJČ)². Pro potřeby tohoto slovníku byl založen počátkem 20. století bohatý lístkový archiv (pro PSJČ byl využit tzv. 1. materiál NLA³), ale i jeho heslář byl ověřován srovnáváním se staršími slovníky (např. se slovníkem Josefa Jungmanna a Františka Š. Kotta). Další dílo, *Slovník spisovného jazyka českého* (dále SSJČ)⁴, už vycházel z hesláře PSJČ a z vlastní doplňkové excerptce (tzv. 2. materiálu). Při sestavování *Slovníku spisovné češtiny pro školu a veřejnost* (dále SSČ⁵) se zase nejprve bral v úvahu heslář SSJČ, a pak byl tento slovník doplněn o další dílčí excerptci – 3. až 5. materiál lístkového archivu Ústavu pro jazyk český (dále ÚJČ). Podle předem stanoveného rozsahu a cíle daného slovníku se tedy každý heslář předchozího slovníku nejprve vyhodnocoval a redukoval – protože zatím měl každý následující obecný výkladový slovník češtiny menší rozsah než ten předchozí –, a poté se doplňoval o další jednotky v souladu se zaměřením každého z nich.

Od konce 60. let 20. století byla ruční excerptce v Ústavu pro jazyk český postupně omezována (jádro slovní zásoby bylo vyexcerptováno díky totální excerptci různých

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 PSJČ vycházel v letech 1935–1957.

3 Novočeský lexikální archiv Ústavu pro jazyk český, srov. i Goláňová 2011.

4 1. vydání SSJČ vyšlo v letech 1960–1971, 2. vydání v roce 1989.

5 1. vyd. SSČ vyšlo v roce 1978, 2. vyd., opravené a doplněné 1994, 3. vyd., opravené 2003, 4. vyd., totožné se 3., vyšlo v letech 2004, 2005 a 2009.

děl už v 1. základním lexikálním materiálu), v 70. letech probíhala výběrově a v 80. letech byla přerušena na nějakou dobu úplně; v té době se excerpovalo pouze pro *Akademický slovník cizích slov* (kolem sto tisíc významů, dále ASCS)⁶ a *Nový akademický slovník cizích slov* (NASCS, doplněno asi 2000 nových slov a významů). Od začátku 90. let se excerpcce prováděná v ÚJČ soustřeďuje především na neologismy, takže kontinuita excerpování běžné slovní zásoby nebyla udržena v úplnosti.

I před námi stála otázka, jaký přístup zvolit pro nově budovanou lexikální databázi (dále LDB) – čili zda bude její heslář vymezen na základě využití dostupných referenčních zdrojů, nové vlastní excerpcce, či materiálových zdrojů odjinud, a kromě toho bylo třeba rozhodnout, zda bude heslář LDB totožný s heslářem uvažovaného budoucího slovníku, či nikoli.

Jelikož byl náš výzkumný záměr zahájen mimo jiné proto, že v ÚJČ ucelená a cílená vlastní excerpcce běžné slovní zásoby z novějších textů než zmíněný lístkový archiv prakticky neexistovala, hodlali jsme využít elektronicky dostupné texty shromážděné v korpusu ÚČNK. K práci s korpusovými texty nás tedy vedlo více důvodů, nejen fakt, že jde o větné kontexty uložené v elektronické podobě, ale i skutečnost, že jsme vlastní soudobou excerpcí (kromě neologické) neměli.

2. Možnosti a kritéria výběru lexikálních prostředků do hesláře LDB Pralex

a) výběr podle frekvence: FSC

Určit vhodná kritéria pro výběr do hesláře dávno před tím, než bude připravena koncepce pro zpracování nového výkladového slovníku⁷, nebylo vůbec jednoduché. Proto vzhledem k uvažované čistě popisné funkci této podkladové, materiálové databáze⁸ jsme se při výběru položek do jejího hesláře přiklonili nejprve k frekvenčnímu hledisku, abychom na začátku prací mohli pracovat se souborem lexémů v současné češtině nejběžnějších. Tento výběr byl primárně determinován požadavkem, aby byly ke všem zpracovávaným položkám databáze k dispozici dostatečné materiálové zdroje, šlo tedy o kritérium doložitelnosti v korpusu SYN2000 (resp. v korpusu FSC2000) s dostatečnou frekvencí (tehdy to byl jediný synchronní lemmatizovaný a morfologicky značkováný korpus psaných textů ÚČNK, který jsme měli k dispozici při zahájení prací).

Při budování výchozího hesláře LDB Pralex jsme proto vycházeli ze seznamu položek uvedených a statisticky vyhodnocených ve *Frekvenčním slovníku češtiny* (dále FSC⁹), který byl zpracován na základě korpusu SYN2000, a z předpokladu, že počtem jednotek v něm zahrnutých se tento soubor nejvíce blíží hesláři jednosvazkového

6 1. vyd. vyšlo v roce 1995 ve dvou dílech (1. díl A-K, 2. díl L-Ž), od roku 1998 (2. vyd.) pak jako slovník jednosvazkový (A-Ž). V roce 2005 vyšel v aktualizované verzi pod názvem *Nový akademický slovník cizích slov* (NASCS). Elektronická verze z nakl. LEDA má název *Velký slovník cizích slov* (2005, VSCS).

7 Přihláška grantového projektu *Vytvoření koncepce nového akademického výkladového slovníku češtiny* byla podána v roce 2011; koncepce by měla vzniknout v letech 2012–2014.

8 Viz zde naši stať *Struktura a zpracování hesel v lexikální databázi Pralex*.

9 FSC vyšel v tištěné i elektronické podobě v roce 2004.

SSČ, a tudíž pokryje zejména jádro současné slovní zásoby¹⁰ – základní slovní fond, který zůstává v jazyce stabilní. Doplnění hesláře na celkový počet databázových položek určený výzkumným záměrem (100 000 až 120 000 jednotek) jsme pak chtěli realizovat dalším výběrem v následující fázi prací, až budou přesněji vymezena kritéria výběru. Původně jsme uvažovali o tom, že by se tento výchozí heslář doplňoval hlavně o víceslovné jednotky; jejich pojetí však bylo zpočátku nevyjasněné, a tudíž jejich zpracování již v první fázi prací bylo problematické¹¹.

Náš první heslář byl tedy v podstatě jen převzat z FSČ. Tento frekvenční slovník zahrnuje tři samostatné soubory – soubor apelativ, který obsahuje 50 tisíc položek, soubor proprií (2 tisíce jednotek) a soubor zkratk (1 tisíc), a je založen pouze na jediném kritériu – na frekvenci slov obsažených v prvním korpusu zpracovaném v ÚČNK (SYN2000). Velkou výhodou pro nás bylo, že FSČ již existoval v elektronické podobě, a tudíž mohl být náš heslář podle něj snadno a rychle naplněn. Nevýhodou ale bylo to, že FSČ obsahuje pouhý seznam lemmat se statistickými údaji a žádná další lexikograficky využitelná data (ani slovnědruhové zařazení). Navíc tento soubor z našeho pohledu vykazoval různé nedostatky vyplývající z jiných kritérií výběru textů pro uvedený korpus, než uplatňují lexikografové¹². Následné analýzy ukázaly, že jako výchozí heslář není FSČ pro naše potřeby dostatečně použitelný bez rozsáhlých úprav a doplňků, protože uvedená frekvence vztahovaná k jedinému korpusu by nebyla dostatečně objektivním kritériem pro sestavení hesláře dnešní češtiny v požadovaném rozsahu.

Kromě toho se nepotvrdila ani naše původní představa, že toto „jádro“ dané frekvenčně bude většinou odpovídat hesláři SSČ.¹³ Bylo tedy evidentní, že při tvorbě vlastního hesláře potřebujeme dále pracovat i s předchozími výkladovými slovníky češtiny.¹⁴

10 Srov. *Encyklopedický slovník češtiny*, s. 244: Slovní zásoba každého jazyka obsahuje základ (jádro) a okrajové vrstvy. Pojetí slovníkového základu může být různé: např. může jít o základní slovní fond, jenž je tvořen slovy, která jsou v jazyce odedávna a vyznačují se stabilitou, anebo podle strukturální teorie centra a periferie; centrum slovní zásoby tvoří slova neodvozená, sémanticky základní, stylově neutrální, produktivní a frekventovaná.

11 Podrobněji o zpracování víceslovných lexikálních jednotek v LDB Pralex viz zde stat' Z. Opavské o zpracování sousloví a E. Kolovecké o zpracování přirovnání.

12 Podle autorů FSČ je „SYN 2000 založen na pečlivě vyvažovaném a proporčním záběru do všech žánrů psané češtiny, pokrývá novodobý jazyk novin, krásné literatury i odborných publikací“. Tím však není řečeno, že jde zároveň i o systematický výběr lexika z různých oborů a oblastí našeho života, potřebný pro další lexikografické zpracování.

13 Analýza FSČ ukázala, že v některých případech jsou na základě frekvence zastoupena celá slovtovorná hnízda či jednotlivé řady, např.: *rod, rodák, rodič, rodička, rodičovství, rodičovský, rodina, rodinný, rodilý, rodiště, rodit, rodný, rodový, rodokmen; relace, relační, relativismus, relativita, relativnost*. Neplatí to však obecně, někdy je z podobné řady nejfrekventovanější podstatné jméno, jindy sloveso nebo přídavné jméno a další slovní druhy vykazují v daném korpusu nízkou frekvenci – FSČ např. zařazuje přídavné jméno *radiologický*, kdežto podstatná jména *radiologie* a *radiolog* v něm uvedena nejsou.

14 Nehledě na to, že jsme korpusový materiál chtěli zařadit rovnou k příslušným významům každého slova, takže i v materiálové části databázových položek se potřebujeme opírat o výkladové definice z některého předchozího výkladového slovníku.

b) výběr podle materiálových či referenčních zdrojů: korpus, nebo výkladové slovníky?

Další možností pro sestavení hesláře bylo využít korpus. Hlavní problém s využitím korpusu pro tvorbu našeho hesláře ale spočíval v tom, že na začátku prací na výzkumném záměru (počátkem roku 2005) existoval jen jediný korpus – SYN2000, a ten nám nemohl poskytnout relativně úplný seznam slov, se kterým bychom mohli pracovat (kromě 50 000 lemmat zařazených do FSČ – další jednotky v něm byly zastoupeny velice málo). Proto jsme potřebovali pracovat i s předchozími výkladovými slovníky. Otázkou bylo, zda se z hlediska hesláře opřít o mladší, výběrový slovník jednosvazkový (SSČ) a pak se více zaměřit rovnou na novější slovní zásobu (od 80. let 20. stol.) včetně víceslovných lexikálních jednotek a terminologie, nebo o nepatrně starší slovník středního rozsahu (SSJČ¹⁵), který pokrývá širší spektrum – zahrnuje slova základová i odvozená, slova aktivně užívaná i běžně známou pasivní slovní zásobu včetně široce zastoupené odborné vrstvy –, a ten pak na závěr doplňovat o novější slova z různých lexikálních vrstev.

Po zpřístupnění dalších korpusů ÚČNK (SYN2005 a SYN20006PUB) se ukázalo, že obsahují doklady na mnoho dalších slov zachycených v SSJČ. Také proto jsme se nakonec rozhodli, že pro nás bude nejvýhodnější při dalším budování hesláře použít jako hlavní srovnávací, referenční zdroj SSJČ, který je mnohem obsáhlejší než SSČ, a doplňovat další jednotky do našeho hesláře hlavně podle něj, a to včetně lemmat jen přihnízdovaných, protože jsme se rozhodli do LDB zahrnout jako samostatné databázové položky i hesla v obou předchozích slovnících (SSJČ, SSČ) buď tradičně nezpracovávaná vůbec, či jen přihnízdovaná a zpracovaná částečně nebo výběrově (např. verbální substantiva, abstrakta zakončená na -ost, adjektiva vztahová a přivlastňovací atd.). Zvláště vyčleňujeme také všechna propria, zkratky a značky, uvedené v SSJČ u jiných heslových slov – doplňujeme je do hesláře jako samostatné položky.

Proto, aby byl náš databázový heslář relativně úplný z hlediska spodní časové hranice, v současné etapě prací do něj postupně doplňujeme (a zároveň vyhodnocujeme z hlediska dnešního úzu, resp. doloženosti v korpusu SYN) téměř všechny lexémy uvedené v SSJČ jako lemma, tedy nejen hesla zpracovaná v tomto slovníku samostatně, ale i všechny varianty uváděné v záhlaví hesla, derivovaná slova přihnízdovaná na konci heslové stati a homonyma uváděná v rámci sdružených heslových statí, u substantiv i přechýlené podoby uváděné v SSJČ za výchozím lemmatem v závorce. Vynecháváme pouze archaismy (tedy slova označená již v SSJČ jako zastaralá či zastarávající) a také slova řídká, pokud ovšem nejsou v korpusu doložena.

Došli jsme totiž k názoru, že chceme-li při výběru do hesláře kompetentně rozhodovat o tom, které slovo je v dnešním úzu ještě stále živé a které už ne, bude pro přípravu nového výkladového slovníku přínosnější, když bude ve stávající LDB Prallex podchycen celý náš rozhodovací a ověřovací proces, než aby se v našem hesláři objevil jen jeho výsledek, tj. hotový výběr pro budoucí slovník (u kterého zatím není vymezeno, jak bude vypadat).

15 Podle údajů uvedených v posledním dílu (po Dodatcích, tj. v 1. vyd. ve 4. dílu na s. 1011, ve 2. vyd. v 8. dílu na s. 569) je uvedeno, že celý SSJČ obsahuje celkem 192 908 lexikograficky zpracovaných slov (ale navíc je v něm mnoho dalších lexikálních jednotek přihnízdovaných).

Na základě všech těchto podkladů shromážděných v LDB Pralex bude pak možné určit kritéria výběru heslových slov pro chystaný výkladový slovník právě v závislosti na tom, jak bude vymezen jeho cíl, typ a rozsah. Počítáme s tím, že všechny podklady zpracované nyní, tj. z hlediska hesláře v ověřovací fázi, využijeme jako dílčí výsledky našeho výzkumu ve formě různých statí o současném stavu české slovní zásoby z hlediska časové osy (na jedné straně dostaneme celkový obraz o slovech zastaralých a zastarávajících, na druhé straně o nových slovech a významech).

Nesnažíme se o pokrytí slovního repertoáru dnešní češtiny v jeho úplnosti (to ani není technicky a časově možné vzhledem k době trvání výzkumného záměru), ale přesto zachycujeme a vyhodnocujeme i různé potenciální derivace registrované v SSJČ, abychom zjistili současný stav z hlediska jejich užívání v dnešním psaném jazyce. Nakonec tedy náš heslář pravděpodobně obsáhne téměř všechna lemmata zachycená v SSJČ; heslář tohoto slovníku tak pro nás představuje východisko a spodní časovou hranici pro posuzování dnešního úzu a frekvence výskytu tohoto souboru lemmat v korpusových textech.

Na druhé straně je také třeba mít na zřeteli, že některá slova, která časově odpovídají výběru a zpracování hesel pro SSJČ, se ale do uvedeného slovníku z různých důvodů nedostala (např. slova nespisovná, úzce odborná, slangová, tabuová apod.), takže ani vymezení heslářem ze SSJČ nemůžeme považovat za rozhodující kritérium. Kromě toho k časové ose patří i evidování některých slov označujících skutečnosti dnes už neexistující (tzv. dobové reálie), které se sice již v aktivním úzu neobjevují, ale přesto se s nimi někdy můžeme setkat i dnes (např. při návštěvě skanzenů a nejrůznějších expozic), a také slov s obnoveným užitím – jde tedy o přesahy různými směry.

Po vyhodnocení hesláře SSJČ budeme postupovat směrem k novějšímu lexiku zachycenému v pozdějších publikovaných slovnících (výkladových i speciálních). Tím náš heslář značně přesáhne předpokládaný počet databázových položek, ale na začátku následujícího úkolu to umožní rychleji provést reprezentativní výběr do budoucího slovníku, protože všechna databázově zpracovaná hesla budou doplněna příklady, nebo alespoň okomentována z hlediska doloženosti v korpusu, se kterým pracujeme. Počítáme s tím, že zcela okrajová část slovní zásoby, pokud byla do hesláře LDB z nějakých důvodů zařazena (např. některé jednotky stylově příznakové či substandardní, nebo z různých okrajových komunikačních oblastí atd.), zůstane zpracována pouze v LDB a do budoucího výkladového slovníku se vzhledem k nutnému omezení jeho rozsahu nedostane. (Může však být zahrnuta do některého slovníku speciálního.) I to je důvod, proč by heslář LDB Pralex měl být obsáhlejší než budoucí slovník – tedy nejen proto, aby bylo z čeho vybírat pro slovníky různého rozsahu a zaměření, ale i proto, že lexikální databáze může plnit základní registrační a informační funkci o dnešním stavu periferie i v té podobě, jakou má zpracování jednotlivých databázových položek už dnes.

3. Zastoupení jednotlivých slovních druhů a typů hesel

Heslář LDB Pralex by měl pochopitelně zahrnovat všechny slovní druhy, i když počítáme s tím, že všechny nebudou do konce roku 2011 zpracovány.

Při zpracování postupujeme po jednotlivých **slovních druzích**, a proto je i jejich doplňování do hesláře v různé fázi: nejprve (od roku 2007, tj. po dvou letech vývoje vlastní databáze) jsme přistoupili ke zpracování **substantiv** (jako samostatné položky databáze rozepisujeme v souladu s naším pojetím všechna homonyma, varianty, u substantiv i přechýlené podoby mužských názvů osob a zdobněliny)¹⁶; od konce roku 2008 jsme též jako samostatná hesla začali rozepisovat všechny slovtvorné řady substantivních složenin zpracované v SSJČ¹⁷. Některé menší slovní druhy (**zájmena, číslovky a citoslovce**) se začaly doplňovat rovnou podle SSJČ jako celek a zpracování těchto dílčích souborů v podobě vymezené výzkumným záměrem už je kompletní¹⁸. V průběhu roku 2009 jsme pak přikročili k systematickému doplňování **adjektiv a adverbii** (od začátku roku 2011 je budeme zpracovávat); do konce roku 2011 by měla být do hesláře LDB Pralex vložena podle SSJČ i všechna **slovesa** a syntémantika.

Z hlediska různých **typů hesel**, které v hesláři LDB Pralex rozlišujeme (hesla jednoslovná – víceslovná – podslovní části a zkratky/značky) jsme vedle jednoslovných hesel postupně započali se systematickým doplňováním hesláře o další typy hesel podle SSJČ: nejprve byly v roce 2008 do hesláře vloženy všechny **frazémy** ze SSJČ uváděné na konci heslových statí jednoslovných hesel za značkou kosočtverce (po tzv. „hřebíku“)¹⁹. V průběhu roku 2009 jsme do hesláře doplnili všechny **podslovní části**²⁰; v roce 2010 jsme se zaměřili zejména na kompletování souboru **zkratk a značek**²¹ a souboru **proprií**²² (tyto dva typy hesel se doplňovaly i z jiných zdrojů). Nefrazeologické víceslovné lexémy – **sousloví** – zatím zpracováváme pouze výběrově²³.

Na rozdíl od výkladových slovníků zatím nemusíme tolik dbát o proporcionalitu hesláře z hlediska tradičního zastoupení různých složek slovní zásoby²⁴ a jejího funkčního rozvrstvení, ani o poměrné zastoupení jednotlivých typů hesel (slova – slovtvorné části slov – spojení slov), spíše se zatím snažíme o úplnost některých dílčích subsystémů, a proto je kompletujeme i z dalších zdrojů.

4. Možnosti využití jiných zdrojů pro doplňování hesláře a další práci s hesly

Předpokládáme, že naplňování LDB Pralex bude pokračovat i po skončení stávajícího výzkumného záměru, který rokem 2011 končí, protože tato databáze bude

16 Srov. zde naši úvodní statí *Struktura a zpracování hesel v lexikální databázi Pralex*; podrobněji viz zde společnou statí J. Světlé a B. Procházkové o zpracování substantiv, H. Pernicové o zpracování jmen přechýlených a E. Čonosové o zpracování zdobnělin v LDB Pralex.

17 Celkem jde o více než 6000 hesel.

18 Blíže ke zpracování zájmen viz zde statí V. Veselého.

19 Celkem tento soubor na začátku obsahoval asi 9000 položek, se kterými se pak dále pracovalo.

20 Celkem jich je založeno skoro 1000. Jde hlavně o první části složených slov, kterých bylo v SSJČ samostatně zpracováno asi 650; postupně se doplňují i druhé části složených slov, blíže viz zde statí E. Podruhové.

21 Výchozí soubor z FSČ obsahoval asi 1000 položek.

22 Z FSČ jich bylo původně vloženo 2000, blíže viz zde statí J. Letafkové.

23 Blíže viz zde statí Z. Opavské *Zpracování víceslovných lexikálních jednotek v LDB Pralex*.

24 Srov. Filipec J., *Teorie a praxe jednojazyčného slovníku výkladového*, In: Filipec, J. – Čermák, F., *Manuál lexikografie*, s. 16.

sloužit jako základna pro další výzkum a popis české slovní zásoby a mohou s ní pracovat i lingvisté z dalších pracovišť.

Abychom při vytěžování **referenčních (slovníkových) zdrojů** postupovali systematicky, po vytěžení SSJČ budeme náš seznam databázových hesel konfrontovat s heslářem ASCS, resp. VSCS pro doplnění dalších slov cizího původu; pro lexikograficky zpracovanou novější vrstvu využijeme publikace *Co v slovnících nenajdete* (1994, CSN) a *Nová slova v češtině. Slovník neologismů 1, 2* (1998, SN 1, 2004, SN 2); odbornou slovní zásobu pak budeme ověřovat a výběrově doplňovat podle jednotlivých oborů a oblastí lidské činnosti, resp. tematických okruhů, z oborových a speciálních slovníků či encyklopedií vyšlých po roce 1990.

Systematicky budeme přistupovat i k vytěžení dalších materiálových a databázových zdrojů pro doplňování hesláře. Z hlediska **materiálových zdrojů** je zatím heslář LDB Pralex průběžně doplňován výhradně jen z korpusu SYN, a to pouze v případě, že při zpracování některých databázových položek narazíme na další lexikální jednotky, které s daným heslem nějak souvisejí²⁵. Později chceme zdrojovou základnu rozšířit – i korpus SYN je od konce roku 2010 co do objemu ztrojnásoben²⁶, takže může pro doplňování hesláře představovat další významný zdroj. Poté můžeme ověřovat a zaznamenávat výskyt lemmat též v ostatních korpusech psaných textů ÚČNK do SYN nezahrnutých; následně můžeme pracovat i s jinými zdroji dostupnými v elektronické podobě na internetu (srov. např. databázi Newton aj.); pro budoucí potřeby se uvažuje i o cíleném sklizení dat na internetu.

Nelze opomenout také dvě **excerpční databáze** ÚJČ, které se postupně budují a ve kterých už je uložen značný objem lexikograficky relevantního materiálu, který je přímo ukládán pomocí výpočetní techniky.

Z hlediska novější slovní zásoby budeme využívat naši **neologickou excerpční databázi (nazvanou NEOMAT)**, která se buduje od začátku 90. let 20. století. Jde o speciální databázi zaměřenou na sledování dynamiky lexikální zásoby a na novou slovní zásobu. 1. část nazvaná **Archiv 1** zahrnuje nový materiál, který byl využit pro tvorbu dvou slovníků neologismů (SN 1 a 2) a obsahuje cca 200 000 elektronicky uložených excerpt; tato část je zpřístupněna veřejnosti. Neologická excerptce pokračuje dále (podle aktualizovaných excerpčních pravidel zaměřených širěji na lexikální dynamiku) a ukládá se v souboru nazvaném **Archiv 2**.

Při ověřování hesel a dalším doplňování hesláře o speciální slovní zásobu budeme též pracovat s **oborovou excerpční databází (ODE)**²⁷, která je určena pro

25 Pro srovnání: v LDB Pralex je v rámci celého písmene **a** založeno 4221 položek (zatím bez sloves, ale včetně zkratk a víceslovných lexikálních jednotek). V prvním korpusu SYN2000 seznam lemmat u písmene **a** nabízí 2953 položek, v souhrnném korpusu SYN (stav koncem r. 2010) jich je na seznamu už 6014. To znamená, že kdyby doplňování hesláře nepodléhalo řízenému výběru, mohl by se počet položek v každém písmeni minimálně zdvojnásobit již na základě souhrnu lemmat z korpusu SYN a dílčích seznamů (využívaných např. pro doplňování zkratk a propříj).

26 Od roku 2011 byl původní korpus SYN (který slučuje korpusy SYN2000, SYN2005 a SYN-2006PUB) přejmenován na orig_syn.

27 Oborovou excerpční databázi (ODE) jsme navrhli a zprovozili v letech 2009–2010; k polovině roku 2011 obsahuje kolem 60 tisíc záznamů.

cílenou excerpci z učebnic základních a středních škol, dále pak z nejrůznějších oborových slovníků či encyklopedií a z dalších textových materiálů popularizačního charakteru a bude vhodně doplňovat a vyvažovat materiály obsažené v korpusech. V první řadě vycházíme především ze zdrojů, které jsme průběžně pořizovali od 90. let a které máme k dispozici v tištěné podobě, abychom pokryli i ty zdroje, které nejsou dostupné v elektronické podobě. Tato ruční excerpce je zaměřena nejen na odborné výrazy z příslušných oborů, ale cíleně též na zachycení dokladů na různé substantivní výrazy charakteristické pro odborné a učebnicové texty (zejména jde o verbální substantiva a názvy abstrakt, tedy hlavně o pojmenování dějů, stavů, procesů, metod a vlastností). Excerptují se též jejich adjektivní a adverbialní deriváty. Na základě této naší vlastní excerpce budou v oborové databázi postupně vznikat dílčí oborové hesláře; do LDB Pralex se pak bude dostávat až výběr z každého oboru, relevantní pro všeobecný výkladový slovník. Databázi ODE využijeme nejen v dalších fázích zpracování hesel v rámci LDB Pralex, ale především při následném slovníkovém zpracování.

Protože se vhodné výrazy excerptují s celým větným kontextem, bude tato oborová databáze sloužit jak k vyvážení nepoměrného zastoupení jednotlivých oborů a oblastí v námi používaných korpusech, tak i k vyvážení nerovnoměrného výběru z hlediska stupně odbornosti korpusových textů a případně k vymezení dalších relevantních komunikačních oblastí. Tato doplňková excerpce napomůže i k doplnění vhodných dokladů na odborné užití běžných lexémů, popř. k podchycení jejich dalších distribučních oblastí, neboť je určena i k zachycení vhodných kontextů napomáhajících k upřesnění výkladových definic termínů a pojmů.

Při následném rozhodování o zařazení hesel do budoucího výkladového slovníku bude možné využít i výsledky **digitalizace lístkového lexikálního archivu** ÚJČ odpovídajícího starším vrstvám českého lexika. Naskenované excerpční lístky mohou být využity jako časově vymezený, starší doplněk současných jazykových korpusů jak ve fázi doplňování hesláře lexikální databáze o další relevantní jednotky, tak i při kompletování příkladové části v korpusu nedoložených hesel o časově relevantní doklady (v rámci probíhajícího výzkumného záměru však zatím o tomto doplňování lexikální databáze ze starších lexikálních archivů neuvažujeme).

Literatura

- Akademický slovník cizích slov. Red. V. Petráčková a J. Kraus. 1. vyd. Praha: Academia 1995, 1. díl A–K, 2. díl L–Ž, 834 s.; 1998 A–Ž, 834 s. (ASCS)
- Encyklopedický slovník češtiny. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: Nakladatelství Lidové noviny 2002. 604 s.
- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSC)
- Goláňová, Hana: Novočeský lexikální archiv a excerpce v průběhu let 1911-2011. Slovo a slovesnost, 2011, roč. 72, s. 20-33.
- Manuál lexikografie. Eds. F. Čermák – R. Blatná. Jinočany: H&H 1995. 283 s.
- Nová slova v češtině 1. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 1998. 356 s. (SN 1)

- Nová slova v češtině 2. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 2004. 568 s. (SN 2)
- Nový akademický slovník cizích slov A–Ž. (J. Kraus a kol.) 1. vyd. Praha: Academia 2005. 880 s. (NASCS)
- PALA, Karel – HORÁK, Aleš – RAMBOUSEK, Adam – RANGELOVA, Albená: Nové nástroje pro českou lexikografii – DEB2. In: Gramatika a korpus, F. Štícha, J. Šimandl (ed.), Praha: ÚJČ AV ČR 2007, s. 197–204.
- Příruční slovník jazyka českého. 9 sv. Praha: Státní nakladatelství – SPN 1935–1957. (PSJČ)
- SOCHOVÁ, Zdeňka – POŠTOLKOVÁ, Běla: Co v slovnících nenajdete. 1. vyd. Praha: Portál 1994. 204 s. (CSN)
- Slovník spisovné češtiny pro školu a veřejnost. Red. V. Mejstřík a kol. 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- SVĚTLÁ, Jindra: Čeština v elektronických médiích z pohledu lexikografie. In: Čeština doma i ve světě, 2006, roč. 14, č. 4, s. 115–121.
- Velký slovník cizích slov. Verze slovníku 2.0. Red. J. Kraus – V. Petráčková. Red. spolupracovaly S. Matoušková, V. Holubová. Voznice: Leda spol. s r. o. 2005. (CD-ROM). (VSCS)

K základnímu zpracování substantiv v lexikální databázi Pralex¹

Jindra Světlá, Barbora Procházková

Ústav pro jazyk český AV ČR, v. v. i.

On Basic Processing of Nouns in the Pralex Lexical Database

The article focuses on how we work with the Pralex lexical database while processing nominal entries, how we describe specific phenomena or a particular issue in an entry form, as well as which nominal data are included (mandatory items). To clarify this, the authors follow the order of particular parts of an entry form. Special attention is paid to problems of homonymy, variants, explanation of meanings and the structure of polysemantic entries.

1. Úvod

První fáze prací (2005–2006) byla v rámci stávajícího výzkumného záměru zaměřena především na návrh a vývoj naší vlastní lexikální databáze (dále LDB). Pro každý slovní druh byla v programu **Praled**² vymezena podoba **heslového formuláře** a bylo určeno, které údaje se v něm budou uvádět a jakým způsobem: některé se do databáze vkládají výběrem z pevné nabídky nebo pomocí zaškrtačovacího pole, jiné se zapisují ručně nebo se kopírují do textového pole z referenčních a materiálových zdrojů. Souběžně s tím jsme se zaměřili na budování hesláře³.

Ve druhé fázi prací (v roce 2007) začalo vlastní naplňování lexikální databáze, jejíž obsahová část byla nazvána **Pralex**, relevantními daty. Jako první bylo na podzim roku 2007 zahájeno tzv. **základní zpracování substantiv**. Nejdříve jsme vymezili obecné zásady pro zpracování tohoto slovního druhu v LDB Pralex, poté byla navržena podrobná struktura příkladové části heslového formuláře (Exemplifikace)⁴ a byla stanovena časová norma na jeden zpracovaný význam. Každý zpracovatel dostal přiděleno některé písmeno abecedy. V průběhu prací byly obecné zásady postupně dále zpřesňovány, aby se dosáhlo požadované jednotnosti a podrobnosti zpracování; na některé dílčí typy substantiv jsme se zaměřili podrobněji⁵.

V rámci základního zpracování substantiv se pohybujeme mezi úrovní celého hesla (na kartě Heslo) a úrovní jednotlivých významů (na kartě Význam). Dále pak postupujeme po jednotlivých položkách heslového formuláře⁶; na **kartě Heslo** se

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Podrobně k etapám vývoje, struktuře heslového formuláře a funkcím LDB Praled viz Světlá 2008.

3 K hesláři viz zde podrobněji stať J. Světlé *Budování hesláře LDB Pralex*.

4 K příkladové části hesla u substantiv viz zde stať M. Koutové.

5 K tomu viz zde stati E. Čonosové o zpracování zdrobnělin, H. Pernicové o zpracování apelativních názvů osob a jmen přechýlených, M. Habrové o zpracování antonym a H. Golášové o zpracování regionálních lexikálních jednotek.

6 K jednotlivým vyplňovaným položkám heslového formuláře viz zde i úvodní stať J. Světlé *Struktura a zpracování hesel v lexikální databázi Pralex*.

zaměřujeme zejména na tyto části (sekce)⁷: Záhloví, Varianty, Statistika a Gramatika; na **kartě Význam** pak na Výklad významu, Exemplifikaci a poznámkový aparát. Pro přehlednost zde při popisu jednotlivých sekcí postupujeme v pořadí podle heslového formuláře, i když se všemi jeho nástroji a položkami můžeme po vyplnění Záhloví v LDB Pralex pracovat v libovolném pořadí.

I když jsou jednotlivé složky popisu takto rozděleny do samostatných nástrojů, často narážíme na nejrůznější přechodné oblasti, protože při lexikografickém popisu vždy existuje určitá míra plynulých přechodů a souvislostí, které je třeba databázově zachytit a někam zaznamenat. Vše bylo třeba nejprve na základě analýz a zkušebních vzorků teoreticky popsat v zásadách pro zpracovatele a poté vyřešit i prakticky; k tomu jsme si postupně vybudovali specifický systém zápisů do jednotlivých položek a formalizovaných odkazů a poznámek pro různé případy, které při zpracování nastaly. Výchozí zásady byly tímto způsobem průběžně doplňovány a zpřesňovány.

2. Údaje vyplňované v Záhloví databázového hesla (na kartě Heslo)

Nástroj Záhloví obsahuje položky Lemma, Typ hesla a Slovní druh/typ, Homonymie a Zdroj pro heslář⁸.

2.1. Lemma u substantiv zapisujeme v reprezentativním tvaru, tj. většinou v nominativu singuláru, pouze ve specifických případech v jiném tvaru – např. v plurálu u jmen pomnožných (*akta, čtvrtohory, dionýsie, narozeniny* aj.). Při práci s lemmaty vycházíme nejprve z podob uvedených ve *Slovníku spisovného jazyka českého* (dále SSJČ). Výjimkou jsou variantní podoby lemmatu – ty uvádíme pro potřeby LDB Pralex jako samostatné databázové položky, aby bylo možné rozdělit všechny relevantní příklady podle formální podoby lemmatu a každou variantu pak samostatně vyhodnocovat na základě její dnešní doloženosti a stylové příznakovosti.

Zároveň s lemmatem u všech hesel v LDB Pralex určujeme **typ hesla** (výběr mezi hesly jednoslovnými a víceslovnými) a **slovní druh**. Pro práci s databází (z hlediska možností vyhledávání a třídění) považujeme za důležitější primárně označit u každého substantiva spíše slovní druh, než koncovku 2. pádu, jak bývá pro úsporu místa zvykem ve slovnících. Proto také gramatické charakteristiky zatím vyplňujeme jen v případě potřeby, tj. tehdy, když gramatický údaj představuje nějaký distinktivní rys v rámci stávajícího databázového popisu (viz též dále v bodu 5).

2.2. Homonymie u substantiv: všechny typy homonym (morfologická, slovnědruhovává, slovtvorná i lexikální) uvádíme jako samostatné položky databáze. Pokud v rámci substantiv v LDB Pralex z nějakého důvodu vyčleníme z původního zpracování podle SSJČ v rámci polysémie zvláště životné a neživotné významy jako homonyma, pak při

7 Z programového hlediska pracujeme s tzv. **nástroji**, protože jednotlivé části heslového formuláře se otevírají pomocí tlačítek jako dílčí okna s různým počtem polí pro zápis dat; těmto jednotlivým polím uvnitř nástrojů pracovně říkáme **položky formuláře**. Každá položka formuláře má svůj název, který dané pole jednoznačně identifikuje. Pokud v textu odkazujeme na tyto názvy položek, resp. názvy nástrojů, uvádíme je s počátečním velkým písmenem.

8 Pro zápis variant, statistik a gramatických charakteristik byly navrženy samostatné nástroje, aby se daly otevírat jen v případě potřeby (nevyplňují se vždy) – viz dále.

číslování těchto substantivních homonym vychází z logické motivace; pokud se jedná o paralelní motivaci (např. člověk a přístroj), přiřazujeme římskou číslici I k osobě (*čistič I*) a číslici II k věci (*čistič II* ve významu ‚stroj, přístroj, zařízení‘).

Při zpracování hesel v LDB Pralex mohou z hlediska homonymie nastat ve srovnání s hesly ze SSJČ, z dalších referenčních zdrojů a/nebo z korpusu následující případy:

1) Dvě homonyma (i více) byla už v SSJČ zpracována jako samostatná hesla; u těch pouze doplňujeme římskou číslici, pokud tam nebyla uvedena⁹.

2) Dvě hesla byla v SSJČ hodnocena jako homonyma, ale byla zpracována:

a) v tzv. sdružené heslové stati (jako heslo a podheslo oddělené pomlčkou), srov. např. *časoměřič* (živ.) – *časoměřič* (neživ.) v SSJČ:

časoměřič -e m. (živ.) (**časoměřička**, -y ž.) *pracovník měřící čas pracovních výkonů*: kurs pro úkoláře a č-e; č-ky a kontrolorky domácích prací; sport. *kdo měří stopkami čas sportovních výkonů*; --- **časoměřič**, -e m. (neživ.) tech. *hodiny (stopky) měřící čas práce strojové i ruční, přestávky i časové ztráty*.

b) v rámci hnízdování, které se týká i slovnědruhových homonym – např. substantiva *červená* a *červené* jsou v SSJČ uvedena u adjektiva *červený*:

červený příd. (...); zpodst. **červená**, -é ž. ob. krev: teče mu z nosu č.; **červené**, -ho s. červené víno; **červené**, -ých ž. mn. kart. jedna z karetních barev (s červenými srdci).

V LDB Pralex taková homonyma rozepisujeme zvlášť, jako samostatné položky databáze, aby bylo možné je podchytit v rámci celého souboru homonym a vyhledávat i jednotlivě, a také náležitě exemplifikovat v příkladové části hesla.

3) V SSJČ bylo zachyceno pouze jedno heslo, ale k dané formě najdeme další homonymum buď v ostatních referenčních zdrojích, nebo jej identifikujeme na základě analýzy materiálu z korpusu. Často se jedná o neologismy, srov. např.: *čajovník I*, kdo připravuje čaje a obsluhuje hosty v čajovně – *čajovník II*, keř n. stromek s kožovitými listy a bílými květy, z jehož mladých listů se připravuje čaj; rod *Thea* – *čajovník III*, čajovníkový strom; kajeput střídavolistý. Podobně *četař I*, poddůstojnická hodnost mezi desátníkem a rotným – *četař II*, účastník četu aj.

4) Někdy postupně dochází k rozpadu polysémie; v jednoznačných případech některé významy původně polysémního hesla vyčleňujeme zvlášť jako homonyma. Často však bývá problémem rozpoznat, zda už k úplnému rozpadu polysémie došlo, či nikoliv – pak jen do poznámky uvedeme návrh na rozdělení, aby se dalo heslo z tohoto aspektu později vyhodnotit.

Propojování homonym: na závěr zpracování hypertextově propojujeme homonymní dvojice (trojice atd.) mezi sebou pomocí seznamu; díky tomu máme při otevření kteréhokoli z nich konkrétní představu, kolik homonym je k dané formě v našem hesláři uvedeno (všechna jsou vždy v daném hesle zapsána na Seznam homonym), a můžeme s nimi pracovat zároveň – můžeme všechna otevřít najednou, nebo mezi nimi přecházet pomocí prokliku.

2.3. V poli Zdroj pro heslář pro potřeby naší další práce s heslem uvádíme informaci, na základě kterého zdroje (referenčního nebo materiálového) jsme dané lemma zařadili do hesláře LDB Pralex. V této položce registrujeme dva různé údaje, resp. jejich kombinaci:

⁹ V označování římskými číslicemi se často tištěná podoba SSJČ liší od elektronické verze, se kterou pracujeme v DEBDictu.

1) první výskyt daného lemmatu ve výkladových slovnících – všeobecných i speciálních – počínaje *Slovníkem spisovného jazyka českého* (dále SSJČ), přes *Slovník spisovné češtiny pro školu a veřejnost* (dále SSČ), *Nový akademický slovník cizích slov* (NASCS), *Co v slovnících nenajdete* (CSN), až po *Nová slova v češtině. Slovník neologizmů 1 a 2* (dále SN 1, SN 2); k ověřování zpracovávaných lexikálních jednotek v uvedených slovnících používáme jejich elektronickou podobu přístupnou v rámci slovníkového prohlížeče DEBDict (Dictionary Editor and Browser), případně na CD-ROM z nakladatelství LEDA¹⁰;

2) výskyt v materiálových zdrojích, se kterými pracujeme, tj. ve *Frekvenčním slovníku češtiny* (dále FSČ), resp. v korpusu SYN¹¹.

V první fázi prací jsme se zaměřili na zpracování těch substantivních hesel, která byla už v korpusu SYN2000 dostatečně doložena (čili na základě frekvence výskytu byla uvedena ve FSČ) a zároveň měla výklad v SSJČ; právě díky údaji označenému v poli Zdroj pro heslář (“fsc+ssjc”) s nimi můžeme pracovat jako s uceleným souborem a snadno je identifikujeme.

Hesla doplněná do hesláře LDB Pralex později (ve druhé fázi prací na substantivech) podle SSJČ a v korpusu doložená označujeme “ssjc+syn”; v korpusu nedoložená jsou označena jen “ssjc”. Od SSJČ jsme pak při doplňování hesláře a zpracování substantiv postupovali chronologicky dále. Pokud byla do hesláře doplněna další hesla z jiných slovníků (a v SSJČ registrována nebyla), odrazí se to na výběru příslušného údaje z nabídky v tomto poli. Přitom platí zásada, že pokud je dané lemma zároveň doloženo v korpusu, vybíráme kombinaci dvou údajů (fsc+slovník, resp. slovník+korpus SYN)¹²; u hesel v korpusu nedoložených uvádíme jen údaj o 1. výskytu ve výkladovém slovníku počínaje SSJČ.

Naopak hesla, u nichž je jako zdroj uvedeno jen “fsc”, nezaznamenává SSJČ ani žádný jiný slovník, ale mají poměrně vysokou frekvenci výskytu v korpusu SYN. Proto se u nich rovněž uvádějí statistické údaje, které máme k dispozici podle elektronické verze FSČ (EFES). Hesla označená v LDB Pralex jen zdrojem “syn” jsou hesla doplněná do hesláře LDB Pralex na základě korpusu SYN (a většinou doložená v malém počtu výskytů), která neregistruje žádný výkladový slovník. K takovým heslům (např. *čórkař*, *zloděj*, *čuplík*, *tlačítko* aj.) vytváří zpracovatel pracovní výkladovou definici sám, pokud mu to nalezené doklady umožní (viz též dále v bodě 7).

3. Variantnost u substantiv

Variantnost je jev, na který v LDB Pralex při zpracování narážíme velice často (u různých slovních druhů). Proto jim byla v průběhu zpracování substantiv věnována soustavná pozornost (teoreticky i prakticky)¹³.

10 V elektronické verzi vyšly v nakl. LEDA dva slovníky – SSČ a VSCS.

11 Od roku 2011 byl původní korpus SYN (který slučuje korpusy SYN2000, SYN2005 a SYN2006PUB) přejmenován na ORIG_SYN.

12 V roletovém menu s nabídkou zdrojů pak vybíráme z těchto kombinací: fsc+ssjc, fsc+ssc, fsc+vscs, fsc+csn, fsc+sn1,2; ssjc+syn, ssc+syn, vscs+syn, csn+syn, sn+syn, scfi+syn.

13 Na velkém objemu variantnosti registrované v LDB Pralex na základě SSJČ, dalších slovníků a dnešního stavu zachyceného v korpusu mají podíl jednak – u pravopisných variant – všechny

Podle našich zásad je zpracování všech variantních podob daného lemmatu v LDB Pralex samostatné (každá varianta tedy představuje samostatnou databázovou položku), ale zároveň paralelní, protože vždy jde o stejný význam/významy¹⁴. Protože jde o samostatná "hesla", všechny informace uváděné v rámci jedné varianty se musí vztahovat k dané podobě lemmatu¹⁵. To platí jak pro metajazyk výkladu, tak i pro doklady uváděné v Exemplifikaci.

Vzhledem k tomuto pravidlu v LDB Pralex někdy potřebujeme upravit metajazyk výkladu převzatý ze SSJČ, jelikož tento slovník (i všechny následující) uvádí pro obě (všechny) varianty společný výklad (v souladu s příslušnou pravopisnou normou). Týká se to např. hláskových variant u přípon cizího původu (srov. např. v SSJČ heslo *čechisace, čechizace* se společným výkladem „počešťování, počeštění, čechizování“ – v LDB upraveno pro samostatnou variantu *čechisace* u výkladového synonyma na „čechizování“).

Propojování variant: abychom v celé databázi signalizovali, že jednotlivé formální podoby patří k sobě, proto všechny ostatní varianty u zpracovávaných "hesel" (databázových položek) zapisujeme na seznam variant a pak všechny varianty hypertextově propojujeme mezi sebou. Zároveň k nim zakládáme zastřešující odkazové heslo. Tímto způsobem u každé variantní podoby lemmatu (a v souhrnném odkazovém hesle) na Seznamu variant vidíme všechny ostatní varianty registrované v LDB Pralex, a zároveň je také můžeme všechny najednou prohlížet, protože je můžeme současně otevírat jako okna a rozmísťovat na obrazovce počítače, jak potřebujeme. Svou grafickou podobou mohou některé varianty přesahovat do jiných písmen abecedy (*čingischán – džingischán; čip – chip I*; hesla začínající na *c-/k-* aj.). Protože každé písmeno dostal ke zpracování jiný autor, i proto je pro nás zpracovatelsky výhodné, když můžeme varianty snadno porovnávat mezi sebou a jejich zpracování kdykoli sjednocovat.

4. Frekvenční údaje u substantiv

Dále pracujeme s nástrojem nazvaným **Statistika**. Výchozí jsou zde statistické údaje FSC, které byly do LDB na počátku prací vloženy programátorem a které nám orientačně slouží jako první informace o tom, zda heslo v korpusu najdeme či nikoli a kolik dokladů bylo již v korpusu SYN2000¹⁶. S těmito údaji pak v některých případech pracujeme dále a komentujeme je podle potřeby.

změny pravopisu, které v češtině proběhly od počátku vydávání SSJČ až po úpravy provedené v posledním vydání PČP (1993) a které výkladové slovníky pochopitelně reflektují, jednak i fakt, že se v korpusu občas vyskytují i grafické podoby, které buď dnešní kodifikaci neodpovídají, nebo – např. u slovtvorných variant – s ní nesouvisí.

14 Pouze ojediněle se stává, že se u polysémního hesla variantnost netýká všech významů stejně.

15 Díky tomuto samostatnému databázovému zpracování různých variantních podob bude snazší vybrat pro slovníkové zpracování ty podoby, které jsou stále živé a frekventované, anebo kodifikované.

16 Statistické údaje byly pro FSC zpracovány na základě korpusu SYN2000: absolutní frekvence – *hodnota FRQ*, průměrná redukovaná frekvence – *hodnota ARF*. Spojením tří korpusů do korpusu SYN se tyto hodnoty mohly značně změnit, avšak potřebné údaje nemáme k dispozici.

Statistiku pochopitelně uvádíme pouze u hesel, která jsou ve FSČ uvedena¹⁷. V některých případech dochází ve FSČ ke společné lemmatizaci (buď dvou základních podob u variant, nebo různých významů u homonym); proto tyto údaje rozepisujeme k příslušným samostatně uvedeným podobám v LDB Pralex. Statistické údaje zůstávají u obou variant či homonym totožné, protože by bylo velice pracné a obtížné přesně spočítat, kolik dokladů přísluší každé podobě lemmatu. Společnou lemmatizaci komentujeme u variant v Pozn. ke statistice takto: *Stat. údaje platí pro +* daná hesla (např. *Stat. údaje platí pro čištění/čistění*); podobnou poznámku vyplňujeme rovněž u homonym: *Stat. údaje platí pro čáka I, II*.

U homonym, která zakládáme navíc na základě korpusu SYN, uvádíme stejné statistické údaje u obou lemmat pouze v případě, že i toto druhé (či další) homonymum je dostatečně doloženo (nad 20 výskytů se dá očekávat, že bylo doloženo už v korpusu SYN2000). Pokud je ale dokladů k dalšímu homonymu málo, statistické údaje podle FSČ u něj neuvádíme vůbec (protože je zřejmé, že se absolutní většinou vztahují jen k jednomu z nich) a místo toho do Celkové poznámky k Exemplifikaci uvedeme konkrétní údaj o počtu dokladů v korpusu SYN, protože ty pravděpodobně pocházejí až z dalších korpusů, které jsou v SYN zahrnuty.

5. Gramatické charakteristiky u substantiv

Sekci gramatických informací vyplňujeme zatím u substantiv jen v případě potřeby, později (v rámci slovníkového zpracování) bude pochopitelně třeba u každého hesla doplnit všechny informace, které se zde tradičně uvádějí – příslušná pole jsou připravena včetně výběru z nabídek, které popis usnadňují. Tento nástroj lze použít na kartě Heslo i Význam, podle toho, zda se údaje vztahují k celému heslu, nebo jen k některému významu (např. jméno pomnožné může být jen jedním z významů polysémního hesla, srov. *cukr – cukry*). Zde uvedeme jen několik příkladů, kdy s tímto nástrojem pracujeme.

Skloňování substantivního hesla je třeba vymežit již nyní např. u zpodstatnělých adjektiv (protože se skloňují jako příd. jm.), nebo pokud např. dané substantivum zůstává v nesklonné podobě, uvádíme u něj *neskl*.

Číslo určujeme např. u pluralií tantum, u nichž celé heslo nebo daný význam bude mít lemma v plurálu; u nich označíme (zvolíme z nabídky) *pouze mn. č.* (např. *čtvrtohory*). U substantiv s neúplně vyvinutou opozicí singulár – plurál, v jejichž paradigmatu převažují buď singulárové, nebo naopak plurálové tvary, rozhodujeme o tom, zda bude lemma v hesláři uvedeno v singuláru či plurálu; k tomu pak z nabídky zvolíme buď formulaci *prav. mn. č.* (např. *čachr*), nebo *prav. jedn. č., pouze jedn. č. atd.*

Rod uvádíme zatím pouze u jednoho typu homonym (pro rozlišení kategorie životnosti a neživotnosti), dále např. u zpodstatnělých adjektiv, u variant lišících se rodem a tvarem v singuláru (*atak/ataka*) nebo u substantiv s kolísáním v rodě (např. u hesla *rez – ž. r. i m. r.*).

17 U ostatních hesel, která nebyla vzhledem k nižší frekvenci do FSČ zahrnuta, v rámci Exemplifikace registrujeme alespoň případy malého počtu výskytů v SYN, což je údaj, který má pro naši další práci s heslem určitou výpovědní hodnotu.

V gramatické poznámce v relevantních případech komentujeme difference dnešního úzu proti stavu popsanému v SSJČ; tvrzení zde uvedené musí být doloženo na konkrétních příkladech v Exemplifikaci (srov. *image: V nepřímých pádech se příležitostně vyskytuje počeštěná psaná podoba, např. 3. p. imidži.*). U obouvidých verbálních substantiv do této poznámky zapisujeme komentář *dok. i nedok. vid* (např. u *definování*).

6. Členění na významy

Na počátku práce s hesly vycházíme z významové struktury hesel v SSJČ (každé heslo v LDB Pralex bylo tedy nejprve pracovně rozčleněno na jednotlivé významy a ke každému výkladu byla pracovně přiřazena výkladová definice podle tohoto slovníku). Tuto výchozí strukturu pak ověřujeme na materiálu z korpusu SYN a vyhodnocujeme dnešní stav. Další významy doplňujeme buď na základě ostatních referenčních zdrojů, s nimiž pracujeme, nebo je přidáváme na základě doloženosti v korpusu.

Provedené **změny ve významové struktuře** hesla oproti SSJČ, ze kterého vždy při zpracování substantiv primárně vycházíme, signalizujeme v **Poznámce k celému heslu**. Jednou z nejčastějších změn může být přidání dalšího významu; přidáme-li další význam oproti SSJČ s oporou o některý výkladový slovník, pak zde zapíšeme *nový význam podle + zkratka příslušného slovníku* (viz např. u hesla *čistírna – nový význam podle SSČ*).

Pro nové významy analyzované jen na základě korpusu použijeme stručnější formulaci *nový význam*, resp. *nové významy* (je-li jich více). V případě rozdělení původního polysémního hesla ze SSJČ na homonyma, nebo při rozdělení jednoho významu na dva v této poznámce uvádíme formulaci *nová struktura*.

7. Výklad významu

Jak již bylo uvedeno výše, u substantiv, která měla v SSJČ vlastní výklad, pracovně (jako výchozí) používáme výkladovou definici z tohoto slovníku, a to včetně tzv. předvýkladu (např. *v kapit., ve středověku, ve Francii, zejm. v soudní praxi aj.*), neuvádíme však původní stylové a oborové kvalifikátory, protože jsou často zastaralé a vše bude třeba znovu vyhodnotit na současném materiálu.

Pro kontrolu a usnadnění další práce s hesly (včetně revize) navíc do **Poznámky k výkladu** u prvního významu vkládáme kompletní heslovou stať ze SSJČ, popř. z dalších slovníků (včetně přihnízdovaných slov), aby zde byla vidět nejen původní struktura celého slovníkového hesla (gramatické informace, oborové a stylové charakteristiky, příkladová část, přihnízdovaná slova atd.), ale i různost zpracování v referenčních zdrojích. Je-li zpracovávané heslo polysémní, pak u dalších významů do této poznámky uvádíme ze slovníků jen údaje k tomuto konkrétnímu významu.

U polysémních hesel, která mají varianty, uvádíme v Poznámce k výkladu u prvního významu celé opěrné heslo ze SSJČ, tedy včetně všech variant uvedených v zá-

hlaví, zatímco u dalších významů pouze první lemma, pod kterým bylo heslo abecedně zařazeno v SSJČ – srov. např. *čokoláda* (ob. *čekoláda, čekuláda*).

Kromě toho navíc do **Poznámky autora** zaznamenáváme všechny neprovedené dílčí návrhy na úpravy definic či změny významové struktury hesla oproti SSJČ, komentáře pracovního charakteru týkající se významových odstínů, posunů ve významu, příp. nesrovnalostí z hlediska dnešního úzu apod.

8. Práce s výkladovou definicí ve vztahu k dokladům uvedeným v příkladové části hesla

Pochopitelně se snažíme řídit lexikografickou zásadou, že výklad musí odpovídat všem dokladům v rámci daného významu, a naopak, že doklady uvedené u příslušného významu mají odpovídat výkladové definici. Protože ale výkladové definice jako východisko přebíráme ze SSJČ, dochází k mnohým diferencím mezi stavem popisovaným v tomto slovníku a dnešním územ tak, jak nám jeho obraz ukazuje korpus SYN.

Při práci s heslem v LDB Pralex máme několik možností zásadních úprav oproti výchozímu členění hesla a výkladovým definicím uvedeným v SSJČ, resp. v dalších slovnících:

1) výklad rovnou upravíme; upravujeme však vždy až po přidání dalšího pole pro výklad významu, aby původní znění ze SSJČ zůstalo pracovní zachováno a aby byla v hesle vidět celá geneze prací na výkladové definici, dokud nebude definitivní (pro veřejnost pochopitelně tento postup zůstane skryt); do přidaného pole zapisujeme upravenou definici nebo další výklad z jiného referenčního zdroje, který více odpovídá korpusovému materiálu, nebo naformulujeme vlastní pracovní návrh definice;

2) pokud nechceme výklad upravit hned, nebo nemáme oporu v žádném jiném slovníku, nebo nevíme, jak výkladovou definici naformulovat, pouze v Poznámce autora signalizujeme, že je třeba výklad změnit, nebo do pole pro výklad zapíšeme otazník (aby bylo vidět, že se s heslem pracovalo a že výklad chybí nebo nesedí);

3) pokud v rámci jednoho významu SSJČ uváděl i významové odstíny nebo terminologické užití, mohou nastat tři případy: buď dnes daný významový odstín v korpusu není doložen (pak se k dnešnímu stavu zatím nevyjadřujeme), nebo je dnes doložen jen několika doklady (pak je ponecháváme zde), nebo je naopak doložen hojně, a pak vyhodnocujeme, zda došlo k jeho lexikalizaci¹⁸;

4) pokud se dnes významový odstín uzualizuje a lexikalizuje, vyčleňujeme jej jako další samostatný význam;

5) další význam (v SSJČ neuvedený) někdy přidáváme s oporou o některý jiný výkladový slovník, ale jen v případě, že byl doložen v korpusu;

6) pokud část dokladů z korpusu ukazuje na další význam dříve neregistrovaný žádným slovníkem, také jej přidáváme do polysémnní struktury hesla – vždy jako poslední¹⁹; v takovém případě naformulujeme vlastní návrh výkladu;

18 K tomu viz též podrobněji zde ve stati M. Koutové.

19 Pořadím významů v polysémnním hesle se zatím nezabýváme – bude se souhrnně vyhodnocovat až v rámci nového slovníkového zpracování (na základě práce s dalším materiálem).

7) pokud najdeme v korpusu doklady, které neodpovídají žádnému dříve registrovanému významu, ale není jich dost pro vymezení dalšího samostatného významu, zatím je bez výkladu registrujeme v poli **Doklady pro nové významy** (jde vlastně o speciální blok Exemplifikace vyčleněný zvlášť).

Analogicky postupujeme u hesel, která byla ve slovnících jen registrována bez výkladu.

9. Příkladová část hesla (Exemplifikace)

Exemplifikace je u substantiv rozčleněna do jednotlivých bloků, do nichž zapisujeme vhodné příklady na syntaktickou a sémantickou spojitelnost heslových slov. Pokud žádné doklady nenajdeme, příslušný blok zůstane prázdný.

U hesel s malou doložeností v korpusu SYN počet výskytů zaznamenáváme v **Celkové poznámce k Exemplifikaci** (např. *čachrářství: v syn 7x*).

Pokud k některému významu hesla nenajdeme v korpusu vůbec žádné doklady, zaškrtneme políčko *Nepodařilo se doložit v syn*; pak se celá Exemplifikace k danému významu automaticky sbalí, aby nezabírala místo.

Nedoložené může být buď celé heslo, nebo jen jeho část (některý význam); nedoloženost se tedy týká jak hesel monosémních (*čertíroh*), tak i polysémních (*čunče*). U hesel ze SSJČ, u nichž už nepředpokládáme možnost dohledání dokladů v jiných materiálových zdrojích, zapisujeme k lemmatu rovnou nulu.

Literatura

- Akademický slovník cizích slov. (V. Petráčková, J. Kraus a kol.) 1. vyd. Praha: Academia 1995, 1. díl A–K, 2. díl L–Ž, 834 s.; 1998 A–Ž, 834 s. (ASCS)
- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSČ)
- Nová slova v češtině 1. Slovník neologizmů. Red. O. Martinová a kol. Praha: Academia 1998. 356 s. (SN 1)
- Nová slova v češtině 2. Slovník neologizmů. Red. O. Martinová a kol. Praha: Academia 2004. 568 s. (SN 2)
- Nový akademický slovník cizích slov A–Ž. Red. J. Kraus a kol. 1. vyd. Praha: Academia 2005. 880 s. (NASCS)
- Slovník spisovné češtiny pro školu a veřejnost. Red. V. Mejstřík a kol. 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovné češtiny pro školu a veřejnost. Red. V. Mejstřík a kol. Voznice: Leda spol. s r. o. 2005. (CD-ROM). (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- SOCHOVÁ, Zdeňka – POŠTOLKOVÁ, Běla: Co v slovnících nenajdete. 1. vyd. Praha: Portál 1994. 204 s. (CSN)

SVĚTLÁ, Jindra: K návrhu, vývoji a funkcím lexikální databáze češtiny. In: Lexikografie v kontextu informační společnosti. Sborník příspěvků z pracovního setkání Praha, 4.–6. 6. 2007. Eds. A. Rangelova – J. Světlá – A. Jarošová. Praha: Ústav pro jazyk český AV ČR, v. v. i., 2008, s. 19–32.

Velký slovník cizích slov. Verze slovníku 2.0. Red. J. Kraus – V. Petráčková. Red. spolupracovaly S. Matoušková, V. Holubová. Voznice: Leda spol. s r. o. 2005. (CD-ROM). (VSCS)

Exemplifikace substantiv v lexikální databázi Pralex (zpracování příkladové části hesla)¹

Marta Koutová

Ústav pro jazyk český AV ČR, v. v. i.

Exemplification of Nouns in the Pralex Lexical Database (Processing of the Example Part of an Entry)

The core of the processing of individual items in the Pralex lexical database is an example part of an entry (exemplification). While processing the nouns, we try to select a relevant evidence from the corpus SYN, i.e. examples of a concrete use of meanings processed in advance according to the *Dictionary of the Standard Czech Language* and, in addition to that, we also try to record evidence of the new meanings. In the exemplification, we focus on a syntactic and semantic linkage. Lemmas are mainly illustrated by means of non-quotational, modified evidences, i.e. example linkage in the form of syntagmas, which are regarded as common collocations of the processed nouns. Sentential records are introduced as well, though in a lower extent.

1. Úvod

Tzv. základní zpracování² všech položek („hesel“) v lexikální databázi (dále jen LDB) Pralex je – vedle uvádění základních údajů platných pro celé heslo a členění na významy – zaměřeno především na příkladovou část hesla. Na rozdíl od *Příručního slovníku jazyka českého* (dále PSJČ), *Slovníku spisovného jazyka českého* (dále SSJČ) a *Slovníku spisovné češtiny pro školu a veřejnost* (dále SSČ), které využívaly lístkový lexikální archiv, se pro naši databázi stal základním materiálovým východiskem Český národní korpus, konkrétně korpus SYN³. Máme tak k dispozici jedinečný zdroj současných psaných textů z oblasti publicistiky, beletrie i odborné literatury. Při zpracování jednotlivých substantiv nám jde o výběr relevantních dokladů, tj. o výběr příkladů na konkrétní užití k předem založeným významům podle SSJČ i o zachycení dokladů na významy nové.

2. Práce s WSE a Bonitem

U substantiv se zaměřujeme na syntaktickou i sémantickou spojitelnost. Oba typy spojitelnosti ilustrujeme primárně pomocí necitátových, tzv. upravených dokladů, tj. příkladových spojení ve formě syntagmat, která představují běžné kolokace zpracovávaných substantiv. V menší míře uvádíme i větné (citátové) doklady. Pro elektronickou lexikální databázi je charakteristický dostatek místa, proto se snaží-

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Viz podrobněji v úvodní stati J. Světlé *Struktura a zpracování hesel v lexikální databázi Pralex*.

3 Od roku 2011 byl původní korpus SYN (který slučuje korpusy SYN2000, SYN2005 a SYN-2006PUB) přejmenován na orig_syn.

me ukázat celou šíři spojitelnosti heslových slov. Je-li korpusových dokladů příliš, vybíráme nejčastější vazebné skupiny, které vystihují spojitelnost substantiva a dokreslují užití heslového slova v textu.

Výběr dokladů je v kompetenci každého zpracovatele. Vodítkem pro výběr je frekvence v korpusu, ale nezastupitelnou roli hraje samozřejmě jazykové povědomí zpracovatelů (a to i v případě výrazů pro autora neznámých, např. odborných pojmů, u kterých musí být zpracovatel rovněž schopen na základě kontextu vybrat adekvátní spojení a rozlišit je od spojení náhodných, neuzuálních).

Zatím pracujeme jen s korpusem SYN, teprve při práci na budoucím výkladovém slovníku budeme dohledávat doklady i z jiných zdrojů. Pro třídění dokladů z korpusu využíváme nástroje Word Sketch Engine (dále WSE) a Bonito. Bonito je korpusový manažer, v němž lze korpus prohledávat. Vyhledávat můžeme buď konkrétní slovní tvar, nebo lemma, nebo v případě potřeby morfologickou značku. WSE je webové rozhraní propojené s Bonitem, dostupné na <http://ucnk.ff.cuni.cz/corpora>. Slouží k zobrazení spojitelnosti daného slova na základě gramatických charakteristik, jimiž je korpus SYN doplněn. Výsledkem po zadání příslušného lemmatu je seznam kontextových slov, která se k danému lemmatu vážou, rozříděných do tabulek podle jednotlivých gramatických kategorií. Pro substantiva jsou to např. tabulky s adjektivními přívlaskty, genitivními, předložkovými a slovesnými vazbami atd. Zobrazení WSE má hierarchickou strukturu, a tak se v příslušné tabulce ukáže jen část syntagmatu (bez příslušného lemmatu). Je možné a často i nutné ověřovat zobrazené výsledky na delším kontextu, který je schován o úroveň níž (každý kontext se musí rozkliknout). Spojení, která vygeneruje WSE, je nutno vyhodnocovat s určitou dávkou obezřetnosti, protože se někdy jedná o chybná spojení – náhodné pozice slov, které tvoří lexikální syntagma, nebo o chybnou lemmatizaci.⁴

WSE používáme pro výběr a podrobné třídění minimálních kontextů k jednotlivým významům. Z Bonita vybíráme tzv. citátové doklady – v průměru 3–5 větných příkladů s uvedením zdroje.⁵ Bonito slouží též pro vyhledávání dokladů na substantiva s malým počtem výskytů, protože ta často nejsou v korpusu lemmatizována, a tudíž je nezachycuje ani WSE.

Při výběru příkladů do jednotlivých bloků si zatím příliš nevšímáme morfologické variantnosti, protože WSE toto v tabulce neukazuje. Pokud na takové příklady narazíme, uvádíme zpravidla i variantní možnosti, ale cíleně je nevyhledáváme, srov. např. u hesla **atlas**: *mapy v atlase, mapy v atlasu*. Případné kolísání v rodě zachycujeme v rámci jednoho hesla, doklady na jednotlivé rody nadepisujeme příslušným titulkem a oddělujeme volnou řádkou, srov. např. u hesla **akné**:

stř.

mírné akné, těžké akné, silné akné;

viditelné akné;

zanícené akné;

4 Např. u lemmatu *kvádru* tabulka „prep v“ zdánlivě zachycuje spojení *muž v kvádru*. Po rozkliknutí však najdeme tuto větu: *Tělo muže v kvádru betonu našli stavební dělníci při výkopových pracích*.

5 U každého větného dokladu uvádíme informace o zdroji v této podobě: opus.rokvyd (rok vydání), opus.txttype (typ textu, tj. např. román, publicistický text, text vědeckonaučné literatury apod.), opus.id (identifikační kód zdrojového textu).

žen.
těžká akné;
cystická akné

3. Práce s korpusovými doklady

Na začátku zpracování substantiv (2007–2009) jsme se zaměřili (vzhledem k napjaté časové normě vymezené na zpracování jednoho významu) na to, abychom do exemplifikace vybrali pouze spojení zcela typická (v rozsahu jen o něco větším, než je obvyklé ve výkladovém slovníku), později, tj. od roku 2010 – v rámci revize základního zpracování – jsme zvýšili podrobnost zpracování příkladové části hesla tak, abychom dostali širší obraz sémantické spojitelnosti daného lexému, a proto vybíráme spojení běžná, v úzu hojně zastoupená, i příklady na přenesené užití. V rámci každého exemplifikačního bloku (ke struktuře exemplifikace u substantiv viz dále v bodu 4) postupujeme částečně na základě frekvence, částečně na základě sémantiky tak, aby byly vedle sebe uvedeny kolokace z jedné lexikálněsémantické skupiny. Pokud nenajdeme žádné relevantní doklady, příslušný blok zůstane prázdný.

V počáteční fázi prací jsme se při výběru tzv. upravených dokladů snažili o vyšší stupeň abstrakce a zapisovali jsme do exemplifikace pouze minimální syntagmata v příslušném tvaru, postupem času jsme však zjistili, že je v některých případech nutné uvádět rozvitější doklady, aby příklad dával smysl, popř. aby se přesně vymezilo, ke kterému významu daný příklad patří, srov. např.:

autodílna: *autodílna v Mariánské ulici* (nikoli pouze „autodílna v ulici“);

houslista: *houslista orchestru Národního divadla* (nikoli pouze „houslista orchestru“);

kachna: *bránit se tiskovým kachnám* (doplnění adjektivního přívlastku ukazuje, že dané spojení nepatří k V1⁶, tj. „vodní pták s nohama opatřenými plovací blánou a se zploštělým zobákem“, ale k V3, tj. „nepravdivá, vymyšlená, senzační zpráva“).

V některých případech je také nutné uvést delší kontext proto, aby odpovídal charakteristikám denotátu:

ozón: *přízemní ozón škodí lidskému zdraví* (zavádějící by bylo pouhé „ozón škodí lidskému zdraví“).

Někdy sice není nezbytně nutné rozvíjet uvádět, avšak pokud vhodně dokresluje dané spojení a v úzu se běžně vyskytuje, můžeme uvést i rozvitější doklady tohoto typu:

kontrola: *jít na kontrolu k lékaři*;

auto: *slabost pro rychlá auta; střet s protijedoucím autem*

Vzhledem k tomu, že LDB Pralex je deskriptivní, snažíme se doklady nalezené v korpusu uvádět v autentické podobě a příliš je neupravovat⁷. To znamená, že např. nepřevádíme negativní tvary sloves a adjektiv na pozitivní, minulý čas na přítomný, trpný rod na činný, plurál na singulár apod. Jedinou abstrakcí na této úrovni hesla je

6 Jednotlivé významy polysémních lexémů označujeme zkratkami V1, V2, V3 a tak dále podle pořadí.

7 Pracovní název „upravené doklady“ znamená hlavně to, že z větných kontextů vybíráme jen minimální syntagma nebo krátký kontext dokládající spojitelnost heslového slova.

převedení finitních tvarů na infinitiv, popř. nutné syntaktické doplnění pomocí neurčitého zájmena. Je však přirozené, že pokud korpus zachycuje jak doklady na bezpříznakové, tak i na příznakové užití, dáváme při výběru z korpusu přednost dokladům v reprezentativním tvaru. Většinou tedy nevybíráme totožné doklady na singulár i plurál, spojení zpracovávaného substantiva s jiným substantivem v mužském i ženském rodě apod., ale vybereme pouze doklad běžnější, zpravidla ten bezpříznakový, srov. např. u hesla **automobil**: *majitel automobilu, řidič automobilu* (není uvedeno i *majitelka automobilu, řidička automobilu*). Pokud však výskyty v korpusu ukazují na typičnost spojení s příznakovým tvarem⁸, zapisujeme do příslušného bloku tuto podobu (výjimečně uvádíme oba tvary, tj. příznakový i bezpříznakový, spíše však vybíráme jen ten běžnější z nich, aby databáze nebyla zahlcena množstvím dubletních dokladů), srov. např. **obyvatel**: *místní obyvatelé, zdejší obyvatelé, tamní obyvatelé* (spojení *místní obyvatel* atd. je sice také doloženo, ale v mnohem menší míře, běžnější jsou zde plurálové tvary, proto v tomto případě singulárové tvary nebyly uvedeny), ale např. *vysoký podpatek* i *vysoké podpatky* – jako relevantní byly vyhodnoceny, a proto uvedeny obě podoby.

Po pravopisné stránce upravujeme doklady nalezené pomocí WSE jen v případech, že se jedná vysloveně o chybu. Chyby a překlepy, které se vyskytnou v citátových dokladech z Bonita, však neopravujeme, stejně tak mezeru před tečkou a dalšími grafickými znaky užívanými v korpusových textech.

4. Práce s významovými odstíny a členění na významy

Vzhledem k tomu, že významovou strukturu hesla zatím pracovně přebíráme ze SSJČ, a k tomu na základě dalších referenčních zdrojů a korpusu SYN registrujeme další významy, musíme se občas vypořádávat s problémem významových odstínů a s přenášením významu.

U některých substantiv v SSJČ výklad uvedený v rámci jednoho významu naznačuje po středníku jeden i více sémantických odstínů. Tyto sémantické odstíny zachycujeme zatím většinou v souladu se SSJČ, tj. rovněž v rámci jednoho významu; tam, kde to jde, uvádíme doklady k nim souhrnně na konci příslušného bloku. Týká se to např. substantiv typu **kachna, králík, kvěťák**, která označují též jídlo, dále hesel, která mají též terminologický význam, např. **mrak** (*radioaktivní mrak, toxický mrak; hřibovitý mrak*) apod. Další dílčí případy jsou např.: **auto** – označuje též hračku (srov. př. *auto na dálkové ovládání*), **sklenice** – označuje nádobu a metonymickým přenesením i její obsah (*vypít celou sklenici*), **autor** – označuje osobu a metonymicky i její dílo (*číst německé autory, vystavovat mladé autory*).

Tzv. dovýklad

Zatím nepracujeme s dovýkladem tak, jak bývá zvykem v rámci slovníkového zpracování, ale jen na úrovni exemplifikace. Pokud potřebujeme v nějakém exemplifikačním bloku v odůvodněných případech uvést dovýklad k některému syntagmatu, aby bylo jasné, oč jde (většinou jde o doplňující vysvětlení vyplývající až z širšího kontextu),

8 Toto je další důvod, proč je nutné rozkliknutím příslušného výrazu ve WSE ověřovat, zda je tvar daného spojení nabízený tabulkou reálný. WSE totiž v tabulce uvádí pouze základní tvar daného výrazu (u substantiv 1. pád singuláru, u sloves infinitiv v pozitivním tvaru apod.).

uvádíme jej kurzivou v závorce za příkladem. Často se to týká spojení terminologického charakteru, srov. např. heslo **kachna**: *přinášení kachny, dohledávka kachny (disciplíny při zkouškách loveckých psů)*. Jindy jde o přenesené užití obrazného charakteru, např.: **podpatek**: *podpatek na italské botě (Apulie)*; **krasavec**: *čtyřkolový krasavec, nablýskaný krasavec, plechový krasavec (o autech)*; **společnice**: *společnice jeho meditací (o dýmce)*.

Upřesnění ale není vždy nutné, obraznost některých spojení je jasná i bez dovýkladu, srov. např. substantivum **kometa**: *kometa extraligy, zářivá kometa zápasu*.

Lexikalizace přeneseného užití

Pokud je nějaké spojení v přeneseném užití, které bylo registrováno již v SSJČ, v korpusu SYN hodně doloženo, a to v různých kontextech, často je to pro nás signálem, že se ustaluje další samostatný význam. Srov. např. heslo **buňka** – z přeneseného užití uvedeného v SSJČ u 2. významu („malá místnost, jedna z mnoha stejných“: *obytné buňky*) se ustálil nový význam V5 „malá místnost, jedna z mnoha stejných; dočasná stavba menších rozměrů sloužící konkrétním účelům“.

U slov cizího původu bývá proces lexikalizace často zachycen už v *Novém akademickém slovníku cizích slov* (NASCS – el. verze implementovaná do DEBDictu je označena jen jako SCS), o který se v takovém případě můžeme opřít. Srov. např. u hesla **analfabet** (SSJČ uvádí v rámci jednoho významu „člověk neznalý čtení a psaní, negramotný; přen. hanl. nevzdělanec, zaostalec“: *být a. v dějepise; hudební a.*; SCS již má rozděleno na dva významy: 1. „člověk neznalý čtení a psaní, negramotný“; 2. hanl. „zaostalý, nevzdělaný, omezený, hloupý člověk vůbec n. v daném oboru“).

Další význam může být vymezen i jinými způsoby, např. specifikací v různých komunikačních oblastech či oborech. Např. u substantiva **delegát** jsme zachytili dva nové významy: 1. „pracovník cestovní kanceláře pečující o rekreanty v místě jejich (zahraničního) pobytu“ (*delegát cestovní kanceláře, delegát Čedoku*), 2. „osoba vyslaná fotbalovým svazem a pověřená dohledem nad regulérností hodnocení utkání“ (*fotbalový delegát; svazový delegát; delegát zápasu, delegát utkání; delegát svazu*). V SCS ani v dalších novějších slovnících tyto významy ještě zachyceny nejsou.

Polysémní hesla členíme na významy primárně ve shodě se SSJČ. Zpracování v SSJČ porovnáváme se zpracováním v dalších slovnících (SSČ, SCS), a pokud uvádějí členění jiné, popř. zachycují oproti SSJČ další významy, v relevantních případech přehodnocujeme významovou strukturu zpracovávaného hesla na základě těchto novějších slovníků a doloženosti v korpusu.

Doklady pro nové významy často poskytuje i korpus SYN (viz např. již zmiňované substantivum *delegát*). V takovém případě nový význam nejprve ověřujeme ve slovníku-suplementu J. Sochové a B. Poštolkové *Co v slovnících nenajdete* (CSN) a dále ve slovnících neologismů (*Nová slova v češtině 1 a 2*), které nejnovější sémantické posuny některých slov již zachytily. Pokud daný význam v žádném slovníku dosud registrován není, na základě dokladů z korpusu ho zpracujeme sami, včetně návrhu pracovní definice (když se nám ve vymezeném čase nepodaří naformulovat vhodnou výkladovou definici, zapíšeme do pole pro výklad otazník). Pokud máme jen několik dokladů, které signalizují významové posuny oproti stavu registrovanému v předchozích slovnících, uvádíme je v poli „Doklady pro nové významy“, aby bylo možné později vyhodnotit, zda jde opravdu o další význam.

5. Struktura exemplifikace u substantiv

Příkladovou část substantivního hesla rozdělujeme do 14, resp. 15 exemplifikačních bloků. V členění příkladů do jednotlivých bloků se projevuje syntaktická spojitelnost daného substantiva⁹, uvnitř jednotlivých bloků je zachycena spojitelnost sémantická.

Do exemplifikačních bloků E1–E13 uvádíme tzv. upravené doklady, tj. minimální nebo někdy i rozvitější syntagmata, která získáváme především pomocí WSE (viz v bodu 1). Pro větné doklady slouží blok E14. Fakultativní blok E15 slouží pro doklady na lemma užité v názvech a přidáváme ho jen v případě potřeby.

Vnitřní členění v rámci jednoho exemplifikačního bloku

V každém exemplifikačním bloku uvádíme nejprve doklady na běžné, nepříznačkové užití. Pokud v korpusu najdeme i doklady na užití odborné, přenesené apod., uvádíme je až po „neutrálních“ dokladech, oddělujeme je volnou řádkou, pokud je jich více, a nadepisujeme příslušným titulkem.

Titulkem „odb.“ označujeme doklady na odborné užití. Srov. např. u hesla **mezera**:

vzduchová mezera;

větrací mezera;

volná mezera, zející mezera;

úzká mezera, široká mezera;

dvoucentimetrová mezera;

odb.

dilatační mezera

Ve výjimečných případech používáme i konkrétnější označení dané oblasti, např. „sport.“ (srov. heslo **celek** ve významu V2 „kolektiv, společnost“: sport. *prvoligový celek; kapitán celku; nejlepší celek v tabulce* atd.). U titulku „odb.“ nejde o kvalifikátor, v žádném případě tím nesignalizujeme, že jde o ověřené termíny, pouze dáváme signál o odborném užití (jde spíše o vymezení sféry, oblasti užití). Je rozdíl mezi hesly či významy výhradně terminologickými, kde by daný titulky musel být na začátku každého bloku, a proto ho neuvádíme nikde, a hesly z běžné slovní zásoby, kde se část dokladů týká odborného užití – ty pak takto vyčleňujeme, abychom si později usnadnili jejich další vyhodnocování.

Pod titulky „přen.“ registrujeme doklady na přenesené užití. Srov. např.:

četník: přen. *četník světa; role četníka Balkánu; Spojené státy nebudou na Haiti dělat četníka;*

kometa: přen. *zářívá kometa zápasu; vyhoupnout se mezi komety letošního roku.*

Někdy v exemplifikaci uvádíme i přenesené užití neustálené, příležitostné, např. metafory typu: *katedrála konzumu, katedrály socialismu.*

Zcela okazionální případy (jednorázové, individuální autorské metafory) ale v upravených dokladech neuvádíme vůbec, mohou však být registrovány ve větných dokladech jako příklad na příležitostné užití daného slova. Srov. např.:

9 Původní model exemplifikace substantivních hesel připravil doc. RNDr. Karel Oliva, Dr.

*Atkins začal vraždit v Detroitu. Jako **kometa děsu a krve** stačil mezi lety 1991 a 1992 brutálně povraždit jedenáct žen, většinou prostitutek a narkomanek.*

Někdy může být spojení užíváno v přímém i přeneseném významu. V takových případech dáváme za podobný doklad do závorky komentář „(i přen.)“, srov. např. u hesla **buben**: *válečné bubny* (i přen.).

Podobně pro frazémy uvádíme „(i fraz.)“ a pro sousloví „(i sousl.)“, srov. např. hesla **podpatek**: *srazit podpatky* (i fraz.), *sklapnout podpatky* (i fraz.), *scvaknout podpatky* (i fraz.); **kůň**: *bílý kůň* (i fraz.), *černý kůň* (i fraz.); **skříňka**: *černá skříňka* (i sousl.).

6. Specifika zpracování jednotlivých bloků

Blok E1 (Adj+SUBST)

První blok zachycuje spojení zpracovávaného substantiva s anteponovaným adjektivním přívlastkem. V tomto bloku zachycujeme odděleně doklady na singulárové a plurálové užití zpracovávaného substantiva podle toho, co u daných spojení převažuje. Srov. heslo **kopec**:

*lesnatý kopec, travnatý kopec, skalnatý kopec, kamenitý kopec;
zalesněný kopec, zasněžený kopec;
hustě porostlý kopec, holý kopec;
vysoký kopec, nízký kopec;
strmý kopec, příkrý kopec;
kuželovitý kopec, homolovitý kopec;
čedičový kopec, vápencový kopec;
protější kopec;
beskydské kopce, šumavské kopce;
okolní kopce, daleké kopce;
pohraniční kopce*

Kromě minimálních syntagmat zde v odůvodněných případech uvádíme i rozvířenější příklady (viz též bod 2). Může se jednat např. o rozvíření příslovcem, srov. *tázavě zdvižené obočí, dokonale vytvarované obočí*. Uvádíme také postupně rozvíjející přívlastky, např. *přední český houslista, špičkový soudobý houslista*.

Nespisovným koncovkám adjektiv v adjektivně-substantivních spojeních se snažíme vyhýbat, ale je-li to vzhledem k podobě lemmatu (např. *brejle, auták*) běžné a je-li takových spojení v korpusu doloženo hodně, registrujeme i podoby s nespisovnou koncovkou, např.: *slepecký brejle; tlustý brejle, silný brejle; černý brejle, tmavý brejle; babičkovský brejle*.

Blok E2 (SUBST+Adj)

Do tohoto bloku dáváme výhradně odborná spojení zpracovávaného substantiva s adjektivem v postpozici typu *konvalinka vonná, kyselina sírová, atribut verbální*.

Nepatří sem spojení zpracovávaného substantiva s postponovaným přívlastkem vyplývající z aktuálního členění věty, jako např. u hesla **ozon**: *troposférický ozon má vůči UV-záření stejné chování jako ozon stratosférický*.

Nedáváme sem ani odborná spojení SUBST+Subst typu *hřib satan*, ta buď zpracováváme samostatně jako víceslovné lexikální jednotky, pokud potřebují výklad, nebo je uvádíme u substantiv v bloku E13 – „Jiná spojitelnost“.

Blok E3 (SUBST+Subst-gen) a blok E4 (Subst+SUBST-gen)

Do bloku E3 vybíráme příklady na spojení daného substantiva s jiným substantivem v genitivu, v bloku E4 je zpracovávané substantivum samo genitivním přívlástkem. V těchto ani v následujících blocích již na rozdíl od bloku E1 doklady na singulár a plurál neuvádíme odděleně. Srov. např. heslo **auto**:

E3: *auto značky Fiat*;
auta policie, auta taxislužby, auta záchranářů;
auta rekreatů, auta rezidentů;
auto podnikatele;
auto snů, auto budoucnosti

E4: *majitel auta, řidič auta, osádka auta*;
výrobce aut, prodejce aut, dovozce aut;
vykradači aut, zloději aut;
prodej aut, dovoz aut, koupě auta;
parkování aut;
pojištění auta;
srážka aut, oprava auta;
odtah auta, odtažení auta, odstavení auta;
krádež auta, krádeže aut, vykrádání aut, pašování aut přes hranici;
půjčovna aut, myčka aut;
karoserie auta, vnitřek auta;
volant auta, kufr auta, kapota auta, korba auta, střecha auta;
výfuky aut, reflektory aut, klaksony aut;
značka auta, SPZ auta, MPZ auta;
kolona aut, fronta aut, šňůra aut;
vrak auta

Při zápisu dokladů v bloku E3 rovněž dbáme na oddělení objektových a subjektivých genitivů – je to zpravidla první kritérium sémantického třídění: *autocenzura textů* x *autocenzura redaktorů*; *makrofotografie hmyzu* x *makrofotografie Jiřího Štreita* apod.

Blok E5 (SUBST+Prep+Subst / SUBST+Subst-jiný pád) a blok E6 (Subst+Prep+SUBST / Subst+SUBST-jiný pád)

Blok E5 zachycuje předložkové přívlastky ke zpracovávanému substantivu. Do tohoto bloku patří rovněž doklady na spojení zpracovávaného substantiva se substantivem v jiném pádě, srov. např. u hesla **dárek**: *dárek ženě, dárky dětem*. Do bloku E6 dáváme spojení, kde se zpracovávané substantivum vyskytuje jako předložkový přívlástek k jiným substantivům nebo se s jinými substantivy pojí jako přívlástek dativní či instrumentální (např. **arterie**: *průřez arterií*).

Vazby se stejnou předložkou a stejným pádem řadíme k sobě, převažuje zde tedy formální kritérium nad sémantickým. Srov. např. blok E6 u hesla **autorita 2**:

*úcta k autoritám, odpor k autoritám, nechť k autoritám;
vzpoura proti autoritám, odpor proti autoritám;
potíže s autoritami, problémy s autoritami, boj s autoritami;
napětí mezi autoritami morálními a politickými;
strach z autorit;
závislost na autoritách;
odvolání na autoritu;
víra v autority, důvěra v autority;
respekt před autoritami, strach před autoritami;
poslušnost vůči autoritám, respekt vůči autoritám, odpor vůči autoritám*

Pokud se však u některých dokladů předložka pojí se stejným pádem, ale sémantika je odlišná, pak rozdělujeme dané doklady na více řádků. Srov. např. u hesla **automat**:

*automat na nápoje, automat na kávu, automat na jízdenky;
automat na mince, automat na karty*

Vazby s různými předložkami uvádíme zvlášť, i když jsou příklady sémanticky příbuzné, srov.:

*automat na výdej jízdenek;
automat pro výdej jízdenek*

Rozkliknutím dokladu ve WSE ověřujeme, zda předložka v daném dokladu nepatří k předložkové vazbě slovesa, a tudíž do některého ze slovesných bloků, např.:

arterie: **arterie ve zdravém stavu* – patří do E8: *udržovat arterie ve zdravém stavu*.

Pořadí předložek není pevně určeno, postupujeme zpravidla podle pořadí, které nám předkládá WSE.

Blok E7 (SUBST-nom+Verbum finitum) a blok E8 (Infinitiv (resp. VF)+ SUBST-acc)

V bloku E7 je zpracovávané substantivum subjektem slovesa, v bloku E8 je dané substantivum objektem slovesa. Srov. např. heslo **autorita 1**:

*E7: chybí mu autorita;
matčina autorita selhala;
něčí autorita vzrostla, něčí autorita klesá;
z jeho očí vyzařovala autorita*

*E8: získat autoritu, vydobýt si autoritu, vybudovat si autoritu, udržet si autoritu, zachovat si autoritu;
mít autoritu, vzbuzovat autoritu, vyzařovat autoritu;
zjednat si náležitou autoritu, upevnit svou autoritu, vynucovat si autoritu, uplatňovat autoritu;*

*ztratit autoritu, postrádat autoritu, obnovit autoritu;
respektovat něčí autoritu, uznávat něčí autoritu, akceptovat autoritu;
nesnášet autoritu;
propůjčit někomu autoritu, dodávat někomu autoritu, posílit něčí autoritu, zvýšit něčí autoritu;
podrývat něčí autoritu, podkopávat autoritu vlád, podlamovat autoritu prezidenta, nahlodávat něčí autoritu;
zpochybnit něčí autoritu, poškodit něčí autoritu, zničit něčí autoritu*

Blok E9 (Infinitiv (resp. VF)+SUBST-jiný pád) a blok E10 (Infinitiv (resp. VF)+Prep+SUBST-v pádě předložky)

Blok E9 zachycuje případy, kdy zpracovávané substantivum je neakuzativním doplněním slovesa, tj. vyskytuje se v genitivu, dativu či instrumentálu. V bloku E10 je zpracovávané substantivum předložkovým doplněním slovesa.

V bloku E9 rozdělujeme doklady podle pádů, dodržujeme pořadí pádů od 2. do 7., srov. např. heslo **kouzlo**:

*být zbaven kouzla;
nevěřit kouzlům;
očarovat někoho mocným kouzlem, vládnout kouzly*

V bloku E10 stejně jako v blocích E5 a E6 postupujeme podle pořadí předložek, které nabízí WSE. Srov. např.:

*věřit na kouzla;
věřit v kouzla;
vyznat se v kouzlech;
chránit před zlými kouzly, ochránit dobytek před kouzly;
dokázat pomocí kouzel očarovat krávu, řešit vše pomocí kouzel*

Slovesa v blocích E8–E10 uvádíme většinou v infinitivním tvaru. Pokud je ale důležitý původce děje, je nutné uvádět daná slovesa ve finitním tvaru, srov. např.:

veřej: *vrata vrzala ve veřejích, dveře skřípou ve veřejích* (tj. nikoli pouze *skřípat ve veřejích, vrzat ve veřejích*).

Blok E11 (SUBST+Subst, případně Subst + SUBST)

V bloku E11 je naším úkolem uvádět typické příklady koordinace. Ukazuje se, že skutečně typických, ustálených, pevných koordinačních spojení, tzv. binomiálů (srov. Čermák, 2007) – *pepř a sůl, otec a matka, klady a záporny*, je poměrně málo a že je obtížné tuto vysokou míru typičnosti rozpoznat.

Zároveň musíme akceptovat omezení daná WSE, kde se uvádějí pouze spojení v poměru slučovacím se spojkami *a* a *nebo*, ostatní slučovací spojovací výrazy chybějí. Do bloku E11 tedy většinou zapisujeme jen koordinace se spojkami *a* a *nebo*, popř. též s předložkou *s*, jako např. u lemmatu **otec**: *otec s matkou; matka s otcem*. Příklady uvedeného typu by rozhodně neměly být uváděny v blocích E5 a E6, ve kterých vyhledáváme přívlastková doplnění (srov. např. *smoking s košilí* – patří do E11 x *smoking s kapsičkami* – patří do E5).

Blok E12 (Ustálená spojení se SUBST – návrhy na samostatné zpracování)

Tento blok není na rozdíl od bloků E1–E11 a E13 vymezen syntakticky, nýbrž sémanticky – zachycuje ustálená spojení se zpracovávaným substantivem; v LDB Pralex se ustálená spojení zpracovávají samostatně, proto se v případě bloku E12 vlastně jedná o návrhy na samostatná zpracování daných ustálených spojení¹⁰.

Blok E13 (Jiná spojitelnost)

Tento blok zatím slouží jako sběrný koš pro případy, které se syntakticky nehodí do žádného jiného bloku. Ukázky různých spojení, která patří do E13, lze nalézt např. u hesla **puchýř**:

nohy plné puchýřů;
puchýře naplněné vodou;
látka vyvolávající na kůži puchýře;
opar ve stadiu puchýře;
mít dlaně samý puchýř

Blok E14 (Větné doklady)

Na rozdíl od předchozích bloků zachycuje blok E14 celé větné doklady z korpusu, tj. příklady rozvitější, často i takové, které díky širšímu kontextu vysvětlují a ilustrují méně jasná spojení z ostatních exemplifikačních bloků (často jde o příznaková užití).

Snažíme se vybírat jazykově zajímavé doklady. Zpravidla sem nedáváme doklady, které nepřinášejí novou relevantní informaci oproti dokladům již uvedeným v předchozích blocích. Výjimkou je případ, kdy je lemma v korpusu málo zastoupené, proto se snažíme z dokladů vytěžit maximum, tj. uvedeme je do E14 i do příslušných exemplifikačních bloků.

Blok E15 (V názvech)

Do bloku E15 dáváme příklady na užití daného substantiva v nejrůznějších názvech. Tento blok v základní nabídce exemplifikační struktury substantiv není, přidáváme jej pouze v případě, že nějaké relevantní doklady v korpusu nalezneme. Srov. např. u hesla **kachna**:

Divoká kachna (drama H. Ibsena z r. 1884);
Past na kachnu (česká filmová detektivka z r. 1978);
Zlatá kachna (název restaurace)

Literatura

ČERMÁK, František: Syntagmatika slovníku: typy lexikálních kombinací. In: Čeština – univerzália a specifika 3. Eds. Z. Hladká – P. Karlík. Brno: Masarykova univerzita 2001, s. 223–232.

ČERMÁK, František – HOLUB, Jan: Syntagmatika a paradigmatica českého slova I. Valence a kolokabilita. 3. vyd. Praha: Karolinum 2005. 191 s.

¹⁰ Blíže viz zde ve stati Z. Opavské.

- ČERMÁK, František: Kolokace v lingvistice. In: Studie z korpusové lingvistiky 2. Kolokace. Eds. F. Čermák – M. Šulc. Praha: Nakladatelství Lidové noviny 2006, s. 9–16.
- ČERMÁK, František: Binomiály. In: Čermák, F.: Frazeologie a idiomatika – česká a obecná. Praha: Karolinum 2007, s. 414–426.
- Nová slova v češtině 1. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 1998. 356 s. (SN 1)
- Nová slova v češtině 2. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 2004. 568 s. (SN 2)
- Nový akademický slovník cizích slov A–Ž. (J. Kraus a kol.) 1. vyd. Praha: Academia 2005. 880 s. (NASCS)
- OPAJSKÁ, Zdeňka – ŠTĚPÁNKOVÁ, Barbora: Práce s lexikálním materiálem a možnosti exemplifikace v lexikální databázi LEXIKON 21. In: Lexikografie v kontextu informační společnosti. Sborník příspěvků z pracovního setkání Praha, 4.–6. 6. 2007. Eds. A. Rangelova – J. Světlá – A. Jarošová. Praha: ÚJČ AV ČR, v. v. i., 2008, s. 33–40.
- Příruční slovník jazyka českého. 9 sv. Praha: Státní nakladatelství – SPN 1935–1957. (PSJČ)
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- SOCHOVÁ, Zdeňka – POŠTOLKOVÁ, Běla: Co v slovnících nenajdete. Novinky v současné slovní zásobě. Praha: Portál 1994. 204 s. (CSN)

Zpracování zdrobňelin v lexikální databázi Pralex¹

Edith Čonosová

Ústav pro jazyk český AV ČR, v. v. i.

Processing of Diminutives in the Pralex Lexical Database

The article deals with processing of diminutives in the Pralex lexical database, in particular with their allomorphy, homonymy and word-forming relations. Moreover, it discusses various definitions of diminutive meanings and provides their detailed classification in particular types.

1. Úvod

Zdrobňeliny jsou podle *Příruční mluvnice češtiny* (§ 196, s. 125) „desubstantivní² odvozeniny, jejichž formant je nositelem přídatného významu znaku menšího rozměru a/nebo pragmatického rysu hodnocení, častěji pozitivního než negativního“. Základním poznávacím znakem, že jde o zdrobňelinu, je existence základového slova v systému jazyka, srov. Doležel (1967, s. 494 nn.): „Podmínkou pro začlenění odvozeného (motivovaného) slova do souboru deminutiv je tedy existence základového substantiva nedeminutivního; v případech, kdy základové slovo této povahy neexistuje, nelze mluvit o deminutivu.“

Hesla tohoto typu zpracováváme v lexikální databázi (dále LDB) Pralex od května roku 2008, tedy už více než dva roky. Soubor hesel, která v současné době obsahují v definici označení „zdrob.“, představuje kolem 4000 heslových položek (a zhruba 5500 významů). Zakládání hesel tohoto typu však ještě není zcela ukončeno, takže konečný součet bude ještě o něco vyšší (celý soubor odhadujeme na více než 5000 hesel). Zpracování zdrobňelin je zaměřeno zejména na příkladovou část hesla, kde uvádíme spojitelnost na základě korpusu SYN³, ale – oproti jiným skupinám substantiv – i na vlastní výklad (viz dále podrobněji).

Nejdříve jsme začali zpracovávat soubor zdrobňelin, které byly už předem zařazeny do hesláře, protože se vyskytovaly ve *Frekvenčním slovníku češtiny* (dále FSČ) i *Slovníku spisového jazyka českého* (dále SSJČ) s výkladem.

Druhá fáze prací začala v září 2008 doplňováním dalších zdrobňelin ze SSJČ, která mají v definici zkratku „zdrob.“ (z těch jsme nejprve zpracovali ty, které měly frekvenci v korpusu SYN alespoň 30 výskytů).

Od prosince 2008 se zpracovávala další hesla, která byla vložena z FSČ, ale v SSJČ neměla výklad a byla jen přihnízdoována. Tento soubor byl na závěr celý zrevidován. Od jara 2010 jsme pak začali zpracovávat další deminutiva doplněná do hesláře LDB

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 V našem případě se budeme soustředit pouze na zdrobňeliny odvozené od substantiv. Nebudeme se zabývat obecným problémem, zda jsou odvozeniny od adjektiv (*malinký*), od sloves (*spinkat*), od adverbí (*pomalíčku*) skutečně deminutiva.

3 Od roku 2011 byl původní korpus SYN (který slučuje korpusy SYN2000, SYN2005 a SYN2006PUB) přejmenován na ORIG_SYN.

Pralex později, a zároveň jsme všechny zdobněliny začali propojovat s jejich základovými slovy pomocí nástroje Slovtvorba. V tomto nástroji jsou dva seznamy, první slouží k propojení zdobnělin směrem k základovým slovům, v druhém seznamu postupujeme opačným směrem.

Zpracování zdobnělin odpovídá obecným zásadám pro zpracování substantiv v LDB Pralex⁴. Pouze v některých částech hesla se ukazují určitá specifika. V tomto článku se zaměříme na homonymii a variantnost, slovtvorné vztahy a typy výkladů u zdobnělin.

2. Homonymie a variantnost u zdobnělin

Zpracování homonym a variant u zdobnělin se v podstatě neliší od zpracování ostatních substantiv. Homonymii označujeme římskou číslicí.

Jedná-li se o homonymii základového slova (např. *plavák I*, *plavák II*, *plavák III*), pak to u zdobněliny označujeme v rámci výkladu významu, srov. **plaváček I** – ,*expr. zdob. k plavák I (kůň se žlutavou, plavou srstí a s tmavými žíněmi)*‘. Pokud je základové slovo navíc polysémní, pak se v první části výkladu za označením homonyma po lomítku označí i číslo příslušného významu, např. **plaváček III** – ,*zdob. k plavák III/1; malý plavák (těleso plovoucí na vodě a sloužící k různ. účelům)*‘. V případě, že homonymie se objeví až u zdobnělin (jako je tomu např. u hesel *lištička I* a *lištička II*, které se vztahují k základovým slovům *lišta* a *liška*), označí se základové slovo v definici jen číslem významu, srov. **lištička I** – ,*zdob. k liška 1; malá n. menší liška (psovitá šelma)*‘ a **lištička II** – ,*zdob. k lišta; malá n. menší lišta (laťka, zprav. různě upravená)*‘.

Pokud je základové slovo variantní, pak tuto variantnost zachycujeme u příslušné zdobněliny ve výkladu významu a v nástroji Slovtvorba, např. **fejetoněk** – ,*expr. zdob. k fejetón*‘ a **fejetoněk** – ,*expr. zdob. k fejeton*‘. Variantní podoby lemmatu (stejně jako u všech ostatních podstatných jmen) v LDB Pralex pracovně uvádíme jako samostatné položky databáze a propojujeme je mezi sebou hypertextově pomocí Seznamu variant.

3. Slovtvorné vztahy u zdobnělin

Zdobněliny jsou slovtvorné deriváty, které se tvoří od základových substantiv pomocí nejrůznějších sufixů. V češtině odlišujeme podle míry zdobnění zdobněliny 1. a 2. stupně. Nejvíce zdobnělin se v češtině tvoří od maskulin (nejčastější sufixy jsou zde -ek a -ík), u feminin je nejběžnějším sufixem -ka. Nejméně zdobnělin tvoří neutra (u nich jsou nejužívanější sufixy -ko a -átka)⁵.

4 V rámci LDB Pralex pracujeme zatím pouze s korpusem SYN; je nám jasné, že takto získané poznatky bude nutno v rámci nového slovníkového zpracování ověřovat a znovu vyhodnocovat na základě všech materiálových zdrojů, se kterými pak budeme pracovat.

5 Srov. Doležel, 1967, s. 494 nn.; jsme si vědomi toho, že tehdejší hodnocení vycházelo z materiálu použitého v PSJČ a SSJČ.

K zachycení slovtvorných vztahů mezi zdobnělinami jednotlivých stupňů a jejich základovými slovy využíváme hypertextové propojení pomocí seznamu v nástroji Slovtvorba. Pokud je například zdobnělina u jednoho základového slova více (např. *osel* – *oslík* – *oslíček*), propojujeme všechny navzájem, takže u hesla *osel* budou jako deriváty zaznamenána hesla *oslík* i *oslíček*, u hesla *oslíček* základová slova *osel* i *oslík* a u hesla *oslík* je základovým slovem *osel* a derivovaným *oslíček*. Tento stav se analogicky odráží i v definici u zdobnělin 1. a 2. stupně, respektive v její první, slovtvorné části: **oslík** V1⁶ – ‚zdrob. k *osel* 1; malý n. menší *osel* (zvíře)‘; **oslíček** V1 – ‚zdrob. k *osel* 1, k *oslík* 1; malý n. menší *osel*, *oslík* (zvíře)‘.

Ne vždy žije v dnešním úzu kompletní řada zdobnělin 1. a 2. stupně; dokladem toho mohou být např. hesla *farářík* a *faráříček* k základovému slovu *farář*. 1. stupeň zdobněliny *farářík* není v korpusu SYN vůbec doložen, existuje tedy spíše jen potenciálně (pokud nebude dohledán v jiných zdrojích), kdežto 2. stupeň *faráříček* zde má 9 dokladů. V takovém případě je otázkou, zda má takováto nepoužívaná či nedoložená zdobnělina 1. stupně figurovat v definici u zdobněliny 2. stupně, srov. definici hesla **faráříček**: ‚expr. zdrob. k *farář*, k *farářík*‘, tj. zda má naše hodnocení vycházet ze současného úzu, nebo ze slovtvorného hlediska.

Na základě materiálu z korpusu SYN nebo z dalších referenčních zdrojů někdy doplňujeme heslář LDB Pralex o další hesla, a to jak o základová slova, tak i o zdobněliny. Např. u hesla ze SSSČ *fejton* s variantou *fejton* a expr. zdrob. *fejtonek*, *fejtonek* jsme v korpusu SYN našli další varianty *fejton*, *fejton* a *fejtonek*, které jsme následně založili jako další položky databáze (v souladu se zásadami pro zpracování variant v LDB Pralex) a jejich slovtvorné vztahy i variantnost jsme zachytili pomocí seznamů, takže vznikly slovtvorné dvojice: *fejton* a *fejtonek*, *fejton* a *fejtonek*, *fejton* a *fejtonek*; a vedle toho variantní čtveřice *fejton*, *fejton*; *fejton* a *fejton* a trojice variantních zdobnělin *fejtonek*, *fejtonek* a *fejtonek*.

4. Typy výkladu významu u zdobnělin

V SSSČ byly některé zdobněliny zpracovány jako samostatná hesla s vlastním výkladem (šlo hlavně o hesla, která se po stránce významové neshodovala se základovým slovem, případně o terminologizované významy⁷), ale většina jich byla jen přihnízdována na konci heslové stati základového slova, a to bez výkladu, pouze se zkratkou „zdrob.“, resp. „expr. zdrob.“.

U substantiv obecně se opíráme o výklad významu v SSSČ, ale u zdobnělin se v mnohých případech (zvláště u polysémních hesel) nedal použít, protože měl většinou jen odkazovou formu k výkladu základového slova. Zdobnělé významy zde byly jen evidovány, jako např. v SSSČ u hesla *růžek* 1 – *zdrob. k roh* 1, 2, 5–8. Pro popis významu u zdobnělin v LDB Pralex jsme proto navrhli jednotný typ výkladové definice, která má většinou dvě části, slovtvornou a výkladovou (výjimku tvoří expresivní zdobněliny).

6 Zkratka V slouží v LDB Pralex k označení jednotlivých významů.

7 Srov. *Směrnice SSSČ*, s. 114.

Zdrobněliny v LDB Pralex můžeme rozdělit do tří skupin: zdrobněliny vykazující rys malosti, dále zdrobněliny vykazující expresivitu a formální zdrobněliny. Ve většině případů se podaří každé konkrétní heslo zařadit k jedné z těchto skupin, ale jsou zde i některé přechodné typy.

U zpracovaných zdrobnělin se vyskytují tři hlavní typy výkladů, založené na následujících rysech:

1. r y s m a l o s t i

V případě, že zdrobnělina označuje skutečnost v jistém smyslu menší než základové slovo, pak definice v LDB Pralex vypadá takto: „**zdrob. k + základové slovo; malý/ menší atp. + základové slovo**“, např. **bambulka**: ‚zdrob. k bambule; malá bambule‘.

Definice se u tohoto typu zdrobnělin skládá ze slovotvorné části před středníkem a pak sémantického výkladu, ve kterém vystupují slova jako *malý, menší, menších rozměrů, drobný, mladý, úzký* apod. U polysémnních slov je pak v závorce výklad či přiblížení toho, o který z významů základového slova se jedná – bývá to spíše zjednodušená, zkrácená definice daného významu, např. **destička**: ‚zdrob. k deska 1; malá n. menší deska (plochý předmět, nejč. čtyřúhelníkovitého tvaru a nevelké tloušťky)‘.

Je-li základové slovo polysémnní, pak podle materiálu nalezeného v korpusu SYN nejprve určujeme, které významy se týkají deminutiva, a pak založíme jen ty doložené. Nezakládáme tedy všechny významy podle základového slova.

Na druhou, výkladovou část definice rezignujeme v případech, kdy je zdrobnělina v SYN zcela nedoložená, nebo je jen velmi málo dokladů, a není se tak o co opřít, případně nelze určit, od kterého z významů základového slova se zdrobnělina tvoří.

Viz např. v SYN nedoložené heslo ze SSJČ: **garniturka**: ‚zdrob. k garnitura‘. Přesto, že základové slovo je v SSJČ polysémnní (má tři významy), necháme vše jen s odkazem na základové slovo, protože jsme nenašli příklady, které by nám pomohly heslo rozčlenit na jednotlivé významy.

Obdobně to řešíme i u expresivních zdrobnělin, např. v SYN je nedoloženo *filosůfek*, proto bude v definici jen ‚*expr. zdrob. k filosof*‘; základové slovo je přitom polysémnní (má čtyři významy).

2. e x p r e s i v i t a

V tomto případě skutečnost označovaná deminutivem není menší, ale jedná se zde o vyjádření citového postoje mluvčího k dané skutečnosti. Definice se v těchto případech tvoří takto: **expr. zdrob. k + základové slovo**. Definice se tedy skládá z kvalifikátorů *expr.* a *zdrob.* (v tomto pořadí) a z odkazu na základové slovo, např. **inženýrek**: ‚*expr. zdrob. k inženýr*‘.

V případě polysémie či nejasného výkladu významu u základového slova pak do závorky uvádíme specifikaci daného významu základového slova, tedy jeho zestručněnou definici, např. u monosémnního hesla: **gáblíček**: ‚*expr. zdrob. ke gábl, ke gáblík (přesnídávka, snídaně na vidličku)*‘. A u polysémnního: **dědoušek** V1: ‚*expr. zdrob. k děd 1 (otec jednoho z rodičů)*‘; V2: ‚*expr. zdrob. k děd 3 (stařec, kmet)*‘.

Přechodné typy:

2a) Existuje množství zdrobnělin, které splňují podmínku rysu malosti, ale ten se zároveň někdy kombinuje s expresivitou. Přestože u nich pocítujeme určitou expresivitu,

sivitu v některých dokladech, nelze jednoznačně určit, že je heslo expresivní vždy. V takových případech označujeme expresivitu pouze na úrovni exemplifikace u konkrétních dokladů, které v daném bloku nadepíšeme titulkem „expr.“. Viz např. u hesla **bratříček** V1: *zdrob. k bratr 1, k bratřík 1; malý bratr (sourozenec mužského rodu)*, pak ale v exemplifikaci v bloku E1⁸ k V1 uvádíme i tyto doklady:

expr.

povedený bratříček, praštěný bratříček;

milovaný bratříček

U těchto příkladů se nejedná o ‚malého bratra‘, ale jde o doklady na citový postoj mluvčího. V SSJČ byla v podobných případech použita další možnost, charakteristika „expr.“ byla vymezena poznámkou „často“ a „zprav.“ Tuto možnost hodnocení zatím na úrovni celého hesla (u lemmatu) nepoužíváme, protože v LDB Pralex jde především o zpracování materiálové části hesla, utřídění a okomentování dokladů. Pochopitelně budeme celá hesla komplexně vyhodnocovat v budoucím slovníkovém zpracování, až budeme pracovat s větším množstvím dokladů i z jiných zdrojů.

Často se také stává, že v bloku upravených dokladů⁹ na minimálních syntagmatech expresivitu tolik nevnímáme a k jejímu rozpoznání je třeba delší kontext. V takovém případě raději použijeme blok Větné doklady a doklad uvedeme tam.

Někdy dochází k posunu v hodnocení expresivity oproti SSJČ: buď bylo heslo v SSJČ hodnoceno jako expresivní, ale dnes již expresivitu nepocítujeme, nebo jde o opačný případ, kdy SSJČ tuto charakteristiku neuvádí, ale my ji na základě dokladů z korpusu doplníme. Heslo *dílnička* je např. v SSJČ hodnoceno jako „expr.“: **expr. zdrob.** *dílnička*. Avšak podle našeho jazykového citění a zjištěných dokladů jsme heslo zařadili spíše do první skupiny zdobnělin s definicí *zdrob. k dílna; malá dílna*.

2b) Pokud je už samo základové slovo v SSJČ označeno jako expresivní, pak vždy předpokládáme, že se daná charakteristika vztahuje i na zdobnělinu, i když to v SSJČ někdy není označeno u obou hesel.

Srov. např. u hesla **fakaně, fakaně** (*faganě), -te s. ob. **expr., zprav. hanl.** dítě; (expr. zdrob. *fakaňátko*, -a s. (6. mn.-ách). Zdobnělina **fakaňátko** je také označena jako „expr.“, což v těchto případech přebíráme, srov. *fakaňátko* – ,expr. zdrob. k *fakaně*.

Podobně mohou být expresivní slova zdobnělá od základu slov obecněčeských nebo hanlivých, ale vždy je třeba na novém materiálu vyhodnotit, zda se expresivita týká jen některých dokladů, nebo celého hesla. Srov. např. u hesla *fald*, kde je v SSJČ charakteristika „ob.“ a některá spojení jsou označena jako „expr.“. Pokud nám doklady z korpusu jednoznačně neukáží expresivitu ve všech případech, pak zdobnělinu jako expresivní neoznačujeme, srov. **faldík** – ,zdrob. k *fald*; **malý n. menší fald** (*záhyb*); **faldíček** – ,zdrob. k *fald*, k *faldík*; **malý n. menší fald** (*záhyb*). Jako expresivní označíme pouze příslušné doklady na úrovni exemplifikace, srov.

faldík E1 k V1 – *expr. roztomilé faldíky*

faldíček E8 k V1 – *expr. roztrást faldíčky*

8 Zkratka E slouží v LDB Pralex k označení exemplifikačních bloků (blíže viz zde ve stati M. Koutové).

9 K upraveným dokladům blíže viz zde ve stati M. Koutové.

3. formální zdrobněliny

Jde o případy, kdy zdrobnělina už nemá ani rys malosti, ani není expresivní, tj. rys malosti i expresivity se dnes stírá. Tyto zdrobněliny jsou pouze formální a v slovní zásobě vystupují jako synonyma slova základového bez jakéhokoli zabarvení¹⁰. Definice u zdrobnělin tohoto typu se proto skládá ze slovotvorné formální části před středníkem a výkladu za ním, který je totožný jako u základového slova. Patří sem např.:

vajíčko V1: ‚zdrob. k vejce 1; oblý útvar ve skořápce‘;

klička II V1: ‚zdrob. ke klika I/3; křivá čára‘;

kravka: ‚zdrob. ke kráva 1; samice tura domácího‘;

babka V10: ‚zdrob. k bába 3; pomocnice při porodu‘;

benjamínek: ‚zdrob. k benjamín; nejmladší člen rodiny n. něj. společnosti‘;

fazónka V1: ‚zdrob. k fazóna; náležitý tvar, forma‘.

Závěrem

Z celkového souboru zdrobnělin v LDB Pralex jich většina je celkem jednoznačně zařaditelných. Kromě toho se ale při práci s materiálem setkáváme s řadou tzv. „problémových“ hesel, která nelze jednoznačně přiřadit k některé ze skupin, a tudíž máme problémy s výkladovou definicí:

Kam např. zařadit heslo *fenka*? Je menší než *fena*, nebo se jedná o synonyma? Váháme zde mezi typem 1 (rys malosti) a 3 (formální zdrobnělina). Dalším prubířským kamenem je třeba heslo *čtvrthodinka*: Je toto heslo ještě expresivní, či se tento příznak už stírá? Objektivně se nemůže jednat o skutečnost menší, než je *čtvrthodina*, takže skupina 1 odpadá, ale zbývá tu skupina 2 (expresivní zdrobnělina) nebo 3 (formální zdrobnělina). U hesel *deštík* a *deštíček* jde zřejmě o přechod mezi hodnocením mluvčího a drobností deště – tedy mezi skupinami 1 a 2. Zda se jedná o rys malosti, nebo o vztah mluvčího, poznáme většinou až z materiálu, a každý příklad se může jevit jinak. V podobných případech není jednoznačné určení možné a vždy do jisté míry záleží na jazykovém cítění a povědomí zpracovatele.

Literatura

- BALÁŽOVÁ, Ľubica: Osobné (depersonálne) deminutíva: špecifická polysémia a jej lexikografické spracovanie. In: Slovo – Tvorba – Dynamickosť. Na počesť Kláry Buzássyovej. Ed. M. Šimková. Bratislava: Veda 2010, s. 209–223.
- DOLEŽEL, Lubomír: Jména zdrobnělá (Kvantitativní rozbor). In: Tvoření slov v češtině. 2. Odvozování podstatných jmen. Red. F. Daneš – M. Dokulil – J. Kuchař. Praha: Academia 1967, s. 494–530.
- Encyklopedický slovník češtiny. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: Nakladatelství Lidové noviny 2002, s. 131–132, s. 283.

¹⁰ Srov. též Štícha, 1978, s. 113 nn.

- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSČ)
- HAUSER, Přemysl: Nauka o slovní zásobě. 1. vyd. Praha: SPN 1980, s. 130–133.
- Mluvnice češtiny 2. Tvarosloví. Red. M. Komárek – J. Kořenský – J. Petr. Praha: Academia 1986, s. 300–303, s. 381, s. 417, s. 435.
- Helcl, Miloš a kol.: Směrnice pro vypracování rukopisu slovníku spisovného jazyka českého. Praha 1957.
- Příruční mluvnice češtiny. Eds. P. Karlík – M. Nekula – Z. Rusínová. Brno: Nakladatelství Lidové noviny 1995, s. 125–128.
- NÁBĚLKOVÁ, Mira: Domiská, knížiská, psiská a iné modifikačné substantívne deriváty v slovensko-českom pohľade. In: Slovo – Tvorba – Dynamickosť. Na počesť Kláry Buzássyovej. Ed. M. Šimková. Bratislava: Veda 2010, s. 396–425.
- RUSÍNOVÁ, Zdeňka: Deminutiva ve frazeologických přirovnáních. In: Sborník prací Filozofické fakulty brněnské univerzity. Linguistica Brunensia, A 46. Brno: Masarykova univerzita 1998, s. 101–106.
- RUSÍNOVÁ, Zdeňka: Deminutiva jako jazykové univerzálie. In: Sborník prací Filozofické fakulty brněnské univerzity. Linguistica Brunensia, A 49. Brno: Masarykova univerzita 2001, s. 137–145.
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- Slovník spisovné češtiny pro školu a veřejnost. Red. V. Mejstřík a kol. 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovné češtiny pro školu a veřejnost. Praha: Leda 2004 (elektronická verze).
- ŠTÍCHA, František: Substantiva deminutivní formy s lexikalizovaným významem. In: Naše řeč, 1978, roč. 61, č. 3, s. 113–127.
- TRÁVNÍČEK, František: Mluvnice spisovné češtiny. Praha: Melantrich 1949, s. 265–271, s. 354.
- Tvoření slov v češtině 1. Teorie odvozování slov. Red. M. Dokulil. Praha: Nakladatelství Československé akademie věd 1962. 263 s.
- ZIMA, Jaroslav: Expresivita slova v současné češtině. Studie lexikologická a stylistická. Praha: Nakladatelství Československé akademie věd 1961. 140 s.

Zpracování antonym v lexikální databázi Pralex¹

Martina Habrová

Ústav pro jazyk český AV ČR, v. v. i.

Processing of Antonyms in the Pralex Lexical Database

In the Pralex lexical database, antonyms are preserved as an explanatory part in the same way as in the *Dictionary of the Standard Czech Language*, except for the entries, which require a modification of the meaning explanation (e.g. some word-formative variants, homonyms etc.). Members of the antonym pairs are hypertextually linked while using The List of Antonyms, which is available for every entry at the level of particular meanings.

I. Zpracování antonym ve slovnících

Ve výkladových slovnících není vymezení antonym na rozdíl od synonym závislé na formulaci výkladové definice, protože opozita vyjadřují nejen významový, ale i logický vztah protikladu; tento protiklad zůstává v průběhu času neměnný, tudíž se s ním dá pracovat bez ohledu na budoucí úpravy výkladů při novém slovníkovém zpracování. Antonymie je tedy pro výkladovou lexikografii důležitou oporou.

Ve výkladových slovnících češtiny bývají antonyma zachycena na konci výkladu. V *Příručním slovníku jazyka českého* (dále PSJČ), *Slovníku spisovného jazyka českého* (dále SSSJČ), *Slovníku spisovné češtiny pro školu a veřejnost* (dále SSČ) i *Novém akademickém slovníku cizích slov* (dále NASCS) jsou definována jako slova opačného (protikladného) významu, opozita (opaky), která slouží k dalšímu zpřesnění výkladu významu slova, podobně jako synonyma. Uvádějí se většinou v závorce s příslušným označením, a to jednak za výkladem významu, popř. za synonymy, jednak za dokladovými spojeními, srov. např. v SSČ **bázlivý** bojácny, ustrašený (op. smělý 1, statečný); **černý 1.** mající barvu sazí (op. bílý 1); tmavý, temný 2, 1 (op. světlý 1, 2): ... č-é pivo (op. světlé).

Systematičtěji jsou antonyma uváděna ve slovnících synonymických. Ve *Slovníku českých synonym* (2000) se definují antonyma (opozita) jako slova s opačným významem a jsou uváděna v relevantních případech u heslového slova, resp. uvnitř příslušných synonymických řad, např. **chytrý** vs. **hloupý**. Pro odlišení jsou zde antonyma označena speciálním znakem ♦ (tzv. hřebík).

Slovník českých synonym a antonym (2007) se podrobněji věnuje i teoretickému výkladu o antonymech. Definuje je jako dvojice slov protikladného významu a charakterizuje jednotlivé typy antonym užitých v tomto slovníku (polární, komplementární, vektorová, konverzivní a kontextová), přičemž však zdůrazňuje, že nejčastěji používá antonyma polární a komplementární.

Zpracováním antonym v českých a slovenských výkladových slovnících se ve svém článku zabývá A. Oravcová (1992). Vychází z excerptu opozit z *Krátkého slov-*

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

níku slovenského jazyka (1987) a s ohledem na dva základní typy klasifikace antonym (klasifikace strukturní a klasifikace na základě logického rozlišení kontrárnosti a kontradiktornosti) navrhuje zásady pro zpracování významových protikladů ve výkladovém slovníku středního typu. Podle ní by se ve slovníku měly zaznamenávat pouze **ostré významové protiklady**, a to mezi krajními členy antonymických řad, protiklady vektorové a komplementární, dále pak kontrasty terminologické a kontrasty mezi neterminologickými lexikalizovanými spojeními. Jako součást výkladu významu slov by se antonyma měla používat pouze v případech, kdy se jedná o stejnou lexikální spojitelnost obou členů opozitní dvojice.

II. Zpracování antonym v LDB Pralex

Antonymům jsme se v rámci lexikální databáze (dále LDB) Pralex začali věnovat až začátkem r. 2010, kdy jsme je začali hypertextově propojovat mezi sebou jako hesla a narazili jsme při tom na řadu nejasností a problémů, jednak proto, že antonyma nejsou ve výkladových slovnících češtiny zachycena v úplnosti, ale také proto, že databázové zpracování má proti slovníkům různá specifika, která je třeba brát v úvahu.

Při zpracování hesel v LDB Pralex se pracovně opíráme o SSJČ, tudíž nás jako východisko pro další práci s antonymy zajímá jejich zastoupení u všech slovních druhů právě v tomto slovníku. V SSJČ se nejčastěji opozita ve výkladu objevují u adjektiv a substantiv, v menší míře u sloves a adverbíí, řidčeji u předložek a velmi zřídka u zájmen (např. **každý** x **žádný**), číslovek (např. **málo** x **mnoho**) a citoslovcí (např. **čehý** x **hot**)².

Zápor může být ve slovnících vyjádřen explicitně antonymem použitým ve výkladu, např. **konečný** ‚(op. nekonečný)‘, **harmonie** ‚(op. disharmonie)‘, **diskrétní** ‚(op. indiskrétní)‘. Může však být vyjádřen i implicitně, jako součást opisu významu antonymního slova, např. **mrtvý 3.** *jsoucí bez pohybu, nesvědčící o životě; nehybný; slepý 1.* *postrádající zrak, neschopný vidět; nevidomý 1; suchý 3.* *ničím nenamazaný, nepotřený; neobsahující žádný tuk.* Negace je tedy ve výkladech významu v SSJČ vyjadřována nejen antonymy, ale také opisem, nebo synonymy. V tomto příspěvku se zaměřujeme pouze na opaky umístěné v hesle za výkladem významu a explicitně označené zkratkou „op.“.

Zatím se v LDB Pralex věnujeme pouze antonymům u substantivních hesel; v současnosti náš heslář obsahuje přes 700 substantivních hesel s opozitem použitým ve výkladové definici (a jejich počet může dále narůstat v důsledku zpracovávání později založených hesel). V roce 2011 se v rámci výzkumného záměru zaměříme i na adjektiva, u nichž jsou opozita ve výkladu rovněž hojně zastoupena.

SSJČ používá antonyma jednak jako součást výkladu významu heslového slova, jednak při výkladu významu typických spojení uvedených v rámci exemplifikace. V druhém případě je v SSJČ opozitum součástí výkladu terminologického nebo ne-

2 Vzhledem k tomu, že v LDB jsou zatím antonyma zpracována pouze u substantivních hesel, jsou příklady k jiným slovním druhům vybírány ze SSJČ.

terminologického spojení. Pokud dané spojení zůstane u substantiva jen v příkladové části hesla, opak u něj neuvádíme; pokud je toto spojení v LDB Pralex zpracováno jako víceslovná jednotka, bude se problematika antonymie řešit v souladu se specifickým zpracováním těchto lexikálních jednotek.

1. Zpracování antonym u substantiv

V LDB Pralex zachováváme stejný zápis opozit jako v SSJČ, tj.: (op. + lemma), protože většinu výkladů přebíráme zatím pracovní z tohoto slovníku. Výkladové definice ze SSJČ byly u slov, která byla na základě frekvence v korpusu SYN2000 zařazena i do *Frekvenčního slovníku češtiny* (dále FSČ), do LDB vloženy programátorem. Došlo však k tomu, že jím navržený program zkopíroval jen tu část výkladu, která byla v SSJČ vyznačena jako souvislý text kurzivou. Takto programově vložená hesla tedy antonyma ve výkladu významu vůbec neměla, protože zkratka „op.“ byla uvedena jiným typem písma než text výkladu. U všech těchto hesel tedy bylo potřeba výklad o opozita doplnit.

Dále jsme zjistili, že propojování mezi antonymy nelze provést zcela mechanicky podle seznamu hesel ze SSJČ, u kterých je opozitum uvedeno, protože i v SSJČ toto nebylo vždy realizováno důsledně u obou členů opozitní dvojice.

Další problém souvisí s tím, že v naší databázi všechny varianty zakládáme jako samostatná hesla, a tudíž upravujeme i jejich výklady v souladu s podobou lemmatu, což se projeví i u antonym ve výkladu (o variantách viz podrobněji dále).

V LDB Pralex zachováváme antonyma jako součást výkladu v nezměněné podobě kromě případů, kdy je potřeba výklad významu upravit (např. u některých slovo tvorných variant) nebo smazat. Pokud někde v heslech vložených programově opozita chyběla, postupně je do výkladové definice podle SSJČ doplňujeme, protože jinak ani dané heslo nelze v databázi vyhledat a nedostane se pak ani do souboru antonym vyhledaného pomocí zadání: „definice – obsahuje: op.“. Jinou možnost, jak s opozity jako s uceleným souborem pracovat, zatím nemáme, proto je nutné je do výkladů doplňovat (tam, kde to je relevantní – s ohledem na variantnost).

Naproti tomu v případech, kdy v SSJČ u jednoho členu dvojice opak chyběl, do příslušného výkladu jej zatím nedoplňujeme vždy, někdy jej pouze uvádíme na Seznam antonym (viz podrobněji v bodě 5). Vyhledat se však tyto dvojice dají na základě toho členu, u kterého opak uveden je.

2. Úpravy výkladových definic oproti SSJČ

V rámci základního zpracování substantivních hesel v LDB Pralex jsme někdy výklady významu ze SSJČ obsahující antonymum upravovali. Např. u variant upravujeme grafickou podobu antonyma (viz podrobněji v bodě 8), v některých případech opozitum z výkladu mažeme, či naopak doplňujeme.

Definice u hesel s antonymem použitým ve výkladu jsme upravovali i v případě homonym (viz podrobněji v bodě 7). U slovnědruhových homonym, která byla

v SSJČ zpracována ve sdruženém hesle, jsme např. opozita uváděná v závorce – viz např.: **bílý** ,(op. černý)‘, **černý** ,(op. bílý)‘, **tepl**o ,(op. chladno)‘ – doplnili o římskou číslici (v souladu se značením homonym v LDB Pralex): ,(op. černý I)‘, ,(op. bílý I)‘, ,(op. chladno I)‘. Obdobně jsme postupovali u homonym v rámci jednoho slovního druhu (v SSJČ sdružených do jednoho heslového odstavce nebo původně zpracovaných jako dva významy jednoho polysémního hesla, které jsme na základě dnešního stavu rozdělili na hesla samostatná), např. **antagonista II** ,(op. synergista II)‘.

3. Doplnování výkladů a přidávání dalších významů

Do hesláře LDB Pralex rozepisujeme na samostatné databázové položky celá slo-
votvorná hnízda ze SSJČ (přihnízďované deriváty, zdobněliny a jména přechýlená)
a slovtvorné řady, které byly v SSJČ zpracovány v rámci heslových odstavců základo-
vých slov. Některá ze slov přihnízďovaných nemají v SSJČ vlastní výklad, proto se
při jejich zpracování opíráme o definice z jiných výkladových slovníků (ve výkladu
významu můžeme použít antonymum v souladu s definicí, která byla vybrána z těchto
zdrojů), nebo vycházíme z upraveného výkladu základového slova ze SSJČ, kde už
byl opak uveden.

Někdy při základním zpracování substantiv podle jiných referenčních zdrojů
nebo na základě korpusu přidáváme další významy; v takovém případě můžeme
rovněž využít antonyma uvedená v dalších slovnících.

Např. substantivum **bezděčnost** bylo v SSJČ přihnízďováno k adjektivu **bezděčný**
bez vlastního výkladu – v LDB jsme použili definici z PSJČ ,(op. schválnost, úmysl-
nost)‘ a do hesláře dodatečně založili heslo **úmyslnost**; u obou hesel jsme na Seznam
antonym zapsali odpovídající opaky.

4. Doplnování hesláře

Pokud opozitum použité v rámci výkladu významu ještě nemáme v hesláři, rov-
nou je doplňujeme (řídíme se zásadou, že všechna slova použitá ve výkladu by měla
být součástí hesláře).

5. Seznam antonym

Antonyma v LDB Pralex mezi sebou hypertextově propojujeme pomocí nástroje
nazvaného SEZNAM ANTONYM, který je k dispozici u každého hesla, a to na úrovni
jednotlivých významů, protože antonymie se týká právě jednotlivých významů, ni-
koli celého hesla, jako je tomu u ostatních seznamů (např. Seznam synonym), které
používáme pro zachycení systémových vztahů mezi hesly (položkami databáze).

Ojediněle bývají za výkladem uvedena dvě i více antonym za sebou, např. u hesla
arze ,(op. teze, klad)‘. V takovém případě bude na Seznamu antonym vyplněno více
položek.

V případě, že SSJČ neuvádí opaky ve výkladu u obou členů opozitní dvojice, v LDB Pralex tento problém řešíme zápisem na seznam u obou antonym, ale do výkladu jej nedoplňujeme, jak již bylo uvedeno výše.

Propojení hesel prostřednictvím antonym je pro nás důležité, protože antonymní dvojice by měly být zpracovány paralelně; většinou jsou členy dvojice od sebe abecedně vzdáleny, neboť antonymie prochází napříč celou abecedou (na rozdíl např. od variantnosti). Hypertextový proklik nám umožňuje pracovat s oběma členy dvojice zároveň.

6. Antonymie u polysémních hesel³

Antonymie se nemusí týkat všech významů polysémního hesla (právě antonymem, podobně jako synonymem, se některé významy od sebe mohou odlišit). Proto je třeba při zpracování polysémních hesel v LDB Pralex ověřovat, kterého významu se opozice týká. Např. u hesla **abstraktum** se opak **konkrétnum** vyskytuje u obou významů, u hesla **adresát** je opozitum **adresant** uvedeno jen u prvního významu, u desetivýznamového hesla **život I** se antonymum v SSJČ vyskytuje pouze u prvního významu. Z tohoto důvodu se antonyma zapisují na seznamy až na rovině jednotlivých významů – viz i polysémní heslo **klid**⁴. Podobným případem je rovněž heslo **klad**, které má v 1. a 2. významu opak **zápor**, kdežto ve 3. významu opak **dvih**, **arze**.

Někdy opak v LDB Pralex přidáváme na Seznam antonym i v případě, že ve výkladové definici daného hesla v SSJČ nebyl uveden. Např. u hesla **teze** SSJČ uvádí u třetího významu antonyma **dvih**, **arze**; první význam má bez opaku. Avšak ve výkladu hesla **antiteze** je u prvního významu uvedeno antonymum **teze**, proto jsme to v LDB vyřešili takto: u hesla **teze** je na Seznamu u prvního významu uvedeno **antiteze**, u vý-

3 J. Filipec (1994, s. 126) upozorňuje na skutečnost, že vedle polysynonymie existuje také polyantonymie, protože členy opozitních dvojic a stupnic bývají nejen monosémické, ale i polysémické. Polyantonymie potvrzuje fakt, že polysémie je jednak projev asymetrie lexému (znaku), jednak významová struktura lexému (hyperlexému), struktura sémémů a tím i relativně samostatných lexikálních jednotek (lexií). Filipec a Čermák (1985) uvádějí příklad antonymní dvojice **silný-slabý**, jejíž členy jsou protikladné z hlediska fyzické a duševní síly, společenského vlivu, účinnosti nějakého prostředku, rozměru do šířky (tloušťky), kvantity, intenzity a obsažnosti.

4 SSJČ: **klid** -u m. (6. j. -u) **1. stav nehybnosti, nečinnosti, odpočinku** (op. *pohyb*): zůstat (ležet) v klidu; polední k.; nedělní, sváteční k.; práv. den pracovního klidu *zákonným předpisem určený den všeobecného odpočinku od práce*; k. řízení *zastavení postupu soudního řízení z návrhu účastníků n. pro jejich nečinnost*; fyz. těleso v klidu *neměnicí polohu* (op. v *pohybu*); těl. cvičení na kruzích v klidu (op. v *hupu*); voj. k. zbraní *dočasně zastavení bojové činnosti, příměří*; v pojišťovnictví uvést pojistku do klidu *na čas ji zastavit* **2. stav nerušenosti; pokoj, ticho** (op. *neklid, ruch*): mít k. k práci; zachovávat, rušit k.; k. lesa; idylický k. *venkova, žít v klidu a míru*; ohrožovat veřejný k. a pořádek; mít od někoho, od něčeho k. *nebýt obtěžován, rušen* **3. stav nevzrušenosti; rozvážnost, chladnokrevnost** (op. *rozčilení, vzrušenost*): zachovat v každé situaci k.; ledový k.; hovořit s klidem; ztratit (při zkoušce) k.; být přiveden z klidu; jednat s klidem Angličana (hovor.); vynikat stoickým klidem; epický k. *nezúčastněnost a nevzrušenost vyprávění*; hovor. s klidem něco udělat jako samozřejmost, bez ostychu, bez ohledu; --> expr. zdrob. ***klideček**, -čku m.

znamu druhého *dvih, arze* (tj. doplnili jsme opak k 1. významu v souladu s heslem *antiteze*).

7. Antonymie u homonym⁵

V LDB Pralex zakládáme jako samostatné databázové položky (hesla) i ta homonyma, která byla v rámci jednoho slovního druhu v SSJČ sdružena do jedné heslové stati (a oddělena pomlčkou), a označujeme je římskými číslicemi. Protože tedy např. homonyma *civil* (m. živ.) a *civil* (m. než.) uvádíme jako dvě položky databáze, pak musíme antonyma *uniforma, kostým* uvedená v SSJČ přiřadit k relevantnímu homonymnímu heslu.

V LDB zakládáme a zpracováváme každý slovní druh zvlášť, což platí i pro slovně-druhovú homonyma, která byla v SSJČ sdružena do jednoho souborného hesla; pracovně je označujeme římskými číslicemi v pořadí podle slovních druhů. V SSJČ byla např. v rámci jedné heslové stati zpracována hesla *bílý* (adjektivum) a *bílý* (substantivum). V LDB jsme založili samostatná hesla *bílý I, bílý II*, u substantiva *bílý I* jsme v souladu s výkladem v SSJČ doplnili u třetího významu antonymum *černý*. Součástí našich úprav je i přiřazení římské číslice k opaku v závorce, proto jsme definici upravili na: ‚(op. černý I)‘. Opozity u adjektiv jsme se prozatím v LDB nezabývali, ale při zpracování adjektiv bude u hesla *bílý II* podle výkladu ze SSJČ k prvnímu významu zapsáno opozitum *černý II* (číslování podle LDB), k druhému významu antonyma *černošský, černý, barevný* a k třetímu významu opak *rudý*.

8. Antonymie u variant⁶

Při zpracování antonym u variant se variantnost promítá do výkladu významu u obou hesel dvojice, včetně antonym v něm uvedených. U variant si všímáme toho, jaké antonymum se k danému lemmatu vztahuje, aby si v rámci metajazyka odpovídaly jednotlivé členy antonymických dvojic. Tento vztah nemusí být symetrický – jeden člen opozitní dvojice může mít varianty a druhý ne, např. *acidita x basicita/bazicita* (tento asymetrický vztah se projeví i na Seznamech antonym).

U pravopisných variant upravujeme odpovídající podobu lemmatu a příslušného antonyma použitého ve výkladu. V SSJČ bývají často v záhlaví hesla uvedeny varianty s odlišnou grafickou podobou, ale výklad významu se vztahuje k oběma podobám, např. *acidosa, acidóza* ‚(op. alkalóza)‘ – u hesla *acidosa* jsme definici v LDB upravili oproti SSJČ na ‚(op. alkalosa)‘.

5 M. Pisárčiková (1989) upozorňuje na úzkou souvislost enantosémie (vnitřněslovní antonymie) s polysémií a homonymií. Dva polární významy jednoho slova jsou v podstatě krajními body polysémie a často bývají příčinou rozpadu slova na homonyma. Užití antonyma ve výkladu významu u homonym tedy rovněž slouží k jejich odlišení, stejně jako u polysémie.

6 Samostatné uvádění variant představuje pracovní řešení v rámci LDB Pralex, aby se při třídění materiálu mohly ke každé formální podobě lemmatu přiřadit příslušné doklady; při následném slovníkovém zpracování budou pochopitelně všechny varianty sloučeny do jediného hesla.

U slovotvorných variant zvažujeme, co bude antonymní dvojice, a co nikoliv. SSJČ např. u hesla **abnormalita** uvádí opak *normálnost* – podle přípon propojujeme dvojice cizí-cizí, domácí-domácí, proto byl výklad významu upraven na: ‚(op. normalita)‘.

Někdy je pro úpravu výkladu hesel s antonymem potřeba porovnat více systémových vztahů mezi hesly v SSJČ s ohledem na varianty a výklad pomocí synonyma. SSJČ např. u hesla **aerob**⁷ uvádí synonymum *aerobiont* a antonymum *anaerob*; u hesla **aerobiont**⁸ pouze synonymum *aerob*, bez opozita; varianty **anaerob**, **anaerobiont**⁹ mají v SSJČ společnou lemmatizaci a ve výkladu opak *aerob*. U hesla **anaerobiont** jsme výkladovou definici v LDB upravili oproti výkladu v SSJČ na ‚(op. aerobiont)‘, vznikly tak antonymické dvojice **aerob** x **anaerob**, **aerobiont** x **anaerobiont**; hesla **aerob** a **aerobiont**, **anaerob** a **anaerobiont** jsme se rozhodli jako varianty nepropojovat.

Tam, kde z hlediska našeho pojetí nejde o antonyma, opozitum z výkladu převzatého podle SSJČ mažeme, např. u řidší varianty – v SSJČ má heslo **starousedlík** řidší variantu **starousedlec** a ve výkladu významu opozitum *novousedlík*. Do hesláře jsme sice založili heslo **starousedlec**, ale nepropojili jsme ho s antonymem **novousedlík**, jelikož se jedná o slovotvornou variantu, které podle přípony netvoří antonymní dvojici s heslem **novousedlík**. Proto jsme u hesla **starousedlec** smazali z výkladové definice ‚(op. novousedlík)‘. Stejný případ se objevil např. u hesel **velkodůl**¹⁰, **velkovýroba**¹¹.

U některých hesel registruje SSJČ více než dvě variantní podoby lemmatu v záhlaví jednoho hesla, avšak antonymum ve výkladu významu má pouze jednu grafickou podobu, např. varianty **antithese**, **antiteze**, **antithesa**, **antitéza** mají ve výkladu v SSJČ opak *teze*; přitom společně lemmatizované varianty **teze**, **these**, **thesis** ve výkladu antonyma nemají. Podle dokladů nalezených v korpusu SYN¹² byla v LDB Pralex založena také hesla **téze** a **antitéze**. Přes Seznamy antonym jsme propojili dvojice: **antithese** x **these**, **antiteze** x **teze**, **antitéze** x **téze**. U hesel **antithesa** a **thesis** jsme na Seznam antonym nezapsali nic, protože LDB neobsahuje hesla, která by svou grafickou podobou vyhovovala jako antonymní dvojice k těmto lemmatům. Rovněž nebyl důvod taková hesla zakládat vzhledem k tomu, že nejsou registrována v žádném výkladovém slovníku češtiny, ani pro ně neexistují žádné doklady v korpusu SYN. Zároveň bylo nutné upravit definice u hesel **antithesa** a **thesis** – ve výkladu jsme smazali opak s jinou grafickou podobou¹³.

7 SSJČ: **aerob** [-a-e], -u m. (6. j. -u) (z řec. zákl.) biol. *organismus, kt. potřebuje k životu vzdušný kyslík; erobiont* (op. *anaerob*).

8 SSJČ: **aerobiont**-u m. (6. j. -u) biol. *aerob*.

9 SSJČ: **anaerob** [an-a-e-], **anaerobiont**, -u m. (6. j. -u) (z řec. zákl.) biol. *organismus nepotřebující k životu vzdušný kyslík* (op. *aerob*).

10 SSJČ: **velkodůl** (*vele- Maj.), -dolu m. (6. j. -e, řidč. -u) *důl velkého rozsahu, s velkou těžbou* (op. *malodůl*).

11 SSJČ: **velkovýroba** (*vele- Staš.), -y ž. *výroba ve velkém množství a rozsahu* (op. *malovýroba*).

12 Od roku 2011 byl původní korpus SYN přejmenován na *orig_syn*; *Český národní korpus – SYN*. Ústav Českého národního korpusu FF UK, Praha. Cit. 22.02.2011, dostupný z WWW: <<http://www.korpus.cz>>.

13 Stejný případ byl podobně vyřešen i u variantních hesel **syntéza/synthesa/synteze/synthese**.

V souvislosti s antonymy bylo potřeba u některých variant založit do hesláře nová hesla, aby se doplnily opozitní dvojice odpovídající si svou grafickou podobou. V LDB Pralex byly např. již založeny varianty **amatérismus/amatérizmus**, avšak u opozita **profesionalismus** chyběla varianta **profesionalizmus**, která nebyla registrována v SSJČ. Založili jsme tedy heslo **profesionalizmus** podle Nového akademického slovníku cizích slov (NASCS), aby vznikla antonymní dvojice **amatérizmus x profesionalizmus**. U hesla **profesionalizmus** jsme upravili výklad ze SSJČ na ‚(op. amatérizmus)‘, u hesla **amatérizmus** na ‚(op. profesionalizmus)‘.

9. Antonyma u víceslovných lexikálních jednotek

Zpracováním antonym u víceslovných lexikálních jednotek se zatím v LDB Pralex souhrnně nezabýváme. (Tento typ hesel dlouhodobě zpracovává Z. Opavská.) Někdy se ale vyskytne přechodný typ, kdy je u jednoslovného hesla (např. **prvočíslo**) ve výkladu v SSJČ uveden opak, kterým je víceslovná jednotka (**číslo složené**). Protože zatím nebylo rozhodnuto, zda toto spojení bude uvedeno v hesláři, zapsali jsme tento opak pouze na Seznam antonym, aby byl registrován.

Literatura

- ČERMÁK, František: Paradigmatika a syntagmatika slovníku: možnosti a výhledy. In: *Manuál lexikografie*. Eds. F. Čermák – R. Blatná. Jinočany: H&H 1995, s. 90–115.
- ČERVENÁ, Vlasta: K lexikálním kontrastům v české slovní zásobě. In: NŘ, 1979, roč. 62, č. 3, s. 240–244.
- ČERVENÁ, Vlasta: K slovtvorným adjektivním kontrastům. In: NŘ, 1979, roč. 62, č. 5, s. 240–244.
- DOLNÍK, Juraj: Antonymá z logického a lingvistického hlediska. In: *Slovenská reč*, roč. 53, č. 2, 1988, s. 83–93.
- DOLNÍK, Juraj: *Lexikológia*. Bratislava: Univerzita Komenského 1990, s. 176–190.
- Encyklopedický slovník češtiny*. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: NLN 2002. 604 s.
- FILÍPEK, Josef: Systémovost antonymie. In: NŘ, 1994, roč. 77, č. 3, s. 124–126.
- FILÍPEK, Josef – ČERMÁK, František: *Česká lexikologie*. Praha: Academia 1985. 281 s.
- Frekvenční slovník češtiny*. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s.
- Krátký slovník slovenského jazyka*. Red. J. Kačala – M. Pisárčiková. Bratislava: Veda 1987. 587 s.
- Nový akademický slovník cizích slov A–Ž*. (J. Kraus a kol.) 1. vyd. Praha: Academia 2005. 880 s. (NASCS)
- ORAVCOVÁ, Anna: Antonymá vo výkladovom slovníku. In: *Slovenská reč*, 1992, roč. 57, č. 5, s. 298–303.

- PALA, Karel – VŠIANSKÝ, Jan: Slovník českých synonym. 3., doplněné vyd. Praha: NLN 2000. 479 s.
- PISÁRČIKOVÁ, Mária: Antonymia. In: Horecký, J. – Buzássyová, K. – Bosák, J. a kol.: Dynamika slovné zásoby súčasnej slovenčiny. Bratislava: Veda 1989, s. 333–337.
- Příruční slovník jazyka českého. 9 sv. Praha: Státní nakladatelství – SPN 1935–1957. (PSJČ)
- Slovník českých synonym a antonym. Red. autorský kolektiv pracovníků *Lingea*. 1. vyd. Brno: Lingea 2007. 573 s.
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)

Zpracování názvů osob a jmen přechýlených v lexikální databázi Pralex¹

Helena Pernicová

Ústav pro jazyk český AV ČR, v. v. i.

Processing of Personal Names and Gender Inflected Names in the Pralex Lexical Database

In the Pralex lexical database, male nouns and their derived feminine forms are treated as individual entries; it means the principle of nesting is not applied. The definition of the derived feminine noun consists of two parts: a purely word-formative part (“derived from + masculine form”) and an explanation. This explanatory part is based on the explanation of the masculine form and appropriate changes of the metalanguage are carried out. In the documentary part of an entry we also pay attention to the frequency of feminine nouns related to their inclusion in dictionaries.

1. Úvod

Osoby se v češtině většinou pojmenovávají ve shodě s přirozeným rodem (názvy osob mužského pohlaví jsou rodu mužského, názvy osob ženského pohlaví jsou rodu ženského). Mezi těmito pojmenováními existuje vzájemný vztah – ve většině případů vytvářejí rodové dvojice. Tyto rodové dvojice jsou buď povahy slootovorné (*řidič – řidička*, *Novák – Nováková*), nebo lexikální (*muž – žena*, *bratr – sestra*). Některé názvy osob jsou tzv. jména vespólná, která bez ohledu na rod označují osoby obou pohlaví, a rodové dvojice tedy netvoří (*závistivec*, *blázen*; *příšera*, *bábovka*; *třeštidlo* aj.)².

Pokud se jedná o **slootovornou dvojici**, primární bývá obvykle pojmenování mužského jedince, pojmenování ženské je od něj odvozeno. V tom případě hovoříme o přechylování. Za nejběžnější jsou v odborné literatuře považovány přechylovací přípony *-ka*, *-ice*, *-yně*.

V ojedinělých případech je základním členem dvojice pojmenování ženské, z kterého je odvozeno pojmenování mužské (*vdova – vdovec*, *pradlena – prادلák*, *hokyně – hokynář*)³. Protože takto vzniklých mužských názvů osob je málo, jsou tato pojmenování pod tlakem systému často pocitována jako primární. Tento tlak může mít i slootovorné důsledky, tendenci dotvořit i k těmto maskulinům další ženský protějšek (*hokyně – hokynář – hokynářka*; *baletka – baleták – baletačka*).

Od mužských podob **zpodstatnělých přídavných jmen** se ženské názvy tvoří konverzí (*vrátný – vrátná*, *účetní m. – účetní ž.*, *rozhodčí m. – rozhodčí ž.*); ojediněle vznikají ženské podoby adjektivní povahy i sufixací z podstatných jmen (*švagr – švagrová*, *správce – správcová*).

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 V Encyklopedickém slovníku češtiny (2002, s. 205) se uvádí: „Po stránce gramatické jsou to většinou mask. (*sobec*, *lenoch*, *tupec*, *hrubec*), řidčeji fem. (*fiflena*, *troska*) n. neutra (*trdlo*, *fintidlo*, *nemehlo*).“

3 Srov. též *Tvoření slov v češtině* 2, s. 551.

Ze slovtvorného hlediska bývá vztah mezi členy těchto dvojic považován za vztah modifikační, to znamená za doplnění významu motivujícího slova o určitý znak (Dokulil 1962, s. 46).

2. Zpracování názvů osob ve výkladových slovnících

Názvy osob jsou ve výkladových slovnících zpracovávány podle příslušných zásad, stanovených pro zpracování substantiv; mezi jednotlivými slovníky nejsou ve způsobu zpracování zásadní rozdíly. Odlišnosti se však vyskytují u přechýlených ženských názvů.

Ve *Slovníku spisovného jazyka českého* (dále SSJČ) je převážná většina ženských přechýlených názvů osob přihnízďována k maskulinu, od kterého jsou odvozeny, a to formou uvedení v závorce za mužskou podobou. Samostatné heslo mají ženské názvy, u nichž bylo jako zvláštní význam zpracováno označení manželky osoby mužské (*učitelka, doktorka, lékárnice, kovářka, chalupnice, domkářka*), a ty názvy, jejichž významová struktura je něčím odlišná (*farářka, kněžka, rodička*) nebo bohatší (*kuchařka, milovnice, dojička*), tedy rozdílná od odpovídajícího pojmenování mužského. Samostatně s plným výkladem jsou v tomto slovníku zpracovány také nečetné ženské názvy, k nimž mužská podoba neexistuje (*švadlena*), je zcela ojedinělá (*šička*), nebo se tehdy neužívala (*modelka*).

Slovník spisovné češtiny pro školu a veřejnost (SSČ) postupoval obdobně, ale přihnízďované ženské názvy uvádí za středníkem na konci heslového odstavce, větší formou přípony (*šéfredaktor + -ka, provokatér + -ka, orientalista + -ka*), někdy je uvedena plná podoba (*tramvajačka*). V některých případech má i přihnízďovaná podoba začleněny své vlastní příklady užití (*šéfka, šéfová*). Samostatných hesel je méně, zejména proto, že význam pro manželky postupně vychází z úzu, a proto už nebyl zařazován. Plné zpracování mají např. hesla *dělnice* (bohatší struktura významu) nebo *chemička* (homonyma ‚žena‘ a ‚továrna‘).

Slovníky cizích slov (ASCS, NASCS, VSCS) používají stejně jako SSJČ metodu přihnízďování v závorce za mužskou podobou a uvádějí přechýlená pojmenování i u některých heslových slov, kdy v SSJČ byla registrována pouze mužská podoba (*alergička, introvertka, klarinetistka, kariéristka, paleontoložka, psychiatřička*).

Slovník neologismů 1, 2 (SN 1, SN 2) uvádí jména přechýlená jako samostatná hesla, naznačuje se zde vztah k základovému mužskému substantivu a uvádí se příklad užití (*florbalistka, -y* ž. přechýl. k florbalista: *vítězství českých florbalistek*). Výkladovou definici zde přechýlené podoby většinou nemají, pokud daná hesla neodkazují na mužskou podobu uvedenou už v prvním svazku.

3. Názvy osob v hesláři LDB Pralex

V lexikální databázi (dále LDB) Pralex, kde nejsme limitováni nedostatkem prostoru, jsme se rozhodli zpracovávat všechna ženská pojmenování jako samostatné položky databáze (hesla) s vlastním výkladem. Princip hnízdování neuplatňujeme

vůbec. Proto jsme se v rámci zpracování substantiv od května 2008 zaměřili též na samostatné zakládání ženských osob do hesláře, pokud tam již nebyly uvedeny podle SSJČ⁴, a jejich jednotné zpracování.

Kromě toho u každého mužského pojmenování, kde je přechylování potenciálně možné, ale ve slovnících nezachycené, ověřujeme doloženost ženské podoby v korpusu SYN⁵. Pokud nějaké doklady najdeme, založíme ženskou podobu. V případě, že přechýlené pojmenování je možné vytvořit a jeho existenci předpokládáme, ale v korpusovém materiálu jsme žádné doklady nenašli, zaznamenáme tuto skutečnost do poznámky u mužského substantiva a ženskou podobu zatím samostatně nezakládáme. (Tuto poznámku o nezaložené ženské podobě máme např. u hesel *onomastik*, *uměnovědec*.)

Soubor názvů osob, který byl součástí prvotního celku substantiv programově vložených do naší databáze z *Frekvenčního slovníku češtiny* (dále FSČ) a SSJČ, se snažíme postupně rozšiřovat, naší snahou je úplnost alespoň v některých sémantických skupinách. Pokud se nám např. z výchozích zdrojů dostal do hesláře pouze název vědního oboru a pojmenování pro osoby věnující se tomuto oboru chybělo, pak – je-li v korpusu doloženo, zakládáme také lemmata pojmenovávající příslušného odborníka a odbornici v daném oboru (*religionistika* – *religionista*, *religionistka*). Stejně tak dbáme na to, aby v případě hesel, u nichž byl z referenčních zdrojů převzat výklad typu „odborník v...“, byl v našem hesláři rovněž obsažen i název příslušného oboru, proto je systematicky doplňujeme. V současné fázi zpracování jsou takto kompletní např. trojice *imunologie* – *imunolog* – *imunoložka*; *homeopatie* – *homeopat* – *homeopatka*; *muzikologie* – *muzikolog* – *muzikoložka*. Stejným způsobem kompletujeme i některé názvy sportů a osob, které se jim věnují (*karate* – *karatista* – *karatistka*; *florbal* – *florbalista* – *florbalistka*; *triatlon* – *triatlonista* – *triatlonistka*), hudebních nástrojů a hráčů / hráček na tyto nástroje (*banjo* – *banjista* – *banjistka*; *klávesy* – *klávesista* – *klávesistka*) apod.

4. Homonymie a variantnost u názvů osob

Některé názvy osob v naší databázi jsou členy homonymních dvojic nebo skupin. Ty byly buď zpracovány jako samostatná homonyma už v SSJČ, např. *protektor* -a m. ‚vlivná osoba poskytující pomoc, přímluvu, ochranu‘ – *protektor* -u, m. (u pneumatiky) obnovený během ojetého pláště, nebo tam byly uvedeny v rámci sdružené heslové stati (*kráječ* -e m. (živ.) ‚kdo odborně provádí krájení‘ – *kráječ*, -e m. (neživ.) ‚nůž n. strojek na krájení‘). V LDB je uvádíme jako samostatné databázové položky. Význam pro osobu označujeme římskou jedničkou, význam pro nástroj (nebo jiný neživotný název) římskou číslicí dvě.

Stejně postupujeme i v případě oprávněného rozdělení polysémního hesla, tj. osamostatnění jednoho z významů a jeho vyčlenění zvlášť jako samostatné hes-

4 Výchozíkem pro heslář LDB Pralex se stal FSČ a SSJČ, viz zde podrobněji ve stati J. Světlé *Budování hesláře LDB Pralex*.

5 Od roku 2011 byl původní korpus SYN (který slučuje korpusy SYN2000, SYN2005 a SYN-2006PUB) přejmenován na *orig_syn*.

lo. Například heslo *antagonista* zpracované v SSJČ jako ,1. protivník, protichůdce, 2. anat. sval, který působí opačně než druhý‘ jsme takto rozdělili na *antagonista I* (živ.) a *antagonista II* (neživ.), heslo *partyzánka* ,1. k partyzán 2. levnější druh cigaret zn. Partyzánka‘ je rozděleno na *partyzánka I* ,přechýl. k partyzán 1; členka bojové organizace k vedení záškodnické války v týlu nepřítele‘ a *partyzánka II* ,levnější druh cigaret zn. Partyzánka‘.

Jiným typem homonymické dvojice v našem hesláři je pár, kdy jedním členem je pojmenování osoby, někdy mužské, častěji ženské, a druhým univerbizovaná podoba slovního spojení, které je dnes často ustálené povahy. Takové dvojice byly ojediněle zachyceny už v SSJČ (*měšťanka* ,žena‘ x ,měšťanská škola‘, *chemička* ,žena‘ x ,chemická továrna‘, *civil* ,člověk‘ x ,civilní, nevojenský život; občanský, nevojenský oblek‘), ve větší míře ale tento typ homonym v předchozích slovnících registrován nebyl, často proto, že univerbáty se vzhledem ke stylovému hodnocení do slovníků dříve příliš nezařazovaly. Další dvojice jsme tedy založili až na základě korpusu SYN (*ekonomka II* ,ekonomická škola‘, *čerpačka II* ,čerpací stanice‘, *státnice II* ,státní zkouška‘).

Jako homonyma rovněž označujeme mužskou a ženskou podobu zpodstatnělých přídavných jmen i další slovní druh – adj. nebo zájmeno (*účetní I* ,kdo vede odborně účetnictví‘ – *účetní II* ,přechýl. k účetní I; žena, která vede odborně účetnictví‘ – *účetní III* adj.; u zájmen: (*můj I* ,můj manžel‘ – *můj II* zájmeno, *jeho I* ,jeho manželka‘ – *jeho II* zájmeno).

Variantnost se u názvů osob projevuje podobně jako u ostatních substantiv a nemá žádná specifika, ale je v našem souboru zastoupena poměrně často. Zpracování probíhá podle našich zásad pro zpracování variant.

5. Práce s výkladovou definicí

Pro výklad **názvů mužských osob** používáme definice ze SSJČ, případně z dalších výkladových slovníků, včetně CSN a SN 1, SN 2. Tento výklad u mužských podob neměníme. Pokud se nám však zdá výklad v některém z dalších výkladových slovníků výstižnější, použijeme i jej nebo na něj upozorníme v poznámce. Pokud byl výklad významu osoby uveden např. politicky tendenčním označením „v třídní společnosti“ (*lidumil, sluha*) nebo „dř.“ (*starosta, radní*), takové označení z výkladu odstraňujeme.⁶

Přechýlená ženská pojmenování zpracováváme od roku 2008, a to jednotně jako ucelený soubor. Naše definice jmen přechýlených se skládá ze dvou částí – čistě slovotvorné a výkladové. Slovotvorná část má formu „přechýl. k + mužská podoba“ a od druhé části, vlastního výkladu, je oddělena středníkem. Tato formální část výkladu má praktický význam – umožňuje souhrnné vyhledávání přechýlených pojmenování v naší databázi, a zároveň je nám oporou pro zápis na seznam Slovotvorba, který slouží k hypertextovému propojení relevantních hesel (viz podrobněji dále).

6 Zpracování mužských názvů osob probíhalo v souladu s našimi zásadami pro zpracování substantiv a účastnili se jej od počátku všichni zpracovatelé podle přidělených písmen. Revize tohoto souboru (všech apelativních názvů osob) však již probíhá jednotně pro celé dvojice – mužské i ženské podoby.

Formu výkladu jsme sjednotili tak, že všude, kde to bylo možné, jsme ji opřeli o výklad mužského základového substantiva uvedený v SSJČ nebo jiném referenčním zdroji. To znamená, že v druhé části definice upravujeme výchozí podobu převzatou ze SSJČ (nebo jiného výkladového slovníku) tak, aby metajazyk výkladu odpovídal novému heslu. Zejména měníme výraz „kdo“ nebo „osoba“ na výraz „žena“ a rovněž v závislosti na konkrétní formulaci původního výkladu upravujeme i koncovku slovesa, případně dalších slovních druhů (přídavných jmen, vztažných zájmen): *cvičitel* ‚kdo je pověřen vedením výcviku‘ – *cvičitelka* ‚přechýl. k cvičitel; žena, která je pověřena vedením výcviku‘. Pokud je ve výkladu mužského substantiva použito jako opěrné slovo odborník, pracovník, zaměstnanec apod., používáme jeho přechýlenou podobu, např. *dietoložka* ‚přechýl. k dietolog; odbornice v dietologii‘; *kuchařka* ‚přechýl. ke kuchař 1; zaměstnankyně obstarávající přípravu pokrmů‘.

Pokud je mužské základové slovo polysémní, zakládáme a zpracováváme u ženských protějšků pouze ty významy, pro které byly v našem korpusu nalezeny doklady. Formou číselného odkazu v první části naší definice jmen přechýlených zaznamenáváme, ke kterému z významů opěrného maskulina se výklad vztahuje (*lukostřelkyně* ‚přechýl. k lukostřelec 2; *sportovkyně pěstující lukostřelbu*‘). V odůvodněných případech zakládáme i význam sice dnes nedoložený, ale z hlediska vývoje dalších významů základní (a označíme jej jako nedoložený v korpusu SYN).

K září 2010 již bylo v LDB Pralex zpracováno přes 3 000 hesel s označením „přechýl“. Nejvíce jsou zastoupeny ženské názvy s příponou *-ka* a jejími variantami. Mnohem méně jsou zastoupeny ostatní přechylovací přípony: např. *-ice* 410x a *-yně* 236x. Zcela ojediněle se objevuje přechýlení příponou *-ička* (*pediatřička, psychiatřička, foniatrička, geriatrička*). Uvedený celkový počet jmen přechýlených znamená, že máme i stejný počet jejich mužských protějšků. Další mužské názvy, ke kterým nebyly ženské protějšky uvedeny v SSJČ, nebo k nim nebyly doplněny příslušné ženské protějšky v LDB Pralex (na základě analýzy korpusu SYN), nelze zatím v rámci daného souboru podchytit, protože nemáme jak je vyhledat.

Jak již bylo uvedeno výše, přechýlená pojmenování a jejich mužské protějšky v LDB Pralex hypertextově propojujeme mezi sebou (prostřednictvím seznamu Slovtvorba). Toto propojení nám umožňuje současné otevření obou hesel (členů dvojice), jejich paralelní zpracování i jejich vzájemné porovnávání mezi sebou, např. z hlediska frekvence výskytu, významové struktury, výkladu významu, šíře spojitelnosti apod. V případě dvojic, u nichž se může jednat o tzv. opačné přechýlení, se zatím přidržujeme pohledu SSJČ, pokud tam daná dvojice byla zachycena (*prostitut – prostitutka*); později nebo zcela nově registrované dvojice (*model I – modelka, striptér – striptérka*) ponecháváme zatím bez propojení (a bez označení „přechýl.“).

6. Členění na významy

Členění na významy převzaté ze SSJČ někdy neodpovídá dnešní situaci – doklady získané z korpusu SYN nás často vedou k novému či jinému rozčlenění významové struktury. „Potenciální“ další významy bývají někdy v předchozích slovnících ve výkladu naznačeny jako významový odstín. Jestliže nám povaha a množství dokladů

ukazují vývoj směrem k samostatnému významu (lexikalizaci), pak původní význam rozdělujeme.

U názvů osob jsme takto někdy vyčlenili význam týkající se sportovní oblasti (např. u hesel *chodec*, *zápasník*, *útočník*, *obránce*). Srov. např. v SSJČ heslo *chodec* ‚kdo jde, chodí někam, koná cestu pěšky; sport. kdo pěstuje chůzi jako sport‘ – v LDB je zpracováno takto: V1 ‚kdo jde, chodí někam, koná cestu pěšky‘, V2 ‚závodník v chůzi‘. Vodítkem nám byla především bohatá a různorodá spojitelnost výrazu ve sféře sportu, a také již existující obdobný způsob zpracování v SSJČ u jiných hesel (*běžec* ‚1. kdo běží‘, 2. ‚závodník v běhu‘) nebo v SSČ (např. *plavec* ‚1. kdo umí plavat‘, 2. ‚závodník v plavání‘).

Stejným způsobem jsme jako zvláštní význam v několika případech vymezili také ‚název příslušníka určitého druhu vojska‘: *chemik* ‚1. odborník v chemii‘, 2. ‚příslušník chemického vojska‘, *průzkumník I* ‚1. kdo nebo co provádí průzkum‘, 2. ‚příslušník průzkumné jednotky‘.

Nové významy názvů osob zakládáme také v řadě jiných případů. Oporou pro výklad nám často bývá zpracování neosémantismů v SN 1 nebo SN 2, jako v případě hesla *populista*, kdy SSJČ ve výkladu tohoto hesla uvádí *populista* ‚1. stoupenec populismu, 2. (v letech 1922–1951) příslušník řecké pravice strany‘ a SN 1 přidává nový význam: *populista* ‚politik praktikující populismus‘.

Pro řadu nově vymezených významů ale nenajdeme oporu v žádném námi používaném referenčním zdroji. V takovém případě se opíráme o vlastní jazykové povědomí, případně vyhledáváme informace v encyklopedických zdrojích a výkladovou definici vytváříme sami. V LDB Pralex jsme takto postupovali například v případě výkladu nového významu hesel *minimalista* ‚2. stoupenec minimalismu (ve výtvarném umění a hudbě)‘ a *absolutista* ‚3. kdo je maximalistický ve svých požadavcích‘.

U některých názvů osob jsme řešili i otázku mluvnického čísla. Výrazy jako *cementář*, *dřevař*, *betonář* aj. dnes většinou už neoznačují individuálního pracovníka, ale podnik, firmu a používají se souhrnně v množném čísle. Toto nové užití řešíme založením a zpracováním dalšího významu (s uvedením příslušného tvaru lemmatu do podheslí: *cementáři*, *dřevaři*, *betonáři*).

7. Příkladová část hesla a doloženost v korpusu

Při práci s korpusem zjišťujeme, že u hesel založených podle SSJČ je poměrně hodně případů **nedoložených** pojmenování osob, jedná se často o povolání, profesu a řemesla, které dnes již neexistují, a proto nejsou v korpusu SYN doloženy ani v mužské, ani ženské podobě (*gilošér* – *gilošérka*; *maltař* – *maltařka*; *moukař* – *moukařka*). Pro některá z mizejících nebo v praktickém životě neprovozovaných řemesel máme ale přesto doklady, protože se o nich píše v souvislosti s řemeslnými jarmarky a trhy, kde bývají prezentována.

Někdy je ve dvojici muž – žena **mužská podoba doložená**, nebo i běžná, ale **ženská podoba** je **nedoložená** (*koželužka*, *lamačka*, *mlatka*), případně je její výskyt omezen na 1 až 3 doklady (*klempířka*, *hornice*, *tesařka*). V uvedených případech je

zřejmé, že více dokladů nelze očekávat, pokud se neobjeví v publicistice nebo umělecké literatuře. Označení „nepodařilo se nedoložit v syn“ má však v naší databázi i poměrně početná skupina ženských názvů osob, kde nulový výskyt může být způsoben omezeností našeho korpusového zdroje a později bude třeba dohledat doklady v jiných zdrojích (*filharmonička, hodinářka, kolařka, křížovkářka, radiotechnička* aj.). Sem patří i skupina mužských názvů pravděpodobně netvořících mužskou a ženskou dvojici, např. některé odborné termíny, zejména z právní a finanční oblasti (*smírce, příkazce; zmocnitel, komitent, mandatář*).

Opačným případem je situace, kdy pojmenování osoby bylo do naší databáze vloženo na základě materiálových zdrojů (z FSC nebo doplněno z korpusu SYN), ale v SSSJ uvedeno nebylo (někdy není uvedeno ani v žádném jiném námi používaném výkladovém slovníku). Jde zejména o velké množství dosud nikde nezachycených ženských názvů. Ty je možné rozčlenit do několika sémantických okruhů: sport (*motýlkářka, bobistka, boxerka*), hudba (*hobojistka, basistka, bluesmanka*), profese, zejména intelektuální povahy (*biofyzička, auditorka, knihkupyně*), církevní oblast (*akolytka, jáhenka, biskupka*), méně často jde o celé dvojice výkladovými slovníky dosud nepodchycených názvů osob (*rozehrávač – rozehrávačka, dráhař – dráhařka, nadhazovač – nadhazovačka*).

Některá jména mají sice v SSSJ přihnízdovanou ženskou podobu, ale ta bývá často zastaralá a dnes už nedoložená (*lakomec – lakomnice, hlupec – hlupkyně, lenoch – lenošnice*), mužská podoba je chápána jako vespolečné podstatné jméno. Někdy jsme však tuto přihnízdomanou ženskou podobu v SYN našli (*vášnivec – vášnivka, zvědavcec – zvědavka, zuřivec – zuřivka, žárlivec – žárlivka*). Ojedinele jsme z korpusu SYN (případně z vlastního povědomí a z internetu) zachytili i další přechýlené podoby ke jménům označujícím nositele vlastnosti. Ty mají obvykle expresivní ráz (*blbka, hňupka, ošklivka, náruživka*).

Literatura

- ADÁMKOVÁ, Petra: Přechylování versus přímé tvoření názvů osob ženského pohlaví. In: Život s morfémy. Sborník studií na počest Zdenky Rusínové. Eds. P. Karlík J. – Pleskalová. Brno: Masarykova univerzita 2004, s. 9–17.
- ČMEJRKOVÁ, Světlá: Rod v jazyce a komunikaci: specifika češtiny. In: Slovo a slovesnost, 2002, roč. 63, č. 4, s. 263–286.
- Encyklopedický slovník češtiny. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: Nakladatelství Lidové noviny 2002. 604 s.
- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSC)
- Nová slova v češtině 1. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 1998. 356 s. (SN 1)
- Nová slova v češtině 2. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 2004. 568 s. (SN 2)
- Nový akademický slovník cizích slov A–Ž. (J. Kraus a kol.) 1. vyd. Praha: Academia 2005. 880 s. (NASCS)

- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- Tvoření slov v češtině 2. Odvozování podstatných jmen. Red. F. Daneš – M. Dokulil – J. Kuchař. Praha: Academia 1967. 780 s.
- Velký slovník cizích slov. Verze slovníku 2.0. Red. J. Kraus – V. Petráčková. Red. spolupracovaly S. Matoušková, V. Holubová. Voznice: Leda spol. s r. o. 2005. (CD-ROM). (VSCS)

Ke zpracování regionálních lexikálních jednotek v lexikální databázi Pralex¹

Hana Golářová

Ústav pro jazyk český AV ČR, v. v. i.

On Processing of Regional Lexical Units in the Pralex Lexical Database

The article deals with the issue of regional lexical units (LU) processed in the Pralex lexical database (LDB), specifically with the processing of regionally marked lexical units in the rising dictionaries while defining regionalisms of different type in the Czech tradition and in the Pralex LDB. It also emphasizes the problems occurring during the processing of this type of LU in the LDB. Furthermore, it examines the shifts in our evaluation of regional LU in comparison with the previous dictionaries (*Dictionary of the Standard Czech Language, Dictionary of Standard Czech, Academy Dictionary of Loanwords in Czech*).

1. Úvod

Oddělení současné lexikologie a lexikografie AV ČR, v. v. i., od r. 2005 pracuje na výzkumném záměru *Vytvoření lexikální databáze českého jazyka počátku 21. století*. Při zpracovávání lexikálních jednotek (dále LJ) v lexikální databázi (dále LDB) Pralex navazujeme na lexikografickou tradici, ale v porovnání s tradičním výkladovým slovníkem má databázový popis slovní zásoby svá specifika a dovoluje nám lépe zpracovávat lexikum jako systém; máme možnost vyhledávat konkrétní jazykové jevy jako celek, můžeme se tedy podrobněji zabývat různými tématy. Náš příspěvek se zaměřuje na problematiku územní příznakovosti, a to zejména u regionálních lexikálních jednotek zpracovávaných v LDB Pralex. Již na začátku je třeba připomenout, že kromě regionalismů a dialektismů se v LDB Pralex u žádných hesel podrobně nezabýváme stylovými charakteristikami, to bude až úkolem budoucího slovníku. Tím se tato skupina LJ odlišuje od zpracování ostatních typů substantiv.

2. K pojetí územní příznakovosti ve slovnících

Teritoriálně diferencované lexikální jednotky jsou příznačné hlavně pro mluvený jazyk, ale díky publicistice a umělecké literatuře se objevují také v písemných jazykových projevech oficiálních i neoficiálních. Podobně se i ve výkladových slovnících češtiny nacházejí výrazy vymezené územně – jedná se o tzv. nářeční výrazy (dialektismy) a oblastní výrazy (regionalismy). Výrazy zařazované do kategorie obecné češtiny necháváme v tomto příspěvku víceméně stranou².

1 Příspěvek vznikl v rámci výzkumného záměru UJČ AV ČR, v. v. i., *Vytvoření lexikální databáze českého jazyka počátku 21. století* (AV0Z90610521, 2005–2011), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Zmíněná vrstva slovní zásoby se původně vyvinula ze středočeského dialektu. V lingvistice bývá interpretována různě: buď jako interdialekt český, nebo také jako vrstva jazykových prostředků užívaných obecně, tj. na celém území České republiky. Někdy se obě pojetí obecné češtiny prolínají. V stávajících výkladových slovnících češtiny se uplatňuje druhé pojetí obecné češtiny, tzn., že se jedná o jazykové prostředky, které se užívají na celém českém území.

Než se dostaneme k úvahám o LJ s regionálním příznakem, vzpomeneme stylistické charakteristiky, kterých užívaly dosavadní, zejména výkladové slovníky češtiny k označení LJ s teritoriálním hodnocením. Nejstarší jednojazyčný výkladový slovník, *Příruční slovník jazyka českého* (dále PSJČ), používal následující zkratky: *dial.* (dialektický)³, *mor.* (moravský), *ob.* (obecný), *lid.* (lidový) a v předvýkladu označení (*na Moravě*)⁴.

Slovník spisovného jazyka českého (dále SSJČ), z něhož při databázovém zpracování češtiny nejvíce čerpáme, označoval nářeční výrazy kvalifikátorem *nář.* a obecné výrazy, stejně jako PSJČ, stylovým hodnocením *ob.* (obecný). **Podrobněji než v PSJČ byl v SSJČ propracován systém značení regionálních výrazů.** Pro oblastní LJ bylo v SSJČ zavedeno označení *obl.* (např. *cácorka 1* ‚konipas‘; *grundle* ‚mřenka‘), kromě toho bylo možno provést přesnější lokalizaci LJ a užít kvalifikátorů *obl. čes.* (oblastní český, např. *drančit* ‚velmi rychle jít n. jet‘) a *obl. mor.* (oblastní moravský, např. *dědina 1* ‚vesnice‘). Dále se někdy oblastní varianty LJ doplňovaly předvýkladem či dovýkladem: (*na Moravě*), např. *stolař I* ‚truhlář‘; a byly víceméně považovány za spisovné⁵. Mimo uvedená značení pro regionální LJ byla v SSJČ nalezena i značka *ob. (mor.)*⁶, mající význam pravděpodobně obecný výraz moravský, např. *stolařka 2* ‚žena truhláře‘; *stolařina*.

Následující slovníky – *Slovník spisovné češtiny pro školu a veřejnost* (dále SSČ)⁷ a *Akademický slovník cizích slov* (dále ASCS)⁸ – v podstatě navázaly na stylové charakteristiky a značky teritoriálních LJ, které byly zavedeny v SSJČ.

3. K definování územně příznakových lexikálních jednotek

Při zpracovávání slovní zásoby s územní charakteristikou narážíme na jeden zásadní problém, a sice jak rozlišit mezi LJ regionálními a dialektickými. Jaké je pojetí regionálních LJ v české lingvistice?

Rozlišení dialektismu a regionalismu se neobejde bez obtíží, rovněž v definicích těchto pojmů bývají posuny v hodnocení⁹. Např. *Encyklopedický slovník češtiny* (dále ESČ) definuje *regionalismus* následovně: „Z hlediska územního rozšíření bývají v rámci d. [dialektismů, pozn. autora] *vyčleňovány regionalismy (oblastní výrazy), tj. d. roz-*

3 Máme na mysli dialektický výraz nebo význam LJ, totéž platí u dalších kvalifikátorů LJ.

4 Zda je v PSJČ heslo s předvýkladem (*na Moravě*) chápáno jako spisovné, nebo nespisovné, není zcela patrné. Pojetí spisovnosti a nespisovnosti nebylo v tomto výkladovém slovníku češtiny podrobně propracováno. Něco málo o tom v předmluvě (1937), srov. PSJČ, 1. díl, s. 7–11.

5 SSJČ, 1960, 1. díl, s. 10.

6 Je možné, že výčet značek užívaných v SSJČ pro regionální výrazy není vyčerpávající.

7 V SSČ je pojetí LJ s regionální charakteristikou poněkud vágní. Vztah *obl.* výrazů ke spisovnosti není zcela zřejmý. Doplnující výklad (*na Moravě*) se v SSČ u regionálních LJ nevyskytuje, LJ jsou označovány pouze jako *obl.*, *obl. čes.* nebo *obl. mor.* (i případ *stolař*), srov. SSČ (1. vydání, 1978, s. 794 a CD, Leda, 2005, kap. 7).

8 V ASCS není také vztah oblastních LJ ke spisovnosti a nespisovnosti příliš jasný. Kromě značek *nář.*, *obl.* a *obl. mor.* se v ASCS u územně příznakových LJ užívá doplňujícího výkladu: *gajdy* (*na vých. Moravě*).

9 Definice dialektismu a regionalismu uvádějí v české tradici PSJČ, SSJČ, SSČ, Příruční mluvnice češtiny (1995, PMČ), ESČ, Manuál lexikografie (1995) a Současná česká stylistika (Čechová, M. a kol. 2003).

šířené přes hranice lokálních dial., příp. interdialektů.”¹⁰. Tento dialektologický pohled na územně příznakové LJ se ovšem kříží s lexikografickým zpracováním centrální slovní zásoby češtiny.

Při rozlišení dialektismu a regionalismu hrají důležitou roli tři aspekty: vztah ke geografickému rozšíření LJ, vztah ke spisovnosti a vztah k vrstvám slovní zásoby. Bere-li v úvahu tyto tři aspekty, nemůžeme z hlediska pojetí územně příznakových jednotek v LDB Pralex hovořit o regionalismech jako o podmnožině dialektismů. Vzhledem ke spisovnosti stojí totiž regionalismus nad dialektismem, protože některé z regionálních výrazů jsou nebo mohou být považovány za spisovné. Kromě toho stojí regionalismus nad dialektismem také tím, že má větší teritoriální rozšíření než úzce lokalizovaný dialektismus.

V rámci LDB Pralex rozumíme pod pojmem **regionalismus lexikální jednotku spisovné nebo nářeční povahy, která se vztahuje k větším územním celkům** (Čechy, Morava, Slezsko; Slezsko + východní Morava apod.), např. *kyselo I: území*¹¹ Čechy; *stolař: území* Morava, Slezsko; *dědina V1* ‚vesnice, osada‘: **region**. Morava, Slezsko; *bývat V2* ‚bydlet‘: **region**. Morava.

Dialektismem chápeme *lexikální jednotku nářeční povahy, která se vztahuje k menším územním celkům* (např.: středomoravská oblast; východomoravská oblast, slovácký úsek apod.), viz např. *bandor V2* ‚brambor‘: **dialekt**. severovýchodočeská oblast; *ogar* ‚chlapec‘: **dialekt**. východomoravská oblast.

Regionálními LJ je nutné se podrobně zabývat také z toho důvodu, že by v novém výkladovém slovníku češtiny měly mít převahu nad LJ dialektickými. Je to dáno jednak tím, že některé dialektismy přehodnocujeme na regionalismy, a také tím, že by do výkladového slovníku češtiny byly zařazovány nářeční LJ méně, než tomu bylo v SSJČ¹².

Soubor územně příznakových LJ zařazených do LDB Pralex bude obsahovat všechny LJ, které jsou v SSJČ označené stylovým hodnocením nář., obl., obl. čes. a obl. mor. (téměř 6 tisíc hesel). LJ se zakládají postupně a zpracovávají se jednotlivě, ale u každé z nich vyhodnocujeme, zda jde v našem pojetí o regionalismus nebo dialektismus, a kromě toho se některé přehodnocují i z hlediska spisovnosti a nespisovnosti.

4. K hodnocení regionálních LJ a k jejich zpracování v LDB Pralex

V elektronické databázi používáme označení zkratkou minimálně, protože máme dostatek místa. Pokud zpracováváme LJ s územním příznakem, můžeme je zařadit do jedné ze čtyř kategorií: **území**, **region.**, **dialekt.**, **region./dialekt.**, přičemž po-

¹⁰ ESČ, s. 110.

¹¹ Podrobnější charakteristikou LJ s územním příznakem v LDB Pralex a jejich zařazením do kategorií „území“, „region.“ a „dialekt.“ se zabýváme níže, v kap. 4.

¹² V SSJČ bylo nalezeno 5971 záznamů LJ s územně příznakovým označením (nář. 5275 LJ, obl. 592 LJ – v tom jsou zahrnuty i výrazy obl. čes. a obl. mor.). SSČ uvádí 27 oblastních výrazů a jen 4 výrazy nářeční (do těchto počtů jsou zahrnuty i přihnízdované lexémy). V nejnovějším slovenském výkladovém slovníku (*Slovník súčasného slovenského jazyka* (A–G), 2006, SSSJ) nejsou vůbec úzce nářeční LJ zařazeny a vyskytují se v něm pouze LJ regionální (označené jako *region.*) a etnografické (označené jako *etnogr.*) (SSSJ, s. 38).

slední možnost je naše pracovní zařazení, abychom mohli takto označené LJ vyhledat v naší databázi a teprve později se rozhodnout, do které kategorie je zařadíme.

U **regionálních LJ** se nabízejí dvě možnosti klasifikace: spisovnou LJ můžeme označit jako kategorii **území** a další údaj si vybereme z nabídky makroareálů (Čechy, Morava, Slezsko); nespisovnou LJ zařadíme do kategorie **region**. a k ní také přiřadíme údaj z nabídky makroareálů (Čechy, Morava, Slezsko). **Oproti starším slovníkům jsme do nabídky v heslovém formuláři doplnili i Slezsko a k tomu je možné připojit i údaj o nářeční oblasti (středomoravská oblast, východomoravská oblast apod.), pokud má výraz větší územní rozšíření**, např. *galán* V2 „nápadník, milý“: Slezsko + východomoravská oblast. Nejsme-li si jisti, zda se jedná o LJ regionální, či spíše dialektickou, pak můžeme LJ zařadit do kategorie **region./dialekt.**, srov. např. *gajdy*, *dudy*.

V některých případech je také dosti problematické rozhodnout, zda zařadit LJ mezi výrazy **spisovné** (v LDB Pralex k tomuto slouží kategorie **území**), či **nespisovné** (možnost zvolit kategorii **region**.). Jak s tímto problémem naložily stávající výkladové slovníky češtiny, o tom byla řeč výše.

Dialektismy označujeme kvalifikátorem¹³ **dialekt**. a kromě makroareálů a oblastí ještě máme možnost zařadit LJ do nářečního úseku, takže na rozdíl od předchozích výkladových slovníků může být v LDB Pralex uvedena přesná lokalizace nářeční LJ. Označení nářečního úseku ovšem používáme jen zřídka, protože tento údaj je vhodný spíše pro dialektology.

Výhodou databázového zpracování slovní zásoby je, že nám umožňuje přesnější třídění lexikálních jednotek. Můžeme si pracovní vytvořit různé kategorie pro naše potřeby. Například, zaznamenáme-li u LJ větší územní rozšíření, můžeme ji zařadit do kategorie **region./dialekt**. Tímto řešením vidíme zřetelněji přechodové pásmo, v němž se vyskytuje určitá skupina LJ s územním příznakem. Vznik této kategorie nám může podstatně pomoci při vyjasňování pojetí regionálních a dialektických LJ.

5. Příkladová část hesla u územně příznakových LJ v LDB Pralex

Při budování lexikální databáze Pralex je pro lexikografický popis jako materiálová základna využíván především korpus SYN¹⁴. V LDB Pralex pracujeme jednak s tzv. upravenými doklady, které řadíme do bloků na základě spojitelnosti a vazebnosti, a jednak s citátovými (větnými) doklady.¹⁵

Upravené doklady představují jakési kolokační schéma zpracovávané LJ. Toto kolokační schéma je dobré mít také u LJ s územní charakteristikou. Např. u oblast-

13 Termín „kvalifikátor“ jsme začali v LDB Pralex používat pro značky, které se v dřívějších slovnících vyskytovaly u LJ se „stylistickým/stylovým hodnocením/charakteristikou“. Inspiraci jsme našli ve slovenské lexikografii pracující s pojmem „funkční kvalifikátor“, např. viz Bosák (2000, s. 75–83), SSSJ (2006, s. 35).

14 Uvedený korpus (obsahuje korpusy SYN2000, SYN2005, SYN2006PUB) není přístupný veřejnosti. Od r. 2011 byl korpus SYN přejmenován na *orig_syn*.

15 Blíže ke zpracování exemplifikace viz zde ve stati M. Koutové.

ních LJ, které se pohybují na hraně mezi spisovností a nespisovností, míváme volnou řádkou odsazeny doklady, které jsou nespisovné (např. *brejle: slepecký brejle, babičkovský brejle* apod.)

Citátové (větné) doklady z korpusu SYN jsou pro nás užitečné z několika důvodů. U LJ s územní charakteristikou mohou sloužit k doplnění výkladu nebo k dokreslení jazyka, prostředí, ze kterého daný výraz pochází. Srov. např. v LDB Pralex heslo **kyselo I**, polévka z chlebového těsta, tzv. kvásku':

opus.rokvyd=1994,opus.txtype=POP,opus.id=zdravina Když kvásek začne fermentovat a kynout , zalijeme ho studenou vodou a dáme stranou . Procezený používáme na přípravu polévky < kyselo> nebo jako kvásek při pečení chleba .

opus.rokvyd=2002,opus.txtype=ENC,opus.id=krkonose Moc lepčí na tom byli tam , hde se deržali oukropu , nebo se spokojili s poliukou hop do vody a hde bylo < kyselo> s bandorem na denním pořádku .

Doloženost LJ s územním příznakem v korpusu SYN je různá: jednak se vyskytují případy lexikálních jednotek, jež mají dost výskytů a jsou zachyceny ve *Frekvenčním slovníku češtiny* (2004, FSČ) (*bečka, brejle, krám, stolař*), ale také LJ, které ani v korpusu SYN žádné doklady nemají a jsou doloženy jen v brněnském dialektologickém archivu (*banče* ‚house‘, *hejhal* ‚hejkal‘), nebo dokonce nebyly nalezeny ani v dialektologickém archivu (*binovač* ‚koledník s pomlázkou‘, *čermák II* ‚čerchman(t)‘). Můžeme říci, že převažují územně příznakové LJ s menším množstvím dokladů (v korpusu SYN do 10 výskytů). LJ označené v LDB Pralex kvalifikátorem dialekt. často v korpusu SYN vůbec doloženy nebývají.

Nižší frekvencí výskytů se tento typ LJ liší od jiných typů substantiv zařazených do LDB Pralex. Proto je nutné výskyt teritoriálních LJ (v SSJČ označovaných zkratkami nař. a obl.) ověřovat v příručkách a jiných referenčních zdrojích. Jinak jsou LJ s územním příznakem v LDB Pralex zpracovávány podle jednotných zásad pro zpracování substantiv.

6. Posuny v hodnocení regionálních LJ v LDB Pralex proti předchozímu stavu (SSJČ, SSČ, ASCS)

Ke změnám v popisu územně příznakových lexikálních jednotek proti hodnocení v předchozích slovnících (SSJČ, SSČ, ASCS) dochází především díky existenci pěti dílů *Českého jazykového atlasu* (1992–2005, dále ČJA). Tato příručka je dnes nepostradatelná při lexikografickém zpracovávání vrstev slovní zásoby spojených s územním příznakem. Při používání ČJA je však nutné brát v úvahu, že ČJA zpracovával výhradně mluvený jazyk na základě materiálu zachyceného přibližně před padesáti lety. Od té doby jsou možné další posuny, neboť teritoriální slovní zásoba je rovněž v pohybu. Dalším důležitým pramenem pro vyhodnocení příznaku u LJ je možnost ověřit si výskyt LJ v brněnském dialektologickém archivu, kde je dobře zachyceno územní rozšíření teritoriálních LJ. Při vyhodnocování územního příznaku u LJ tedy vycházíme z ČJA, z dialektologického archivu, z hodnocení LJ ve stávajících slovnících a ze současného úzu zjišťovaného zatím pouze v korpusu SYN.

K posunům v hodnocení LJ s územní charakteristikou dochází nejčastěji na ose od nářečních LJ, přes regionální LJ (nespisovné i spisovné) a obecněčeské LJ (substandardní) až po výrazy hodnocené dnes už jako neutrální. Posun na ose v opačném směru může nastat také, např. jde o přehodnocení LJ v SSJČ označené jako ob. na kvalifikátor územní (*dudák II* ‚dudlík‘), ale je méně častý. Je dán spíše tím, že dříve neexistoval ČJA a nebyl přístup do referenčních zdrojů, kde by byla možnost si územní výskyt LJ ověřit. V SSJČ mnohdy sloužil kvalifikátor *nář.* jako sběrný koš pro různé příznaky, podobně tomu bylo u kvalifikátoru *ob.*

V následující tabulce se zaměříme především na LJ s regionálním příznakem. Ukážeme si posun hodnocení LJ z příznaku nářečního přes označení oblastní do pozice neutrální; dále posun LJ z pozice neutrální do hodnocení LJ sice jako spisovné, ale s regionálním příznakem; potom posun LJ hodnocené dříve jako *nář.* do příznaku regionálního.

	SSJČ	SSČ	ASCS	LDB Pralex
				LJ bez územní charakteristiky
<i>burčák</i>	<i>nář.</i>	<i>obl.</i>	0	0 (jen poznámka k územnímu hodnocení)
				LJ regionální spisovná
<i>kyselo I</i> ‚polévka‘	bez charakteristiky	0	0	území Čechy
<i>brejle</i>	bez charakteristiky	bez charakteristiky	bez charakteristiky	území Čechy
<i>krám V1</i> ‚obchod‘	<i>ob.</i>	<i>hovor.</i>	<i>hovor.</i>	území Čechy
				LJ regionální nespisovná
<i>děvucha</i>	<i>nář.</i>	0	0	region. (Slezsko)
<i>fešák V2</i> ‚nápadník, milý‘	<i>nář.</i>	neuvádí tento význam	neuvádí tento význam	region. (Morava + Slezsko)
<i>gajdy</i> ‚dudy‘	<i>nář.</i>	0	(na vých. Moravě)	region./dialekt (vých. Morava + Slezsko)

Závěr

V této stati jsme se soustředili zejména na LJ s regionální charakteristikou zpracovávané v LDB Pralex. Zdá se, že se objevila nová řešení, která jsou možná právě díky databázovému zpracování slovní zásoby. Je přirozené, že zůstává mnoho otázek otevřeno – ať už je to chápání pojmu regionalismus a vymezení přesnějších krité-

říí pro jeho definování, vztah regionálních LJ k vrstvám národního jazyka, zejména k obecné češtině, otázka spisovnosti a nespisovnosti u oblastních LJ, zařazování územně příznakových LJ do výkladového slovníku češtiny, resp. do LDB Pralex, aj. K řešení problému spisovnosti a nespisovnosti teritoriálních lexikálních jednotek by mohla přispět i veřejná diskuse. Zajímavým jevem jsou posuny LJ mezi vrstvami národního jazyka, které by si jistě zasloužily zvláštní pozornost, neboť ukazují jazyk jako dynamický a živý mechanismus.

Literatura

- BOSÁK, Ján: Funkčné kvalifikátory v nových výkladových slovníkoch. In: Nová slovní zásoba ve výkladových slovníkách. Sborník příspěvků z konference (Praha 31. 10. – 1. 11. 2000). Eds. O. Martincová – J. Světlá. Praha: Ústav pro jazyk český AV ČR 2000, s. 75–83.
- Český jazykový atlas. 5. sv. Red. J. Balhar a kol. 1. vyd. Praha: Academia 1992–2005.
- Čeština – řeč a jazyk. Ed. M. Čechová. 2., přepracované vyd. Praha: ISV nakladatelství 2000. 407 s.
- Encyklopedický slovník češtiny. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: Nakladatelství Lidové noviny 2002. 604 s.
- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSC)
- GOLÁŇOVÁ, Hana: Územně příznakové jednotky v Lexikální databázi češtiny 21. století. In: Jazyková interakce a jazykové rozhraní a strategie “cutting-edge”. Sborník příspěvků z 8. mezinárodní konference Setkání mladých lingvistů, konané na Filozofické fakultě Univerzity Palackého ve dnech 14.–16. května 2007. Ed. Vl. P. Polách. Olomouc: Univerzita Palackého v Olomouci 2008, s. 104–109.
- JANČÁK, Pavel: K jazykově zeměpisné charakteristice česko-moravských a česko-slezských protikladů v slovní zásobě. In: Naše řeč, 2001, roč. 84, s. 175–83.
- KLOFEROVÁ, Stanislava: Nářeční slovní zásoba a status spisovnosti. In: Spisovnost a nespisovnost dnes. Sborník příspěvků z mezinárodní konference Spisovnost a nespisovnost v současné jazykové a literární komunikaci (Šlapanice u Brna 17.–19. ledna 1995). Ed. R. Šrámek. Brno: Masarykova univerzita 1996, s. 142–144.
- Manuál lexikografie. Eds. F. Čermák – R. Blatná. Jinočany: H&H 1995. 283 s.
- Nový akademický slovník cizích slov A–Ž. (J. Kraus a kol.) 1. vyd. Praha: Academia 2005. 880 s. (NASCS)
- Nárečia a národný jazyk. Materiály z medzinárodnej vedeckej konferencie (Budmerice 24.–26. septembra 1997). Ed. A. Ferenčíková. Bratislava: Veda 1999. 332 s.
- Příruční mluvnice češtiny. Eds. P. Karlík – M. Nekula – Z. Rusínová. 1. vyd. Praha: Nakladatelství Lidové noviny 1995. 800 s.
- Příruční slovník jazyka českého. 9 sv. Praha: Státní nakladatelství – SPN 1935–1957. (PSJČ)
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)

Slovník spisovné češtiny pro školu a veřejnost. Verze slovníku 1.0. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. vyd.). Red. spolupracovaly S. Matoušková, V. Holubová. Voznice: Leda spol. s r. o. 2004. (CD-ROM).

Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)

Slovník súčasného slovenského jazyka (A-G). Red. K. Buzássyová – A. Jarošová. Bratislava: Veda 2006. 1134 s.

Základní charakteristika kompozit v lexikální databázi Pralex¹

Eva Podruhová

Ústav pro jazyk český AV ČR, v. v. i.

Basic Characteristics of Compounds in the Pralex Lexical Database

The paper deals with compounds in the Pralex lexical database, especially with their general characteristics and problems with their processing in the database. Apart from the compounds, the author focuses on the particular components and their origin, meanings and mutual relations. Compounds with a quantitative meaning as well as compounds containing such component are not handled in here.

1. Úvod

Tématem příspěvku jsou kompozita v lexikální databázi (dále LDB) Pralex. Konkrétně se jedná o charakteristiku vybraných typů složených slov a nastínění problematiky jejich zpracování v databázi. Kromě samotných složených slov je položen velký důraz také na jejich komponenty, tj. první a druhé části kompozit. Na nich je totiž možné ukázat různé systémové vztahy a vazby, které mezi těmito komponenty fungují podobně jako mezi běžnými lexikálními jednotkami slovní zásoby češtiny (synonymie, antonymie, homonymie, vztahy významové souvislosti atd.). Vedle pojmu první část složených slov je v textu užíván také poměrně rozšířený termín prefixoid, který označuje přední komponent, jenž si částečně zachovává původní význam slova, od něhož byl odvozen.²

Text příspěvku je založen na důkladné analýze všech typů kompozit a jejich podslovních částí převzatých do LDB Pralex ze *Slovníku spisovného jazyka českého* (dále SSSJČ), s výjimkou číslovkových komponentů a složenin obsahujících takový komponent.³ Při popisu vybraného materiálu jsou konkrétní složeniny chápány jako izolované lexikální jednotky vícedílného charakteru, přičemž nejsou rozlišována kompozita vzniklá přímou a nepřímou kompozicí, resp. kompozicí ve vlastním smyslu a derivací.⁴ Za účelné je naproti tomu považováno třídění kompozit podle slovního

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Kromě prefixoidů se obecně mluví o tzv. afixoidech, jejichž vymezení v české odborné literatuře je dílem O. Martincové a N. Savického (1987). Dále viz např. práce P. Mittera. „*Termín afixoid označuje morfém přechodného rázu buď mezi kořenem a prefixem (tzv. prefixoid), nebo mezi kořenem a sufixem (tzv. sufixoid). Morfémy, které si zachovaly svůj věcný význam a povahou se blíží spíše kořeným morfémům, se označují jako radixoidy.*“ (Mitter 2005, s. 57)

3 Máme na mysli pouze komponenty s významem konkrétní kvantifikace (př. *jedno-*, *dvanácti-*, *osmašedesáti-* ad.), ne neurčitou kvantifikaci (*mnoho-*, *málo-* ad.). Analýza kompozit s číslovkovým komponentem by byla náročná z důvodu velkého množství takových složených slov a podle našeho názoru by její výsledky nebyly obzvláště přínosné pro celkovou charakteristiku kompozit v databázi Pralex.

4 Nepřímá kompozita vzniklá derivací od jiných složenin se za kompozita ve vlastním smyslu obvykle nepovažují, srov. např. Trávníček (1951, s. 240), *Příruční mluvnice češtiny* (1995, s. 110) aj.

druhu, podle původu jejich konstituujících členů (domácí, nebo přejaté) nebo podle pozice, kterou jednotlivé komponenty zauímají v systému slovní zásoby češtiny.⁵ Po stručné charakteristice práce s kompozity a jejich komponenty v lexikální databázi Pralex jsou v druhé části textu probírány jednotlivé příklady systémových vztahů, které fungují zejména na úrovni prefixoidů a odrážejí systém vztahů běžných v lexikální zásobě češtiny. Tento výklad je nutně stručný a má sloužit pouze pro ilustraci některých typů kompozit, s nimiž je možné se při práci s materiálem v lexikální databázi Pralex setkat.

2. Práce s kompozity a jejich částmi v lexikální databázi Pralex

Systematické doplňování kompozit do hesláře LDB Pralex začalo přibližně v srpnu 2008 zakládáním podslovních částí, které byly převzaty ze SSJČ, v němž byly zpracovány jako samostatná hesla. Do hesláře databáze byly doplněny všechny bez výběru (včetně zastaralých) právě proto, aby se tento soubor dal analyzovat z různých aspektů a vyhodnotit jako celek. Nejprve se jednalo o prefixoidy, později byly doplněny také zadní, opěrné členy kompozit a některé předpony. Tato hesla mají v záhlaví uveden Typ ‚Podslovní‘ a Typ podsl. ‚První část slož. slov‘, resp. ‚Druhá část slož. slov‘.⁶ Naše definice kopíruje první část výkladu ze SSJČ až po slovo, ze kterého je komponent odvozen: např. u hesla *xylo-* ‚první část složených slov spojující část druhou s významem řec. slova *xylon* *dřevo*‘ nebo u hesla *vzducho-* ‚první část složených slov spojující část druhou s významem slova *vzduch*‘.

K databázovým položkám představujícím části složených slov byly podle SSJČ zakládány také varianty a homonyma, podobně jako u ostatních substantiv a dalších slovních druhů v LDB Pralex. Dále jsme však s těmito hesly nepracovali ani jsme nevymezovali jednotlivé významy podle SSJČ. Kompletnímu zpracování těchto hesel musí předcházet jejich vyhodnocení z hlediska současného stavu podle nově zpracovaných kompozit. To znamená, že u nich není vyplňována ani exemplifikace, neboť za konkrétní příklady jsou považována až kompozita vzniklá pomocí daných podslovních částí, a bude tedy relevantní až pro nové slovníkové zpracování. V současné době je v LDB Pralex založeno téměř 1000 podslovních částí (cca 870 prvních a 72 druhých částí složených slov).⁷

5 Konkrétně jde o jejich podíl na tvorbě slov v oblasti terminologie nebo běžné slovní zásoby, o jejich produktivitu a frekvenci jejich používání, resp. o aktuálnost vzniklých kompozit, doloženost v korpusu apod.

6 Další dva typy podslovních částí představují předpony a přípony.

7 Ne všechna kompozita v databázi Pralex mají založen odpovídající první nebo druhý komponent, což představuje problém při komplexní práci s kompozity. Tento stav vyplývá z faktu, že ani v SSJČ nejsou všechna složená slova zastřešena adekvátním heslem a jsou často „utopena“ mezi ostatními slovy začínajícími stejným řetězcem písmen. V Pralexu se tedy nachází o něco více kompozit, než jsme schopni podchytit nebo spočítat, a je možné je najít jen náhodou nebo při důkladné analýze hesláře. Např. máme založeno heslo *sněho-* a 11 jeho kompozit, avšak hesla *mrazo-*, *mrazu-* zde zařazena nejsou, ačkoliv kompozit s těmito předními členy najdeme v hesláři 7.

Zakládání hesel do hesláře LDB Pralex dále pokračovalo doplňováním kompozit utvořených pomocí daných prefixoidů ze SSJČ. Jejich postupné zpracovávání v současné době probíhá. Pokud jde o substantivní kompozita, zpracování jejich exemplifikace je v zásadě totožné se základním zpracováním substantiv obecně, složená slova se totiž co do syntaktické spojitelnosti nijak neliší od ostatních slov běžné slovní zásoby. Jediný rozdíl, který lze vysledovat, spočívá v nízké frekvenci výskytu většiny těchto složených podstatných jmen v korpusu SYN⁸. Mnoho z nich zůstává nedoloženo, další velká skupina slov má v korpusu velmi nízkou četnost a jen u malého procenta těchto složenin⁹ najdeme dostatek dokladů na to, aby mohly být plně exemplifikovány.¹⁰ Zpracování kompozit z dalších slovních druhů v LDB Pralex bude probíhat souběžně se zpracováním ostatních slov hesláře.

Celkové složení kompozit v lexikální databázi ze slovnědruhového hlediska odráží rozvrstvení slovních druhů složených slov v lexikální zásobě obecně. Největší procento představují substantiva a adjektiva, poměrně méně jsou zastoupena adverbia (často pouze jako deriváty od složených adjektiv). Složené číslovky a zájmena tvoří zvláštní, nepříliš obsáhlou, ale kompaktní a sourodou skupinu, zatímco slovesa se vyskytují zcela výjimečně.

Podle množství kompozit v LDB Pralex by se tedy dalo soudit, že složená slova představují v lexikálním systému češtiny poměrně velkou a důležitou skupinu slov¹¹, avšak jde především o lexikální jednotky stojící mimo centrum slovní zásoby. Tento závěr je ovšem poněkud komplikován tím, že heslář LDB nebyl zatím doplněn o kompozita z jiných slovníků, než je SSJČ, tudíž odráží ponejvíce dobu vzniku tohoto lexikografického díla. Na základě *Frekvenčního slovníku češtiny* (dále FSČ) se do hesláře dostala pouze nejvíce frekventovaná kompozita z korpusu SYN, což ovšem představuje jen minimální množství z celkového objemu kompozit v našem hesláři.

3. Kompozita v databázi Pralex z hlediska původu jejich komponentů

Složení analyzovaného materiálu (tj. ze SSJČ a FSČ) v LDB Pralex můžeme popsat také z hlediska původu jednotlivých komponentů, které se podílejí na vzniku kompozit. První části složených slov jsou jak domácí, tak přejaté, přičemž domácí komponenty tvoří téměř dvě třetiny materiálu a přejaté zbytek.

Cizí slovní základy (především prefixoidy) jsou často součástí terminologie různých oborů (např. chemie, fyziky, technických, přírodovědných oborů ad.),

8 Od začátku roku 2011 je původní korpus SYN pro naše potřeby přejmenován na orig_syn, aby se původní název SYN zachoval pro potřeby zřizovatele.

9 Nejčastěji jde o ta slova, která byla vložena do hesláře podle *Frekvenčního slovníku češtiny*, tj. na základě jejich frekvence.

10 S tím souvisí také problematičnost jejich vyhledávání v korpusu, protože tato málo frekventovaná slova nejsou obvykle lemmatizována.

11 Tento fakt potvrzuje také analýza neologického lexikálního materiálu, např. slovníku *Nová slova v češtině 1 a 2*, ve kterých je podíl složených slov přibližně 30 % a 50 %. Protože však kompozita v LDB nijak neklasifikujeme, je obtížné zjistit jejich celkový počet. Dá se odhadnout, že mnoho dalších nalezneme i v korpusu, zatím však nebyly cíleně vyhledávány.

takže i slova složená pomocí těchto komponentů náleží z velké části do kategorie termínů a profesionalismů. Nejčastěji se jedná o kompozita s oběma členy cizího původu, méně se objevují hybridní složeniny. Vzhledem k úzké oblasti jejich používání a jejich značné specifičnosti¹² nenacházíme u těchto případů v korpusu dostatek materiálu pro zpracování jejich exemplifikace. Poměrně obsáhlou skupinu kompozit zde představují tzv. vícekomponentová kompozita, resp. rekompozita, tedy slova složená z více než dvou částí. Nejvíce kompozit je utvořeno pomocí komponentů **agro-**, **elektro-**, **foto-**, **fyto-**, **mikro-**, **paleo-**, **radio-**, **stereo-** a **zoo-**. Jako příklady můžeme uvést kompozita *agrometeorolog*, *antropogeograf*, *bioklimatologie*, *dynamometamorfóza*, *elektrokardiogram*, *aerofotografie*, *fytosociologie*, *chromoxylografie*, *kilowatthodina*, *makrofotografie*, *paleoetnologie*, *radiotelegram*, *stereomagnetofon*, *zoopaleontologie*.

Mezi složenými slovy s oběma **komponenty domácího původu** najdeme mnoho nedoložených substantiv, adjektiv i dalších slovních druhů.¹³ Nízká frekvence užívání domácích složenin a v mnoha případech jejich zastaralost spočívá v zastaralosti samotných podslovních částí kompozit. Pokud je totiž jeden komponent pocitován jako zastaralý, budou pravděpodobně jako zastaralá vnímána také slova, na jejichž vzniku se podílí.¹⁴ Zvláště patrné je to u předních komponentů adjektivního charakteru, jejichž stylové hodnocení se přenáší na hodnocení celého kompozita. Z velké části sem patří slova, která se do SsJČ dostala z krásné literatury 19. století, a nezdá se, že se jedná o autorská užití slov již dříve málo používaných. Jako příklady můžeme jmenovat prefixoidy **blaho-**, **darmo-**, **hojno-**, **jaro-** (ve významu jarý), **lepo-**, **libo-**, **sivo-**, **skvělo-** (odvozeno od skvět se), **snivo-**, **širo-**, **trpko-** atd. Nejvíce se tyto komponenty podílejí na tvorbě adjektiv, méně často i substantiv (př. *bystrovtip*, *čaromoc*).¹⁵ Následující příklady pak ukazují, že v některých slovech dochází také ke kombinaci dvou méně užívaných komponentů, což stylový příznak zastaralosti ještě více posiluje: adj. *jarošumný*, *lepokadeřný*, *libodechý*, *sivobradý*, *skvělobleský*, *snivolící*, *trpkobolný*, subst. *darmotlach*, *libomrav*, *pestrokřídlec*, *truchlozvěst*.

4. Variantnost u kompozit

Variantnost složených slov je v souboru kompozit ze SsJČ v lexikální databázi Pralex poměrně vysoká (přibližně 25 %). Na první pohled je zřejmé, že vyplývá především z variantnosti grafické podoby jednotlivých částí kompozit – jak jejich prefixoidů, tak

12 Velmi často se jedná o pojmy a termíny, kterým běžný uživatel jazyka (laik) bez slovníku nemůže rozumět.

13 Tento stav je však způsoben jinými důvody, než je vysoká odbornost daných slov a jejich částí, jak tomu bylo u kompozit cizího původu.

14 Výjimky tvoří pouze slova pevně zakořeněná ve slovní zásobě, u nichž vlastní význam jednotlivých částí už není pocitován.

15 Příklad *blaho-*: V hesláři máme přes 20 kompozit, z nich používáme v současné češtině cca 5 (*blahobyť*, *blahodárný*, *blahopřání*, *blahosklonný*, *blahoslavený*). Ostatní charakterizujeme jako zastaralá (subst. *blahocit*, *blahomrav*, *blahoslov*, *blahovůle*, *blahozvěst*; adj. *blahodajný*, *blahodyšný*, *blahojasný*, *blahonosný*, *blahoplodý*, *blahosnivý*, *blahověrný*, *blahozvučný*).

zadních členů a derivačních sufixů.¹⁶ První typ představují kompozita, jejichž přední členy vykazují variantnost na švu, resp. u kterých může být variantní konekt (př. **boho-**, **bohu-**, **boha-**; **díku-**, **díků-**; **chvalo-**, **chvály-**; **mnoho-**, **mnoha-**; **polo-**, **polou-**). Dále kompozita, u kterých je více či méně redukováno slovo vstupující do výsledného slova jako jeho první člen nebo u kterých jsou základem pro první člen různé slovní druhy, nejčastěji substantivum a adjektivum (**rusko-**, **ruso-**; **vele-**, **velko-**; **lesko-**, **lesklo-**; **žido-**, **židovsko-**) a kompozita, jejichž první část je odvozena od již variantních slov (**kdoví-**, **kdožví-**; **leda-**, **ledas-**; **nebe-**, **nebes-**; **také-**, **taky-**). U podslovních částí přejatých z latiny nebo řečtiny se objevují typické problémy grafického přepisu těchto slov do češtiny, tedy s-z, o-ó, t-th apod. (př. **fero-**, **ferro-**; **fysio-**, **fyzio-**; **chromo-**, **chrómo-**; **meso-**, **mezo-**; **teo-**, **theo-** atd.). Podobně je tomu také u druhých částí složených slov a sufixů (př. **-lysa**, **-lýza**; **-ismus**, **-izmus** a další, z domácích např. **-koli**, **-koliv**).

5. Homonymie u předních členů kompozit

Přestože homonymních prefixoidů převzatých ze SSJČ není v databázi Pralex mnoho,¹⁷ je zajímavé se na tento jev podívat trochu podrobněji. Stejně jako při ostatních krocích naší analýzy se také zde jeví jako velmi účelné dělení podslovních částí podle jejich původu na domácí a přejaté. Přejaté komponenty pocházejí z velké většiny ze dvou jazyků, a to z latiny a řečtiny.¹⁸ U vybraných komponentů tedy vstupují do vztahu homonymie domácí a cizí slovní základy a konstituují se přitom tři typy: 1) domácí – domácí, 2) domácí – přejatý, 3) přejatý – přejatý.

V prvním typu jsou to např. komponenty **leto-** (I k *léto*, II k *let*), **pravo-** (I k *pravý*, II k *právo*), **rudo-** (I k *rudý*, II k *ruda*), **sebe-** (I jako ‚největší míra vlastnosti‘, II jako ‚zájmeno se‘), **stavo-** (I ke *stavět*, *stavba*, II ke *stav*), **vlno-** (I k *vlna I*, II k *vlna II*). U druhého typu můžeme uvést např. **holo-** (I k *holý*, II k *holos* řec. ‚celý, úplný‘), **nitro-** (I k *nitro*, II k *nitron* řec. ‚dusík‘). Nejvíce zastoupen je třetí typ, v jehož rámci lze rozlišit dva podtypy:

a) dva lexémy na úrovni jednoho cizího jazyka, např. **chloro-** (řec. I k *chlóros* ‚zelený‘, II k *chlór*), **chromo-** (řec. I k *chróma* ‚barva‘, II k *chróm*), **pedo-** (I k *pais* řec., lat. ‚dítě‘, II k *pedon* řec. ‚půda‘),

b) lexémy na úrovni dvou různých cizích jazyků, př. **homo-** (I k *homo* lat. ‚člověk‘, II k *homos* řec. ‚stejný‘).

Speciální typ tvoří komponenty, u kterých zůstává jako I člen s původním významem slova, ze kterého byl komponent odvozen, zatímco lexém II nese význam již složeného slova, obvykle frekventovaného v běžné slovní zásobě. Příklady ukážou více: **aero-** (I k *aér* řec. ‚vzduch‘, II k *aeroplán*), **auto-** (I k *autos* řec. ‚sám‘, II k *au-*

16 To je dáno také tím, že SSJČ se řídil starším vydáním *Pravidel českého pravopisu* (k poslední změně došlo až v roce 1993), takže v databázi vlastně zachycujeme dynamiku vývoje variantnosti od počátku 20. století.

17 27 komponentů je homonymních jako dvojice, jeden komponent jako trojice.

18 Ostatní jazyky v podstatě nejsou zastoupeny.

tomobil), **foto-** (I k *fós, fotos* řec. ‚světlo‘, II k *fotografie*), **gramo-** (I ke *gramma* řec. ‚záznam‘, II ke *gramofon*), **moto-** (I k *motus* lat. ‚pohyb‘, II k *motocykl, motorový*), **radio-** (I k *radius* lat. ‚paprsek‘, II k *rádio, rádiový*), **strato-** (I ke *stratum* lat. ‚vrstva‘, II ke *stratosféra*).

Na úrovni celých kompozit (substantiv nebo jiných slovních druhů) se s homonymií prakticky nesetkáme. U homonymních základů, které tvoří komponenty domácího (I) a cizího (II) původu (př. **holo-** I, II), to může být způsobeno faktem, že přejaté prefixoidy mají tendenci se spojovat s přejatými druhými částmi složených slov a domácí naopak s domácími, takže zřídka dochází k významovému křížení. U kombinací dvou stejnojjazyčných homonymních lexémů máme však také pouze několik příkladů: *fotokartograf* (I ‚konstruktér map na základě fotogrammetrie‘, II ‚přístroj k transformaci leteckých snímků pro mapování‘), *pedologie* (I ‚půdoznalství‘, II ‚nauka o dítěti po všech stránkách‘), *vlnoměr* (I ‚přístroj, kt. se měří zkadeření ovčí vlny‘, II ‚měřič kmitočtu vysokofrekvenčních oscilací‘). Zdá se tedy, že homonymie kompozit je zde opět založena na homonymii jednotlivých komponentů těchto slov, zde konkrétně jejich prvních členů.

6. Vztahy významové souvislosti u kompozit a jejich členů

Vztahy významové souvislosti mezi kompozity jsou (podobně jako vztahy homonymie) založeny na vztazích, které se konstituují již v systému podslovních komponentů podílejících se na vzniku těchto složených slov. Pro mnoho z doložených prvních i druhých částí kompozit jsou charakteristické zejména vztahy synonymie a antonymie, které fungují napříč dělením na komponenty domácí a přejaté, takže vytvářejí dvojice (eventuálně trojice) různého složení. Zvláštní skupinu však tvoří dvojice, v nichž je jeden člen domácí a druhý přejatý, internacionální. Např. synonymních dvojic takového složení máme v databázi obsaženo přes šedesát. Obvykle jde o případy, kdy internacionální člen funguje v oblasti terminologie a tvoří kompozita s oběma cizími slovními základy, zatímco domácí člen se kombinuje výhradně s domácími komponenty a vytváří slova frekventovaná v běžné slovní zásobě. Jen zřídka přitom dochází ke kombinaci či k zaměnitelnému postavení těchto dvou členů. Některé z daných prefixoidů mají potenciál vytvářet hybridní složeniny, ale ani zde není běžné, že bychom měli dvě synonyma – hybridní a přejaté nebo hybridní a domácí. Synonymie a antonymie u těchto slov se tedy projevuje silněji na úrovni komponentů než na úrovni výsledných kompozit. Př. komponent **dendro-** a **stromo-** (*dendrolit, dendrolog, dendrometr; stromořadí, stromoměr*) nebo **geo-** a **země-** (*geodesie, geofyzika, geografie; zeměkoule, zeměpis, zemětřesení, zeměžluč*). Hybridní kompozita pak tvoří nejvíce komponenty, které jsou všeobecně rozšířené i mimo úzké hranice jednoho oboru a které jsou stále produktivní např. při tvoření nových slov v češtině. Př. **elektro-** (termíny: *elektrodynamika, elektrochirurgie, elektrolyt, elektronarkóza, elektroterapie*; hybridní kompozita: *elektroléčba, elektrooprava, elektropotřeby, elektroprodejna, elektroprůmysl, elektropřístroj, elektrospotřebič, elektroúdržbář, elektrovrtáčka* aj.), dále **auto-**, **bio-**, **foto-**, **hydro-**, **kino-**, **moto-**, **radio-** nebo komponenty s intenzifikačním významem jako

makro-, maxi-, mikro-, mini-, super-, ultra- ad. (*makrosvět, maxisukně, mikrospánek, miniponorka, superpočítač, ultravlastenec*).¹⁹

Kromě vztahů synonymie a antonymie vstupují podslovní komponenty i celá složená slova do různých dalších významových vztahů a vytvářejí různé sémantické řady a sémantická pole. Nezřídka bývá jeden komponent zapojen do více takových vztahů. Přitom i zde platí fakt, že na úrovni komponentů se tyto vztahy manifestují zřetelněji než na úrovni hotových kompozit, u kterých je často významová souvislost odvozena od významové souvislosti jejich předních nebo zadních členů. Lexikální materiál zahrnující kompozita z LDB Pralex je bohatý a obsahuje různé významové typy a kategorie lexémů. Z toho důvodu nelze zmínit všechny aspekty, které vyplývají z práce s tímto materiálem, proto pro účely svého příspěvku vybíráme jen několik významových skupin komponentů, na nichž chceme ilustrovat šíři a rozmanitost vztahů, které mezi nimi existují.

I. Komponenty odvozené od názvů zemí, regionů, národů apod.

Př. **anglo-, čecho-, česko-, franko-, indo-, moravsko-, německo-, řecko-, římsko-, slovensko-, srbo-** apod.

Tyto části složených slov se podílejí především na vzniku adjektivních kompozit, v nichž se spojují s druhým členem podobného významu (dvouslovné názvy: *česko-krumlovský, německobrodský*; kompozita složená ze zeměpisných jmen dvou různých zemí, měst apod.: *anglosaský, indoevropský, řeckořímský, srbochorvatský*), nebo jsou výsledná kompozita odvozena od sousloví s příslušným adjektivem (*římskokatolický, řeckokatolický, českobratrský*). Některé z uvedených komponentů se spojují také se základy cizího původu, nejčastěji označujícími osoby (*angloman, frankoman, čechofil, srbofil*).

II. Komponenty s významem světových stran

Př. **jih-, severo-, východo-, západ-**.

Uvedená čtveřice prefixoidů představuje uzavřenou skupinu komponentů, která je v současné době pouze omezeně produktivní. Substantiva tvoří minimálně, konkrétně názvy světových stran (*jihozápad, severovýchod* a vícekomponentová slova jako *jihojihovýchod, severoseverozápad, západojihozápad, východoseverovýchod*), výjimečně se objevují i složeniny s jiným významem (*Jihoslovan, Východočech, jihozemec*). Adjektiva utvořená pomocí těchto komponentů představují odvozeniny od názvů kontinentů, zemí, regionů apod. a patří mezi častá a běžně užívaná slova slovní zásoby (*jihoafrický, severomoravský, západoidický, východořímský*). Dalšími příklady jsou slova jako *západokřesťanský, jihopolární, severotočnový, jihozemský*.

III. Komponenty **bůhví-, čertví-, kdoví-, kdožví-, pánbůhví-**

Jedná se o skupinu prvních členů složenin, která z principu netvoří substantiva, ale spojují se pouze se členy představujícími adjektiva, zájmena, číslovky a příslovce

¹⁹ Nové materiálové zdroje ukazují, že produktivita těchto komponentů s intenzifikačním významem nápadně vzrostla zejména během posledních dvou desetiletí. Už SSJČ zachycuje tento potenciál u několika z nich, např. viz heslo *mini-*: „V publicistickém stylu a v slangu je v poslední době tvoření složenin s *mini-* hojně, proto je uvádíme jen v přísném výběru.“

neurčitého významu. Vznikají tak kompozita s více než dvěma částmi (už přední člen těchto kompozit je složený), zásadně domácího původu. Repertoár celkem dvanácti zadních, opěrných členů kompozit se opakuje u všech prefixoidů (*-co*, *-čí*, *-jak*, *-jaký*, *-kam*, *-kde*, *-kdo*, *-kdy*, *-kolik*, *-kolikátý*, *-odkud*, *-proč*), ovšem ne všechny první části složených slov v našem hesláři jej zcela vyčerpávají (*bůhví-* tvoří kompozita s jedenácti z nich, naopak nejméně zastoupen je komponent *kdožví-*).²⁰

IV. Komponenty odvozené od názvů částí těla

Př. *břicho-*, *hlavo-*, *ruko-*, *srdce-*, *tělo-*, *vlaso-*, *zubo-*, *jazyko-*.

Tato skupina zahrnuje poměrně velký počet komponentů, které jsou odvozeny od frekventovaných substantiv pevně zakořeněných v lexikální zásobě češtiny. Přesto je jejich potenciál k vytváření nových kompozit omezený a již vzniklá složená slova se z hlediska stylistické charakteristiky přesouvají do oblasti zastaralých nebo velmi málo užívaných slov. Časté jsou příklady kompozit s druhým členem verbálního charakteru, které představují vedle slov běžně užívaných (jako jsou např. *břichomluvec*, *hlavolam*, *jazykolam*, *jazykověda*, *rukojet*, *rukopis*, *tělocvik*) i slova zastaralá (*rukotlesk*, *srdcebol*, *zuboměr*, *vasostřih*, *jazykozpyt*). Podobně je tomu i u kompozit se základem odvozeným z dalších slovních druhů. Běžně užívaná jsou např. slova jako *hlavohrud*, *hlavonožec*, *rukodělný*, *srdceryvný*, *tělovýchovný*, *zubodásňový*, *jazykohltanový*. Nápadná je u těchto komponentů jejich spojovatelnost s adjektivními druhými členy, značně bohatá a nápaditá např. u slova *srdce* (*srdcebolný*, *srdcejemný*, *srdcejímaný*, *srdcelomný*, *srdcerozrývající*, *srdcervavý*, *srdcervoucí*, *srdceryvný*, *srdcetklivý*).²¹

V. Komponenty s významem ‚starý‘, ‚mladý‘ apod. různého původu

Př. *paleo-*, *neo-*, *staro-*, *mlado-*, *novo-*.

Poslední skupinu slov, na kterou jsme se zaměřili, tvoří komponenty s významem stáří pocházející jak z češtiny, tak z latiny a řečtiny. Jmenované prefixoidy mají adjektivní povahu, avšak často se podílejí na vzniku složených substantiv. Přejaté první části složených slov *paleo-* (z řečtiny) a *neo-* (z latiny) se kombinují s druhým zadním členem cizího původu a tvoří především odborné termíny a professionalismy, např. názvy vědeckých disciplín, filosofických směrů apod. (*paleobotanika*, *paleoetnologie*, *paleografie*, *neodarwinismus*, *neoimpresionismus*). Prefixoid *paleo-* je přitom velmi produktivní při tvorbě vícekomponentových složených slov se všemi částmi cizího původu (např. *paleoantropologie*, *paleoklimatologie*, *paleozoologie* ad.), naopak netvoří hybridní kompozita. Domácí prefixoidy *staro-* a *mlado-* se pak spojují především s domácími slovními základy a tvoří substantivní kompozita (*starobinec*, *staročestina*, *starohory*, *staropraven*, *starousedlík*, *mladočech*). Prefixoid *mlado-* je přitom v materiálu zastoupen méně a slova obsahující tento komponent jsou povětšinou málo frekventovaná (*mladolamač*, *mladomanžel*, *mladokomik*), zatímco kompozit s prefixoidem *staro-* máme podstatně více, i když i mezi nimi převažují málo používané složeniny (*starodůchodce*, *staromanžel*, *staromilství*, *starovlastenec*).

20 Stav v našem hesláři je odvozen od SSJČ.

21 Jedná se o komponenty, od kterých máme v databázi velké množství slov, ale většina zůstává v kategorii nedoložených nebo velmi málo doložených slov (na základě korpusu).

Prefixoid **staro-** vytváří také hybridní kompozita s druhým členem cizího původu (*starokonzervativec, staropenzista, staroromantismus, staroutrakvismus*). Konečně komponent **ново-** představuje zvláštní případ prefixoidu, jelikož se v něm stýkají dva stejné slovní základy, pokaždé odvozené z jiného jazyka: **ново-** I z latiny („novus“) a **ново-** II z domácího adjektiva ‚nový‘. Protože význam obou komponentů je stejný, je otázkou, zda se v tomto případě vlastně jedná o homonymii. Z uvedeného vyplývá fakt, že u tohoto prefixoidu není snadné určit, zda tvoří, nebo netvoří hybridní kompozita: mezi druhými částmi složených slov jsou zastoupeny jak přejaté, tak domácí slovní základy. Podle předchozích zkušeností lze předpokládat, že přejatý komponent **ново-** I se spojuje s přejatými zadními členy složených slov a domácí komponent **ново-** II se členy domácího původu, avšak tento závěr může být potvrzen pouze důkladnou slovotvornou analýzou každého jednotlivého příkladu. Kompozita utvořená pomocí prefixoidu **ново-** patří mezi poměrně frekventovaná slova, jde např. o adjektiva odvozená od víceslovných názvů zemí, regionů, obcí atd. (*novodvorský, novoguinejský, novohradský*) a další adjektiva (*novodobý, novogotický, novoroční, novorozený, novozákonní*) nebo substantiva (*novofašismus, novokněz, novokřtěnc, novomanžel, novostavba, novověk*).

Závěr

Jak je zřejmé z předchozího výkladu, problematika kompozit je velice zajímavá a také podrobná analýza složených slov v lexikální databázi Pralex, která byla provedena pro účely tohoto příspěvku, ukazuje mnohé tendence uplatňující se jak v oblasti kompozit obecně, tak i v celkové slovní zásobě češtiny. Složená slova jako celek však bude možné vyhodnotit až tehdy, když budou všechna hesla kompletně zpracována a doložena na základě dnešního úzu. To bude součástí následujícího úkolu, který se zaměří na nové slovníkové zpracování lexikální zásoby současné češtiny.

Literatura

- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSČ)
- MARTINCOVÁ, Olga – SAVICKÝ, Nikolaj: Hybridní slova a některé obecné otázky neologie. In: Slovo a slovesnost, 1987, roč. 48, s. 124–139.
- MITTER, Patrik: Vymezení některých prefixoidů a radixoidů v české lingvistické literatuře. In: Eurolingua 2004. Opera Academiae Paedagogicae Liberecensis. Series Bohemistica. Ed. O. Uličný. Liberec: Technická univerzita 2005, s. 57–62.
- Příruční mluvnice češtiny. Eds. P. Karlík – M. Nekula – Z. Rusínová. 1. vyd. Praha: Nakladatelství Lidové noviny 1995. 800 s.
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- TRÁVNÍČEK, František: Mluvnice spisovné češtiny. 1. Hláskosloví. Tvoření slov. Tvarosloví. 3., opravené a doplněné vyd. Praha: Slovanské nakladatelství 1951. 611 s.

Ke zpracování zájmen v lexikální databázi Pralex¹

Vojtěch Veselý

Ústav pro jazyk český AV ČR, v. v. i.

Processing of Pronouns in the Pralex Lexical Database

This article deals with a processing of pronouns in the Pralex lexical database. Pronouns can be homonymous with other word types, e.g. with nouns or adverbs. Grammatical data for polysemic pronouns are introduced at the level of a full entry or at the level of individual meanings. Within every lexical unit, the author records a pronoun type as well as its primary syntactic function. Sometimes, identification of a word type can be problematic, e.g. the pronouns *kteřý* (*who*) and *jaký* (*which*), which are used to link a subordinate clause, but have no relative function at all. When defining an explanation of a meaning, we rely on *Dictionary of the Standard Czech Language* and *Dictionary of Standard Czech for Schools and the General Public*. An example part of entries is divided in the so-called modified and sentential records.

1. Úvod

Zájmena definujeme jako tu část deiktických výrazů, která má morfologickou povahu jmen, tj. skloňuje se. Tato kritéria splňují i některé zájmenné číslovky, ty však v lexikální databázi (dále LDB) Pralex klasifikujeme právě jako číslovky a ne jako zájmena. Deiktické výrazy vymezujeme v souladu s *Mluvnicí češtiny 2* (1986, s. 81; dále jen MČ 2) jako výrazy ukazující do situace nebo odkazující do kontextu.

Soubor zájmen zpracovaných v LDB Pralex představuje přibližně 200 databázových položek (lexémů), resp. 450 lexií. Asymetrii mezi počtem lexémů a jim odpovídajících významů způsobuje fakt, že některá hesla jsou monosémní, jiná polysémní. Sémantická struktura polysémních zájmen může být relativně jednoduchá (hesla se dvěma či třemi významy, např. *kdokoli, který* aj.), anebo poměrně složitá (zájmena s více než čtyřmi významy, např. *jaký, takový* aj.). Většina zájmen má jednoslovnou formu, víceslovnost je jevem spíše výjimečným².

2. Homonymie u zájmen

Poměrně častým jevem u zájmen je slovnědruhovú homonymie. Zájmena bývají homonymní se substantivy (*já, můj*), zájmennými příslovci (*copak*), částicemi či ci-

1 Příspěvek vznikl v souvislosti s budováním lexikální databáze Pralex v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Víceslovné heslo v platnosti zájmena zatím označujeme jako víceslovnou lexikální jednotku (VLJ) bez bližšího zařazení, protože rozhodnutí, zda patří k souslovím nebo k frazémům, je často problematické. Např. spojení *ten samý* můžeme v LDB Pralex hodnotit spíše jako zájmenné sousloví, jiné případy naopak jako frazémy. Výrazy typu *bůhví co* lze klasifikovat jako frazémy, ale vzhledem ke dvojí možné grafické podobě (*bůhví co* x *bůhvíco*) tyto výrazy kolísají mezi jednoslovností a víceslovností; proto lze též uvažovat též o pojetí nefrazeologickém.

toslovci (*což*). Ne vždy se homonymie týká celého paradigmatu: s neohebnými slovními druhy bývá homonymní většinou jen tvar nominativu, popř. i akuzativu (srov. *copak, čehopak* atd.). Jednotlivá homonyma odlišujeme v rámci celé lexikální databáze římskými číslicemi. Pořadí homonym v LDB Pralex nekoresponduje vždy s jejich centrálností/periferností v jazykovém systému ani se směrem sémantické motivace. Pracovně v LDB Pralex homonyma číslujeme tak, aby to odpovídalo tradičně uváděnému pořadí slovních druhů. Např. nesklonné substantivum *já* je derivováno od zájmena *já* a jeho role v systému jazyka je možno říci periferní; v *Slovníku spisovného jazyka českého* (1960–1971, dále SSSJČ) i v *Slovníku spisovné češtiny pro školu veřejnost* (1994, dále SSČ) je k zájmenu *já* přihnízďováno. V lexikální databázi je pracovně naše pořadí homonym opačné: *já I* (subst.), *já II* (zájm.).

3. Variantnost u zájmen

U polysémních zájmen se variantnost většinou týká všech významů. Někdy je ale omezena jen na některé lexie, např. *takovýhle* je variantou k *takový* jen pro významy ‚soudí toho způsobu, druhu, té podoby n. těch vlastností‘ a ‚vynikající jakostí, velikostí n. množstvím; tak krásný, dobrý, velký, četný‘; význam ‚mravně narušený, hodný opovržení, bezvýznamný, špatný‘ se podařilo doložit jen u zájmena *takovýhle*.

4. Gramatická charakteristika zájmen v LDB Pralex

U zájmen (na rozdíl od ostatních slovních druhů) uvádíme souhrnné gramatické informace umístěné v SSSJČ před prvním významem (a platící tedy pro celé heslo); gramatický popis je u některých zájmen v SSSJČ velmi podrobný, je proto vhodné jej využít (viz např. u zájmen *já* nebo *můj*).

Pokud má zájmené lemma v nominativu dva a více heteronymních tvarů, uvádíme je v položce Paradigma.³ To je případ většiny rodových zájmen, existují však mezi nimi i výjimky: např. zájmena *jeho* a *její* vyjadřují kategorii nominativu pouze jedinou formou, pole Paradigma proto zůstává prázdné.

Úplné paradigma zájmena *ten* je např. tvořeno osmi subparadigmaty definovanými hodnotami sémů jmenného rodu a jmenného čísla; jde o subparadigma singuláru maskulina inanimata, subparadigma plurálu feminina atd. (počítáme se čtyřmi jmennými rody: mužským životným, mužským neživotným, ženským a středním). Každé z těchto subparadigmat je reprezentováno jediným tvarem nominativním; jejich úhrn vypadá takto: *ten, ten, ta, to, ti, ty, ty, ta*. Homonymní tvary uvádíme v poli pro Paradigma pouze jednou, tedy: *ten, ta, to, ti*. Homonymii vždy okomentujeme v poznámce: tvar *ten* platí pro j. č. rodu m. živ. i neživ.; tvar *ta* platí pro j. č. rodu ž. a mn. č. rodu stř.; tvar *ty* platí pro mn. č. rodu m. neživ. a ž. Jestliže je zájmeno polysémní a paradigma tvarů není pro všechny lexie stejné, popisujeme paradigma až na úrovni jednotlivých významů.

3 Paradigma v LDB Pralex nezachycuje úplné zájmené paradigma, ale pouze heteronymní tvary nominativní.

5. Gramatické údaje na rovině jednotlivých významů

U zájmen určujeme tyto druhy: osobní, osobní zvrtné, přivlastňovací, přivlastňovací zvrtné, tázací, vztažné, neurčité, totalizační, záporné. Tento soubor koresponduje s druhy zájmen uváděnými v MČ 2. Některé lexie mohou náležet k více druhům zároveň; např. *čí 1* je současně zájmenem přivlastňovacím a tázacím.

Kromě druhu určujeme u zájmen syntaktické funkce (adjektivní, substantivní) a obecné sémantické rysy, které daná lexie může vyjadřovat (deiktický, kvantifikační). Deiktický rys je přítomen v sémantické struktuře všech zájmen, neplatí však, že by byl na tento slovní druh omezen: kategorie deiktických výrazů je širší než kategorie zájmen. Kvantifikační rys je naopak vyjadřován jen menšinou zájmen. Kvantifikátory zde chápeme v souladu s naším článkem publikovaným ve Slově a slovesnosti (Veselý V., 2008). Soubor syntaktických funkcí byl převzat z MČ 2 (s. 84), neobsahuje však funkci cirkumstantivní a verbální. Absence těchto funkcí je dána tím, že u zájmen a číslovek v LDB Pralex vyznačujeme **pouze primární** syntaktické funkce.

Primární syntaktickou funkci definuje Komárek (2006, s. 30) takto: „(...) v primární syntaktické funkci se uplatňují sémantické vlastnosti slovního druhu v maximální míře; mluvíme zde proto o primárním významu slovního druhu a míníme tím slovnědruhový význam nezabarvený sémantikou jiného slovního druhu.“ Naše pojetí se od pojetí Komárkova poněkud liší: primární syntaktická funkce musí být u konkrétní lexie vždy realizovatelná, v opačném případě její místo obsadí funkce, která je pro daný slovní druh sekundární. Dobře je to vidět na lexémech *ten* nebo *všechn*. Primární syntaktická funkce většiny jejich lexíí je adjektivní (*ten člověk, všichni lidé*), avšak lexie reprezentované formami *to* a *všechno*⁴ adjektivní funkci plnit nemohou: pro ně je primární funkce substantivní. Nasvědčuje tomu i fakt, že zájmeno v této funkci ztrácí jednu z vlastností pro tzv. adjektivní zájmena typických: kongruenční vyjadřování jmenného rodu.

Složitější je určování primární syntaktické funkce u tázacích, resp. vztažných zájmen *který* a *jaký* . Komárek (2006, s. 86) postuluje pro všechny instance těchto zájmen primární syntaktickou funkci adjektivní. Argumentuje tím, že v závislých relativních větách uvozených zájmeny *který* nebo *jaký* předává relativum závislé větě svůj adjektivní slovnědruhový a sémantický charakter, a věta závislá jako celek tak plní funkci přivlastku. Pro nás je však důležitější úloha relativa **uvnitř** věty závislé – ta je v těchto případech primárně substantivní, nikoli adjektivní. Např. *učíme se o těch národech, které kdysi žily v naší vlasti* (příklad ze SSJČ) nelze změnit na ** učíme se o těch národech, které národy kdysi žily v naší vlasti*, nýbrž pouze na *učíme se o tom, které národy kdysi žily v naší vlasti*, což je ovšem již souvětí s jinou syntaktickou strukturou i jiným významem. Závislá věta v prvním příkladu je přivlastkem substantiva, na němž závisí, a relativum ji uvozuující je funkčním ekvivalentem substantiva.

Syntaktické hodnocení zájmen *který* a *jaký* je komplikováno hodnocením jejich lexikálněsémantické struktury. SSJČ – a stejně tak i SSČ – řadí zájmena ve výše exem-

4 Srov. omezující morfologické poznámky a výklady významu v SSJČ: *ten 10*. jen *to* ‚všeobecně souhrnně odkazuje na něco, o čem byla řeč n. co je známé‘; *všechn 4*. jen *s. vše, všechno, všechno*, ‚vyjadřuje úplnost, celost něčeho; nic nevyjímající‘.

plifikovaném typu mezi relativa. Nesporné jsou kontexty, v nichž zájmeno plní funkci tázací a syntakticky se primárně chová jako adjektivum: *Jaký byl koncert?* (SSČ), *Jaké bude počasí?* (SSJČ). Méně jasné je, jak uvedená zájmena hodnotit v tzv. nepřímých otázkách. Tento pojem užívá SSJČ pro gramatickou specifikaci zájmen tázacích. Např. *kteřý 1:* táz. (v ot. přímých i nepřímých), ptáme se jím obecně na vlastnost někoho, něčeho. Co to však jsou nepřímé otázky? Příklady uváděné v SSJČ na to jasnou odpověď nedávají. Lze je v zásadě řadit do tří skupin:

- 1) *nevím, co vlastně chceš; nevím, v k-ém roce jste se narodil,*
- 2) *ptal se, co (je) nového; ptal se, s kým jsem jednal; ptal se, jaký je program,*
- 3) *psal, jaký měl úspěch.*

Příklady prvního typu lze hodnotit jako nepřímé otázky ve smyslu pragmatickém. Výpovědi uvedené sub 2) jsou pouhým popisem tázacího aktu, jejich ilokuční funkce tázací není. Příklad posledního typu za nepřímou otázku považovat nelze. Komárek (tamtéž) se k řešení SSJČ přibližuje tím, že zájmena všech uvedených typů pokládá za interogativa. Všimá si mimo jiné toho, že v povrchové struktuře stojí vždy po verbu dicendi nebo cogitandi, popř. po substantivu nebo adjektivu, které je výsledkem jeho nominalizace. Např. strukturu *Otázka, kterého malíře považuje za nejlepšího...* Komárek traktuje jako výsledek transformace: jejím východiskem byly postupně struktury *Ptal se, kterého malíře považuje za nejlepšího* (nepřímá otázka) a *Ptal se: „Kterého malíře považujete za nejlepšího?“* (přímá otázka). Celou řadu tzv. nepřímých otázek však takto interpretovat nelze; z výše uvedených příkladů ze SSJČ se to týká bodu 1) a zejména pak bodu 3). SSČ vymezuje významy tázacích/vztažných zájmen podobně jako SSJČ, nepoužívá však pojem „otázka nepřímá“ a problematické příklady zásadně neuvádí.

V LDB Pralex jsme lexikálněsémantickou strukturu tázacích/vztažných zájmen definovali s ohledem na jejich roli v povrchové syntaktické struktuře. Je třeba rozlišovat tři syntaktické typy vět, v nichž se tato zájmena objevují.

- A) Zájmeno přímo vyjadřuje otázku a zároveň ji uvozuje (před ním ovšem ještě může stát předložka).
Příklady: *Které šaty si vezmeš? Jaké bude počasí? Kdo to přišel?*
Primární syntaktická funkce zájmena: u tzv. adjektivních zájmen (*jaký, který, čím*) je adjektivní, u tzv. substantivních zájmen (*kdo, co*) substantivní; druh zájmena: tázací.
- B) Zájmeno uvozuje vedlejší větu a zároveň vyjadřuje její vztah zpravidla k určité osobě nebo věci uvedené ve větě řídící; zájmeno samo k této osobě nebo věci ukazuje.
Příklady: *Vzal si šaty, které nosí v neděli. Přítel, jakých je málo.*
Primární syntaktická funkce zájmena: vždy substantivní; druh zájmena: vztažné.
- C) Zájmeno uvozuje vedlejší větu, nevyjadřuje však její vztah k určité osobě nebo věci uvedené ve větě řídící ani samo k ničemu takovému neukazuje.

Příklady: *Ptal se, co je nového. Ptal se, kterého malíře považuje za nejlepšího. Otázka, kterého malíře považuje za nejlepšího... Nevím, co vlastně chceš.*

Primární syntaktická funkce zájmena: platí totéž co v A); druh zájmena: tázací? neurčitý?

Zájmeno typu C) nelze považovat za vztažné, ale ani za tázací v pravém slova smyslu (ne vždy signalizuje otázku). Zdá se nám, že svým významem stojí kdesi mezi zájmeny tázacími a neurčitými. Na blízkost obou druhů upozorňuje i Komárek (2006, s. 73): „Navrhujeme pojem neurčitosti rozšířit o slova tázací (interogativa): obojí totiž ukazuje k něčemu neurčitému, ne úplně identifikovanému v tom úseku skutečnosti, ke kterému má být výpověď referenčně vztažena: *Kdo přišel? – Někdo přišel.* Rozdíl mezi nimi je na první pohled v tom, že tázací zájmeno obsahuje v sobě apel k odstranění neurčitosti, k úplné identifikaci.“ Zdá se, že zájmena plnicí funkci C) tento apel implikují různě silně: silně (*Chtěl bych vědět, co bude dnes k večeři.*), slaběji (*Nevím, co vlastně chceš.*) nebo téměř vůbec (*Psal, jaký měl úspěch.*). Z nedostatku vhodnější klasifikace řadíme v LDB Pralex tento typ pod zájmena tázací.

6. Výklad významu

Při zpracování zájmen se opíráme především o výklad významu uvedený v SSJČ. Pokud zde výklad významu chybí, využíváme i další referenční zdroje. Např. uvádí-li SSČ u některého hesla více významů než SSJČ, opíráme se při zpracování hesla o něj.⁵ Podrobnější strukturaci významů v SSČ než v SSJČ najdeme např. u hesla *my*: SSČ na rozdíl od SSJČ uvádí také význam ‚označ. nejméně dva mluvčí zároveň jako autory společného projevu (zprav. písemného)‘. Pokud výklad významu z těchto slovníků nevyhovuje, upravujeme jej, popř. navrhujeme výklad vlastní (srov. níže výklad významu u zájmena *tvůj*).

Podobně jako např. u zdobnělin v LDB Pralex výklad významu obsahuje nejprve část derivační, realizovanou formou odkazu k motivujícímu heslu, a poté část popisnou, tak i výklad významu některých zájmen, konkrétně zájmen nereflexivních přivlastňovacích, je zřetelně dvousložkový. Například u zájmena *tvůj* byly výklady ze SSJČ sjednoceny tak, aby se u všech tří významů opakovala stejná část derivační: *tvůj 1*: ‚k ty; patřící, náležející osobě, k níž se v řeči obracíme; takový, kt. oslovená osoba má v majetku, ve vlastnictví, v držení‘; *tvůj 2*: ‚k ty; vyjadřuje jiný než vlastnický, zprav. však blízký, těsný vztah k této osobě (někdy s citovým zabarvením)‘; *tvůj 3*: ‚k ty; pocházející od tebe; tebe se týkající‘.

Pro srovnání uvedme i příklad naší pracovní výkladové definice používané v LDB Pralex u zdobnělin⁶:

lístek 1: ‚zdrob. k list 1; malý n. menší list rostliny‘ (upravený výklad ze SSJČ).

5 Opačný případ (který je nepoměrně častější, protože SSČ je slovník jednosvazkový, a tedy stručnější) je ponechán bez komentáře.

6 Ke zpracování zdobnělin viz zde stat' E. Čonosové.

Přes strukturální podobnost je mezi oběma typy zásadní rozdíl: zatímco u zdobnělin jde o derivaci slovotvornou, přivlastňovací zájmena jsou od zájmen osobních derivována jenom ve smyslu sémantickém. Lexikální význam zájmen přivlastňovacích je možno chápat jako lexikální význam zájmen osobních obohacený o další sémantické složky. U reflexivního zájmena *svůj* derivační část výkladu významu popositelně chybí.

7. Příkladová část hesla u zájmen v LDB Pralex

V exemplifikaci (stejně jako v hesláři) uvádíme zejména tvary spisovné; pouze okrajově, jen v odůvodněných případech, i tvary nespisovné (u *nějaký* 3: *Já jsem nějaké Štrosmajer... apod.*).⁷ Exemplifikaci u zájmen členíme na tzv. upravené a větné doklady. Fakultativně používáme i blok nazvaný Návrhy na VLJ (víceslovné lexikální jednotky). V něm uvádíme víceslovné lexikální jednotky registrované v některém z referenčních zdrojů, avšak doposud nezaložené v hesláři LDB Pralex.

V bloku větných (citátových) dokladů uplatňujeme řazení podle pádu. Objevují se zde zejména doklady, které z různých důvodů není vhodné uvádět v bloku upravených dokladů – např. zájmena plnicí sekundární syntaktické funkce (*málokterý* po elipse řídicího substantiva apod.). U některých lexií je blok upravených dokladů ponechán prázdný a vyplněn je právě jen blok větných dokladů. Systematicky takto postupujeme např. u zájmen tázacích nebo vztažných. Přednost větným dokladům dáváme i u lexií, které jsou v korpusu málo doloženy.

Blok upravených dokladů: na rozdíl od substantiv upravené doklady nemají explicitně stanovenou strukturu (členění do více bloků s titulky); tento blok členíme jen pomocí volných řádků na jednotlivé odstavce.

V bloku upravených dokladů je exemplifikační struktura specifická a) pro jednotlivé syntaktické typy, b) pro jednotlivé druhy zájmen.

Specifičnost typu a): U zájmen s primární syntaktickou funkcí substantivní uvádíme zejména spojení sloveso + zájmeno (*zabili ho*), substantivum + předložka + zájmeno (*většina z nás*), u zájmen s primární syntaktickou funkcí adjektivní především spojení zájmeno + substantivum (*takový člověk*).

Specifičnost typu b): Např. u osobních zájmen najdeme typy kolokací, které se u jiných druhů tzv. substantivních zájmen vůbec neobjevují: zájmeno + číslovka (*my dva*), zájmeno + substantivum (*já hlupák*).

U zájmen osobních uvádíme v jednom odstavci také jejich spojení s různými předložkami (toto je zde oproti jiným druhům zájmen navíc); předložky tu slouží především jako indikátor pádu. Cílem je ukázat relativně složitou morfologii osobních zájmen v jednotlivých pádech.

U rodových zájmen uvádíme v různých odstavcích jednotlivé rodové podoby. Tyto odstavce (v ideálním případě jsou čtyři) pak dále členíme podle typu kolokací,

⁷ Nespisovné modifikace lemmatu (*nějaký* > *ňáký*) nejsou v našem pojetí uváděny jako samostatné položky databáze, některé se však do hesláře dostaly na základě *Frekvenčního slovníku češtiny* (2004) – např. právě *ňáký*.

do nichž zájmeno vstupuje, např. tedy sloveso + zájmeno, substantivum + zájmeno atd. V každé takové skupině jsou doklady uváděny v pořadí podle pádu zájmena, a to od 1. do 7. pádu.

Třídění dokladů, dosud probíhající vertikálně (do jednotlivých bloků, resp. odstavců), má i rozměr horizontální: v jednom řádku jsou uváděny kolokace sémanticky blízké. Doklady jsou tříděny podle následujícího algoritmu: jm. rod > kolokační typ > pád > lexikální sémantika. Nevýhodou toho, že lexikální sémantika stojí v algoritmu až na posledním místě, je fakt, že se od sebe odtrhávají kolokace sémanticky příbuzné (např. doklady *můj otec, moje matka* jsou uvedeny v různých odstavcích). Přednost před hlediskem sémantickým tedy dostal zřetel formální – zohlednit oboje současně nebylo možné.

Při exemplifikaci rodových podob čtyřrodých zájmen motivovaných rodem přirozeným jsou upravené doklady rozděleny do bloků a, b, c atd. Týká se to osobních a přivlastňovacích zájmen 3. osoby označujících živé bytosti. Nominativní formy *on, ona, ono, oni, ony, ona*, podobně jako formy *jeho, její, jejich*, považujeme za tvary jednoho zájmena. U těchto hesel uvádíme v upravených dokladech jen kolokace, v nichž zájmeno označuje živou bytost / živé bytosti. Případy, kdy zájmeno označuje věci (a jeho forma je nutně dána kontextovou shodou, tj. není motivována rodem přirozeným), uvádíme ve větných dokladech.

Příklad: blok upravených dokladů u zájmena *tento* 2 (význam: ,odkazuje na osobu n. věc, o kt. právě byla n. je řeč n. kt. je známá‘):

tento muž, tento pán, tento chlap, tito lidé;
tento politik, tento umělec, tento novinář, tento architekt;
tito psi, tito koně;

tento názor, tento postoj;
z tohoto důvodu;
v tomto případě;
v tomto smyslu, v tomto ohledu;
tímto způsobem;
tyto věci, tyto skutečnosti;
tato žena, tato dívka, tato slečna;
v této souvislosti;
při této příležitosti;

toto téma;
v tomto městě, na tomto místě

8. Zájmena jako komponenty frazémů

Zájmeno se často stává komponentem frazémů; zájmena ale hypertextově propojujeme pouze s těmi frazémy, v nichž je zájmeno reálným komponentem. Např. heslo *můj II* (zájmeno) je takto provázáno s frazémy *jen přes mou mrtvolu, moje maličkost, to je má řeč, ty jsi mé živé svědomí aj.*

V některých frazémehch zájmena plní jinou funkci: a) potenciálně alternují s jinými výrazy téže větné platnosti, nebo b) specifikují valenční pozice predikátového slovesa nebo jiného nositele valence. Příklady ad a) /*to/ je můj člověk, leží /mu/ to na srdci jako skála*, ad b) *brát někoho za slovo, být někomu k ruce, být na něco krátký*. Funkci b) plní pouze několik zájmen neurčitých.

Určení hranice mezi kolokacemi, které lze interpretovat jako spojení autonomních lexikálních jednotek, a kolokacemi, které mají povahu frazému, je mnohdy problematické. Demonstrovat to lze na spojeních *toto ono, tento ono, tohle a onohle*. SSJČ postuluje pro lexém *tento* následující lexii:

6. nespis. spoj. *toto ono, tento ono, zř. tohle a onohle má význam zájm. neurč. něco I 1, leccos, všelicos (místo urč. vyjádření n. při hledání výrazu): musí to být jako by stříhl, žádné toto ono, takové kmotrovské dohadování (Šlej.); začal (krejčík) kritisovat, že prý tohle a onohle (K. Čap.)*

SSČ takovou funkci zájmena *tento* výslovně neuvádí. Uvedená spojení mají nepochybně některé znaky frazémů:

- a) sémantická nesamostatnost komponentů: význam je nutno připisovat až spojení jako celku; nelze identifikovat diskrétní antecedenty obou demonstrativ;
- b) morfologicko-syntaktická anomálnost: v kolokaci *tento ono* vyjadřují zájmena disjunktní hodnoty jmenného rodu.

V korpusu ovšem najdeme příklady, kdy antecedenty obou zájmen jsou sice neznámé, ale tvoří disjunktní množiny charakterizované odlišným predikátovým znakem: opus.rokvyd=1996,opus.txtype=COL,opus.id=uzhorela *A někdo z rodiny nákup vybaluje a zajímá se, proč jste koupili tohle a nekoupili onohle, a vůbec kde že je tamto?*

To naznačuje, že si obě zájmena zachovala jistou lexikální samostatnost. V uvedeném kontextu lze navíc zájmena *tohle – onohle* interpretovat ve smyslu bližšího a vzdálenějšího ukazování; přibližují se tak k významům *tenhle 4, onen 2* (význam zájmena *onenhle* vykládá SSJČ odkazem na *onen*), srov. *tento 4*:

4. *ukazuje na bližší ze dvou osob n. věcí (v situaci n. v souvislosti kontextu), zejm. ve dvojicích tento onen, tento tamten, tenhle tamten: tento je tlustý, onen hubený; toto stanovisko je správné, ono nesprávné; tento papír, materiál je kvalitní, tamten špatný; tahle kniha se mi líbila, ale tamta za moc nestála; Morava severní na dvě rozdělená, z této i z oné strany Dunaje (Šaf.); užívají (kacíři) dvou věcí: výmluvnosti a přisnosti. Oné na lid příliš prostý, této na muže příliš bázlivé (Vanč.); ust. spoj. ... (zkrátit V. V.)*

Srov. *onen 2*:

2. ve spoj. *ten onen, tento onen ukazuje na jinou z několika osob n. věcí; jeden druhý: truhlice byla brzo u toho, brzo u onoho rychtáře v úkrytu (Jir.); ukazovala rukou hned v tuto, hned v onu stranu; prohlížel si hned to, hned ono; přemýšlí, jak to či ono udělat*

Případy, kdy je lineární sled zájmen takto „přerušeno“ dalšími výrazy, by bylo složitě zpracovávat jako frazémy i technicky. Přiklonili jsme se proto k řešení SSJČ: *toto ono, tento ono i tohle a onohle* pokládáme za spojení autonomních lexikálních jednotek.

Literatura

- BĚLIČ, Jaromír: Bez něj je to těžké. In: Naše řeč, 1977, roč. 60, č. 2, s. 57–67. Čeština – řeč a jazyk. Ed. M. Čechová. 2., přepracované vyd. Praha: ISV nakladatelství 2000. 407 s.
- DANEŠ, František – HAUSENBLAS, Karel: Přivlastňovací zájmena osobní a zvrtná ve spisovné češtině. In: Slavica Pragensia, 4. Acta Universitatis Carolinae. Philologica. Věnováno akademiku Bohuslavu Havránkovi k sedmdesátým narozeninám. Red. V. Barnet. Praha: Univerzita Karlova 1962, s. 191–202.
- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSC)
- HIRSCHOVÁ, Milada: Netytické případy užití výrazů takový, tak. In: Naše řeč, 1988, roč. 71, č. 2, s. 57–61.
- HLAVSA, Zdeněk – SVOZILOVÁ, Naděžda: K povaze významu u zájmen. In: Slovo a slovesnost, 1969, roč. 30, s. 120–124.
- KOMÁREK, Miroslav: Systém autosémantických slovních druhů ve spisovné češtině. In: KOMÁREK, M.: Příspěvky k české morfologii. Olomouc: Periplum 2006, s. 13–94.
- MACHÁČKOVÁ, Eva: K přivlastňovacím zájmenům vztažným. In: Naše řeč, 1990, roč. 73, č. 5, s. 229–234.
- Mluvnice češtiny 2. Tvarosloví. Red. M. Komárek – J. Kořenský – J. Petr. Praha: Academia 1986. 536 s.
- Příruční mluvnice češtiny. Eds. P. Karlík – M. Nekula – Z. Rusínová. Praha: Nakladatelství Lidové noviny 1995. 800 s.
- SGALL, Petr: K tvarům mně – mě, mi – mne. In: Naše řeč, 1961, roč. 44, č. 5–6, s. 163–168.
- SCHNEIDEROVÁ, Eva: K užívání zájmena ten (v přivlastkové pozici) v mluvených projevech. In: Naše řeč, 1993, roč. 86, č. 1, s. 31–37.
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- Slovník spisovné češtiny pro školu a veřejnost. Red. V. Mejstřík a kol. 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- VESELÝ, Vojtěch: Lexikální sémantika neurčitých kvantifikátorů v současné češtině. In: Slovo a slovesnost, 2009, roč. 70, č. 3, s. 175–192.
- VOBORSKÁ, Milada: Zápis morfologických dat – návrh řešení pro lexikální databázi LEXIKON 21. In: Lexikografie v kontextu informační společnosti. Sborník příspěvků z pracovního setkání Praha, 4.–6. 6. 2007. Eds. A. Rangelova – J. Světlá – A. Jarošová. Praha: Ústav pro jazyk český AV ČR, v. v. i., 2008, s. 41–46.
- ZIMOVÁ, Ludmila: O vyjadřování podmětu osobními zájmeny 1. a 2. osoby. In: Naše řeč, 1988, roč. 71, č. 1, s. 22–32.

Zpracování vlastních jmen v LDB Pralex¹

Jana Letafková

Ústav pro jazyk český AV ČR, v. v. i.

Processing of Proper Names in the Pralex Lexical Database

The purpose of this study is to introduce a processing of proper names in the Pralex lexical database (LDB). The author gives her reasons for inclusion of proper names in the Pralex LDB and for the criteria of their selection. The proper names are divided in three categories in here: personal names, geographical names and the other proper names. The study is mainly focused on the processing of personal names, especially surnames, and their variation. Moreover, it also deals with their homonymy and polysemy.

1. Úvod

Vlastní jména (*propria*) tvoří v jazyce zvláštní systém, který je odlišuje od jmen obecných. Považují se za doplňkovou zásobu jazyka (Dokulil 1970) či se řadí mimo slovní druh (Trost 1958). Jejich odlišnost je dána především oslabenou sémantikou a rozdílnými funkcemi, které v jazyce plní. Podstatná je jejich **individualizační, identifikační a diferenciací funkce**, předmět pouze označují, nikoliv charakterizují. Podle *Encyklopedického slovníku češtiny* (2002) je jejich funkcí označovat jedince, jednotlivinu nebo jako individuum chápané množství, odlišovat je od ostatních jedinců nebo jednotlivin dané třídy a identifikovat je jako jedinečné objekty. Od obecných jmen se liší především absencí generického významu (neobsahují pojmové sémy, nýbrž pouze gramémy). P. Trost (1958) za vlastní jméno považuje slovo, které má identifikovat na základě individuální konvence, ne na základě všeobecného významu. M. Knappová (1980) je hodnotí jako lexikálně neplnohodnotná, defektní. Řadí je do oblasti periferní slovní zásoby. Z formálního hlediska jsou vlastní jména pouze substantivy, nebo jde o slova užitá v jejich funkci, anebo často tvoří víceslovná substantivní pojmenování.

K vlastním jménům je proto třeba přistupovat jinak než k běžné slovní zásobě. Specifické vlastnosti, které je odlišují od apelativ, vyžadují i rozdílné způsoby zpracování ve výkladových slovnících a vedou ke vzniku speciálních onomastických slovníků. Ve výkladových slovnících jsou často uváděna jen formou souhrnných přehledů v příloze. Tak je tomu ve *Slovníku spisovné češtiny pro školu a veřejnost* (1978, dále SSČ), který obsahuje přílohu rodných jmen, příjmení a jmen zeměpisných. Některá *propria* se však nacházejí také v těle tohoto slovníku: „V abecední části slovníku se uvádějí z vlastních jmen a jejich odvozenin jen apelativizované, tj. tvořící součást běžné lexikální zásoby a frazeologie.“² Čtyřsvazkový výkladový *Slovník spisovného*

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je Doc. RNDr. Karel Oliva, Dr.

2 *Slovník spisovné češtiny pro školu a veřejnost, Zásady zpracování slovníku*. Praha, Academia, 2. vyd. 1994, str. 641.

jazyka českého (1960–1971, dále jen SSJČ) řadu proprií zařadil do těla slovníku, tvoří zde tedy samostatná hesla s údaji gramatickými a/nebo věcnými. Uplatňují se ve frazeologii, v terminologii, dochází u nich k přenesenému užití či apelativizaci: „Vlastní jména osobní a místní budou do slovníku pojata v těch případech, kdy se u nich vyvinulo také užívání ve významu obecném a přeneseném (*Jidáš, babylon*), jinak jen tehdy, vyskytuje-li se u nich nějaký normativní problém při skloňování nebo při odvozování (*Boleslavě–Boleslavi, Aleše–Alše, haažský–haagský*), ev. pochybnosti o jejich pravopisné (*Alláh, Anči*) nebo hláskové (*Alsasko*) stránce. U jmen národů, zemí a států přihlížíme však také k jejich politickému, hospodářskému a kulturnímu významu v dějinách i pro dnešek.“³

V rámci současného výzkumného záměru se soustředujeme na budování lexikální databáze (dále LDB) Pralex, která by měla obsahovat okolo 100 000 až 120 000 položek a která bude východiskem pro budoucí zpracování moderního výkladového slovníku češtiny, jehož předností má být především bohatá příkladová část. Vzhledem k uvedenému rozsahu jsme se rozhodli do LDB Pralex zahrnout i mnohá propria. Argumenty pro zařazení proprií do LDB Pralex vycházejí především z toho, že se v první etapě prací stal základem pro heslář *Frekvenční slovník češtiny* (dále FSČ), zahrnující i samostatnou část *Frekvenční slovník proprií* s 2000 vlastních jmen, a z toho, že byla propria uvedena i v SSJČ, který se pro nás stal dalším východiskem⁴. Jedním z podstatných důvodů pro zařazení do našeho hesláře se stal i fakt, že se od proprií derivují další slova – např. přídavná jména včetně přivlastňovacích, jména obyvatelská aj., a taktéž skutečnost, že bohatě vstupují do frazeologie (přirovnání, pranostiky). Zajímavou problematiku představuje i proprializace apelativ a apelativizace proprií. V neposlední řadě jsou propria i přitažlivým tématem pro širokou veřejnost, možná právě pro převažující encyklopedičnost zpracování.

2. Výběr proprií do hesláře a jeho postupné doplňování

Jak jsme uvedli výše, základem pro heslář LDB Pralex se stalo 2000 proprií z *Frekvenčního slovníku proprií*. Ten obsahuje nejčastější vlastní jména česká i cizí z korpusu SYN2000, uspořádaná podle průměrné redukované frekvence. Výrazně se v něm projevuje konkrétní dobové zakotvení korpusu a vliv médií.⁵ Z antroponym jsou zde nejčastěji rodná jména a jejich přivlastňovací podoby, dále příjmení, ale můžeme zde najít i hypokoristika, jména biblická a mytologická, jména literárních a pohádkových postav. V hojně míře se vyskytují též jména příslušníků národů a jména obyvatelská. Z toponym jsou nejvíce zastoupeny názvy států a českých i cizích měst. V neposlední řadě přináší FSČ i značné množství chrématonym různého druhu (od názvů firem, podniků a sportovních klubů až po názvy značek výrobků).

3 Helcl, M. a kol.: Směrnice pro vypracování rukopisu slovníku spisovného jazyka českého. Praha 1957, str. 3.

4 K hesláři viz zde podrobněji ve stati J. Světlé *Budování hesláře LDB Pralex*.

5 Frekvenční slovník češtiny, str. 12.

Dalších více než 3000 proprií ze SSJČ uvedených v těle slovníku jako samostatná hesla jsme do hesláře vložili ručně a postupně doplňujeme i další propria ze SSJČ uvedená uvnitř heslové stati apelativ. Jsou to propria různého druhu, z různých tříd a podtříd. Každá třída proprií je specifická a je třeba k ní přistupovat diferencovaně, proto jsme se rozhodli, že je budeme zpracovávat zvlášť. Ke každé třídě proprií jsme navrhli a použili vlastní formalizované výkladové formulace (v tomto případě se tedy výkladová definice neopírá o SSJČ jako u substantiv).

Nejprve však bylo nutné alespoň pracovní takové třídy proprií stanovit. Proto jsme nejdřív (zvlášť k heslům ze SSJČ) vkládali místo přesných výkladových formulací jen pomocná klíčová slova, která nám usnadňují vyhledávání dílčích skupin (např. země, hydro, obyv, mytol atd.), než bude pro každou třídu proprií naformulována celá výkladová definice.

Doplňování do hesláře se však neomezilo pouze na FSC a SSJČ. Systematicky se v průběhu roku 2008 až 2010 ručně vkládala další propria na základě vybraných příruček a seznamů.

Jako první podtřidu u jmen osob jsme začali doplňovat a zpracovávat **rodná jména a hypokoristika**. I ta byla doplňována podle speciálních seznamů. V první etapě práce byla do hesláře zapsána všechna jména obsažená v českém občanském kalendáři (včetně informace o jmeninách). Pro doplnění rodných jmen jsme dále postupovali podle příručky M. Knappové *Jak se bude vaše dítě jmenovat?*. Základem nám bylo její třetí vydání z roku 1999, které obsahuje přes 1700 heslových podob mužských a ženských rodných jmen, které splňují jazykové podmínky pro zápis jmen do matriky.⁶ Jsou zde uvedena především jména v ČR známá a obvyklá, často užívaná dnes i v dřívějších dobách, a dále nejznámější jména z našich, evropských i světových dějin, z kulturní historie i současného dění. Celkově jsme takto rozšířili heslář LDB Pralex o dalších asi 1300 mužských a ženských rodných jmen. Do našeho hesláře nebyla zařazena zvlášť tzv. polohešla, jak je Knappová označuje, uvádějící jen některé údaje. Často se totiž jedná o okrajová jména či cizojazyčné ekvivalenty českých či počeštěných podob (např. *Idzi, Lajos*).

V rámci jmen osobních v současné době probíhá také zakládání a zpracování **mužských a ženských příjmení**. Východiskem se pro nás stala příloha SSČ. Tento soupis obsahuje 1500 nejčastějších mužských příjmení občanů ČR. Zvlášť k nim rozepisujeme i příjmení ženská (tedy jako samostatné položky databáze).

Současně se zakládají a zpracovávají i vybraná **jména světců a světic**. Jako východisko pro doplňování je nám příručka Václava Rameše *Po kom se jmenujeme?* (2005), opíráme se ale též o knihy *O světcích a patronech* (Ravík 2006), *Český kalendář světců* (Pracný 1994), *Ikonografie a atributy svatých* (Ramešová 1991) aj.

Dále se uvažuje o postupném doplňování dalších podtříd jmen osobních, zejména v závislosti na tom, jaké referenční zdroje pro výběr do hesláře máme k dispozici. Po dokončení zpracování rodných jmen budou adekvátním způsobem zpracována například **jména antická a mytologická**.

6 Na začátku prací jsme neměli k dispozici poslední vydání z roku 2006, kde byly provedeny různé změny a doplňky; podle tohoto vydání budeme postupovat při revizi zpracovaných hesel.

V našem hesláři jsou též **zeměpisná jména**, nejprve založená podle FSC a SSJČ, později doplněná o všechna česká města a o většinu obcí ČR⁷. Postupně probíhá doplňování zeměpisných jmen v SSJČ přihnízdovaných k apelativním heslům.

V rámci proprií si všímáme i víceslovných jednotek, které se v SSJČ objevují na konci hesla po frazémeh nebo jen mezi příklady k významu. Takové propriální názvy (např. s komponentem hora, jezero aj.) zatím uvádíme v LDB Pralex u příslušných substantiv, resp. adjektiv, která jsou jejich komponentem, v bloku V názvech. Později bude rozhodnuto, které typy budou v LDB Pralex zpracovány samostatně. Výjimku tvoří víceslovné názvy měst a obcí, ty jsou v našem hesláři již založeny jako samostatné položky.

3. Zpracování proprií v LDB Pralex

Propria pracovním členíme na tři základní oblasti: jména osob, jména zeměpisná a ostatní. Nejprve se zde zaměříme na některé otázky související s proprií obecně, poté na zpracování prvního typu – jmen osob v LDB Pralex.

3.1. Homonymie a polysémie u vlastních jmen

K homonymii a polysémii vlastních jmen se jazykovědci staví různě. Někteří přímo odmítají o homonymii v souvislosti s vlastními jmény uvažovat (vycházím ze stati E. Pokorné). Pro potřeby budoucího slovníku je však potřebné homonymii a polysémii u proprií rozlišovat. Naše pojetí navazuje na M. Knappovou (1980), která homonymii chápe jako vztah dvou jmen náležejících v promluvě k různým onomastickým polím (*Orava* – příjmení, název řeky a hradu). Polysémii pak jako vztah dvou jmen náležejících do stejného, např. antroponymického pole (*Karel* – jméno, příjmení).

Pokud se tedy stejná grafická podoba vyskytne u různých tříd proprií, například rodné jméno a jméno zeměpisné, v LDB Pralex řešíme tyto případy v rámci homonymie, např. mužské jméno vs. obec (*Vlastislav I*, rodné jméno mužské; nositel tohoto jména; *Vlastislav II*, obec v Ústeckém kraji, okres Litoměřice; podobně u hesel *Miroslav I*, *Miroslav II*).

Taktéž pokud se vyskytne stejná grafická podoba uplatněná v různých významech pro různý rod, jsou taková lemmata zpracována jako dvě samostatná hesla, označena římskými číslicemi jako homonyma, analogicky se zpracováním substantiv (např. hesla *Milota I*, rodné jméno mužské; nositel tohoto jména; *Milota II*, rodné jméno ženské; nositelka tohoto jména).

3.2. Apelativizace a proprializace

Vztah apelativizovaných proprií a proprializovaných apelativ k výchozímu propriu, resp. výchozímu apelativu v LDB Pralex evidujeme hypertextovým propojením mezi příslušnými hesly prostřednictvím seznamu Obecná/vlastní jména (např. *Adam*, *adam*). Obě taková hesla se zakládají a zpracovávají samostatně, jejich pro-

⁷ Názvy měst a obcí ČR jsme zakládali podle soupisu na adrese www.statnisprava.cz a kontrolovali podle informačního portálu www.obce.cz. K nim byla rovnou založena (podle příručky A. Polívkové *Naše místní jména a jak jich užívat*) a hypertextově propojena příslušná adjektiva od nich tvořená.

pojení nám umožňuje jejich vzájemné porovnávání. Již teď se ukazuje, že zvláště apelativizace proprií bude zajímavým tématem k dalšímu lingvistickému zkoumání. V korpusovém materiálu se řada již apelativizovaných proprií, tak jak je uvádí SSJČ, sice chová jako apelativa, tj. mají schopnost zobecnění, charakterizují předmět, který označují, ale uživatelé je často stále píšou s velkými písmeny, tedy jako propria. Řešení takových hesel bude v budoucím výkladovém slovníku jistě velmi podnětné.

3.3 Propria jako komponenty frazémů

Vlastní jména, která jsou zpracovávána jako samostatná hesla v rámci souboru vlastních jmen, se také často objevují jako komponent různých frazémů (a méně často sousloví).

3.4. Zpracování jmen osobních

První velkou oblastí vlastních jmen, kterou jsme se začali zabývat, byla jména osob. Zde je na místě připomenout, že LDB Pralex má být materiálovou základnou pro nový výkladový slovník. Nemůžeme si proto činit nároky na tak podrobné zpracování, jaké mají propria ve speciálních onomastických slovnících, a také naše cíle a východiska jsou pochopitelně jiná. Je však zřejmé, že specifika vlastních jmen jsou patrná i v naší databázi. Výklady významů ustupují do pozadí a převládá encyklopedičnost informací.

Jako východisko při přípravě zásad zpracování jmen osob nám posloužila rodná jména, u nichž je možné rozlišit v rámci polysémie různé významy (rodné jméno, příjmení, jméno světce, biblické postavy atd.). Museli jsme si také vymezit hranice homonymie a polysémie a určit pořadí významů.

Už v počátku prací jsme se rozhodli, že nebudeme významové definice přebírat ze SSJČ, jak je tomu u substantiv, ale postupně vymezujeme **formalizované výklady podle tříd a dílčích typů**.

U jmen osob vypadají naše definice a pořadí zápisu do LDB Pralex takto:

1. *rodné jméno mužské; nositel tohoto jména* a paralelně k tomu *rodné jméno ženské; nositelka tohoto jména*; 2. *mužské příjmení; nositel tohoto příjmení*, resp. *ženské příjmení; nositelka tohoto příjmení*; 3. *domácká podoba k ...* (např. heslo *Anina – domácká podoba k Anna*); 4. *jméno světce*, resp. *svěťce; jeho/její výroční den*; dále pak např. *biblická postava; mytologická postava; pohádková postava; jméno zvířete* atd. (zde bude buď tato obecná formulace, nebo konkretizace podle běžnosti daného jména, např. *jméno psa, kočky* apod.).

4. Zpracování rodných jmen

Zpracování rodných jmen v LDB Pralex probíhá od června 2009. Jelikož má být tato lexikální databáze především materiálovou základnou pro budoucí výkladový slovník, odpovídá tomu i výběr a zpracování jednotlivých hesel⁸. V současné době se

8 Celkem LDB Pralex ke konci roku 2010 obsahuje téměř 2200 rodných jmen zpracovaných jako samostatné položky databáze; přibližně u dvou pětín z nich registrujeme další významy (2–4). Zpracována zatím nejsou poslední písmena abecedy (V–Ž).

zaměřujeme především na příkladovou (exemplifikační) část hesla. Základním materiálovým zdrojem je nám Český národní korpus, z kterého využíváme především korpus SYN (slučující korpusy SYN2000, SYN2005 a SYN2006PUB)⁹. Nezbytným pomocníkem je korpusový manažer Bonito a program Word Sketch Engine.

4.1. Exemplifikace u rodných jmen

Pro rodná jména většinou není relevantní struktura exemplifikace používaná u substantiv, ale přesto rozlišujeme i zde mezi větnými (citátovými) a tzv. upravenými doklady¹⁰. V bloku upravených dokladů jsme se rozhodli uvádět zejména spojení dokladující rodinné vztahy; označení funkce, pozice; církevní funkce; osobnosti vládnoucí vrstvy s titulem; společenské vztahy a profese. K tomu viz např. hesla: *Alois, Ferdinand, Kateřina, Leopold* aj.

Členění bloku **Větné doklady** jsme si interně vymezili takto:

E1a: doklady související s výkladem jména, jeho původem, svátkem apod.;

E1b: doklady související s nominací (spojení typu: jmenovat se ..., říkali mu ..., dát někomu jméno); oslovení (5. p.); jméno a příjmení: pochází, narodil se, vzal si, oženil se (rodinné události a vztahy); rodák/rodačka ...; v případě problematického či variantního skloňování zařazujeme doklady na různé pádové zakončení (např. u hesla *Andrea*);

E1c: zajímavosti či podrobnosti kolem významných osobností;

E1d: postavy z literárních aj. děl, filmové a divadelní role atd.

V případě doloženosti v korpusu uvádíme v exemplifikaci zvlášť **Pranostiky** (pokud se však pranostika vztahuje ke světci/světici, zařazujeme ji k tomuto významu), viz např. heslo *Agáta, Medard*. Případně též evidujeme další specifika v bloku **V názvech** (např. u hesel *Amos, Markéta*) nebo v encyklopedické poznámce (data svátků u jmen osob apod., např. u hesla *Kryštof, Lambert*).

4.2. Variantnost u rodných jmen

Tak jako u jiných slovních druhů a typů hesel platí i u vlastních jmen zásada, že všechny varianty jsou uváděny jako samostatné databázové položky. Hypertextové propojení variant pomocí seznamu nám umožňuje jejich paralelní zpracování a vzájemné porovnávání.

Při uvádění variant v LDB Pralex se u rodných jmen opíráme o příručku M. Knapkové *Jak se bude vaše dítě jmenovat?* (1999). Propojena jsou hesla, která jsou u Knapkové uvedena v jednom heslovém odstavci za sebou, mají stejnou grafiku – kapitálky, tučně (např. *Abdon, Abadon*), nebo hesla, u kterých je poznámka „variantou je“ (např. *Albert*, variantou je *Albertýn, Albertin*); jedná se především o pravopisné varianty typu *Albertýn/Albertin, Alfons/Alfonz, Absolón/Absolon*, či o tvarové varianty *Aleška/Aleša, Anděla/Andělína, Astrid/Astrida*.

9 Od začátku roku 2011 je původní korpus SYN pro naše potřeby přejmenován na orig_syn, aby se původní název SYN zachoval pro potřeby zřizovatele.

10 Srov. též úvodní stať J. Světlé *Struktura a zpracování hesel v lexikální databázi Pralex* a stať M. Koutové o exemplifikaci substantiv.

Za varianty nepovažujeme tyto dvojice hesel: rodné jméno a jeho domácké podoby; rodné jméno a apelativum; rodné jméno a jeho cizojazyčné ekvivalenty (např. *Antonín x Antony x Antonio, Anděl/Andělín x Angel, Anna x Anne x Annie, Albert x Alberto, Alan/Alen x Alain, Anatol x Anatolij*); u cizích jmen české podoby jména, u českých jmen cizí podoby jména, která jsou uvedena tučně (např. *Artur x Artuš, Antonie x Antonína, Anděla/Andělína x Angela, Apolena x Apolonie*); staročeské podoby jmen (např. *Adéla*, její staročeskou podobou je *Adléta*), starší či zastaralé podoby jmen (*Dušan x Duchoslav, Adrian/Adrián x Hadrián*). Variantami v našem pojetí nejsou ani jména odvozená z téhož základu, např. *Adalbert x Albert x Albrecht*; jména, která vznikla zkrácením jiného jména (často s poznámkou „osamostatnělou zkráceninou je ...“, např. *Alexej x Alex, Barbora/Barbara x Bára, Dorota x Dora*; nebo s poznámkou „původem osamostatnělá zkrácenina jména ...“, např.: *Haidi*, původem osamostatnělá zkrácenina německého j. *Adelheid (Adelaida)*).

Naopak propojeny jsou také ty varianty, které uvádí ve společném hesle SSJČ, a dále taková hesla, jejichž jedna varianta je ze SSJČ, druhá je přidána ze seznamu (*Jak se bude vaše dítě jmenovat?*). A taktéž zakládáme a propojujeme varianty, na které narazíme při vyhledávání v korpusu na základě společné lemmatizace (většinou jde o pravopisné varianty typu *Alfréd/Alfred, Agáta/Agata*). Uvedme příklad. Do našeho hesláře se z FSC dostala hesla *Andrej, Andreas, André, Andrew a Ondřej*. V souladu s M. Knappovou (*Jak se bude vaše dítě jmenovat?*)¹¹ tato hesla nechápejme jako varianty; vztah mezi českými a cizími ekvivalenty zatím nikde nezachycujeme.

5. Zpracování polysémních hesel u jmen osob

Jak již bylo naznačeno výše, vztahy v rámci jedné třídy proprií vyhodnocujeme jako polysémii. V takovém případě jednotlivé významy řadíme v jednotném pořadí (tedy – v relevantních případech – nejprve rodné jméno, pak biblické jméno vč. jmen proroků, a dále antické jméno / mytologická postava, jméno světece/světece, domácká podoba vč. zdrob. a expr. zdrob., příjmení, pohádková postava, jméno zvířete atd.).

Rodná jména mají často stejnou grafickou podobu jako další podtřídy jmen osob. Velice frekventované jsou v našem materiálu z korpusu v tomto smyslu kombinace rodného jména s příjmením nebo se jmény světců či světic. Právě proto jsme se rozhodli, že je budeme zakládat a zpracovávat paralelně s rodnými jmény v rámci polysémie. Naopak jména antická a mytologická zpracováváme zatím jen tehdy, když jsou dalším významem rodného jména. Jako soubor budou zpracována jednotně až později (po dokončení souboru rodných jmen).

11 Srov. z této příručky heslo: **ANDREJ**, slovanská podoba řeckého j. **ANDREAS** a z francouzštiny přejatého **ANDRÉ**, znamená „statný, mužný, odvážný“; českou podobou jména je Ondřej. – *Dom. podoby*: And(y), Andráš, Androuš(ek) apod. – *Slov., rus., bulh., sch.* Andrej, *něm., hol.* Andreas, *angl.* Andrew, *pol.* Andrej, *lat., franc.* André, *ital.* Andrea, *špaň.* Andrés, *dán., švéd.* Anders, *mad.* Andrés, *Andor*, *Endre*, *fin.* Endre – Andrej stále oblíben, Andrea běžný, André se vyskytuje zřídka. – *Jmeniny*: 11. října.

Jména světců a světic

Jednak jsou často dalším významem rodných jmen¹², jednak bývají zakládána též jako samostatná víceslovná hesla s označením „sv.“, např. *sv. Havel*, *sv. Alžběta*, a to v případě, že se jedná o jméno, pro nějž máme dostatek podkladů v korpusu či v odborné referenční literatuře, o kterou se opíráme (viz výše)¹³. Výklad významu pak obsahuje i bližší informace o světci či světici a století, ve kterém žili (např. heslo *sv. František z Assisi – světec; vyznavač a zakladatel františkánského řádu, 13. století*).

U jednovýznamových jmen světců a světic uvádíme definici podle referenčního zdroje a v exemplifikaci doklady z korpusu o životě, skutcích či zázracích konkrétního světce. Víceslovné jazykové jednotky propojujeme s příslušnými hesly jednoslovnými (viz např. hesla *sv. Markéta z Antiochie a Markéta*, *sv. Kateřina Švédská a Kateřina*, *sv. Filip Neri a Filip*, *sv. Havel a Havel*).

U polysémních jmen světců a světic, která jsou dalším významem rodných jmen, používáme tuto výkladovou formulaci: *jméno světce/světice; jeho/její výroční den*. V encyklopedické poznámce pak uvádíme plné jméno každého světce a datum, na které připadá jeho výročí. Do exemplifikace vybíráme doklady vztahující se především k výročnímu dni daného světce a spojení typu *kostel sv. Ambrože*, *řád sv. Huberta* apod. V relevantních případech zakládáme další blok s nadpisem **Postava svatého/svaté** (např. u hesla *Ambrož, Barbora, Mikuláš*), **Pranostiky a V názvech**. Oporou při vyhledávání pranostik v korpusu je nám kniha Z. Vašků *Velký pranostikon*.

Příjmení

Nejprve zakládáme jako hesla a zpracováváme ta česká příjmení, na která narazíme v korpusu při zpracování rodných jmen (tj. mají stejnou formální podobu), a dále pak heslář doplňujeme systematicky podle seznamu ze SSČ. Označení pro rodinu či některé její členy (typ *Novákoví*) přidáváme k mužským příjmením do samostatného exemplifikačního bloku nazvaného *jméno rodiny, její členové*. Varianty zatím nepropojujeme. Obourodá příjmení typu *Krejčí* jsou v LDB Pralex rozepsána zvlášť a označena jako homonyma.

Hypokoristika

Zpracování hypokoristických hesel v LDB Pralex probíhalo paralelně se zpracováním rodných jmen a je už dokončeno – viz k tomu příspěvek R. Volejníka. Hypokoristická podoba rodného jména nabyla v některých případech charakteru základní podoby rodného jména, a pak tedy existují vedle sebe dvě totožné grafické podoby, z nichž jedna má charakter základní podoby jména a druhá hypokoristické podoby jména (srov. např. *Dora, Bára*). V korpusovém materiálu bylo někdy velmi těžké přiřadit doklady k příslušnému významu, proto byly v takovém případě všechny doklady umístěny k prvnímu z významů a v Celkové poznámce k exemplifikaci uvádíme: *společné doklady pro V1 a V2*.

12 Tak je tomu u více než 500 propriálních hesel.

13 Takto bylo zatím zpracováno téměř 300 hesel.

Závěr

Vlastní jména jako součást české slovní zásoby tvoří obsáhlý soubor, který není v intencích výzkumného záměru ani v silách našeho kolektivu zpracovat celý. Proto jsme se zaměřili především na ty třídy proprií, které byly nejvíce zastoupeny ve FSČ (jména osob, českých měst a obcí; paralelně nás čeká jak zpracování dalších jmen osob, tak i zeměpisných jmen a s nimi souvisejících jmen obyvatelských a jmen příslušníků rodů, dynastií, kmenů a národů).

Zpracování vlastních jmen v LDB Pralex je pro budoucí slovníkové zpracování důležité tím, že se stává východiskem nejen pro řadu dalších hesel odvozených, ale i pro zpracování frazémů, jejichž komponenty jsou v některých případech právě propriální. Všechna hesla, která spolu nějak souvisí, jsou propojována hypertextovými odkazy. Díky encyklopedickému charakteru těchto hesel od počátku pracujeme i s výkladovou definicí, kde jsme rovnou navrhli vlastní jednotné formulace pro dílčí typy. Některé typy proprií jsou již v podstatě zpracovány slovníkově, i když do slovníku bude moci být zahrnut jen užší výběr. Kromě toho i příkladová část propriálních hesel zpracovaných v LDB Pralex představuje cenný zdroj informací pro budoucí uživatele lexikální databáze i výkladového slovníku, neboť větné doklady mají výrazně encyklopedický charakter.

Na řadu koncepčních otázek jsme si již odpověděli, je jisté, že s dalšími třídami a podtřídami proprií přijdou otázky další, které se budeme snažit vyřešit v souladu s teoretickými závěry lingvistických studií, jež se vlastními jmény zabývají, i v souladu s potřebami všeobecného výkladového slovníku. Praktickou inspirací nám bude zpracování vlastních jmen v předchozích českých i zahraničních výkladových slovnících i v různých příručkách, které máme k dispozici v tištěné podobě.

Literatura

- DOKULIL, Miloš: O vyjadřování jednosti a jedinečnosti v českém jazyce. In: Naše řeč, 1970, roč. 53, č. 1, s. 1–15. Encyklopedický slovník češtiny. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: Nakladatelství Lidové noviny 2002. 604 s.
- FILIPEC, Josef – ČERMÁK, František: Česká lexikologie. Praha: Academia 1985. 281 s.
- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSČ)
- HŮRKOVÁ, Jiřina: Antická jména. Jak je skloňovat a číst. Praha: AMU 2005. 165 s.
- KNAPPOVÁ, Miloslava: Jak se bude vaše dítě jmenovat? Praha: Academia 1978, 2. vyd. 1985, 3. vyd. 1999, 4. vyd. 2006.
- KNAPPOVÁ, Miloslava: Lexikální a slovtvorná charakteristika českých rodných jmen. In: VIII. slovenská onomastická konferencia. Banská Bystrica – Dedinky 2.–6. júna 1980. Zborník materiálův. Red. M. Majtán. Bratislava – Banská Bystrica – Prešov: Jazykovedný ústav Ľudovíta Štúra SAV v Bratislave – Pedagogická fakulta v Banskej Bystrici – Pedagogická fakulta Univerzity Pavla Jozefa Šafárika v Prešove 1983, s. 307–313.
- KNAPPOVÁ, Miloslava: Významové aspekty vlastních jmen. In: Slovo a slovesnost, 1980, roč. 41, s. 57–60.

- KOPEČNÝ, František: Průvodce našimi jmény. 2., přepracované a rozšířené vyd. Praha: Academia 1991. 259 s.
- Manuál lexikografie. Eds. F. Čermák – R. Blatná. Jinočany: H&H 1995. 283 s.
- PRACNÝ, Petr: Český kalendář světců. Praha: EWA Edition 1994. 333 s.
- POKORNÁ, Eva: Vymezení vlastních jmen, gramatika a materiálové a metodologické práce v české odborné literatuře; Vymezení vlastních jmen a některá dílčí hlediska na vlastní jména v zahraniční literatuře. (rpk., nedatováno, 14 s.)
- RAMEŠ, Václav: Po kom se jmenujeme? Encyklopedie křestních jmen. 2. vyd. Praha: Libri 2005. 500 s.
- RAVIK, Slavomír: O světcích a patronech. Praha: Levné knihy KMa 2006. 670 s.
- REMEŠOVÁ, V.: Ikonografie a atributy svatých. 2. vyd. Praha: ZVON 1991. 85 s.
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- Statistický lexikon obcí České republiky. Praha: Ottovo nakladatelství 2005. 1358 s.
- ŠRÁMEK, Rudolf: Propria ve frazémeh a příslovích. In: Parémie národů slovan-ských. Sborník příspěvků z mezinárodní konference konané v Ostravě ve dnech 20.–21. 11. 2002 u příležitosti 150. výročí úmrtí F. L. Čelakovského a vydání jeho „Mudrosloví“. Ostrava: Filozofická fakulta Ostravské univerzity 2003, s. 61–66.
- TROST, Pavel: Zur Theorie des Eigennamens. In: Omagiu lui Iorgu Iordan, cu prilejul implinirii a 70 de ani. Red. A. Rosetti. Bucuresti: Editura Academii republicii populare Romine 1958, s. 867–869.
- VÁŠKŮ, Zdeněk: Velký pranostikon. Praha: Academia 2002. 370 s.

Zpracování hypokoristik v lexikální databázi Pralex¹

Radek Volejník

Ústav pro jazyk český AV ČR, v. v. i.

The Treatment of Hypocorisms in the Pralex lexical database

The article deals with the treatment of hypocorisms in the Pralex lexical database. After the introduction, which outlines the aims and starting points of this treatment (record sources, the software used etc.), the types of hypocorisms, their definitions and expressivity are provided. These are followed by the sources for the index of the hypocorisms (*Dictionary of the Standard Czech Language* etc.). After the issues of homonymy and variability are mentioned, it is proceeded to the actual treatment of hypocorisms in the Pralex LDB; the form for the data entry (the items filled in on the Entry and Meaning cards) is described in more detail and the order of the meanings in proper-name entries and the explanatory definitions of the hypocorisms are elucidated. The core of the article is the description of the work with exemplifications (sentential examples) and the problems of searching for them in the Bonito program. In the end, also the records or the lack of records of hypocorisms in the *Frequency Dictionary of Czech* and in the SYN corpus are discussed.

1. Úvod

V rámci výzkumného záměru *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století (2005–2011)*, na jehož realizaci se podílejí pracovníci oddělení současné lexikologie a lexikografie Ústavu pro jazyk český AV ČR, v. v. i., byly zpracovány rovněž hypokoristické podoby rodných jmen (neboli hypokoristika).

Východiskem pro práci na této lexikální databázi, která je určena ke shromažďování a dalšímu zpracování lexikálních dat, se stal zpracovatelský software **Praled**.² Cílem naší práce je komplexní příprava podkladů pro budoucí zpracování moderního výkladového slovníku českého jazyka. Jako hlavní materiálovou základnu využíváme Český národní korpus. Exmplifikační část hesel naplnujeme doklady z korpusu SYN, který je složen z původních korpusů SYN2000, SYN2005 a SYN2006PUB³. K získávání dokladů z korpusu používáme korpusový manažer Bonito (<http://ucnk.ff.cuni.cz/bonito>) a on-line vyhledávač Word Sketch Engine⁴. V Bonitu je snazší vyhledávání větných dokladů a manipulace s nimi, Word Sketch Engine (WSE) naproti tomu lépe využívá statistické údaje a napomáhá tím k získání tzv. upravených dokladů (viz níže).

Zpracování hypokoristik probíhalo v roce 2009 a 2010; při zpracování se přihlíželo ke zpracování hypokoristik v dosavadních výkladových slovnících, zejména ve *Slovníku spisovného jazyka českého* (dále SSJČ).

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století (AV0Z90610521)*, jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Ten vznikl podle našeho zadání ve spolupráci s Centrem zpracování přirozeného jazyka z Fakulty informatiky Masarykovy univerzity v Brně a jeho vývoj průběžně pokračuje dál na základě našich dalších požadavků.

3 Od roku 2011 byl původní korpus SYN přejmenován na orig_syn.

4 Ukázka např.: http://www.fi.muni.cz/~thomas/corpora/CQL/CQL_pic2.htm.

2. Přehled typů hypokoristik

Pro potřeby lexikální databáze (dále LDB) Pralex přebíráme rozdělení M. Knappové (1989), která preferuje především slovo-tvorné hledisko. Valná většina hypokoristik totiž vzniká derivací; odvozují se jednak z oficiálních podob rodných jmen, jednak z jiných hypokoristik. Některé typy odvozování tak lze vyložit dvojím způsobem. Např. podobu *Jiříček* je možno chápat jako odvozenou z hypokoristika *Jiřík* pomocí přípony *-ek*, nebo jako derivát vzniklý ze základní podoby *Jiří* připojením sufixu *-íček*. U hypokoristik v LDB Pralex používáme pokud možno první postup; reflektován je také ve výkladových definicích. (Při odvozování také samozřejmě dochází k samohláskovým a souhláskovým alternacím.)

Jako odvozovací základ pro tvoření hypokoristik slouží **začátek, střed i konec jména, k němuž se připojuje odvozovací sufix**, který může být i nulový; hypokoristikon pak vzniká:

a) **odtržením zakončení (koncového morfému) fundujícího jména, u složených jmen jejich druhé části, a připojením sufixu** (též sufixu nulového): *František – Franta I⁵, Přemysl – Přemík, Tomáš – Tom, Jiřina – Jiřinka, Kristýna – Kristla, Marie – Maru*;

b) **vynecháním první, popř. i druhé slabiky, u složených jmen první části fundujícího slova, a připojením sufixu** (též sufixu nulového): *Ignác – Náca, Jakub – Kuba, Jaroslav – Slávek, Leopolda – Poldýna, Matylda – Tylida, Tereza – Réza*;

c) **ze střední části jména, zpravidla odtržením první slabiky a zároveň zakončení (koncového morfému) fundujícího slova, k střední části je připojován sufix**: *Alois – Lojza I, Antonín – Tonda I, Cecílie – Cíla, Terezie – Réza*;

d) **vynecháním větší či menší středové části fundujícího jména a event. připojením sufixu**: *Gabriel – Gaba I, Václav – Venda I, Bedřiška – Beda II, Františka – Fána II*;

e) **z jiného základu, obvykle převzatého z cizího jazyka** (zdomácnělé podoby): *Jan – Honza, Josef – Pepa*. K tomu lze doplnit případy, kdy se z cizího jazyka spolu s rodným jménem přebírá již **hotové hypokoristikon** (popř. jen graficky upravené), např. *Alexej – Aljoša, Alfréd – Fred* (a z něj pak opět derivací případně vznikají další).

Dále se hypokoristika tvoří **na základě různých zvukových variací**, např. opakováním první slabiky jména: *Ladislav – Lála* (tento způsob je méně častý). Podle Knappové se ve funkci hypokoristika užívá rovněž základní podoba jména (např. podoba formálně podobného jména nebo cizí ekvivalent základní podoby) a také domácká podoba jiného jména. To jsou ovšem mezní případy, které mají víceméně okazionalní povahu a z lexikografického hlediska nejsou podstatné.

A samozřejmě je nutné také počítat s tím, že jedno hypokoristikon se může užívat pro více rodných jmen, ať už v rámci homonym, či mimo ně (např. *Jarek – domácká*

5 Římská číslice odpovídá značení homonym v LDB Pralex.

podoba k Jaroslav, k Jaromír, k Jarmil ap., **Milka** – domácí podoba k Emílie, k Milada, k Miloslava ap.). (Podrobnější rozbor této problematiky zde pro její rozsáhlost vynecháváme.)

3. Vymezení hypokoristik a jejich expresivita

Hypokoristika bývají chápána a vymežována různě. Stručný přehled v tomto směru podává Pastyřík (1994), který sám za hypokoristikon považuje „expresivní obměnu neutrálního jména (propria i apelativa) v neoficiálním prostředí“. Pro naše účely bylo však přínosnější chápání hypokoristik pouze jako součásti propriálního systému. A z podstaty svého významu jsou hypokoristika pouze jednoslovná propria.

Zásadním rozhodnutím při zpracování hypokoristik bylo především to, že **všechny formálně rozdílné grafické podoby budou zpracovány jako samostatné položky databáze** („hesla“). Také expresivní složku hypokoristického významu bylo nutné trochu přehodnotit.

Původně jsme vycházeli z označení expresivity u příslušných lemmat v SSJČ. Už na počátku zpracování hypokoristik se však ukázalo, že rozhodování o jejich expresivitě bude často obtížné. Expresivita totiž velmi záleží na kontextu, v němž se dané lemma nachází, a rovněž na věku osob, které jsou takto osloveny (dětí i dospělých), přičemž v komunikaci s dětmi je expresivita hypokoristik více výrazná. Hypokoristika, která by formálně měla být expresivní, byla často při oslovování dospělých doložena jak v expresivním, tak v neutrálním kontextu. Neutrální kontext postačil k tomu, aby chom expresivitu nezapisovali do definice daného lemmatu (jinak by byla expresivní téměř všechna deminutivní hypokoristika). V SSJČ je mnoho hypokoristik chápáno jako expresivních, v LDB Pralex jich však na základě uvedeného postupu bylo takto označeno v definicích pouze 16 (tím spíše, že se to často z kontextu nedalo vůbec určit). Zkratku **expr.** mají tedy v definici pouze lemmata, pro která se v korpusu SYN našly jen takové doklady, u nichž bylo vždy zřejmé, že se jedná o děti.

4. Soubor hypokoristik v LDB Pralex

Soubor hypokoristik pro LDB Pralex byl sestaven z několika různých zdrojů. Největší část, 724 lemmat, byla převzata ze SSJČ, kde byla zpracována jako samostatná hesla (např. *Jenda, Pepíček, Kačka, Nataška*). Dále bylo asi 20 lemmat doplněno z Frekvenčního slovníku proprií, který tvoří součást *Frekvenčního slovníku češtiny* (dále FSČ), přičemž dalších 64 lemmat z FSČ je zároveň uvedeno i v SSJČ (např. *Kubíček, Pavlík, Hanka, Lída*); 43 lemmat bylo přidáno jen na základě výskytu v korpusu SYN (tj. v žádném výkladovém slovníku nebyla registrována, např. *Jířa I a II, Pétá I a II, Gábina, Jani*) a 4 lemmata zřejmě pocházela ze zvláštního seznamu (*Daníček, Filípek, Bohuše, Halka*). Vcelku heslář LDB Pralex obsahuje zejména hypokoristika původu domácího, resp. slovanského, v mnohem menší míře pak hypokoristika přejatá (uvedená hlavně ve FSČ, např. *Bob, Charlie, Eddie, Harry, Heinz, Jim*). Východiskem pro

výběr a zpracování hypokoristik v LDB Pralex se stal SSJČ především z toho důvodu, že pouze v tomto slovníku je hodně hypokoristik zachyceno a také zpracováno jako samostatná hesla. Naproti tomu seznam rodných jmen ve *Slovníku spisovné češtiny pro školu a veřejnost* (SSČ) uvádí jen základní, oficiální podoby jmen, nikoli tedy tvary domácké a zdrobnělé.

5. Homonymie a variantnost u hypokoristik

Pokud se vyskytla stejná grafická podoba uplatněná v různých významech pro různý rod, byla taková lemmata zpracována jako dvě samostatná hesla a označena římskými číslicemi jako **homonyma** (např. *Saša I* a *Saša II*). Dohromady bylo zpracováno 177 hesel takto označených. Mezi nimi bylo 81 párů (maskulinum a femininum – heslo s maskulinním lemmatem je bez ohledu na etymologii vždy označeno římskou jedničkou). Šlo např. o tyto páry: *Fána I, Fána II, Jožka I, Jožka II, Pétá I, Pétá II, Vlád'a I, Vlád'a II*. Kromě nich byla ještě jako homonyma označena hypokoristika *Honda, Frantík, Jura, Karlík, Kuba, Lenora, Líba, Mařena, Mařenka, Mates, Ruda, Sába, Slávinka, Slávka, Svatka* (tato hesla jsou ovšem ve vztahu s homonymy, jež nepatří k hypokoristikům).

Celkem 186 hypokoristik bylo možno chápat jako **variantu jiné hypokoristické podoby**. Uvedené číslo však bylo získáno ručním součtem hesel, neboť pomocí funkce vyhledávání jich bylo nalezeno jen 159 (oba údaje v sobě zahrnují i významy rodného jména).

Každá varianta hypokoristika byla hypertextově propojena s ostatními variantami téhož propria – a to bez ohledu na to, zda byly doloženy v korpusu SYN, nebo nikoliv. Jako varianty byly u hypokoristik propojeny: **a) varianty, které uvádí jako varianty už SSJČ**: např. *Béda – Béda – Beda*; **b) varianta ze SSJČ a varianta z korpusu SYN** (i pro variantu z korpusu SYN bylo pochopitelně založeno samostatné heslo): např. *Dášenka – Dášeňka, Eda – Éda*. Jako varianty nejsou chápány ani propojovány hypokoristika a jejich cizojazyčné ekvivalenty (kromě těch, jež uvádí SSJČ).

6. Zpracování hypokoristik v LDB Pralex

Samotné zpracování hypokoristik v LDB Pralex probíhalo od července 2009 do dubna 2010. Zakládání těchto hesel do hesláře však začalo dříve; mnohá z nich (hlavně ta, jež mají jako první význam rodné jméno a jako druhý hypokoristikon, např. *Iva, Květa, Táňa*) byla založena již v rámci zpracování jmen osob. Celkem jsem zpracoval 853 hesel s 1337 významy, z toho je přes 180 variant (stav k 15. 4. 2010).

V rámci mně přidělených hypokoristik jsem také v relevantních případech zpracovával další významy příslušných lexémů. Nejčastěji se jednalo o **rodná jména** (46 hesel – z toho 13 mužských a 33 ženských rodných jmen), **mužská příjmení** (164 hesel, např. *Hanka, Kuba, Michálek*), **jména zvířat** (cca 181 hesel, např. *Ferda, Honzík, Líza, Pepina*) a **jména pohádkových bytostí** (44 hesel, např. *Honza, Mája, Ma-*

řenka, Pepek). Podrobnostem o těchto a dalších významech se zde věnovat nebudu (přibližné číslo u počtu významů jmen zvířat je tu proto, že na začátku zpracování byly definice významů konkrétnější, např. *jméno kočky*, a teprve později se přešlo na obecnější formulaci *jméno zvířete* – pokud tedy bylo z kontextu zřejmé, o jaké zvíře se jedná)⁶.

7. Formulář pro zápis dat – vyplňované položky na kartách Heslo a Význam

Formulář pro zápis dat do LDB Pralex se primárně skládá z karty **Heslo** a z karty **Význam** (karta Heslo je vždy pouze jedna, karet Význam může být více – v závislosti na tom, kolik má daný lexém významů – a označují se V1, V2, V3 atd.). Tyto karty obsahují množství položek, které byly rovnou navrženy i pro slovníkové zpracování všech hesel⁷.

Protože současné zpracování hesel v LDB Pralex v rámci probíhajícího výzkumného záměru je zaměřeno především na vytěžení jazykových dat z korpusu SYN, tj. především na získání adekvátního počtu příkladů pro dokladovou část hesla, nebylo zatím třeba vyplňovat všechny navržené položky na kartách Heslo a Význam. Zatím se vyplňovaly následující položky:

a) **položky na kartě Heslo: lemma; homonymie** (v relevantních případech); **Zdroj pro heslář; Typ hesla:** *jednosl.*; **Slovní druh/typ:** *vlastní jméno*; **VAR.:** zápis variant na seznam a jejich vzájemné hypertextové propojení (v relevantních případech);
 b) **položky na kartě Význam: Gram. – Poznámka ke gramatice** (v případě neskloňnosti: *neskl. + rod*); **Výklad významu; Původ definice:** *vlastní návrh*; **Nepodařilo se doložit v syn** (zaškrtnout v relevantních případech); **Poznámka k výkladu:** sem vkládáme kopii příslušné heslové stati ze SSJČ (pokud v něm je); **Poznámka k exemplifikaci celkově** (v případě nízkého výskytu daného lemmatu v korpusu SYN: např.: *syn 1x*); **Exemplifikace** (blok Větné doklady, popř. blok V názvech).

8. Řazení významů v propriálních heslech a jejich výkladové definice

Z hlediska **pořadí významů** při zpracování proprií v LDB Pralex platí: pokud bylo v korpusu SYN k téže grafické podobě lexému nalezeno více významů, pak jednotlivé významy v rámci polysémie řadíme v tomto pořadí: rodné jméno; biblické jméno vč. proroků; antické jméno/mytologická postava; jméno světce/svěťice; domácká podoba; příjmení; jméno pohádkové bytosti; případně i běžné jméno zvířete; ostatní případy (uvádím zjednodušené formulace výkladových definic).

Pro výkladovou definici hypokoristika byl vybrán formalizovaný zápis *domácká podoba k...* (např. heslo *Anina – domácká podoba k Anna*). Pokud byla domácká podoba zároveň zdrobnělinou rodného jména, pak se zápis obměnil takto: *domácká*

⁶ Je ovšem otázkou, do jaké míry je vhodné zachycovat i jména zvířat, protože z hlediska jejich výskytu a rozšíření v úzu se nemůžeme spolehnout pouze na jejich frekvenci v jednom korpusu.

⁷ Blíže viz zde stat' J. Světlé *Struktura a zpracování hesel v lexikální databázi Pralex*.

podoba a zdrob. k... (např. heslo *Adélka – domácí podoba a zdrob. k Adéla*), popř. *domácká podoba a expr. zdrob. k...* (v případě průkazné expresivity nalezených dokladů, např. heslo *Magdalénka – domácí podoba a expr. zdrob. k Magdalena, k Magdaléna*). Při souběhu hypokoristik a deminutiva (pokud se deminutivum vztahovalo k jiné výchozí podobě – většinou k jiné zdrobnělině) uvádíme nejprve domácí podobu a pak teprve po středníku *zdrob. k...*, resp. *expr. zdrob. k...* (např. u hesla *Ančička – domácí podoba k Anna; expr. zdrob. k Anča, k Ančí*).

9. Exemplifikace u hypokoristik

Jako u všech hesel v LDB Pralex, tak i u hypokoristik bylo záhodno, aby vybrané příklady v příkladové části hesla (Exemplifikaci) co nejlépe dokládaly význam lemmatu (samotné změny významu se zde však až na expresivitu příliš nepředpokládaly). Proto bylo třeba určit adekvátní rozsah kontextu exemplifikačních dokladů. Ukázalo se, že ve většině případů postačí kontext jedné věty, resp. souvětí. Někdy však bylo nutné kontext rozšířit (především tehdy, když význam hypokoristika nebyl z věty zřejmý, nebo byla-li věta či souvětí v korpusu neúplná).

U hypokoristik se bloky Exemplifikace stejně jako u rodných jmen (a jmen vůbec) dělí na **Větné doklady (E1)**, **Upravené doklady (E2)** a fakultativní blok **V názvech (E3)**, přičemž důraz byl kladen na první skupinu. Větné doklady však nemělo smysl členit tak podrobně jako u rodných jmen. Např. kontexty týkající se svátku v kalendáři (jež jsou relevantní u rodných jmen) nebyly u hypokoristik téměř nikdy nalezeny, protože svátek se nevztahuje k hypokoristiku, ale k oficiální podobě rodného jména.

Pokud tedy nebylo dokladů mnoho, všechny příklady byly ponechány v jednom bloku **větných dokladů** (jen jako **E1**). Při větším množství a při důrazu na specifické doklady jsme blok E1 dále členili na **E1a** a **E1b** (optimální počet byl 3–5 dokladů v každém bloku). V rámci bloku E1a byl kladen důraz hlavně na **a) doklady se slovesem jmenovat se, říkat někomu (nějak) apod.; b) doklady v oslovení nebo jiném než 1. p.; c) doklady s příjmením**. Do bloku E1b se v případě doložení uváděly pouze ty doklady, které se vztahovaly **k nějaké známé osobnosti nebo postavě ve hře, filmu apod.**

Při souběhu téže grafické podoby rodného jména a hypokoristika bylo často obtížné rozdělit doklady na jednotlivé významy. Proto byly v takových případech všechny doklady přiřazeny **k prvnímu z významů** (tj. k rodnému jménu) a tamtéž do Celkové poznámky k Exemplifikaci se zapsalo: *společné doklady pro V1 a V2* (takto zpracováno např. u hesla *Dora*).

Upravené doklady nebylo víceméně nutné uvádět; hypokoristika se totiž svými vazebnými a syntaktickými vlastnostmi neliší od rodných jmen, s nimiž mají stejný denotát (pomineme-li tedy významové rozdíly a expresivitu).

Do bloku **V názvech** se vkládaly jednak doklady, u nichž byla hypokoristika částí víceslovné lexikální jednotky (např. *Chlum Svaté Máří, Pražské Jezulátko*), jednak názvy lidských výtvorů, jako např. tituly uměleckých děl (např. *Anička Skřítek a Slaměný Hubert*, drama *Maryša*, pohádka *Květuška a její zahrádka*), výrobků (kojenecká voda *Lucka*) apod.

10. Vyhledávání v Bonitu

Korpusový manažer Bonito z Ústavu Českého národního korpusu slouží k vyhledávání dokladů v korpusech (těch je přes Bonito v současné době on-line k dispozici 14, my z nich používáme jen korpus SYN). Před samotným vyhledáním a vložením dokladů do položky **Exemplifikace** je třeba v menu Bonita nastavit určité parametry: jako výchozí vybrat **korpus SYN** (lze ho takto nastavit i trvale); v nabídce Zobrazení jako relevantní informace pro nás zaškrtaváme:

1) v podnabídce **Zdroje: opus.rokvyd, opus.txttype, opus.id**; tyto značky jsou specifické pro každý dokument v korpusu a po svém označení v nabídce se stávají pevnou součástí každé vybrané exemplifikace uvedené ve Větných dokladech;

2) v podnabídce **Kontext**: vybíráme rozsah kontextu, v němž bude zasazen doklad hledaného lemmatu: možností je více, ale pro naše účely většinou postačuje tato: **1 věta vlevo a 1 věta vpravo** (věta je označena jako s, neboli sentence);

3) v podnabídce **Rozsah**: počet a způsob výběru řádků obsahujících dané lemma: pro účely hypokoristik to bylo **300 náhodných**.

Samotné vyhledávání lze v Bonitu provést jednoduchým zadáním lemmatu do vyhledávacího řádku – tímto způsobem se dají nalézt všechny výskyty konkrétního tvaru daného lemmatu. Např. při zadání propria **Mirek** vyhledá Bonito věty obsahující lemma Mirek v nominativním tvaru. Totéž nastane při použití vyhledávací formule [**lc="Mirek"**] – Bonito opět vyhledá pouze konkrétní tvar. Naproti tomu při zadání formule [**lemma="Mirek"**] Bonito nalezne všechny pády a tvary tohoto hypokoristika, i když ne úplně důsledně. Při nízkém nebo nulovém výskytu dokladů (za použití druhé formule) je proto nutná kontrola zadáváním jednotlivých pádů daného hypokoristika (pochopitelně postačí jen tvary singuláru). Samozřejmě je při výběru dokladů nutno dávat pozor na občasnou tvarovou a pádovou homonymii.

Protože u hypokoristik nebylo tak důležité vyplňovat položku Upravené doklady, nepoužíval jsem program Word Sketch Engine.

11. Doloženost/nedoloženost hypokoristik ve FSC (=SYN 2000) a v korpusu SYN

Celkem 584 lemmat původem ze SSJČ bylo doloženo v korpusu SYN (29 hesel z nich má jako první uveden význam rodného jména, např. *Arne, Přemek, Táňa, Vendula*), naproti tomu 142 lemmat ze SSJČ nebylo v korpusu SYN doloženo vůbec (např. *Čeněček, Lubomírek, Rézička, Zuzanečka*). Mezi lemmaty převzatými z FSC bylo 64 hypokoristik, jež jsou i v SSJČ (nezapočítáno do uvedených 584), např. *Petřík, Vojta, Hanka, Míša II*) a 12 hypokoristik cizího původu (např. *Bobby, Eddie, Jeff, Jimmy*). 15 lemmat z FSC má jako první uveden význam rodného jména – jde o 11 podob ženských a 4 mužské (např. *Magda, Naďa, Ivo, Mirek*). Malá část lemmat byla doplněna do hesláře na základě jejich výskytu v korpusu (např. *Gábina, Jani, Janka, Věrunka*) – často šlo o homonymii v rámci mužského a ženského rodu (např. *Fanča I, Ola I, Honza II, Venda II*) nebo o nalezené varianty (např. *Éda, Pepik, Josefka, Magdalénka*).

Z hlediska nulové doloženosti vypadá situace následovně: celkem 142 hesel se v korpusu SYN nepodařilo doložit vůbec (šlo o jednovýznamové lexémy, např. *Bobíšek*,

Veník, Fančička, Tončinka), a dalších 163 hesel se nepodařilo doložit v hypokoristickém významu daného lexému (např. *Antoníček, Tománek, Flórinka, Žanynka*).

Závěr

Hypokoristická hesla jsou v LDB Pralex zpracována vesměs v souladu se zadáním výzkumného záměru a podle stanovených zásad pro zpracování proprií a hypokoristik. Jádrem dosavadní práce tkví především ve velkém množství lexikograficky relevantních dokladů uvedených v exemplifikační části hesla, a také v systémových návrzích definic. Hypokoristika však nejsou v LDB Pralex zachycena kompletně všechna, na mnohé známé tvary se z časových a jiných důvodů zatím nedostalo (tím spíše, že některá ani nejsou zachycena v SSJČ, např. *Jája, Kája, Víta*). V budoucnu bude třeba u všech hesel doplnit další adekvátní položky na kartách Heslo a Význam. Doklady z korpusu SYN rovněž ukázaly, že v současné době je pro méně hypokoristik relevantní expresivita, resp. že expresivita není oproti SSJČ tak stabilní součástí hypokoristického významu.

Literatura

- BAUEROVÁ, Anna: Tisíc jmen v kalendáři. Praha: Šulc a spol. 1992. 231 s.
- BAYEROVÁ, Naděžda: Hypokoristika v oblasti slezských nářečí (jejich slovtvorba a jejich podoby užívané při oslovování v komunikaci rodiny). In: Sborník prací Filozofické fakulty Ostravské univerzity 158/1996. Jazykověda. Linguistica, č. 2. Ed. J. Damborský. Ostrava: Filozofická fakulta Ostravské univerzity 1996, s. 109–138.
- DAVÍDEK, Václav – DOSKOČIL, Karel – SVOBODA, Jan: Česká jména osobní a rodová. Praha: Rodopisná společnost v Praze II 1941. 109 s.
- Frekvenční slovník češtiny. Red. F. Čermák – M. Křen. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSC)
- HÁJKOVÁ, Kristýna: Tvoření hypokoristik. Bakalářská práce. Brno: FF Masarykovy univerzity 2006.
- CHEJNOVÁ, Jindra: Mužská a ženská hypokoristika tvořená sufixy opačného rodu. In: Opera studiosorum Universitatis Reginaegradecensis, 2. Ed. Kňákalová. Hradec Králové: KČJL VŠP 1997, s. 12–16.
- KLÍMOVÁ, Květoslava – KNESELOVÁ, Helena: Propria v systému mluvnickém a slovtvorném. Sborník příspěvků z mezinárodní konference „Onomastika a škola“ konané v Brně ve dnech 10.–11. 2. 1998. Brno: Masarykova univerzita 1999, s. 11–15.
- KNAPPOVÁ, Miloslava: Jak se bude vaše dítě jmenovat? 3. vyd. Praha: Academia 1999. 358 s.
- KNAPPOVÁ, Miloslava: K proměnám systému hypokoristik. In: Onomastické práce, sv. 4. Sborník rozprav k sedmdesátým narozeninám univ. prof. PhDr. Ivana Luttera, CSc. Eds. L. Olivová-Nezbedová – R. Šrámek – M. Harvalík, Praha: Ústav pro jazyk český AV ČR 2000, s. 209–215.

- KNAPPOVÁ, Miloslava: Lexikální a slovotvorná charakteristika českých rodných jmen. In: VIII. slovenská onomastická konferencia. Banská Bystrica – Dedinky 2.–6. júna 1980. Zborník materiálov. Red. M. Majtán. Bratislava – Banská Bystrica – Prešov: Jazykovedný ústav Ľudovíta Štúra SAV v Bratislave – Pedagogická fakulta v Banskej Bystrici – Pedagogická fakulta Univerzity Pavla Jozefa Šafárika v Prešove 1983, s. 307–313.
- KNAPPOVÁ, Miloslava: Rodné jméno v jazyce a společnosti. Praha: Academia 1989. 204 s.
- KOPEČNÝ, František: Průvodce našimi jmény. 2., přepracované a rozšířené vyd. Praha: Academia 1991. 259 s.
- Manuál lexikografie. Eds. F. Čermák – R. Blatná. Jinočany: H&H 1995. 283 s.
- OBERPFCALCER, František: Domácí tvary křestních jmen. In: Naše řeč, 1931, roč. 15, s. 171–176, 199–204.
- Onymické systémy v regionech. Sborník příspěvků z V. semináře „Onomastika a škola“, konaného 12.–14. ledna 1993 v Hradci Králové. Eds. R. Šráme – J. Bartůňková – Koblížek, V. Hradec Králové: Gaudeamus 1995. 259 s.
- OSOLSOBĚ, Klára: Hypokoristika v korpusu soukromé korespondence KKS. In: Sborník prací filozofické fakulty brněnské univerzity. Linguistica Brunensia, A 53. Brno: Masarykova univerzita 2005, 125–136.
- PASTYŘÍK, Svatopluk: Aktuálně nad hypokoristiky a nad souborem královéhradeckých diplomových prací hypokoristiky se zabývajících. In: Češtinář, 2001–2002, roč. 12, č. 3, s. 65–70.
- PASTYŘÍK, Svatopluk: Co je to „hypokoristikon“? In: Jazyková a mimojazyková stránka vlastních mien. 11. slovenská onomastická konferencia (Nitra 19.–20. mája 1994). Zborník referátov. Ed. E. Krošláková. Bratislava – Nitra: Vysoká škola pedagogická Nitra, Fakulta humanitných vied, s. 104–107.
- PASTYŘÍK, Svatopluk: Hypokoristické podoby vlastních jmen a jejich současné zkoumání. In: Čeština doma a ve světě, 2007, roč. 15, č. 3–4.
- PASTYŘÍK, Svatopluk: Hypokoristika a kategorie rodu. In: Acta Onomastica, 1999, roč. 39, s. 55–58.
- PASTYŘÍK, Svatopluk: Hypokoristika a současný mluvený jazyk. (Otázky, otazníky a problémy.) In: K diferenciaci současného mluveného jazyka. Sborník prací z mezinárodní vědecké konference (Ostrava, září 1994). Ed. D. Davidová. Ostrava: Filozofická fakulta Ostravské univerzity 1995, s. 93–96.
- PASTYŘÍK, Svatopluk: Hypokoristika jako osobitý druh antroponym. In: Acta Onomastica, 2006, roč. 47, s. 392–396.
- PASTYŘÍK, Svatopluk: K projektu výzkumu hypokoristik našich rodných jmen. In: Onomastický zpravodaj, 1993–1994, s. 111–113.
- PASTYŘÍK, Svatopluk: O expresivitě hypokoristik a jejím zaznamenávání. In: Spišovnosť a nespisovnosť dnes. Sborník příspěvků z mezinárodní konference Spišovnosť a nespisovnosť v současné jazykové a literární komunikaci (Šlapanice u Brna 17.–19. ledna 1995). Ed. R. Šrámek. Brno: Masarykova univerzita 1996, s. 154–156.
- PASTYŘÍK, Svatopluk: Před lexikografickým zpracováním hypokoristických podob rodných jmen. In: Acta onomastica, 2001, roč. 41–42, s. 138–142.

- PASTYŘÍK, Svatopluk: Studie o současných hypokoristických podobách rodných jmen v češtině. Hradec Králové: Gaudeamus 2003. 140 s.
- PLESKALOVÁ, Jana: Hypokoristika v češtině. In: Přednášky a besedy z XL. běhu LŠSS. Ed. E. Rusínová. Brno: Masarykova univerzita 2007, s. 150–156.
- RAMEŠ, Václav: Po kom se jmenujeme? Encyklopedie křestních jmen. 2. vyd. Praha: Libri 2005. 500 s.
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- ŠRÁMEK, Rudolf: Úvod do obecné onomastiky. Brno: Masarykova univerzita 1999. 191 s.
- VOLEJNÍK, Radek: Hypokoristické formy rodných jmen a mládeže v Přelouči. In: Acta Onomastica, 2004, roč. 45, s. 84–104.
- Zásady zpracování slovníku. In: Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač. 1. vyd., Praha: Academia 1978, s. 779–799.

Zpracování přirovnání v LDB Pralex¹

Eva Kolovecká

Ústav pro jazyk český AV ČR, v. v. i.

Processing of Similes in the Pralex Lexical Database

V tomto článku se zabýváme zpracováním přirovnání v LDB Pralex. Uvádíme zde postup prací od vložení přirovnání do databáze a práce s referenčními zdroji, přes veškeré úpravy lemmat až ke konečnému zpracování. Dále se zabýváme jak určováním reprezentativního tvaru přirovnání, tak zpracováním fakultativních komponentů, substitutů, variant a transformací. Zaměřujeme se také na práci s výkladovými definicemi. V závěru podrobně rozebíráme zpracování příkladové části hesla.

1. Úvod

Podle *Encyklopedického slovníku češtiny* **přirovnání je** zpravidla frazém kolo-kační (řídčeji i propoziční) založený vždy na explicitním vyjádření podobnosti mezi členem různě doplňovaným kontextem a zobecněným a obvykle známým modelem, prototypem určité vlastnosti, činnosti, entity apod.

2. Postup prací

V lexikální databázi (dále LDB) Pralex pracujeme převážně se souborem frazémů ze *Slovníku spisovného jazyka českého* (dále SSJČ), v případě potřeby i s přirovnáními z dalších slovníků, *Slovníku spisovné češtiny pro školu a veřejnost* (dále SSČ) a *Slovníku české frazeologie a idiomatiky* (dále SČFI).

Nejprve bylo nutné uvedené frazémy založit do hesláře. Úkolem programátora bylo vložit do LDB Pralex všechny frazémy, které jsou v SSJČ uvedeny po značce pro frazémy (tzv. hřebík, v el. verzi jsou označeny [x]). Tento počáteční soubor pak bylo třeba upravit tak, abychom s ním mohli dále pracovat.

Nejprve bylo nutné oddělit od sebe frazémy, které byly v SSJČ původně uvedeny v jednom hesle (např. frazémy *běžel, jako by ho honili, jako by ho někdo honil, jako když ho čerti honí* byly rozděleny na tři samostatná hesla – *běžel, jako by ho honili; běžel, jako by ho někdo honil; běžel, jako když ho čerti honí*, která pak byla dále upravena).

Dále jsme odstraňovali chyby způsobené málo explicitním zadáním úkolu programátorovi (aby jako samostatné položky databáze založil vše, co je uvedeno za hřebíkem), protože po frazeologickém oddělení se v SSJČ uváděly i termíny nebo příklady na významové odstíny – i ty se nám tak dostaly do hesláře. Hesla, která k frazémům nepatřila, jsme tedy museli zase odstranit.

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je Doc. RNDr. Karel Oliva, Dr.

Rovněž bylo potřeba v mnohých případech rozhodnout, jaká podoba či která část frazému ze SSJČ je pro nás relevantní (pokud se výraz měnil na reprezentativní tvar), takže jsme mnohá lemmata pro potřeby LDB upravovali či doplňovali (viz dále). Poté jsme museli určit, zda jde vůbec o frazém. Další fází bylo ověřit, zda u frazému došlo k transformacím (slovesný x jmenný) registrovaným ve slovnících.

Následně jsme uvažovali, kterou část frazeologie začneme komplexně zpracovávat. Hledali jsme ucelený soubor, na kterém bychom navrhli zásady pro zpracování frazémů odzkoušeli. Vzhledem k tomu, že zpracování frazémů probíhá paralelně se zpracováním substantiv, jako nevhodnější jsme vybrali soubor přirovnání se substantivním komponentem. Zpracování tohoto typu frazémů nám tedy poskytuje i částečnou zpětnou vazbu zpracované příkladové části substantiv.

Zpracování přirovnání v LDB Pralex začalo v červnu 2009. Jednalo se o frazémy se spojkou *jak/jako*, později byla doplněna lemmata obsahující i další komparátory uvádějící přirovnání (*jako by, jako že, jako když, až, že by, co, div, jen, než* apod.) – ty však zpracováváme doplňkově, nikoli soustavně. Cílem bylo na uvedeném souboru odzkoušet celkové zpracování frazémů, včetně vyhledávání dokladů v korpusu a vyplňování exemplifikace, a celý soubor pak zpracovat jako celek. Tento soubor původně obsahoval asi 700 hesel vložených do databáze na základě SSJČ, postupně byly zakládány některé další jednotky nalezené v SSČ nebo SČFI, případně v korpusu SYN². Nová hesla byla zakládána pouze v případě potřeby tak, aby byly doplněny všechny doložené varianty a soubor byl z tohoto aspektu úplný.

Jak již bylo uvedeno výše, při zpracování přirovnání jsme postupovali podle předem vytvořeného seznamu substantivních komponentů, které se v přirovnáních založených do našeho hesláře podle SSJČ vyskytují. V současné době je dokončeno zpracování přirovnání obsahujících substantivní komponent. Celkem je zpracováno 1853 hesel / 1967 významů, z toho: se spojkou *jako* 1299/1378, dalších 207/218 se spojkou *jak*, 105/121 s komparátorem *jako by*, 2 – *jako že*, 78 – *jako když*, 11 hesel/14 významů má komparátor *až*, 126/131 – *co*, 16 – *než*, 9 – *jen*. Z celkového počtu je 410 / 449 nedoložených.

3. Práce s referenčními zdroji

SSJČ uvádí přirovnání u jednoslovných hesel většinou po značce pro frazémy na konci heslové stati, někdy ale i před touto značkou nebo bez ní; někdy s výkladem, jindy bez výkladu. V případě, že je přirovnání v SSJČ uvedeno bez jakékoli signalizace, že jde o ustálenou lexikální jednotku, ověřujeme ho rovnou v dalších slovnících, srov. v SSJČ např.:

čumák -u m. (6. mn. -cích) **1.** *čenich*: psí č.; ob. expr. studený jak psí č. **2.** zhrub. část obličej s nosem a ústy; *huba*: dát někomu přes č.; [x] strkat do všeho č. *být zvědavý, plést se do všeho*. (přirovnání nijak neoznačené podtrženo autorkou příspěvku)
Srov. zpracování téhož frazému v SSČ:

2 Od roku 2011 byl původní korpus SYN (který slučuje korpusy SYN2000, SYN2005 a SYN-2006PUB) přejmenován na *orig_syn*.

čumák -u m (6. mn. -cích, -ách) čenich (syno) [x] studený jako psí čumák, přen.
 expr. (o člověku) *chladný, uzavřený.*

V SSČ je v tomto konkrétním případě přirovnání už uvedeno za značkou pro frazémy, a to v prvním významu bez výkladu (týká se zejména jídla, nosu apod.), a teprve v druhém významu (o člověku) je uveden výklad.

Porovnání SSJČ a SSČ provádíme vždy a pokud je výraz v SSČ přehodnocen, opřeme se o něj. Často se ukazuje, že dále potřebujeme konzultovat i SČFI, který věnuje přirovnáním celý první svazek³. V případě potřeby tedy porovnáváme výchozí zpracování v SSJČ se zpracováním v obou těchto dalších slovnících, i když v SČFI může být uvedena jiná podoba frazému, srov.:

být^o (studený) j. psí čumák/nos Nom někdo studený j. p. č. 1. (Člověk...), 2. (Část těla, ruka, nos člověka) (...) Zde se na rozdíl od SSJČ a SSČ jedná o frazém verbální s vypustitelnou syntagmatickou variantou (*studený*) a transformací (nominalizace).

Na základě analýzy výše uvedených referenčních zdrojů rozhodujeme o podobě lemmatu a jeho variant či transformací a podle možností a doloženosti v korpusu se o ně opíráme i při výkladu nebo při zpracování exemplifikace (podrobněji viz dále).

4. Proces lemmatizace

Jak již bylo řečeno, u každého přirovnání nejprve rozhodujeme o podobě lemmatu. Naše výkladové slovníky řeší tento problém různě. Bez problémů není ani hledání opory v *Slovníku české frazeologie a idiomatiky. Přirovnání* (SČFIP) (případně v *Slovníku české frazeologie a idiomatiky. Výrazy slovesné* (SČFIS) a v *Slovníku české frazeologie a idiomatiky. Výrazy neslovesné* (SČFIN) při uvádění přirovnání jako transformace. V SČFIP a SČFIS se navíc přístup k tomuto typu frazémů liší.⁴ V našich zásadách pro zpracování frazémů jsme tedy přijali vlastní řešení, které se týká pře-

3 *Slovník české frazeologie a idiomatiky. Přirovnání* (SČFIP) zachycuje přirovnání jako samostatná hesla s podrobnou strukturou heslové stati. Lemma je uváděno v reprezentativní podobě s variantami po lomítku. Morfologicky obměnitelný komponent se zde označuje symbolem °, vypustitelné varianty syntagmatické jsou uváděny v kulatých závorkách. Poté následuje stylová a gramatická charakteristika, transformace, kontext a valence, výklad významu a další údaje o úzu, etymologii aj. Na konci heslové stati jsou uvedena synonyma, opozita a cizojazyčné ekvivalenty. Pro komparátory *jako a jak* používá SČFIP souhrnné zkratku „j.“.

4 V SČFIP se každý heslový frazém uvádí ve svém neutrálním tvaru, pokud ho má (tj. nominativu, infinitivu apod.); pokud ho nemá, zastupuje ho obvykle jeho nejčastější tvar (u sloves je to obvykle 3. sg.), srov. *být/ležet j. lazar a je tam (tepló) j. v lázni.* (SČFIP, s. 13). V SČFIS se v případě, že je 2. sloveso gramaticky obměnitelné, uvádí 1. sloveso v infinitivu a 2. sloveso, případně jiné komponenty (zejm. zájmena) jsou ponechány v typické podobě + v hranatých závorkách. SČFIS, s. 15: „Frazémy obsahující dvě slovesa (druhé bývá v závislé klauzi) však nemohou mít tvar dvou infinitivů (kromě prostých kombinací V-V, např. *dát se podat*). Pro jejich neutrální podobu se volí tudíž jedna z typických a častých podob druhého slovesa, popř. i dalších komponentů (první sloveso je v infinitivu), a jejich gramatická obměnitelnost se v tomto případě signalizuje hranatými závorkami; zároveň se takto naznačuje i to, že podoba těchto hesel je konstruovaná a že je proto třeba ji lišit od čistě propozičních, textových forem frazému.“ – Např.: *dělat°*, co [je] v [jeho] silách = *dělá/děláme/děje...*, co je/bylo/bude... v jeho/našich... silách, *cpát°* se, až [mu] pupek prská = *cpe se/cpou se, až mu/jim pupek prská.*

devším uvádění fakultativních komponentů, substitutů, variant, transformací, výkladu významu a exemplifikace. Pro potřeby LDB Pralex jsme také stanovili následující grafické znaky užívané v lemmatu: kulaté závorky pro fakultativní komponenty, šikmé závorky pro substituty.

Při práci s lemmaty v LDB Pralex převádíme přirovnání vyskytující se ve tvaru 3. sg. na infinitiv všude, kde je to možné (např. *je hluchý jako poleno* bylo převedeno na *být hluchý jako poleno*). Pokud to možné není, volíme nejčastější tvar, což bývá právě 3. sg. (např. *to se prodá jako med; bylo jich jako klasů na poli*). Jde zejména o případy, kdy je v levé části přirovnání pevný subjekt, včetně stálého „to“ (jde o propoziční frazém), nebo kdy pravou část přirovnání tvoří klauze s gramaticky obměnitelným slovesem či zájmenem.

Pokud k úpravám lemmatu vedou doklady z korpusu, projeví se to např. vymezením některého komponentu jako fakultativního nebo jako substitutu. Na základě nalezených dokladů a SČFIP tak někdy dochází ke změně tvaru lemmatu převzatého ze SSJČ.

V případě potřeby komentujeme lemma v poznámce k lemmatu, např. u hesla *být jako růže mezi trnům* je v Pozn. k lemmatu uvedeno: „na základě dalších dokladů případně doplnit do lemmatu fakultativní komponent: *být (někde) jako růže mezi trnům*“.

5. Práce s fakultativními komponenty

Pokud přirovnání obsahuje fakultativní komponent, uvádíme jej v souladu s praxí SSJČ i SČFI v kulatých závorkách, např. *mít paměť (děravou) jako řešeto*. Jde o syntagmatickou variantu frazému, kterou ponecháváme v jednom hesle, kdežto paradigmatické varianty rozepisujeme v LDB Pralex na dvě i více samostatných položek databáze.

Např. s fakultativním komponentem je v SSJČ uveden frazém (*být jako*) *mouchy snězte si mě*; ten zůstal v LDB Pralex zachován. SČFIP v tomto případě uvádí souhrnnou podobu frazémů takto: *být°/sedět° (někde) j. mouchy vemte/snězte si mě*.

Jiným případem je souhrnná podoba frazémů v SČFIP (*to je°*) *jasný/jednoduchý/lehký j. facka/pár facek*. V LDB Pralex byly k uvedeným přirovnáním založeny samostatně ještě následující položky jako transformace: *jasné jako facka, jednoduché jako facka, lehké jako facka; to je jasné jako facka, to je jednoduché jako facka, to je lehké jako facka*.

V některých případech je v SČFI určitý výraz uveden v kulatých závorkách, přestože jde o frazémovou transformaci⁵. Jde o případ, kdy závorka u fakultativního komponentu je ve slovníku použita pro úsporu místa, nejde však o variantu, nýbrž o transformaci. V takovém případě při zpracování v LDB Pralex převažuje trans-

5 Srov. SČFIP, s. 13 a 18: „Varianty syntagmatické (v kulaté závorce) označují část hesla v určitých kontextech vypustitelnou; někdy tak naznačují i jeden z druhů transformací (...)“

„Pro úspornost záznamu všechny výrazy, jejichž levá strana obsahuje sloveso (zpravidla *být*) a adjektivum, přičemž je toto sloveso vypustitelné (označuje se kulatou závorkou jako varianta), mají automaticky již v záhlaví naznačenou adjektivizaci, která se pak zvlášť mezi vlastními transformacemi neuvádí, srov. (*to je*) *jasný/jednoduchý/lehký j. facka* → Adj *jasný/jednoduchý/lehký j. facka* a podobně (*být*) *utahaný j. kotě* atd.“

formační hledisko: souhrnnou podobu z SČFI rozepíšeme na dvě samostatná hesla a zpracujeme je zvlášť (a tudíž je propojíme mezi sebou jako transformace).

6. Práce se substitutem (zástupcem třídy)

Komponenty přirovnání, které mají v lemmatu zástupnou funkci a platí pro širší třídu nebo skupinu slov, jsme se rozhodli uvádět v šikmých závorkách⁶ (např. /dělat/, /někdo/, /něco/, /mu/ apod.). Při užití přirovnání ve větě se za substitut dosazuje konkrétní výraz, takže lemma v LDB Pralex vypadá např. takto: /*dělat*/ *něco*, jako když *bičem mrská*. V příkladech v exemplifikaci pak uvádíme konkrétní realizaci, např. *nadávat*, *vypočítávat něco*, *obléknout se*, *pracovat*, *odpovídat*, *hovořit*, ve větších dokladech ze SYN pak:

To nadával velice rychle , jako < když bičem mrská >

a tu nový host vypočítával její předky , jako < když bičem mrská >

Copak vy se obléknete , jako < když bičem mrská > , ale mně to trvá už déle .

a on si mohl laskavost dovolit , protože pracoval , jako < když bičem mrská >

Působila jako vzorná malá holčička , která umí odpovídat , jako < když bičem mrská >.

Zájmena *někdo*, *něco* označující valenční pozice u verbálních frazémů, uvádíme bez šikmých závorek, viz např. /*dělat*/ *něco* jako když *bičem mrská*. Někdy je obtížné rozhodnout, zda se jedná o variantnost či substitut, zvláště v případě, že je v korpusu málo dokladů; pak tuto skutečnost signalizujeme v poznámce autora k výkladu významu. Např. u hesla *chcípnout jako pes u cesty* je v Pozn. autora uvedeno: „*znovu ověřit kontexty v SYN; rozhodnout „zajít jako pes“ apod.“*

Výkladové slovníky se substitutem nepracují, naše řešení vychází z pojetí SČFIP, kde se na **zástupná slovesa**⁷ upozorňuje v poznámce, např. *udělat něco j. nic (...)* \diamond^8 *Udělat* lze často konkretizovat: *ujít*, *snést*, *napsat*, *postavit* apod.; na **lexikálně obměnitelné komponenty**⁹ (zvl. zájmena, příslovce) se upozorňuje značkou \circ , např. *je \circ to tu \circ j. v blázinci*, *je \circ to \circ j. balzám na rány*.

Naproti tomu SČFIS volí někdy jiné řešení: v případě, že jde o závislou klauzi, označuje gramaticky obměnitelná slovesa či lexikálně obměnitelná zájmena hrnatými závorkami, např. *cpát \circ se, až [mu] pupek prská = cpe se/cpou se, až mu/jim pupek prská*.

6 Šikmé závorky pro substitut jsme navrhli proto, že v lemmatu nebude nikdy lomítka v jiné funkci (varianty důsledně rozepisujeme na dvě hesla).

7 SČFIP, s. 13: „Část frazému (zvl. slovesná) je u některých hesel zobecněná, zástupná, srov. např. hesla *dělat něco j. nikdy (v životě)* a *udělat něco j. nic*, kde *dělat*, *udělat* zastupuje celou řadu různých akčních sloves, konkretizovatelných podle potřeb kontextu (na to obvykle upozorňuje v hesle poznámka ...) a naproti tomu např. *dělat/dřít j. mourovatej*, kde jde o vlastní, konkrétní význam *dělat* apod.“

8 Značka pro frazémy.

9 SČFIP, s. 14: „Stejně [tj. značkou \circ – pozn. aut.] jsou vyznačeny i části výrazu obměnitelné lexikálně, tedy takové, které je možné zaměnit ve stejné platnosti dalšími (obvykle to jsou příslovce nebo zájmena), srov. *je \circ to tu \circ j. v blázinci* (tu: v kanceláři, v bytě apod.) a *je \circ to \circ j. balzám na rány* (to: ta odměna, to pohlazení aj).“

7. Pojetí a zpracování variant u přirovnání v LDB Pralex

V našem pojetí variant vycházíme z vymezení F. Čermáka (1985), podle něhož je varianta taková ustálená nebo neustálená podoba, tj. relativně malá formální nebo sémantická modifikace frazému, která nemění jeho funkci a v zásadě ani podstatu významu. Varianta je tedy vždy obměna téže jednotky, která si zachovává svou funkci, transformací však může vzniknout jednotka další.

Zakládání variant na základě referenčních a materiálových zdrojů

a) Varianty ze SSJČ: Má-li některé přirovnání založené podle SSJČ více variant, zpracováváme je všechny najednou, tj. rozepisujeme všechny na samostatné položky databáze bez ohledu na jejich doloženost v korpusu SYN (např. k *mít peněz jako babek* existují varianty *mít peněz jako cuků*, *mít peněz jako smetí*, *mít peněz jako želez*, *mít peněz jako slupek*, *mít peněz jako šlupek a mít peněz jako třísek* – všechny budou v LDB Pralex založeny a zpracovány samostatně).

Poznámka: V SSJČ nejsou některé varianty zařazeny u daného substantiva ve frazeologické části hesla, ale jsou jen uvozeny komentářem typu „ve frazeologických spojeních“, „v ustálených spojeních“ nebo jsou uvedeny jen v exemplifikaci bez speciálního označení. Zpočátku jsme si jich všímali méně, nyní už pracujeme soustavně i s nimi.

b) Varianty z jiných referenčních zdrojů: Další varianty doplňujeme podle SČFIP, SSČ, případně *Nového akademického slovníku cizích slov* (NASCS). Při vyhodnocování lemmat nalezené varianty nejprve zapisujeme na seznam VARIANTY a jako samostatná hesla je zakládáme až později, po verifikaci materiálem (zakládáme tedy jen ty doložené v korpusu SYN). Důvodem pro založení může být např. to, že transformační frazém se vztahuje právě k příslušné variantě, nebo tuto variantu bylo třeba založit jako „zástupce“ pro jiné nezaložené varianty (např. v hesle *stál, jako když do něj hrom uhodí* je na seznam uvedeno několik samostatně nezaložených variant, např.: *stál, jako když do něj hrom udeří*, *stál, jako když do něj blesk uhodí* atd.).

c) Méně často zakládáme další varianty na základě vlastního povědomí a zároveň doloženosti v korpusu SYN.

d) Jako samostatná hesla zakládáme i ustálené varianty, které se objeví navíc v korpusu SYN, čili až při práci s exemplifikací. Při zpracování příkladové části hesla na základě korpusu znovu kontrolujeme všechny variantní podoby hesla. Pokud najdeme doklady k variantě, kterou máme uvedenou pouze na seznamu VARIANTY podle referenčních zdrojů, ale v hesláři zatím není, zpracujeme ji jako nové heslo (a hypertextově propojíme s komponenty i ostatními variantami). Např. u frazému *děvče jako jahoda* byla nejprve (v 1. fázi) zapsána na seznam, avšak nebyla založena varianta *dívčina jako malina* s Pozn. k var.: (*jen zde podle scfip, nezaloženo*). Při zpracovávání exemplifikace byla však i tato varianta doložena, a tudíž jsme ji založili dodatečně.

Široká variantnost se ukazuje např. u komponentů *palice/kladivo* (viz např. heslo *být jako palicí praštěný*).

Řešení pro dílčí typy variant: ustálené varianty syntagmatické uvádíme v rámci jednoho lemmatu (viz výše v bodu 5); **paradigmatické zpracováváme** jako samostatné položky databáze, u **slovosledných variant typu je to/to je** zatím

v hesláři uvádíme jen jednu podobu lemmatu (přednost dáváme podobě v korpusu frekventovanější). Pokud jsou doloženy obě podoby, uvádíme je v exemplifikaci v samostatných blocích, které nadepíšeme příslušnou podobou lemmatu. **Neustálené varianty** uvádíme ve zvláštním bloku příkladové části hesla. Případné nespisovné varianty přirovnání zakládáme jen tehdy, pokud je tato nespisovná podoba komponentu uvedena samostatně ve Frekvenčním slovníku češtiny nebo SSJČ; podle SČFIP zakládáme variantu pouze v případě doloženosti této podoby v korpusu.

8. Pojetí a zpracování transformací v LDB Pralex

Z hlediska pojetí transformací vycházíme rovněž z vymezení F. Čermáka (1985, s. 223), podle něhož je transformace v idiomatice a frazeologii (IF) takový proces (strukturní nebo gramatické) změny, jehož rezultat (nazývaný též transformace) se od své báze liší funkcí, a tedy postavením v textu, avšak je své bázi formálně a sémanticky v různém stupni podobný.

Transformace mohou být buď ustálené, např. *jehla v kupě sena* (← *hledat jehlu v kupce sena*), nebo neustálené (srov. poloustálené¹⁰), nejčastější v nominalizacích: *postavení něčeho na hlavu* (← *postavit něco na hlavu*). V LDB Pralex uvedené typy transformací rozlišujeme tak, že **transformace ustálené** zakládáme jako samostatná hesla, **transformace neustálené** však uvádíme pouze v příkladové části základového hesla.

Stejně jako u variant u každého hesla zpracováváme všechny ustálené transformace k danému přirovnání. Nejprve ty, které zachycují předchozí slovníky, tj. jak transformace s formou přirovnání (např. k frazému *hubený jako lunt* byly založeny a zpracovány transformace *být (hubený) jako lunt* (SČFIP + srov. SSČ) a */někdo/ hubený jako lunt* (SČFIP)), tak transformace, které již formu přirovnání nemají (např. k heslu *(je to) jako by hrách na stěnu házel* byla v LDB Pralex zpracována i nepřirovnávací transformace *házet hrách na stěnu*).

Na závěr práce s původním přirovnáním zpracováváme případně i další ustálené transformace na základě analýzy materiálu v korpusu SYN.

9. Práce s výkladovými definicemi

Řídíme se zásadou, že výklad významu musí odpovídat konečné podobě lemmatu; definici tedy doplňujeme jako celek. V SSJČ většinou chyběl komponent, u kterého nedocházelo k posunu významu, např. frazém *zdravý jako buk* SSJČ vykládal pouze: *„dokonale“*. V LDB Pralex v tomto případě zatím výklad přeformulujeme takto: *„dokonale, naprosto zdravý“*¹¹.

Výklad má také odpovídat dokladům. Pokud ale korpusové doklady ukazují posuny ve významu proti způsobu zpracování v SSJČ (resp. v SČFI), heslo upravujeme takto: buď rovnou upravíme výklad významu, nebo zatím v poznámce autora k vý-

10 „Řada transformací je **poloustálená**, tj. od své báze pravidelně vytvořitelná ad hoc, srov. např. adjektivizaci frazému *bát se vystrčit hlavu* : *bojící se vystrčit hlavu* apod.“ (SČFIS, s. 627)

11 Jsme si vědomi toho, že frazémy jako celek nabývají i dalších významů a že náš doplněný výklad je zatím ryze pracovní – v budoucím slovníku je třeba se jim věnovat podrobněji.

kladu významu signalizujeme nutnou budoucí úpravu (např.: *upravit výklad – nesouhlasí s doklady*).

Pokud v SSJČ výklad chybí úplně, opíráme se o další slovníky (SSČ, SČFI) – např. *být jako břitva* (SSJČ 0, výklad ze SSČ); pokud není o co se opřít a sami nejsme schopni naformulovat novou definici ve vymezeném čase, zatím pracovně zapisujeme jen otazník, aby bylo vidět, že bude třeba jej pak doplnit.

Pokud doklady ukazují na posuny ve významu (např. u abstrakt), signalizujeme to buď v rámci exemplifikace, nebo založíme další heslo/význam (např. *utíkat jako jelen x čas letí jako jelen*).

10. Zpracování příkladové části hesla u přirovnání v LDB Pralex (nástroj EXEMPLIFIKACE)

Členění exemplifikačních bloků a výběr dokladů

Struktura exemplifikace pro všechny frazémy obecně byla navržena takto:

E1: Větné doklady

E2: Upravené doklady

E3 (fakultativně): V názvech.

U přirovnání však využíváme jen blok Větných dokladů (blok upravených dokladů zde není relevantní, protože forma přirovnání je většinou už předem daná a zajímavé je hlavně jeho zapojení do větného kontextu.

Protože LDB slouží především jako materiálová báze pro budoucí výkladový slovník, v příkladové části hesla uvádíme zejména doklady z korpusu SYN ilustrující konkrétní užití frazému v psaném textu; na rozdíl od SČFIP (Čermák 1983, s. 20) tedy nevybíráme pouze příklady typické (tento požadavek platí až pro výběr dokladů pro slovníkové zpracování).

Do jediného bloku větných dokladů vybíráme doklady, které odpovídají formě a významu daného lemmatu v LDB Pralex, včetně příkladů na negaci (pokud nedochází k posunu významu) a včetně pozičních variant (pokud nebudou zpracovány jako samostatné položky). Např. frazémy *nemastný neslaný/neslaný nemastný* budou zpracovány jako dvě položky databáze (a propojeny jako varianty), ale u výrazů *je to/to je jako ...* budou vyplněny dva samostatné bloky (E1a, E1b) v exemplifikaci u jediného hesla. V LDB Pralex jde konkrétně např. o hesla: *to je jako den a noc; to je jasně jako facka*. Pokud najdeme pouze doklady se slovosledem oproti výchozí podobě ze slovníků inverzním, nadepíšeme blok konkrétní podobou v Názvu 2 (viz např. u přirovnání *to je jednoduché jako facka*)¹². Vzhledem k tomu, že jsme tuto zásadu přijali až v průběhu práce, není zatím provedena u všech dokončených hesel.

V jediném bloku exemplifikace dále zpracováváme doklady, kde je daná jednotka v kontextu rozvíta o další, neustálené komponenty. Doklady s užitím příslušného substantiva v plurálu zařazujeme na konec příkladové části hesla.

V případě, že některé bloky potřebujeme dále členit, číslováme je E1a, E1b atd. a k tomu v příslušném bloku vyplníme podle potřeby Název 2 (sem uvádíme podobu lemmatu nalezenou v dokladech).

¹² Název 2 je volné pole v oddílu exemplifikace, kam je možné vpisovat libovolný text dle potřeby.

Takto členíme především doklady s **fakultativním komponentem**: v prvním bloku (E1a) uvádíme doklady na plnou podobu, v dalším bloku (E1b) pak doklady s vynechaným komponentem, např. u hesla *mít paměť (děravou) jako řešeto* je v bloku E1a k V1 Větné doklady blok nadepsaný titulkem: *mít paměť děravou jako řešeto* s konkrétním větným dokladem:

Má < paměť děravou jako řešeto > a pořád se hádá .

a v bloku E1b k V1 Větné doklady titulek s nadpisem: *mít paměť jako řešeto* s příklady:

„Panebože , na mou čest , to tedy nevím , jak se z tebe mohlo stát takový velký zvíře !“ divil se .

Máš < paměť jako řešeto > , jako kluk .

Jen utrousí : „ Mám < paměť jako řešeto > . Už mi měkne mozek .

Výše uvedené členění používáme také při dokládání příkladů na **substitut**. Např. u hesla */spadnout/ jako pytel brambor* je blok E1 k V1 Větné doklady nadepsaný titulkem *jiná slovesa*, konkrétně ve větných dokladech je pak uvedeno: *padnout, žuchnout, rozsypat se*. (V tomto případě v SYN nebyl nalezen ani jediný doklad s původním slovesem *spadnout*.)

Co se týče doložení **aktualizací** v příkladové části hesla, problémem je už samo jejich vymezení. Ve větných kontextech najdeme často obměny, např. nahrazení ustáleného slovesa synonymem. Aktualizaci zařazujeme do samostatného bloku a označujeme ji nadpisem AKT/VAR v Názvu 2, např. u hesla *tváří se jako by kousl do kyselého jablka* je v bloku E1a k V1 Větné doklady příklad:

Stevardka , která se stále tváří , jako by < kousla > do kyselého jablka

a v bloku E1b k V1 Větné doklady AKT/VAR pak konkrétní příklady na aktualizaci:

Palmu mávl rukou a tvář se mu zkřivila , jako by < kousl > do kyselého jablka .

„Jako kdyby < kousla > do kyselého jablka ,“ řekl jsem .

Praxe ukazuje, že hranice mezi ustálenými variantami a drobnými textovými obměnami se vymezují obtížně.

V případě potřeby dalšího upřesnění zapisujeme rovnou komentář (tučně) dovnitř bloku jako text titulku nad větným dokladem, kterého se to týká, nebo rovnou uvedeme konkrétní podobu přirovnání v textu – viz např. u hesla */něco/ jako břitva* je v E1a k V1 Větných dokladech v Názvu 2 popisek „*o ostré věci*“, v E1b „*o ostré chuti*“ a v E1c „*o sportu, pointě*“. Zatím jde o naše pracovní řešení.

Vyhledávání a výběr vhodných dokladů v korpusu

Doklady většinou vyhledáváme postupně po jednotlivých komponentech. Při používání filtrů se jako nejpraktičtější ukazuje nastavení větší vzdálenosti mezi slovy (oboustranně), např.: při doložení hesla *vypadat jako jezevec* postupujeme takto: zadáme 1) Nový dotaz: [lemma=“jezevec“], 2) P filtr: od -9 do 9 (příp. od -15 do 15) [lemma=“jako“], 3) P filtr: [lemma=“vypadat“], atd.

Pokud je forma frazému pevná, je možné zadávat do pole pro vyhledávání rovnou více lemmat (např. [lemma=“psí“] [lemma=“čumák“]) a u filtrů menší vzdálenost mezi slovy. Např. u hesla *studený jako psí čumák* 1) Nový dotaz: [lemma=“psí“] [lemma=“čumák“], 2) P filtr: od -5 do 5 [lemma=“jako“], 3) P filtr: [lemma=“studený“]. Tento postup je však nutné používat velmi opatrně, protože u přirovnání při kon-

krétním užití v kontextu často dochází k vkládání nejrůznějších komponentů oproti formě uvedené v hesláři.

V poslední fázi výběru podle filtrů (tak do 50 výskytů) frazémy procházíme všechny a třídíme „ručně“. V takovém „předtříděném“ vzorku se objevují i aktualizace, popř. transformace nalezených frazémů. V případě pochybností o členění exemplifikace jsou dotazy či připomínky signalizovány v poznámce k exemplifikaci, např. u hesla *utíkat jako jelen* je v Pozn. k exemplifikaci celkově zapsáno: *syn jen tento doklad, prohlédnout doklady v 1b + přesunout/vymazat?*

Co se týče výběru vhodných dokladů, zásadou je (jak už bylo uvedeno výše), že vybrané doklady by měly odpovídat výkladu významu; často však neodpovídají jak formálně, tak významově, a pak je nutné problém řešit na základě dokladů z korpusu buď na úrovni lemmat/variant/transformací, nebo změnou výkladu, či členěním exemplifikace a využitím Názvu 2.

Pokud máme k dispozici dostatečné množství korpusového materiálu, vybíráme k dané formě uvedené v lemmatu např. i takové doklady, které ukazují jak morfologické obměny (rodu, čísla, času), tak nahrazení substitutu komponenty dle kontextu, změnu slovosledu, aktualizace (př. */růst/ jako houby po dešti*). Vybíráme 5–10 dokladů; v případě menšího počtu dokladů v korpusu (často je dokladů jen velmi málo, 1–3) použijeme všechny (opět jen pokud odpovídají výkladu významu a jsou lexikograficky relevantní) a komentujeme jejich počet v celkové poznámce k exemplifikaci (např. *syn 1x*). Pokud je naopak dokladů moc, vybíráme z nich doklady vhodné (viz výše).

Z hlediska rozsahu pro dostatečnou ilustraci zapojení příslušného výrazu do kontextu většinou stačí jedna věta, někdy je však třeba kvůli srozumitelnosti vybrat kontext delší (dvě věty i více), např. u hesla *chodit jako bludná ovce* konkrétně: *„Zatímco lékaři vás oficiálně považují za uzdravené, pro zaměstnavatele jste nemocní, nevykonní a nevyhovující. A tak chodíte < jako bludná ovce > od jednoho přijímacího pohovoru k druhému, fámy o vaší nestálosti a nespolehlivosti přibývají.“*

Při práci s korpusem se ukazuje, že mnoho přirovnání zůstává nedoložených, a naopak nová (v našem hesláři na základě SSJČ nezachycená) se v něm objevují. Skutečnost, že se heslo nepodaří doložit, může být způsobena jak zastaráváním některých přirovnání, tak i nedostatečností lexikálního materiálu v daném korpusu psaných textů využitelného pro dokládání frazémů obecně.

11. Celkové vyhodnocení zpracování souboru přirovnání v LDB Pralex

V předchozím textu jsme popsali zpracovaná **přirovnání se substantivním komponentem**. Nyní se naopak zaměříme na ta přirovnání, která sice doposud zpracována nebyla, v našem hesláři jsou však na základě SSJČ uvedena.

První skupinou jsou **přirovnání obsahující adjektivní či slovesný komponent**, která však neobsahují žádný komponent substantivní (např. */chodit/ jako bezhlavý, jak to leží a běží*). V hesláři totiž ještě nejsou v této fázi založena všechna adjektiva a slovesa, a tudíž uvedená přirovnání nemůžeme propojovat s příslušnými komponenty a vyhodnocovat je spolu s nimi.

Pokud však přirovnání s adjektivním či slovesným komponentem obsahuje zároveň i komponent substantivní, zpracováno bylo. Takovým příkladem jsou třeba frazémy *bledý jako/jak stěna* a *žít si jak prase v žitě*, se zpracovanými variantami (*žít/mít se*) a oběma komparátory (*jak/jako*).

Zpracována rovněž nejsou přirovnání, která mají velmi širokou variantnost (někde dokonce kolem šedesáti variant); zatím jsme nerozhodli, jakým způsobem budeme takové případy zpracovávat a zapisovat na seznam VARIANTY (např. již uvedené přirovnání *být jako palicí praštěný*).

Rovněž frazémy, které SSJČ označuje jako přísloví (např. *jak nabyt, tak pozbyt*) či pořekadlo (např. *jak k jídlu, tak k dílu*), zatím ponecháváme stranou.

Další nezpracovanou skupinou jsou ta spojení, která jsme vyhodnotili jako k frazémům nenáležící. V hesláři LDB Pralex sice zatím existují, jsou však pracovníě označena nulou za lemmatem (např. *jako zlá saň na mne vylítla 0*). Víceslovné lexikální jednotky, které sice obsahovaly komponent *jak/jako*, který zde však neplnil funkci komparátoru, byly z oddílu „frazémy“ zatím přesunuty do oddílu „nezařazeno“. Jde zejména o různé formule (např. *jak jdou pořád časy?*).

Zhruba čtvrtinu přirovnání se zatím nepodařilo doložit, často se jedná o zastaralé či zastaralá přirovnání (např. *jíst jako sedm drahých let, kluk zdravý jako tuřín, kouká, jako by se octa napil, koukat jako drahé časy*).

Literatura

- ČERMÁK, František: Česká přirovnání. In: Slovník české frazeologie a idiomatiky. Přirovnání. Red. F. Čermák – J. Hronek – J. Machač. 1. vyd. Praha: Academia 1983, s. 463–492.
- ČERMÁK, František: České frazémy a idiomy verbální. In: Slovník české frazeologie a idiomatiky. Výrazy slovesné. Red. F. Čermák – J. Hronek – J. Machač. 1. vyd. Praha: Academia 1994, s. 597–630.
- ČERMÁK, František: Frazeologie a idiomatika. In: Filipec, J. – Čermák, F.: Česká lexikologie. Praha: Academia 1985, s. 166–236.
- ČERMÁK, František: Základy a zásady frazeologie a idiomatiky. In: Čermák, F.: Frazeologie a idiomatika česká a obecná. Praha: Karolinum 2007, s. 23–72.
- Encyklopedický slovník češtiny. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: Nakladatelství Lidové noviny 2002, s. 356. (EŠČ)
- JANOVEC, Ladislav: Frazémové aktualizace v nové slovní zásobě. In: Svět za slovy a jejich tvary, svět za spojením slov. Sborník příspěvků ze 7. mezinárodní konference Setkání mladých lingvistů, konané na Filozofické fakultě Univerzity Palackého ve dnech 15.–17. května 2006. Ed. V. P. Polách. Olomouc: Univerzita Palackého 2008, s. 103–106.
- JANOVEC, Ladislav: Z nové české frazeologie – poznámky k internacionalizaci frazeologie. In: Varia XI. Zborník materiálů z XI. kolokvia mladých jazykovedcov (Spišská Nová Ves 28.–30. 11. 2001). Ed. M. Šimková. Bratislava: Slovenská jazykovedná spoločnosť při SAV 2004, s. 22–25.

- JANOVEC, Ladislav: Zpracování frazeologie v české lexikální databázi. In: *Varia XV. Zborník materiálov z XV. kolokvia mladých jazykovedcov (Banská Bystrica – Tajov 7.–9. 12. 2005)*. Zost. A. Gáliková – A. Chomová. Bratislava: Slovenská jazykovedná spoločnosť pri SAV – Katedra slovenského jazyka a literatúry FHV UMB v Banskej Bystrici 2008, s. 232–236.
- Nový akademický slovník cizích slov A–Ž. (J. Kraus a kol.) 1. vyd. Praha: Academia 2005. 880 s. (NASCS)
- Slovník české frazeologie a idiomatiky. Přirovnání. Red. F. Čermák – J. Hronek – J. Machač. 1. vyd. Praha: Academia 1983, s. 9–27.
- Slovník české frazeologie a idiomatiky. Výrazy neslovesné. Red. F. Čermák – J. Hronek – J. Machač. 1. vyd. Praha: Academia 1988, s. 9–23.
- Slovník české frazeologie a idiomatiky. Výrazy slovesné. Díl I (A–P). Red. F. Čermák – J. Hronek – J. Machač. 1. vyd. Praha: Academia 1994, s. 9–33.
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)

Zpracování víceslovných lexikálních jednotek v lexikální databázi Pralex (se zaměřením na zpracování nefrazeologických víceslovných lexémů)¹

Zdeňka Opavská

Ústav pro jazyk český AV ČR, v. v. i.

Processing of Multi-Word Lexical Units in the Pralex Lexical Database (with Focus on the Processing of Non-Phraseological Multi-Word Lexemes)

The first part of this paper deals with the concept of multi-word lexical units and the delimitation of their basic types in the Pralex lexical database, including the analysis of some terminological and conceptual issues. After a short preview of the approaches to the lexicographical treatment of multi-word lexical units in contemporary monolingual dictionaries, the second part of this paper deals with the database processing of multi-word lexical units, both with the general rules of their treatment in the Pralex lexical database and with the specific rules of treatment of one of two basic types of multi-word lexical units, i.e. multiple-word namings.

1. Víceslovné lexikální jednotky v lexikální databázi Pralex

Jednou z oblastí, které při práci na lexikální databázi (dále LDB) Pralex věnujeme zvláštní pozornost, je oblast víceslovných lexikálních jednotek (VLJ). V dosavadních českých výkladových slovnících byla tato problematika přes určité shodné rysy řešena vždy specificky, v rámci koncepce daného slovníku (podrobněji viz dále). S růzností přístupů, s různými klasifikacemi a odlišnými delimitacemi jednotlivých typů víceslovných lexémů, stejně tak jako s růzností používané terminologie se setkáváme i v řadě teoretických článků, příspěvků a studií, které byly problematice víceslovných lexémů věnovány².

Pro zpracovatele LDB Pralex to tak znamená řešit řadu lexikograficko-lexikologických problémů spojených s touto oblastí. Volbu způsobu zpracování VLJ v LDB Pralex rovněž ovlivňuje i databázová podoba díla, díky níž lze uvažovat i o takových řešeních, která se v tradičních tištěných výkladových slovnících běžně nepoužívala nebo která ani nebylo možné použít z praktických důvodů nedostatku místa (jde zde např. o otázku zpracování VLJ jako samostatných hesel nebo o způsob exemplifikace VLJ). Na druhé straně právě tato odlišnost staví před zpracovatele databáze řadu nových problémů nebo i problémů již dříve známých, které je však potřeba pro LDB Pralex uchopit a řešit nově, z jiného úhlu pohledu.

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století (AV0Z90610521)*, jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Vzhledem k tomu, že česká a slovenská lexikografie a lexikologie často řeší v oblasti VLJ shodné či velmi obdobné problémy, opíráme se v tomto příspěvku o výsledky bádání jak českých, tak i slovenských autorů.

2. Pojetí VLJ v LDB Pralex. Charakteristika jejich základních typů

V LDB Pralex vycházíme z tradičního rozdělení VLJ do dvou základních skupin, a to na nefrazeologické a frazeologické/idiomatické lexikální jednotky. Zároveň si však v souladu s upozorněním řady lingvistů (např. Čermák 2007, s. 32–33, Jarošová 2000a, 2007, Martincová 2005, s. 212) plně uvědomujeme, že pro oblast VLJ jsou příznačné plynulé přechody mezi jednotlivými typy VLJ a že hranice mezi nefrazeologickými a frazeologickými jednotkami je neostrá. Srov. k tomu např. charakteristiku rozsáhlé oblasti kolokací jako kombinačního kontinua, v němž nalezneme prototypická jádra určitých kategorií i poměrně velké oblasti přechodných jevů tvořících periferii (Jarošová 2007) nebo konstatování F. Čermáka o neostrosti hranic frazeologie: „Každý frazém a idiom lze vždy komplexně vymezit více kritérii zároveň, především stupněm jeho anomálie. Neostrosta hranic IF [idiomatiky a frazeologie – pozn. ZO] a pozvolný přechod do pravidelného jazyka vede k tomu, že u některých typů klesá i relevantní počet kritérií a aspektů anomálnosti výrazu. Periferii IF, tj. pásmo přechodu do pravidelného jazyka, tvoří oblast kvazifrazémů a kvaziidiomů.“ (Čermák 2007, s. 42).

Vzhledem ke zmíněné komplikované problematice VLJ i k obecně známým problémům při jejich popisu a vymezení (viz např. různá pojetí jednotlivých typů VLJ u jednotlivých autorů, rozdílné klasifikace VLJ, různorodost a nejednotnost při užívání lingvistických termínů týkajících se této oblasti) se nejprve věnujeme otázce vymezení VLJ v LDB Pralex a dotkneme se i problému užívané terminologie. V druhé části příspěvku se zaměříme na zásady zpracování jednoho z typů VLJ v LDB Pralex, totiž na zpracování nefrazeologických lexikálních jednotek, tj. víceslovných pojmenování (sousloví).

2.1. Terminologické přístupy

V současnosti lze v české literatuře vysledovat v oblasti VLJ dvě hlavní terminologické linie. Jednu tvoří déle používaná, „tradiční“, avšak různými pojetími zatížená terminologie používající pojmy/termíny jako **sdružená pojmenování, víceslovná pojmenování, sousloví, frazeologické jednotky, frazémy, ustálená slovní spojení, ustálená spojení slov, víceslovné lexikální jednotky, víceslovné lexémy**. Autoři uvádějí ve svých pracích různá kritéria pro vymezení jednotlivých typů VLJ, jejich pojetí se různě překrývají a kříží (srov. např. Bečka 1934, Man 1953, Kuchař 1963, Havránek – Jedlička 1963, Machač 1967, Šmilauer 1982, Filipec – Čermák 1985, Hauser 1986, Příruční mluvnice češtiny 1995, Čeština – řeč a jazyk 2000, Čermák 2001b, Martincová 2005, Čermák 2007, 2010; Slovník slovanské lingvistické terminologie 1977–1979, Lotko 1998).

Druhá – v české literatuře novější – terminologická linie vychází ze syntagmatického, kombinačního pohledu a týká se zkoumání kolokací³ (tj. kombinací, syntagmat jazykových prvků lexikální povahy, viz Encyklopedický slovník češtiny 2002, dále

3 Termín kolokace zavedl J. R. Firth (1957, cit. podle Čermák 2006, s. 10); do české lingvistiky ho zavedl a vymezil F. Čermák (zejm. 2001a, 2001b, 2006, 2010; ESČ, heslo kolokace). Rovněž termín kolokace je v literatuře různě vymezován. Podrobný a velmi instruktivní přehled různých pojetí kolokací podává A. Jarošová (2007).

ESČ, s. 218). V české lingvistice propracoval klasifikaci **kolokací** F. Čermák (zejm. 2001b; 2006; 2010, s. 222–225, 314–318; ESČ, heslo kolokace⁴). Čermákova klasifikace je založena na základních distinkcích systém – text, pravidelný – nepravidelný a vymezují se v ní následující skupiny: 1) systémové – pravidelné: **termínové kolokace** (víceslovné termíny), **propriální kolokace** (víceslovná propria), 2) systémové – nepravidelné: **idiomatické kolokace** (idiomy a frazémy), 3) textové – pravidelné: **běžné kolokace**, 4) textové – nepravidelné: **individuální metaforické kolokace**. Na přechodu mezi textovými a systémovými kolokacemi stojí **běžné uzuální kolokace**⁵. V české literatuře byl kolokační přístup k víceslovným lexémům uplatněn zejména ve studiích z korpusové lingvistiky (viz např. sborník Kolokace 2006); v oblasti terminologie byl pak aplikován v monografii I. Bozděchové (2009).

V LDB Pralex jsme se přiklonili k takovému pojetí VLJ, které pro odlišení nefrazeologických lexikálních jednotek od frazémů užívá lexikálně-sémantická a nominační kritéria. Neopíráme se tedy o protiklad pravidelnost :: nepravidelnost, anomálnost jako o základní distinkci pro delimitaci jednotlivých typů VLJ. Z tohoto důvodu – a také i v souladu s předcházejícími lexikografickými díly – používáme tradiční terminologii, i když jsme si zároveň vědomi jak terminologických problémů, které tento přístup s sebou přináší (srov. bod 2.2), tak i výhrad, které byly vůči těmto „tradičním“ termínům vzneseny (srov. Čermák 2001b).

2.2. Vymezení VLJ v LDB Pralex a některé problémy s tím spjaté

V LDB Pralex v rámci **víceslovných lexikálních jednotek, víceslovných lexémů** (v literatuře označovaných též jako **ustálená slovní spojení, ustálená spojení slov**) vymezujeme dvě hlavní skupiny: 1) **víceslovná pojmenování, sousloví** (v literatuře též **sdužená pojmenování**), 2) **frazémy**; srov. Filipec – Čermák 1985, s. 105: „Ustálená spojení slov dělíme na široce chápaná sdužená pojmenování, sousloví ... a s nimi paralelní vlastní frazémy.“

Zvláštní skupinu tvoří **víceslovná sysémantika a víceslovná zájmena, číslovky a citoslovce**. Jejich pojetí bude třeba v LDB Pralex dále propracovat (srov. např. otázku hodnocení těchto spojení jako gramatických frazémů⁶); touto skupinou se zde vzhledem k její specifičnosti dále nezabýváme.

Poznámky k používané terminologii

1) V literatuře se objevují i jiná vymezení a různá hierarchie uvedených „tradičních“ termínů. Vedle *ustáleného slovního spojení* chápaného jako nadřazený pojem

4 Autory hesla jsou F. Čermák a Z. Hladká.

5 Do této klasifikace nejsou explicitně zahrnuty neterminologické kolokace; tento typ je však zahrnut v Čermákově typologii a klasifikaci pojmenování (Čermák 2010, s. 152–153), v níž se na základě rysu definičnosti (distinkce Ne/Definice) dělí textové a systémové nominace (a lexémy) na terminologická a neutrální, tj. běžná neterminologická pojmenování.

6 J. Filipec (Filipec – Čermák 1985, s. 106) označuje sekundární předložky, spojivá, částicová, citoslovená a zájmenná spojení jako „specifická ustálená spojení“ a uvádí je v rámci výkladu o sdužených pojmenováních. Naopak v Čermákově frazeologické koncepci se víceslovné předložky, spojky, částice, zájmena a číslovky označují jako gramatické frazémy a řadí se k periférii frazeologie, ke kvazifrazémům (Čermák 2007, s. 315).

pro víceslovné lexikální jednotky různého druhu může být tento výraz vymezen i úžeji, viz: „Slovní zásoba ... je souhrn lexikálních jednotek, tj. slov, ustálených slovních spojení a frazémů.“ (Filipec 1984, s. 140). Rovněž termín *víceslovné pojmenování* je v literatuře vymezován jednak úžeji, tj. jako synonymum pro sdružená pojmenování, sousloví (např. Kuchař 1963, Martincová 2005, Příruční mluvnice češtiny 1995) – k tomuto užšímu vymezení se přikláníme –, jednak šířeji, a to a) jako označení pro víceslovné lexémy / ustálená slovní spojení vůbec (např. Hauser 1986, Machač, rkp.), b) jednak jako termín zahrnující i víceslovné výrazy neustálené („pojmenování víceslovné: ne/ustálený víceslovný jazykový výraz, kombinace dvou/více lexémů, zvl. pro jediný pojem“, Čermák 2001a, s. 272). Stejně tak i *sousloví* může být někdy chápáno i šířeji, tj. nejen jako jeden z typů VLJ, srov. „sousloví – starší název pro víceslovný lexém, resp. víceslovné pojmenování a kolokaci“ (Čermák 2001a, s. 284).

2) Terminologie, označovaná v tomto příspěvku jako „tradiční“, tedy užívá pro označení nefrazeologických jednotek několik výrazů, a to *víceslovná pojmenování*, *sdružená pojmenování*, *sousloví* (v Čermákově kolokační klasifikaci této skupině odpovídá termín *systémové pravidelné kolokace*). V LDB Pralex užíváme pro označení nefrazeologických lexémů z praktických důvodů (a na základě dosavadní lexikografické praxe) jednoslovný výraz *sousloví*; v tomto příspěvku užíváme zejména při zdůraznění nominativního aspektu i synonymní termín *víceslovná pojmenování*.

Víceslovné lexikální jednotky vymezujeme v souladu s Macháčem (1967, s. 138) jako „komplexy dvou nebo více slov reprodukováných a reprodukovatelných v konkrétních promluvách jako hotové již celky v ustáleném významu“. K charakteristickým rysům⁷ VLJ náleží **ustálenost** (tj. ustálenost formy, ustálenost významu)⁸, **reprodukovatelnost a reprodukovatelnost spojení jako celku**⁹, chápána z jiného úhlu jako **dispozičnost**¹⁰ (jednotka je součástí jazykového systému, je součástí inventáře slovní zásoby). Víceslovná pojmenování a frazémy se vyznačují rovněž rysem **nominativnosti**¹¹, tj. tyto jednotky mají nominativní, pojmenovací funkci.

V LDB Pralex vycházíme při delimitaci základních typů VLJ, tj. víceslovných pojmenování a frazémů, z nominativní funkce VLJ. Toto hledisko doplňujeme souborem

7 Srov. Machač (1967), Filipec – Čermák (1985), Příruční mluvnice češtiny (1995), Manuál slovníku neologizmů (1996), Horecký (1997), Jarošová (2000a, 2000b, 2000c, 2005), ESČ.

8 Srov. ESČ (s. 516): „ustálenost: opakovaným územ fixovaná existence dané formy n. kombinace forem v relativně pevném významu, pevné funkci a formě“.

9 Srov. ESČ (s. 516): „Svou ustáleností se jednotka stává součástí jazykového systému, ze kterého je pak jen reprodukována, ne pokaždé znovu vytvářena, a je umožněno její opakované užití se stejnými výše uvedenými vlastnostmi.“

10 S rysem dispozičnosti pracuje zejména slovenská literatura, např. Jarošová (2000a, 2000b), Dolník (1997).

11 Je otázkou, zda se tímto rysem vyznačují i víceslovná synsémantika. Srov. k tomu ESČ (s. 318): „Pojem pojmenování je tradičně spojován (výhradně n. přednostně) s tzv. pojmenovacími slovními druhy, majícími pojmový význam, tj. se subst., adj., slovesy a adv. V širším smyslu však lze pojmenovací funkci přiznat všem slovním druhům, tedy i synsémantikům (pojmenování relací a modifikací).“

dalších kritérií. **Víceslovná pojmenování (sousloví)** vymezujeme v souladu s Machačem (1967) a Martincovou (2005) jako komplex dvou nebo více slov, s primární nominativní funkcí, reprodukovatelný a reprodukovaný v konkrétních promluvách jako celek¹². Charakteristickými rysy víceslovných pojmenování¹³ jsou primární nominativní funkce, vztah ke konkrétnímu (určitému) denotátu, nociónálnost a významová určitost. Víceslovné pojmenování může být jediným pojmenovacím prostředkem k označení dané skutečnosti, nebo může být synonymem jednoslovného, nebo i jiného víceslovného pojmenování; víceslovné pojmenování může být dále součástí pojmoslovné soustavy, součástí řady druhových názvů. V protikladu k výše uvedeným vlastnostem víceslovných pojmenování se **frazémy** vyznačují sekundární pojmenovací funkcí, denotativní vágností, významovou neurčitostí, expresivitou, emocionálností, hodnotící funkcí, přeneseností, obrazností a idiomaticití¹⁴. Pro oba typy VLJ však platí, že u dané lexikální jednotky nemusí být uvedené rysy přítomny v plné šíři, některé z nich mohou u lexikální jednotky chybět. Rovněž tak v oblasti VLJ existují takové lexémy, které se vyznačují jak souslovnými, tak frazémovými rysy. Tyto lexikální jednotky svým charakterem náležejí do přechodné oblasti, průnikové periferie sousloví a frazémů, a činí proto z hlediska zařazení lexému k souslovím nebo frazémům značné problémy (k některým aspektům tohoto problému viz dále).

Jiný přístup k odlišení frazémů/idiomů od nefrazeologických jednotek nalezneme v pracích F. Čermáka (zejm. Filipec – Čermák 1985; Čermák 1995, 2007, 2010; hesla frazém a idiom, frazeologie a idiomatika v ESČ). F. Čermák vede rozdíl mezi nefrazeologickými lexémy a frazémy na základě strukturní distinkce pravidelnost :: nepravidelnost (anomálnost), tj. nefrazeologické a frazeologické jednotky od sebe odlišuje podle toho, „zda jsou budovány podle sémantických a formálních (gramatických) pravidel nebo nikoliv“¹⁵. Frazém/idiom je na základě uvedeného komplexního kritéria anomálnosti vymezen jako „jedinečná ustálená kombinace minimálně dvou prvků, z nichž některý (popř. žádný) nefunguje stejným způsobem v žádné jiné kombinaci n. více kombinacích, resp. vyskytuje se v takové funkci pouze ve výrazu jediném, popř. několika málo“¹⁶. Na tomto teoretickém základu je postavena koncepce čtyřdílného *Slovníku české frazeologie a idiomatiky*¹⁷ (dále SČFI). Jak bylo uvedeno výše, v LDB Pralex vycházíme při rozlišení víceslovných pojmenování a frazémů z jiných kritérií, než je kritérium pravidelnosti :: anomálnosti; z tohoto důvodu také některé lexikální jednotky hodnotíme z hlediska (ne)příslušnosti k frazémům jinak, než je tomu v SČFI (jde zejména o ty případy, které jsou v Čermákově koncepci řa-

12 Viz Machač (1967, s. 138 a 141), Martincová (2005, s. 212).

13 Srov. Machač (1967), Filipec – Čermák (1985), *Příruční mluvnice češtiny* (1995), *Manuál slovníku neologizmů* (1996), Jarošová (2000a, 2000b, 2000c, 2005).

14 Srov. Machač (1967), Horecký (1997), Jarošová (2000a, 2000b, 2000c, 2005), ESČ (s. 140), Janovec (2005).

15 Čermák (2010, s. 153).

16 ESČ (s. 140, autorem hesla je F. Čermák).

17 *Slovník české frazeologie a idiomatiky* obsahuje následující díly: *Přirovnání* (1. vyd. 1983; 2., přepracované a doplněné vyd. 2009), *Výrazy neslovesné* (1. vyd. 1988; 2., přepracované a doplněné vyd. 2009), *Výrazy slovesné* (1. vyd. 1994; 2., přepracované a doplněné vyd. 2009), *Výrazy větné* (1. vyd. 2009).

zeny k periférii frazeologie, k tzv. kvazifrazémům/kvaziidiomům, např. *kočičí hlavy, mrtvý chod, provádět analýzu*).

Z hlediska delimitace sousloví :: frazém činí problémy zejména ty lexikální jednotky, které stojí na pomezí víceslovných pojmenování a frazémů. Typickým případem jsou lexémy s konkrétním denotátem (rys charakteristický pro víceslovná pojmenování), které se však vyznačují rysem obraznosti (rys pokládáný za příznačný pro frazémy), např. *černá díra, slepé střevo, babočka paví oko, rajske jablko, vraní oko, vlčí mák, zlatý déšť, slzičky panny Marie, vousy sv. Ivana, kočičí hlavy, bílá technika, modrá knížka, červené barety, modré přilby, modrá armáda, černé zlato, červená růže*. Ve zkoumaném souboru jsou zastoupena terminologická a terminologizovaná pojmenování (*černá díra, slepé střevo, babočka paví oko, rajske jablko, vlčí mák, vraní oko*), neoficiální či lidová pojmenování (*zlatý déšť, slzičky panny Marie, vousy sv. Ivana*) i obrazné neterminologické názvy (*kočičí hlavy, modrá knížka, bílá technika, červené barety, modré přilby, modrá armáda, černé zlato, červená růže*).

Přechodný ráz uvedených výrazů se projevuje i v jejich odlišném hodnocení v odborné literatuře i ve slovníkovém zpracování. Např. výraz *vlčí mák* řadí k souslovím Havránek – Jedlička (1963), Slovník slovanské lingvistické terminologie (1977–1979), Šmilauer (1982), naopak v rámci frazeologických spojení ho uvádí Hauser (1986). Výraz *slepé střevo* je ve *Slovníku spisovné češtiny pro školu a veřejnost* (dále SSČ)¹⁸ uveden v exemplifikaci u hesel *slepý* a *střevo* s oborovým kvalifikátorem *anat.*, naproti tomu v SČFI je pokládán za frazém. Obdobně kolísá hodnocení výrazu *modrá armáda*: v SSČ je tento lexém uveden v exemplifikaci u hesla *armáda* pouze s označením *přen.* (není tedy uveden jako frazém v případně speciálně vyčleněné frazémové části), naproti tomu *Nový akademický slovník cizích slov* (2005, dále NASCS) i SČFI tento výraz za frazém pokládají. Výraz *červená růže* (= Česká strana sociálně demokratická) je ve slovníku *Nová slova v češtině 2* (SN 2) hodnocen jako víceslovné pojmenování; O. Martincová však ve své studii o víceslovných pojmenováních (Martincová 2005) upozorňuje na jeho frazémový charakter.

V LDB Pralex jsme se rozhodli v případě obrazných výrazů s konkrétním denotátem opřít se především o rozdíl primární a sekundární pojmenovací funkce (srov. Machač 1967). K víceslovným pojmenováním řadíme tedy obrazné terminologické i neterminologické výrazy s primární pojmenovací funkcí, a to včetně neoficiálních, lidových označení rostlin (a živočichů), např. *černá díra, slepé střevo, babočka paví oko, rajske jablko, vraní oko, vlčí mák, zlatý déšť, slzičky panny Marie, vousy sv. Ivana, kočičí hlavy, bílá technika, modrá knížka, červené barety, modré přilby*. Hledisko primární nominativní funkce může být v některých případech podpořeno i přítomností jiných rysů typických pro víceslovná pojmenování, např. existencí řady druhových názvů (např. *bílá technika – černá technika – hnědá technika; červené barety – zelené barety – modré barety / modré přilby*). Obrazné výrazy se sekundární pojmenovací funkcí hodnotíme jako frazémy, např. *modrá armáda* (= železničáři), *černé zlato* (= 1. uhlí, 2. ropa), *červená růže* (= Česká strana sociálně demokratická).

18 1. vyd. 1978; 2., opravené a doplněné vyd. 1994; 3., opravené vyd. 2003; 4. vyd. (obsahově nezměněné vůči 3. vyd.) 2005.

Přehled vytyčených rysů charakteristických pro víceslovná pojmenování a frazémy

Víceslovná pojmenování	Frazémy
ustálenost (u některých typů může být velmi nízká, např. u některých typů jmenných multiverbátů)	ustálenost (v mnohem větší míře se zde však projevuje variantnost lexikálních jednotek; specificky pak aktualizace)
reprodukovanosť, dispozičnost	reprodukovanosť, dispozičnost
primární nominativní funkce	sekundární nominativní funkce
jeden (konkrétní) denotát; denotativní a sémantická určitost	neurčitý, vágní denotát; sémantická a denotativní vágnosť
především nociónální pojmenování	především pojmenování s emocionálním, hodnotícím, expresivním příznakem
obraznosť, přenesenosť – v našem pojetí může být přítomna	obraznosť, přenesenosť, idiomatičnost – charakteristický rys pro většinu frazémů
monokolokabilní komponent – v našem pojetí může být přítomen	monokolokabilní komponent
součást pojmoslovné soustavy nebo součást řady druhových názvů	–
synonymnosť: a) s jednoslovnou LJ, včetně univerbátů, b) s jinými víceslovnými pojmenováními	a) synonymnosť s jinými frazémy, b) tzv. ekvivalence se slovem
v některých případech gramatická omezení (např. u adj. nemožnosť stupňování)	omezení gramatických, textových funkcí frazému (viz Čermák)

3. Zpracování VLJ v LDB Pralex (s přihlédnutím k jejich zpracování v českých výkladových slovnících)

Jednou z prvních otázek, kterou bylo potřeba z hlediska zpracování VLJ v LDB Pralex řešit, bylo, zda VLJ budou, či nebudou součástí hesláře databáze, neboli zda VLJ budou, nebo nebudou zpracovány v podobě samostatných hesel. Volba přístupu ke zpracování VLJ má totiž výrazný vliv jak na obsah a na podrobnosť lexikálního popisu, tak i na volbu lexikografických technik; jiný lexikograficko-technický postup budeme volit v případě, že daná lexikální jednotka bude v LDB zpracována jako samostatné heslo, a bude tedy možné využít všech možností, které při databázovém zpracování heslová stať poskytuje, včetně vzájemného propojení víceslovných a jednoslovných hesel, a jiné řešení zvolíme v případě těch VLJ, které budou uvedeny v rámci jednoslovných hesel – jejich komponentů. Ve druhém případě bude podrobnosť zpracování nižší, může se omezit i na pouhou registraci, tj. uvedení VLJ v příkladové části heslové stati jednoslovného hesla.

3.1. Zpracování VLJ v dosavadních slovnících

V tradici českých výkladových slovníků bylo zpracovávat většinu VLJ v rámci jednoslovného hesla. Základním východiskem pro zpracování slovníku tedy bylo slovo – jednoslovné heslo. Zobecněně je možné říci, že slovníky *Slovník spisovného jazyka českého*¹⁹ (dále SSJČ), *Slovník spisovné češtiny pro školu a veřejnost* (SSČ), *Akademický slovník cizích slov* (1995, ASCS; v 2. vydání 2005 *Nový akademický slovník cizích slov*; NASCS) uváděly **a) VLJ na různých místech příkladové části hesla (exemplifikace)**, a to s doplňujícím výkladem či bez výkladu, v případě víceslovných termínů s uvedením oboru (výjimkou byly víceslovné termíny uvedené v rámci terminologického významu dané jednoslovné lexikální jednotky), **b) VLJ jako jeden z významů** polysémního hesla (často šlo zároveň o upozornění na monokolokabilitu některého z komponentů spojení), **c) VLJ jako samostatné heslo** – toto řešení se užívalo jen v případě víceslovných přejímek a citátových výrazů. Způsobem a umístěním VLJ se jednotlivé slovníky liší, přesto však u nich nalezneme i určité shodné body.

Ke společným rysům ve zpracování VLJ patří uvádění frazémů v exemplifikaci ve zvláštním oddílu za speciálním znakem ♦ (tj. za tzv. hřebíkem), uvádění tzv. frazeologických kvazihomonym (tj. VLJ fungujících jak v přímém, tak přeneseném – frazeologickém – významu) v exemplifikaci formou „daný víceslovný výraz – čárka – přen. + výklad“²⁰ a zpracování víceslovných přejímek a citátových výrazů v podobě samostatných hesel. V řadě jiných zpracovatelských zásad se slovníky od sebe liší. Tak např. frazémová část označená ♦ byla v SSJČ umístěna na konci dokladové části daného významu (a někdy tak byly frazémy uvedeny u každého z nich), v SSČ však byla tato část uvedena souhrnně až po posledním významu. Jiným případem je způsob uvádění víceslovných termínů. Zatímco SSČ a NASCS je uváděly v příkladové části hesla, aniž by předem určovaly jejich umístění, SSJČ se pokusil vymezit v exemplifikaci i terminologický oddíl. Ten byl umístěn na samotný konec příkladové části významového odstavce²¹.

V české slovníkové produkci existují rovněž lexikografická díla, která zpracování VLJ věnují podstatně větší prostor. Zatímco uvedené všeobecné výkladové slovníky zpracovávají formou samostatných hesel pouze určitý typ VLJ, české neologické slovníky *Nová slova v češtině 1* (1998, SN 1), *Nová slova v češtině 2* (2004, SN 2) a slovník – suplement *Co v slovnících nenajdete* (1994, CSN) vedle víceslovných přejímek podávají jako samostatná hesla i další typy VLJ (a to jak víceslovná pojmenování, tak frazémy), byť VLJ jsou u nich po formální stránce podřazeny pod jednoslovný komponent. Speciální *Slovník české frazeologie a idiomatiky* (SČFI) zpracovává – v souladu se svým zaměřením a koncepcí – frazémy jako samostatná hesla (výjimkou jsou verbonominální spojení s abstraktem, jež jsou řazena ke kvazifrazémům a jsou uváděna sdruženě pod abstraktním substantivem).

3.2. Zpracování VLJ v LDB Pralex

Odpověď na otázku, jakým způsobem budou VLJ v LDB Pralex zpracovány, byla ovlivněna jak poznatky a zkušenostmi z dosavadní lexikografické praxe, tak i mož-

19 1. vyd. 1960–1971 ve 4 svazcích; 2., nezměněné vyd. 1989 v 8 svazcích.

20 Tento způsob zavedl SSJČ a přejímá ho rovněž SSČ.

21 Pokud heslo obsahovalo jak samostatný frazémový, tak terminologický oddíl, ocitly se v takovém případě víceslovné termíny až za frazémovými jednotkami.

nostmi, které databáze nabízí. Dosavadní české obecné výkladové slovníky zpracovávaly na základě své koncepce většinu VLJ v rámci jednoslovných hesel (viz výše); s tím také souvisejí i některé nevýhody, např. zdánlivá (či i skutečná) nejednotnost zpracování dané VLJ u jednotlivých komponentů (u jednoho z komponentů mohla být daná VLJ vyložena, u druhého nikoli, výklady téže VLJ se mohly v heslech jednotlivých komponentů lišit) či různost zpracování frazeologických lexikálních jednotek (frazémy byly v heslové stati jednoslovného hesla uváděny na různých místech a s různým označením – za tzv. hřebíkem, před ním, s označením *ve frazeol. spojení* či *v ustál. spoj.* nebo se speciálním označením typu „čárka přen.“; týž frazém mohl být uváděn pod jednotlivými komponenty v odlišné (variantní) podobě nebo mohly být u jednotlivých komponentů uváděny místo jednoho frazému různé frazémové transformace).

Oproti praxi dosavadních obecných výkladových slovníků zpracováváme v LDB Pralex značnou část VLJ jako samostatná hesla (a využíváme tak i jedné výhody databázového zpracování, totiž dostatku prostoru pro zpracování lexikálních jednotek). V tomto ohledu tak v mnohém navazujeme na způsob zpracování VLJ ve speciálních výkladových slovnících – SN 1, SN 2 a SČFI.

V LDB Pralex však nezpracováváme jako samostatná hesla všechny VLJ (vzhledem k velkému množství VLJ ve slovní zásobě to ani z praktických důvodů není možné). Přístup k samostatnému/nesamostatnému zpracování VLJ v LDB Pralex se liší podle typu VLJ. **Frazémy** jsou zpracovány jako samostatná hesla en bloc²². U **víceslovných pojmenování (sousloví)** uplatňujeme při jejich zpracování především lexikografické kritérium tzv. potřeby výkladu²³. Samostatně zpracováváme ta sousloví, která podle našeho názoru výklad potřebují proto, že význam dané VLJ nevyplývá ze „slovníkových“ významů jejich komponentů. Ta sousloví, která z lexikografického hlediska „výklad nepotřebují“, uvádíme v exemplifikaci u jednoslovného hesla v příslušném exemplifikačním bloku²⁴ a v této fázi prací daná spojení nijak neoznačujeme. Výhodou zvoleného postupu je možnost vybírat si (v rámci stávajícího výzkumného záměru) z rozsáhlého souboru sousloví k samostatnému zpracování jen jejich část. Na druhé straně je však třeba si uvědomit, že důsledkem použití lexikografického kritéria „potřeby výkladu“ je značná závislost způsobu zpracování sousloví na výkladu významů jejich komponentů – jednoslovných hesel²⁵.

Při zpracování VLJ jako samostatných hesel – vedle tradičních lexikografických položek, jako je např. výklad významu a exemplifikace – klademe důraz i na podchytení systémových vztahů jak mezi danou VLJ a jejími komponenty, tak mezi VLJ navzájem či mezi VLJ a jednoslovným lexémem. Tak např. u víceslovného hesla uvádíme formou hypertextových odkazů soupis většiny jednoslovných komponentů,

22 Ve zpracování frazémů jako samostatných hesel, a to včetně transformací, se sblížíme s praxí SČFI; viz též zde stať E. Kolovecké o přirovnáních.

23 Toto kritérium je používáno – avšak z jiných důvodů – v koncepci slovenského Slovníku současného slovenského jazyka (SSSJ). Srov. též praxi předcházejících všeobecných výkladových slovníků, které uváděly doplňující výklad u VLJ tehdy, pokud šlo o nesoučtovost významu komponentů.

24 K exemplifikaci viz zde podrobněji ve stati M. Koutové.

25 K tomuto aspektu lexikografického zpracování srov. Jarošová (2000b, s. 53).

u jednoslovných hesel naopak uvádíme přehledy frazémů a samostatně zpracovaných sousloví, která daný jednoslovný lexém obsahují. U sousloví uvádíme přehled jeho variantních podob, které jsou zpracovány jako samostatné položky databáze, např. *lymeská borelióza / lymeská borelióza; soap opera / soap-opera*, a zachycujeme rovněž derivační vztahy mezi souslovími, např. *dentální hygienista – dentální hygienistka; akrobatické lyžování – akrobatický lyžař – akrobatická lyžařka*. U frazému zachycujeme formou soupisu jednak jeho paradigmatické varianty, jednak jeho ustálené transformace.

Hypertextovými odkazy propojujeme mezi sebou i samostatně zpracovaná tzv. frazeologická kvazihomonyma. Jde o ty lexikální jednotky, které fungují jak ve svém přímém, tak přeneseném významu, tj. jde o výrazy s dvojí denotací. Tyto výrazy jsou v LDB Pralex zpracovány jednak jako sousloví, jednak jako frazémy, např. *Adamo-vo jablko* ‚grapefruit‘ x ‚lidský ohryzek‘, *fata morgana* ‚atmosférický úkaz‘ x ‚klamná, nereálná představa, vidina‘, *advocatus diaboli* (církv.) x ‚člověk v dobrovolné roli obhájce nepopulárního opatření, šetření ap. n. druhého‘²⁶. V LDB Pralex odlišujeme členy těchto dvojic jednak označením typu VLJ, tj. sousloví :: frazém, jednak uvedením arabské číslice v lemmatu (číslici jedna připojujeme k sousloví, číslici dvě k frazému)²⁷.

3.3. Zpracování víceslovných pojmenování (sousloví) v LDB Pralex

Pod **víceslovná pojmenování (sousloví)** zahrnujeme nefrazémové VLJ různého typu. Řadíme sem jak víceslovná neterminologická, tak terminologická pojmenování (víceslovné termíny), víceslovné přejímky a citátové výrazy, dále pak víceslovná propria²⁸. K víceslovným pojmenováním zařazujeme rovněž jmenná a slovesná multiverbální pojmenování²⁹. K souslovím řadíme i víceslovnou (grafickou, pravopisnou) variantu v případě kolísání lexému mezi jednoslovností a víceslovností (jde zejména o přejímky).

V LDB Pralex u sousloví nevycházíme na rozdíl od frazémů z žádného předem stanoveného souboru³⁰. Při výběru a zpracování sousloví se proto opíráme o naše

26 Viz SČFI, Výrazy neslovesné, heslo *ďáblův advokát*.

27 Na rozdíl od homonym, která v rámci LDB Pralex označujeme římskými číslicemi.

28 Víceslovná propria je třeba řešit v kontextu zpracování proprií jako celku. Zde se jimi dále nezabýváme.

29 Toto řazení je v souladu s pojetím J. Filipce v České lexikologii (Filipec – Čermák 1985). K nefrazémovým jednotkám řadí jmenná multiverbální spojení Martinová (2005), jmenná a slovesná multiverbální spojení (resp. jmenná a slovesná spojení s kategoriálním slovem) Jarošová (2000a, 2000b). Naproti tomu v Čermákově frazeologické koncepci jsou verbonominální spojení s abstraktem řazena ke kvazifrazémům (viz část Frazeologie a idiomatika in Filipce – Čermák 1985, dále např. Čermák 2007). V mluvnici Čeština – řeč a jazyk jsou multiverbální spojení (podle příkladů slovesná) vyčleněna jako speciální skupina v rámci pojmenování majících víceslovnou formu.

30 U frazémů je tomu jinak, srov. výchozí soubor pro přirovnání, vytvořený na základě SSJČ (blíže viz zde příspěvek E. Kolovecké). Nicméně v pozdější fázi zpracování v LDB Pralex nebo při samotném zpracování hesel v budoucím výkladovém slovníku češtiny je možné u některých typů víceslovných pojmenování o takovýchto soupisech uvažovat. Jde zejména o víceslovné výrazy cizího původu (jejich východiskový heslář by bylo možné stanovit na základě jejich samostat-

databázové nebo dřívější slovníkové zpracování jejich jednoslovných komponentů. Vzhledem k tomu, že se základní zpracování hesel v LDB Pralex v letech 2007–2010 zaměřuje především na substantiva, jsou našimi současnými východisky pro zpracování sousloví a) korpusové doklady vybrané při zpracování daného substantiva, b) předchozí slovníkové zpracování tohoto substantiva (a to především zpracování hesla v SSJČ; při řešení otázky samostatného zpracování daného sousloví – včetně výkladu významu – vstupuje do analýzy rovněž i zpracování daného spojení pod dalšími komponenty v SSJČ a jeho zpracování v dalších výkladových slovnících, ojedinele pak jeho ověřování i v dalších slovníkových a encyklopedických či materiálových zdrojích).

Základní způsob zpracování sousloví v LDB Pralex je (jak jsme uvedli výše) primárně založen na lexikografickém kritériu „potřebnosti výkladu“ vycházejícím z nesusoučtovosti významu jednotlivých komponentů. (Někdy se k tomuto kritériu připojují i hlediska další, zejména zásady zpracování variant v LDB Pralex, viz dále.) Část lexémů z rozsáhlé množiny sousloví je tedy na základě zmíněného lexikografického kritéria zpracována jako samostatná hesla, část je uvedena v exemplifikaci u jednoslovných hesel a část (zatím nerozhodnutých případů) je uváděna v bloku Návrhy na samostatné zpracování VLJ u jednoslovných hesel.

V LDB Pralex na základě kritéria potřebnosti výkladu jako samostatná hesla zpracováváme:

1) Terminologická i neterminologická **víceslovná pojmenování** obrazného i neobrazného typu, např. *stolní tenis, lehká atletika, absolutní bod mrazu, absolutní nula, cedulová banka, elektronická peněženka, Morseova abeceda, barnumská reklama, lymeská borelióza, rajské jablko, černá díra, černá skříňka, bílá káva, černý kašel, bílá technika, černá technika, hnědá technika, černý šerif, turecký med, svatojánská muška, zlatý dešť, australský čajovník / čajový strom / čajovníkový strom* („kajeput střídavolistý“), *hvězdicový anýz* („badyán“), *židovský brambor* („topinambur“), *vousy svatého Ivana* („kavyl Ivanův“).

Jedním z problémů tohoto řešení bylo, jakým způsobem podat dvouslovná terminologická (nomenklaturní) pojmenování rostlin a živočichů se strukturou S-S_{Nom'}, např. *babočka admirál, tetřev hlušec, hadilov pisař; trnovník akát, prha arnika, artyčok ostník, ryzec černohlávek, čirůvka májovka*. Nejprve jsme tyto názvy uváděli jen v exemplifikaci u substantiv³¹. Avšak vzhledem k tomu, že se u uvedených botanických a zoologických nomenklaturních názvů nemůžeme opřít o součtovost jednotlivých komponentů – substantivum fungující jako substantivní přívlástek je v daném spojení monokolokabilní (např. *artyčok ostník*), případně první či druhý komponent funguje jako jednoslovné synonymum daného sousloví (např. *bedrník anýz :: anýz, trnovník akát :: akát :: trnovník*), rozhodli jsme se v takovém případě tyto jednotky zpracovávat v LDB Pralex jako samostatná hesla.

ného zpracování v dosavadních výkladových slovnících) a o víceslovné termíny. Ty se budou moci spolu s jednoslovnými termíny doplňovat do hesláře systematicky na základě uvážlivého výběru z terminologických slovníků, encyklopedií apod.

31 Příklady tohoto typu jsme uváděli v bloku nazvaném Jiná spojitelnost. K zásadám zpracování exemplifikace u substantiv viz zde podrobněji ve stati M. Koutové.

Při zpracování sousloví se však ukazuje, že ne vždy je uvedené lexikografické kritérium potřebnosti výkladu jednoznačné či dostačující. Např. u sousloví *přízemní ozon*, *extrémní sport* (*extrémní sporty*), *elektronické bankovníctví* aj. můžeme být na pochybách, zda význam komponentů dostatečně pokrývá význam daného spojení, a tedy zda tyto výrazy zpracovat, či nezpracovat samostatně³². Tuto problematiku bude třeba plně dořešit v rámci koncepce budoucího výkladového slovníku, a to jak z hlediska zpracování VLJ, tak z hlediska zpracování terminologie obecně, neboť v mnoha případech jde zároveň o terminologická či terminologizovaná spojení. Nicméně vzhledem k tomu, že LDB Pralex slouží jako materiálová báze pro budoucí výkladový slovník, máme dvojí možnost, jak tyto problémové případy v LDB Pralex podchytit. V případě, že se na základě dostupných informací rozhodneme dané sousloví nezpracovávat jako samostatné heslo, ale přesto se nadále domníváme, že jde o „kandidáta“ na samostatné zpracování, zapisujeme toto sousloví do bloku Návrhy na samostatné zpracování³³ v exemplifikaci u jednoslovného hesla (komponentu). Pokud se rozhodneme danou VLJ zpracovat samostatně, můžeme naopak případné pochybnosti o platnosti samostatného zpracování uvést do tzv. Poznámky autora u daného hesla.

2) Jako samostatná hesla zpracováváme v LDB Pralex **víceslovné výrazy cizího původu**, např. *direct banking*, *heavy metal*, *science fiction*, *chargé d'affaires*, *curriculum vitae*, *corpus delicti*, *alma mater*, *ex offio*, *a priori*, *a posteriori*.

3) **Derivovaná víceslovná pojmenování**³⁴ k VLJ zpracovaným samostatně podle bodu 1 uvádíme rovněž jako další samostatná hesla, protože u nich také platí kritérium nesoučtovosti významu komponentů, tedy např. *adrenalinový sport* → *adrenalinový sportovec*, *stolní tenis* → *stolní tenista* → *stolní tenistka*.

Vzhledem k obecné zásadě o samostatném zpracování variant v LDB Pralex zpracováváme formou samostatných databázových položek (hesel) i **variantní podoby sousloví**.

Může jít **a) o hláskové aj. varianty**, např. *lymeská borelióza* / *lymská borelióza*, *svatojanská muška* / *svatojánská muška*, *dentální hygienička* / *dentální hygienistka*. Jiným způsobem jsou zpracovány syntagmatické varianty sousloví. Ty nezpracováváme samostatně, ale – obdobně jako u frazémů – v rámci jednoho hesla. Fakultativní komponenty jsou uváděny v lemmatu v kulatých závorkách (např. *Lormova (doteková) abeceda*); exemplifikace pro jednotlivé syntagmatické varianty je zachycena v samostatných exemplifikačních blocích.

b) Druhým případem, kdy varianty zachycujeme formou samostatně zpracovaných hesel, je zpracování víceslovné pravopisné/grafické varianty k jednoslovnému heslu, např. *home banking* k *homebanking*, *soap opera* k *soap-opera*, *ping pong* k *ping-pong*.

32 Kritérium „potřebnosti výkladu“ je navíc konfrontováno se způsobem zpracování synonymních pojmenování. Např. dospějeme-li k závěru, že výrazy *elektronické bankovníctví*, *elektronický banking* budou uvedeny na základě součtovosti výkladu významu v exemplifikaci u jejich komponentů, tj. v heslech *elektronický*, *bankovníctví*, *banking*, pak naproti tomu synonymní výrazy *e-bankovníctví*, *e-banking* budou zpracovány formou samostatných hesel.

33 Podrobněji k exemplifikačnímu bloku „Návrhy na samostatné zpracování“ viz dále.

34 V pojetí derivovaných sousloví vycházíme ze studie J. Horeckého (1957) a O. Martincové (2000, 2005).

Naproti tomu pouze v **exemplifikaci jednoslovného hesla** (tj. u komponentu sousloví) a zatím bez speciálního odlišení od jiných spojení uvádíme:

1) **sousloví**, pro která neplatí kritérium potřebnosti výkladu, např. *nákladní auto*, *osobní auto*, *terénní auto*, *závodní auto* – u hesla *auto*; *konvalinka vonná* – u *konvalinka*.

2) **jmenné multiverbální jednotky, slovesné multiverbální jednotky**, případně i některá další **neobrazná verbonominální spojení**, např. *učitelská obec*; *proces globalizace*; *oblast energetiky*; *provádět analýzu*, *provést analýzu*, *podrobit analýze*; *zaměřit pozornost na něco*, *věnovat pozornost něčemu*, *odvrátit pozornost od něčeho*, *budit pozornost*; *získat souhlas*, *dát souhlas*, *vyslovit souhlas*. Nicméně vzhledem ke specifickému charakteru těchto spojení (spojení stojící na periférii sousloví a frazémů³⁵; sémantická „průhlednost“ spojení; sémantická vyprázdňenost řídicího členu; částečná synonymičnost víceslovného a jednoslovného výrazu) můžeme v budoucím výkladovém slovníku uvažovat o odlišení tohoto typu spojení od spojení jiných, zejména pokud jde o nejvýraznější typy V-S, A-S. Určitým vodítkem pro zpracování v tomto duchu nám mohou být specifické způsoby zpracování těchto jednotek v některých českých i slovenských slovnících (SN 1, SN 2, SČFI; Krátky slovník slovenského jazyka, SSSJ).

V LDB Pralex počítáme též s pracovním lexikograficko-technickým řešením pro ta sousloví, u nichž předpokládáme „potřebnost výkladu“, ale nemáme k dispozici potřebné údaje (např. výklad významu v LDB Pralex u zatím nezpracovaného komponentu), nebo u nichž váháme s rozhodnutím, zda je samostatně zpracovat, či nikoli. Taková **sousloví** uvádíme jako **návrhy na samostatné zpracování** do speciálního bloku exemplifikace u jednoslovného hesla; u substantiv jde o blok E12. Týmž způsobem zapisujeme do bloku E12 i návrhy na zpracování těch frazémů, které zatím nejsou v hesláři jako samostatná hesla uvedeny. Blok určený pro zápis lexémů navržených na samostatné zpracování je sice po formální stránce začleněn do exemplifikace u daného heslového významu substantiva, po stránce obsahové a funkční však stojí mimo ni. Z hlediska lexikografa-zpracovatele jde o blok čistě pracovní a uvedení určité lexikální jednotky v něm je pouze dočasné. Po vyhodnocení buď danou víceslovnou lexikální jednotku zpracujeme jako samostatné heslo, nebo spojení zapíšeme do příslušného exemplifikačního bloku u substantiva podle syntaktické spojitelnosti.

Ve srovnání s výkladovými slovníky SSJČ, SSČ a NASCS je způsob zpracování sousloví (i VLJ vůbec) v LDB Pralex do jisté míry jednodušší – sousloví zpracováváme buď jako samostatná hesla, nebo jsou uvedena v exemplifikaci u jednoslovného komponentu. V LDB Pralex tedy nezpracováváme sousloví (a obdobně i frazémy) v podobě jednoho z významů heslové stati. Sousloví či frazémy, které byly v SSJČ zpracovány takto, tj. jako samostatný význam jednoslovného hesla, převádíme v LDB Pralex do podoby samostatných víceslovných hesel. Protože však u hesel ze SSJČ jednotlivé významy ze struktury jednoslovného hesla nevypouštíme, v LDB Pralex je u takového významu na původním místě poznamenáno „zpracováno jako samostatné heslo“ (srov. zpracování u hesel *bob I* – 4. význam: *vlčí bob*; *cibule* – 3. význam: *mořská*

35 Srov. zpracování verbonominálních spojení s abstraktem v SČFI. K problematice frazeologické platnosti verbonominálních spojení viz např. Čermák (2007), Pisárčiková (1977), Kralčák (2003), Mlacek (2007, s. 60–61).

cibule; chléb – 5. význam: *svatojánský chléb*). Zároveň je však zpracování sousloví, která splňují lexikografickou podmínku „potřebnosti výkladu“, v LDB Pralex oproti SSJČ, SSČ a NASCS bohatší a podrobnější, protože jsou zpracovávána jako samostatná hesla s vlastním výkladem významu i exemplifikací³⁶. Základní struktura exemplifikace u sousloví je v databázi rozdělena do tří exemplifikačních bloků s názvem Větné doklady, Upravené doklady a V názvech³⁷. Do exemplifikace je možné podle potřeby přidávat další exemplifikační bloky; při zápisu dokladů do databáze však dodržujeme zásadu, aby byly upravené a větné doklady od sebe odděleny, tj. aby nebyly uvedeny v témž exemplifikačním bloku.

Databázové zpracování nám navíc dovoluje pracovat s hypertextovými odkazy a zdůraznit tak různé formální i systémové vztahy (u sousloví např. mezi samostatně zpracovanými souslovími a jejich jednoslovnými komponenty, mezi motivující/základovou VLJ a od ní derivovaným víceslovným pojmenováním, mezi frazeologickými kvazihomonymy navzájem; další hypertextové provázání umožňuje i zápis synonymem, příp. antonym do příslušných seznamů).

Literatura

Akademický slovník cizích slov. 2 sv. Red. V. Kraus – J. Petráčková. Praha: Academia 1995.

BEČKA, Josef: Slovesná rčení a sousloví I. In: Naše řeč, 1934, roč. 18, č. 1, s. 8–17.

BEČKA, Josef: Slovesná rčení a sousloví II. In: Naše řeč, 1934, roč. 18, č. 2, s. 38–44.

BOZDĚCHOVÁ, Ivana: Současná terminologie (se zaměřením na kolokační termíny z lékařství). Praha: Karolinum 2009. 270 s.

Čeština – řeč a jazyk. Ed. M. Čechová. 2., přepracované vyd. Praha: ISV nakladatelství 2000. 407 s.

ČERMÁK, František: Frazeografie. In: Manuál lexikografie, Eds. F. Čermák – R. Blatná. Jinočany: H&H 1995, s. 116–136.

ČERMÁK, František: Jazyk a jazykověda. 3., doplněné vydání. Praha: Karolinum 2001a. 341 s.

ČERMÁK, František: Syntagmatika slovníku: typy lexikálních kombinací. In: Čeština – univerzália a specifika 3. Eds. Z. Hladká – P. Karlík. Brno: Masarykova univerzita 2001b, s. 223–232.

ČERMÁK, František: Kolokace v lingvistice. In: Kolokace. Studie z korpusové lingvistiky 2. Eds. F. Čermák – M. Šulc. Praha: Nakladatelství Lidové noviny 2006, s. 9–16.

ČERMÁK, František: Frazeologie a idiomatika česká a obecná. Praha: Karolinum 2007. 718 s.

ČERMÁK, František: Lexikon a sémantika. Praha: Nakladatelství Lidové noviny 2010. 357 s.

DOLNÍK, Juraj: Jazykové principy vo výstavbe frazém. In: Frazeologické štúdie II. Ed. P. Ďurčo. Bratislava: Esprima 1997, s. 36–44.

36 Obdobně je tomu u frazémů, viz příspěvek E. Kolovecké.

37 K charakteristice jednotlivých typů exemplifikace obecně a u substantiv viz příspěvek M. Koutové.

- Encyklopedický slovník češtiny. Eds. P. Karlík – M. Nekula – E. Pleskalová. Praha: Nakladatelství Lidové noviny 2002. 604 s.
- FILIPEC, Josef: Pojetí formy v slovní zásobě a některé další otázky s tím spojené. In: Obsah a forma v slovní zásobě. Materiály z vedeckej konferencie o výskume a opise slovní zásoby slovenčiny (Smolenice 1.–4. marca 1983). Ed. J. Kačala. Bratislava: Jazykovedný ústav Ľudovíta Štúra SAV 1984, s. 138–141.
- FILIPEC, Josef – ČERMÁK, František: Česká lexikologie. Praha: Academia 1985. 281 s.
- HAUSER, Přemysl: Nauka o slovní zásobě. 2. vyd. Praha: SPN 1986. 195 s.
- HAVRÁNEK, Bohuslav – JEDLIČKA, Alois: Česká mluvnice. 2., upravené a rozšířené vydání. Praha: Státní pedagogické nakladatelství 1963. 561 s.
- HORECKÝ, Ján: Dve osobitosti v menách osôb. In: Slovenská reč, 1957, roč. 22, č. 6, s. 352–358.
- HORECKÝ, Ján: Návrh na vymedzenie frazémy. In: Frazeologické štúdie II. Ed. P. Ďurčo. Bratislava: Esprima 1997, s. 78–81.
- JANOVEC, Ladislav: Nevětné frazémy v nové slovní zásobě In: Neologizmy v dnešní češtině. Eds. O. Martincová – A. Rangelova – J. Světlá. Praha: Ústav pro jazyk český AV ČR 2005, s. 147–158.
- JAROŠOVÁ, Alexandra: Lexikalizované spojenie v kontexte ustálených spojení. In: Princípy jazyka a textu. Materiály z medzinárodnej vedeckej konferencie konanej 9.–10. 3. 2000 na Katedre slovenského jazyka Filozofickej fakulty Univerzity Komenského Ed. J. Dolník. Bratislava: Univerzita Komenského 2000a, s. 138–153.
- JAROŠOVÁ, Alexandra: Spracovanie ustálených spojení vo výkladovom slovníku. Návrh. In: Nová slovní zásoba ve výkladových slovníkách. Sborník z příspěvků z konference (Praha, 31. 10. – 1. 11. 2000). Eds. O. Martincová – J. Světlá. Praha: Ústav pro jazyk český AV ČR 2000b, s. 43–54.
- JAROŠOVÁ, Alexandra: Viacslovný termín a lexikalizované spojenie. In: Človek a jeho jazyk 1. Jazyk ako fenomén kultúry. Na počesť profesora Jána Horeckého. Ed. K. Buzássyová. Bratislava: Veda 2000c, s. 481–493.
- JAROŠOVÁ, Alexandra: Obrazná terminológia alebo odborná frazeológia? Ako riešime túto a podobné otázky v pripravovanom Slovníku súčasného slovenského jazyka. In: Frazeologické štúdie IV. Eds. M. Jankovičová – J. Mlacek – J. Skladaná. Bratislava: Veda 2005, s. 131–145.
- JAROŠOVÁ, Alexandra: Problém vymedzenia kolokácií. In: Jazykovedný časopis, 2007, roč. 58, č. 2, s. 81–102.
- Kolokace. Studie z korpusové lingvistiky 2. Eds. F. Čermák – M. Šulc. Praha: Nakladatelství Lidové noviny 2006. 452 s.
- KRALČÁK, Lubomír: K otázke hraníc medzi frazémami a analytickými verbonominálnymi spojeniami. In: Frazeologické štúdie III. Eds. J. Mlacek – P. Ďurčo. Bratislava: Stimul 2003, s. 116–123.
- Krátky slovník slovenského jazyka. Red. J. Kačala – M. Pisárčiková – M. Považaj. 4., doplněné a upravené vyd. Bratislava: Veda 2003. 985 s.
- KUCHAŘ, Jaroslav: Základní rysy struktur pojmenování. In: Slovo a slovesnost, 1963, roč. 24, s. 105–114.
- LOTKO, Edvard: Slovník lingvistických termínů pro filology. 1. vyd. Olomouc: Univerzita Palackého v Olomouci 1998. 113 s.

- MACHAČ, Jaroslav: K lexikologické problematice slovních spojení. In: Slovo a slovesnost, 1967, roč. 28, č. 2, s. 137–149.
- MACHAČ, Jaroslav: Slovní spojení (v jednojazyčném výkladovém slovníku). (rkp., nedatováno, 10 stran)
- MAN, Oldřich: Ustálená spojení a frazeologické jednotky, jejich podstata a hranice. In: Lexikografický sborník. Materiály z I. celoštátní konferencie čs. lexikografů, konané v dnech 5.–7. júna 1952 v Bratislave. Ed. Š. Peciar. Bratislava: Vydavateľstvo Slovenskej Akadémie vied 1953, s. 101–110.
- Manuál slovníku neologizmů. Praha: Ústav pro jazyk český AV ČR 1996. Interní tisk. 55 s.
- MARTINCOVÁ, Olga: Derivatívnosť u nových víceslovných pojmenovaní. In: Človek a jeho jazyk 1. Jazyk ako fenomén kultúry. Na počesť profesora Jána Horeckého. Ed. K. Buzássyová. Bratislava: Veda 2000, s. 382–387.
- MARTINCOVÁ, Olga: Nová víceslovná pojmenovaní. In: Neologizmy v dnešnej češtině. Ed. O. Martincová – A. Rangelova – J. Světlá. Praha: Ústav pro jazyk český AV ČR 2005, s. 212–231.
- MLACEK, Jozef: Slovenská frazeológia. 2., doplnené vydanie. Bratislava: Slovenské pedagogické nakladateľstvo 1984. 156 s.
- MLACEK, Jozef: Štúdie a state o frazeológii. Ružomberok: Katolícka univerzita v Ružomberku 2007. 375 s.
- Nová slova v češtině 1. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 1998. 356 s. (SN 1)
- Nová slova v češtině 2. Slovník neologizmů. (O. Martincová a kol.) Praha: Academia 2004. 568 s. (SN 2)
- Nový akademický slovník cizích slov A–Ž. (J. Kraus a kol.) 1. vyd. Praha: Academia 2005. 880 s. (NASCS)
- PISÁRČIKOVÁ, Mária: Synonymické vzťahy pri ustálených spojeniach typu dat súhlas. In: Slovenská reč, 1977, roč. 42, č. 3, s. 144–153.
- Příruční mluvnice češtiny. Eds. P. Karlík – M. Nekula – Z. Rusínová. Praha: Nakladatelství Lidové noviny 1995. 800 s.
- Slovník české frazeologie a idiomatiky. 4 sv. Red. F. Čermák a kol. 2. přepracované a doplněné vyd. Voznice: LEDA 2009. (SČFI)
- Slovník slovanské lingvistické terminologie. 2 sv. Red. A. Jedlička. Praha, Academia 1977–1979.
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek. 2., nezměněné vyd. Praha: Academia 1989. (SSJČ)
- Slovník súčasného slovenského jazyka. A–G. Red. K. Buzássyová – A. Jarošová. Bratislava: Veda 2006. 1134 s. (SSSJ)
- SOCHOVÁ, Zdeňka – POŠTOLKOVÁ, Běla: Co v slovnících nenajdete. Novinky v současné slovní zásobě. Praha: Portál 1994. 204 s. (CSN)
- ŠMILAUER, Vladimír: Nauka o českém jazyku. 6. vyd. Praha: Státní pedagogické nakladatelství 1982. 335 s.

Typologie abreviačních útvarů v lexikální databázi Pralex¹

Albena Rangelova

Ústav pro jazyk český AV ČR, v. v. i.

Typology of Abbreviated Forms in the Pralex Lexical Database

The paper deals with abbreviations, i.e. shortenings, acronyms and abbreviated expressions; it interprets and internally divides this special group of lexical units. The chosen typology was applied for a development of the database Praled. Following the particular types of units, the author monitors possible processing ways of various types of abbreviated forms, while building the Pralex lexical database.

1. Úvod

Při vývoji lexikální databáze (dále LDB) Praled byly zkratky, značky a zkratkové útvary různého typu zahrnuty do jednoho souhrnného výchozího souboru, jehož předběžná analýza nás vedla k vymezení centrální části skupiny (typ Zkratky/značky) a jevy přechodové, které byly zpracovány v rámci jiných typů dle své podstaty. Návrhy na zpracování tohoto typu lexikálních jednotek vycházejí z analýz jejich lexikologického pojetí a z rozborů jejich lexikografického ztvárnění ve výkladových slovnících všeobecných i speciálních – *Slovník spisovného jazyka českého* (SSJČ), *Slovník spisovné češtiny pro školu a veřejnost* (SSČ), *Akademický slovník cizích slov* (ASCS), *Nová slova v češtině. Slovník neologizmů 1* (SN 1), *Nová slova v češtině. Slovník neologizmů 2* (SN 2), pro srovnání bylo přihlédnuto též k specializovaným příručkám – slovník *Zkratky, značky, akronymy* (Kos 1983) a *Slovník zkratek* (dále SZ ED)².

2. Lexikologické pojetí zkratek a zkratkových útvarů

Pojetí zkratky v **českých lexikologických pracích** se shoduje v celkovém teoretickém vymezení; liší se v různých aspektech vnitřního třídění. Zkratky jsou vymezovány jako zvláštní lexikální útvary, sekundární pojmenování, která existují na pozadí pojmenování plných; jejich základní strukturální rys je fragmentárnost (viz níže), přičemž platí, že čím menší je závislost zkratky na plném pojmenování, tím pevněji se zkratka začleňuje do lexikálního systému a tím více ztrácí charakter sekundárního prostředku (zástupky). Závislost na plném znění se může částečně až úplně ztratit (*rodokaps* ← román do kapsy, *karma* ← Karel Macháček apod.).

V české jazykovědné tradici jsou zkratky a značky vnímány jako jedna kategoriální skupina, i když jsou důsledně rozlišovány. Jako zastřešující se uplatňují ozna-

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. Karel Oliva, Dr.

2 Tyto publikace byly zvoleny jako základní referenční zdroje pro zpracování zkratek a značek v LDB Pralex.

čení abreviační útvar, méně často abreviatura, nejčastěji ovšem zkratka, pro případy, kdy je zapotřebí souhrnně pojmenovat jednotky tradičně rozlišované (např. zkratky, značky, zkratková slova, popř. slova od zkratek odvozená). Jak již bylo řečeno, pojetí zkratky v české tradici je poměrně jednotné, uplatňuje se kombinace procesuálního a rezultativního přístupu. I když jsou různé definice jemně odlišné, shodují se v tom, že zkracování je okrajový způsob tvoření slov (Hauser 1980, s. 155), že tvoření zkratek stojí na okraji slovo tvorby a zasahuje do ní vytvářením zkratkových slov (MČ 1³, 1986, s. 517). Nejužší pojetí je prezentováno v PMČ 1995⁴, s. 224), kde se z kategorie zkratek vylučují tzv. zkratky textové (*apod., např. aj.*) a standardizované značky různých oborů (*kg* – kilogram, *A* – ampér). Z jiného pohledu lze zkratky vymezit negativně jako neslova (v PMČ, s. 224, jsou takto označeny pouze zkratky iniciálové).

Pojetí zkratek v dřívější literatuře podává Hrbáček (1979, s. 9–16)⁵, výstižná je též jeho celková charakteristika zkratek (toto pojmenování užívá jako zastřešující označení; vedle toho též abreviační útvar, abreviatura). Za hlavní rys zkratek považuje fragmentárnost; podle povahy fragmentů je dále vnitřně člení. Je-li fragment pouze písemným faktem, jde o zkratky grafické – *sl.* (= slečna)⁶, *soudr.* (= soudruh, soudružka), *s. / str.* (= stránka); je-li fragment též zvukovým faktem, jde o grafiko-fónické zkratky – podle povahy fragmentů pak lze tuto skupinu rozdělit na a) iniciálové zkratky (*JZD* = jednotné zemědělské družstvo) a b) větší než iniciálové („morfémové“) – *PeDF* (= pedagogická fakulta), *MěNV* (= městský národní výbor); sem patří též zkrácená slova – *limo* (= limonáda) a zkratková slova – *Čedok* (= Česká dopravní kancelář)⁷.

MČ 1 věnuje zkratkám pozornost v rámci výkladu o slovo tvorbě: „Při tvoření zkratek [...] nastává tzv. krácení mechanické (na rozdíl od krácení sémantického, jak se obvykle označuje elipsa a univerbizace). [...] slova nebo víceslovné výrazy se redukuje na některé hlásky (písmena) nebo skupiny hlásek (písmen), které svým nápadným složením nebo pravopisem a současně vžitou podobou signalizují v povědomí adresáta projevu význam nezkráceného výrazu...“ (MČ 1, s. 517). Zkratka a plná podoba mají tedy v principu stejný význam; zkratka je sekundární znak, který je z hlediska slovo tvorného fundován plným slovem, výrazem (souslovím, vlastním jménem atp.).

Podobnými rysy se vyznačují též značky, i když mají řadu specifických rysů. Původně byl výraz *značka* vyhrazen jenom pro ideogramy (např. %, §): „... značka se

3 Část F. *Tvoření zkratek a zkratkových slov*, autor M. Churavý, s. 517–526.

4 Část *Tvoření zkratek a zkratkových slov*, tvoření zkracování, s. 224–225.

5 Autor se zkratkami v češtině zabýval soustavně (viz Hrbáček 1967, 1968), zde odkazujeme ke kapitole *Dosavadní zpracování problematiky jazykových zkratek* z jeho monografie (1979).

6 Pro označení korelace mezi a zkratkou a jejím plným zněním užíváme znaménko „=“, i když jsme si vědomi, že nejde nutně o vztah ekvivalence.

7 Hrbáček vyčleňuje jednu širší kategorii, neodděluje zkratky (jen) fónické jako MČ 1. Ve své starší práci (Hrbáček 1968) vede autor hranici mezi zkratkami (iniciálovými a zkratkovými slovy) a zkrácenými slovy: zkracování slov vidí jako operaci technickou, která nezasahuje do struktury lexika, kdežto zkratky vznikají v procesu univerbizace, tj. vždy na pozadí víceslovného pojmenování (a jen těmi se zabývá), nicméně ve své monografii (Hrbáček 1979) již zkrácená slova do obecného třídění zkratkových/abreviačních útvarů zařazuje.

neskládá z písmen, na př. §, ♂.“ (Cejp 1940, s. 39). Dnes se uznává existence značek písmenných, jež jsou pokládány za součást grafických zkratk (MČ 1, s. 518).

Speciální pozornost **vztahu značky a zkratky** věnuje Hrbáček (1979, s. 19–21), a to na základě shrnutí všeobecně uznávaných rysů značek: a) značka je druh jazykového znaku s neprůhlednou motivací; b) značka je znak buď jazykový, nebo neязыkový, se zvýrazněnou funkcí označování, k čemuž bývá formálně přizpůsobován.

Vztah zkratka – značka je vystižen následovně: „Jazykové značky jsou ve velké většině zkratky, které mají povahu samostatných znaků, tj. nemají funkci jen zkracovat slova, nýbrž krátkou výraznou formou vyjadřují pojem.“ (Hrbáček 1979, s. 20). Dle jeho názoru zkratka je pojem strukturní, kdežto značka pojem funkční, a z toho plyne, že zkratky mohou vystupovat ve funkci značky; u značek je zvýrazněna funkce označovací, grafický rozdíl se projevuje tím, že značka je bez tečky (srov. též PČP, s. 36–37). Pravděpodobně na základě této funkční charakteristiky řadí někteří lingvisté mezi značky také zkratky iniciálové: „Za značky považujeme i tzv. iniciálové zkratky“ (Havránek – Jedlička 1981, s. 74).

Věcné okruhy, ke kterým nejčastěji zkratky náležejí, jsou různorodé: názvy a kódy států, organizací, institucí, vědecké a jiné tituly; názvy svátků a akcí atp. Oblastí užití, bez rozlišení zkratk a značek uvádí jako první Cejp (1940) – obchodní, reklamní, inzertní; vědecké (podle oborů), úřední (vojenské, dopravní, akademické); sportovní. Z dnešního pohledu jde o následující oblasti:

- tituly a zkratky u podpisů (*doc., Ing. / ing., Mgr. / mgr.*);
- zkratky společenského styku (*p., s., br.*);
- zkratky jmen podniků, institucí, politických stran (*AV ČR, FF UK*);
- zkratky jmen států, územních celků/částí (*ČR, MLR*);
- zkratky jmen paktů, smluv, mezinárodních spolků (*ASEAN, G7*);
- zkratky jmen osobních – iniciály, monogram, šifra (*TGM*);
- české názvy svátků, akcí, výročí (*MDŽ*);
- značky v širokém smyslu oborové⁸: měř, měnových jednotek, značky odborné (např. matematické a fyzikální veličiny, chemické prvky a symboly, hudební značky, astronomická tělesa a seskupení atp.);
- značky písmenné a ideogramy (platící i mimo jazykové projevy): poznávací značky států/uskupení na dopravních prostředcích (auta, vlaky, letadla), kódy měn, výstražné dopravní značky, některé orientační značky (srov. zde pozn. 14).

Gramatické začlenění⁹ zkratk do textů probíhá zpravidla podle kategorií operného členu nezkráceného výrazu či nezkráceného slova. Velmi často mají zkratky substantivní platnost, nověji adjektivní (*aids poradna*), ale vyskytuje se též platnost příslovečná – *p* (= piano). U zkratkových slov začlenění do slovnědruhových kategorií probíhá buď pouhým zeslovněním (viz níže u Dolníka 2003, též PMČ, s. 224, srov. též starší práci Helcl 1949), tj. morfologické zformování probíhá dle fonetické realizace zkratky: *AIDS / aids* → *AIDSu / aidsu*, *DIK / dik* → *DIKové / dikové*, *DJ i dýdžej* → *Dje i dýdžeje*. Druhá možnost zeslovnění zkratky je použití gramatikalizačních pří-

8 Srov. Pohnert 1973, Michajlišin 1999 aj.

9 V tomto příspěvku se výslovnosti abreviačních útvarů nebudeme zvlášť zabývat, odkazujeme k příspěvkům Romportl 1951, Daneš 1989. Zpracovatelský software Pralex ovšem obsahuje nástroje pro zachycení tohoto aspektu abreviačních útvarů.

pon, např. *esemeska / SMSka / smska, eseróčko / eseročko, Djovat / dýdžejovat* (srov. též Ološtiak 2005, Furdík 2004, Staněk 1995, Vlková 1984 aj.).

Velmi inspirativní pojetí nabízí Dolník 2003, který vidí tvoření zkratk (abreviací) jako dílčí případ jednoho ze způsobů rozvíjení slovní zásoby. Abreviace spolu s univerbizací jsou projevy redukce jako postup přetváření – vedle tvoření a přebírání (viz Dolník 2003, s. 155). Autor sice kombinuje procesuální a rezultativní pohled na zkratkové útvary, ale uplatňuje je samostatně. Podává procesuální definici abreviace: „... pretváranie jedno- alebo viacslovného lexikálneho výrazu na hlásky (písmená), skupinu hlások (písmen), slabiky alebo na kombinácie morfémy s inými segmentmi“ (Dolník 2003, s. 164). Jeho vnitřní typologie je v podstatě stejná jako v české lingvistice – podle povahy fragmentů liší: a) grafické zkratky: *resp.*, *tzv.*, *dr.*; značky: *Cu*, *Ca*, *W*; b) graficko-fónické zkratky: *OSN*, *KDH*, *SAV*, *VŠMU*; c) fónické zkratky čili zkratková slova: *bankomat*, *hifi*, *motorest*.

Samostatně definuje Dolník jednotlivé abreviační postupy – jsou to: a) výběr iniciálních písmen či skupin: *doc.*, *j. č.*, *a. s.*; b) přetváření kontrakcí (stahování): *dr.* (doktor), *fa* (firma); c) přetváření stylizací, tj. výběrem charakteristických konsonantních písmen v jejich původním sledu: *rkp.* (rukopis), *mgr.* (magister), *rtg.* (röntgen); d) supletivismus – matematické znaky (symboly), supletivní abreviace: operátor „krát“ se vyjadřuje písmenem *x*, „plus“ speciálním znakem *+*.

Nejnovější polské práce jsou, co se týče typologie abreviačních útvarů, velice jednotné: rozlišuje se „skrót“ a „skrótowce“, jako zastřešující je užíváno označení „skrócenie“ (Sawicka 1998, Markowski 2006, Jadačka 2006, Müldner-Nieckowski 2007).

„Skrót“ představuje písmeno nebo skupinu písmen (s tečkou či bez), zastupující plný výraz v textu. Tyto jednotky jsou vždy vyslovovány jako plný výraz, výjimky lze pozorovat v běžném jazyce – *itp.* i *itepe* (= i tym podobne); *itd.* i *itede* (= i tak dalej); sem tedy náleží jak zkratky (*doc.* = docent, *p.* = pan, *m.in.* = między innymi), tak značky v českém pojetí (*zł* = złoty). Jako „skrótowce“ jsou označovány takové výrazy, které sestávají z písmen, hlásek, slabik nezkrácených výrazů – mohou fungovat v textu samostatně jako výrazy (mají přečtení jiné, než je plný výraz); jsou děleny dále na „literowce“ – *NBP* [en-be-pe] (= Narodowy Bank Polski), „gloskowce“ – *MEN* [men] (= Ministerstwo Edukacji Narodowej), „grupowce/sylabowce“ – *POLFA* (= Polska Farmacja); vedle toho se vyčleňují útvary smíšené („mieszane“) – *CBOS* [cebos] (= Centrum Badania Opinii Społecznej) a „struktury słoeniowe“: *wideopirat*, *speckomisja* (viz podrobněji Sawicka 1998, Rudnik-Karwatowa 2008).

Společné rysy zkratk a značek dovolují vymezit jednu společnou kategorii: v obou případech jde o sekundární pojmenovací prvek, fungující ve většině případů na pozadí nezkráceného výrazu (pojmenovací jednotky); dalšími rysy jsou stručnost a úspornost, s tím souvisí jejich častá víceznačnost (homonymie). Jako souhrnné označení pro všechny jednotky s těmito rysy budeme v této stati užívat synonymně označení abreviační útvar, abreviatura a zkratka (v širokém smyslu slova). Pro procesuální aspekt tvoření zkratk budeme spolu s MČ 1, Dolníkem 2003, Ološtiakem 2005 užívat označení abreviace, zkracování¹⁰.

10 Trochu jinak užívá tato označení *Encyklopedický slovník češtiny* (2002), viz heslo *Zkratka (abreviace)*, s. 565.

Základní typologické rozlišení uplatňujeme v souladu s MČ 1, Dolníkem 2003, kde je základní rozlišení provedeno podle povahy fragmentů a realizace a užívání těchto jednotek v textu, s podrobnějším dalším členěním podle strukturních vlastností jednotek:

- zkratky čistě grafické (někdy nazývané též textové¹¹) – *např.* (= například), *tj.* (= to jest), *čs.* (= československý), *pp* (= pánové), *pí* (= paní), *kpt.* (= kapitán) atd.; do této skupiny patří též značky – *Kč* (= koruna česká), *mm* (= milimetr);
- zkratky graficko-fónické (iniciálové i kombinované), a to hláskované – *MDŽ* (= Mezinárodní den žen), *MLR* (= Maďarská lidová republika), i akronyma¹² *AMU* (= Akademie muzických umění); vyskytují se též zkratky kombinované (často se jako zkratka nečtou a tím se podobají zkratkám textovým), např. *MěNV / MěstNV* (= Městský národní výbor) atd.;
- zkratky fónické, které jsou v plném smyslu zkratkovými slovy – *alobal*, *limo*, *Máňa*, *Lída*.

3. Zpracování zkratkových útvarů ve výkladových slovnících

V prvním velkém lexikografickém díle – *Příručním slovníku jazyka českého* – jsou zkratkové útvary uváděny výjimečně, a to zejména zkratková slova (*Čedok*, *Četka*). Textové ani iniciálové zkratky se neuvádějí, ba ani ve výkladu slov (např. *kilogram*, *metr*, *litr*) se nepřipomíná jejich značka.

SSJČ je mnohem „všímavější“ ke zkratkám a abreviačním útvarům vůbec – některé jsou zařazeny do hesláře, některé se uvádějí v heslové stati. Ustálené zkratky textové se zpracovávají samostatně (*aj.*, *atd.*) anebo jako významy u jednoho ze svých komponentů, např. *t.m.*, *m.p.* jsou zařazeny jako významy zkratky *m.*: [...] 2. t. m. tohoto měsíce 3. m. p. manu propria [-nú -iá] (lat.) vlastní rukou (za podpisem na opisech, vyhláškách ap.); iniciálové zkratky (*ČTK*, *ROH*, *SEATO*), iniciálová zkratková slova (*radar*) a značky (*km*, *mA*, římské číslice) uvádí SSJČ jako samostatná hesla nebo samostatné významy. Dále některé ustálené značky (*kA*, *kg*, *kgm*, *kHz*, *Cal*, *km*, *kV*, *kW*, *kWh*) připomíná za výkladem příslušných nezkrácených slov. Značky chemických prvků následují po výkladu (např. u slova *chlór* ... dusivý jedovatý plynný prvek ze skupiny halogenů (zn. Cl), příp. jsou jeho součástí (*astatin* ... nestálý radioaktivní prvek, zn. At, objevený v posledních letech). Některé značky jsou uvedeny jenom v exemplifikaci slova, např. u slova *metr* za spojením: m. délkový (m), čtvereční (m²), krychlový (m³).

V SSČ je uplatněno nové řešení – samostatný seznam zkratek a značek (cca 650 jednotek), v němž je výběr nejužívanějších textových a iniciálových zkratek, domácích i cizích, zkratky názvů států, mezinárodních i domácích organizací a institucí, zpravodajských agentur a politických stran, často se objevujících v našem denním

11 Viz Filipec (bez vročení), SSČ, s. 597 aj.

12 Označení akronyma používáme v užším smyslu (viz i Horecký 1980) pro dílčí případ iniciálových zkratek, jejichž prvky (iniciály slov plného znění) vytvářejí snadno přečitelný celek (čteme je jako slovo), srov. též Kochová 2003, 2005.

tisku. Zahrnuty jsou značky měn, měrové jednotky platné podle československých norem, značky hudebních pojmů; naopak stranou zůstávají chemické prvky, matematické a fyzikální veličiny, mezinárodní a státní poznávací značky motorových vozidel. Ve výběru jsou zpracována též zkratková slova, která jsou však zařazena do hesláře slovníku.

ASCS postupuje obdobně jako SSIČ, s mnohem podrobnějším propracováním homonymie u zkratk a značek cizího původu – především z odborné sféry (viz příklad zpracování *p, p., P, P.*). Jednotky jsou zařazeny do hesláře, přičemž se uvádějí též znaky z jiných abeced (na jejich místo příslušné podle fonetické realizace znaku v češtině), viz např. **π** [pí] <ř.> mat. zn. pro Ludolfovo číslo – značka je zařazena po poslední jednopísmenné zkratce z řady písmene *p, P* (tj. po **P**. nejč. církl. zkr. pro pátera).

Jako samostatné jednotky v hesláři jsou zařazovány abreviační útvary též v slovnících neologismů, přičemž souvislost mezi zkratkou a jejím plným zněním je signalizována v poznámce u plného znění, např. ze SN 1:

DPH [dépéhá] zkr. ž. i s.

ekon. *daň z přidané hodnoty*: plátci DPH; vyúčtování DPH

daň z přidané hodnoty ekon. *částka vybíraná při každém prodeji zboží, vypočítaná z rozdílu mezi cenou prodávaného zboží a cenou nákupu zboží*: poradenství v oblasti daně z přidané hodnoty

Zkr. *DPH* (viz tam).

Pokud jde o jednotku původem z jiného jazyka, její plné znění se uvádí v poznámce na konci heslové stati, příklad ze SN 2:

HR [há-er, ejč-ár] zkr. <z angl.>

oblast lidských zdrojů: manažeři HR; lidé, kteří pracují v oblasti HR ...

Z angl. *human resources* 'lidské zdroje'.

4. Pojetí abreviačních útvarů pro LDB Pralex

V návaznosti na českou lexikologickou tradici jsme pro potřeby vstupní analýzy materiálu pro LDB Pralex zvolili širší pojetí, tj. abreviační útvary pojímáme jako zvláštní lexikální jednotky, sekundární pojmenování, která existují na pozadí pojmenování plných (viz výše). Čím menší je závislost zkratky na plném pojmenování, tím pevněji se začleňuje do lexikálního systému. Oslabuje se její motivační spojení s plným zněním (tj. závislost na plném znění se může částečně až úplně ztratit). Vztah mezi zkratkou/značkou a plným zněním může být rozvolněný také tím, že se zkratka/značka paralelně vztahuje ke dvěma jednotkám stejného významu, např. *cf. / cfr. (confer / conferatur a srovnej)*, častý je tento jev u chemických značek – *Ca* (kalciium a vápník).

Takovýto široký záběr dovoluje analyzovat celou skupinu jinak kategoriálně různých jednotek, které vznikají při zkracování nebo se zkracováním úzce souvisí. Tento přístup umožňuje provést následný lexikografický popis dle specifik vnitřní kategorizace skupiny a reflektovat přitom rysy společné pro všechny její prvky. Jádro skupiny tvoří jednotky tradičně označované jako zkratky a značky (dle PMČ „neslova“, v intencích základní typologie – zkratky grafické a graficko-fónické).

Pro lexikografický přístup je třeba nejdříve odlišit zkracování a zkratky **čistě operativní** od zkratk **standardizovaných** (uzualizovaných). Zkratky operativní jsou produktem aktuální potřeby časové ekonomie a mají dosah pouze pro daný text. Zkratky standardizované jsou prvky lexikálního systému jako celku, a tudíž mají dosah pro praktický výběr jednotek do hesláře lexikografického díla – v našem případě do hesláře LDB Pralex.

4.1. Modelová stránka – LDB Praled a programování potřebných nástrojů

Při volbě základního/kmenového typu popisované jednotky v LDB Praled, jímž je typ **zkratka/značka** jsme zvolili užší pojetí zahrnující jen jádro skupiny abreviačních útvarů. Při programování formuláře a nástrojů pro lexikografický popis jsme vycházeli z formuláře pro zpracování substantiv¹³, který byl následně modifikován pro specifické potřeby popisu zkratk a značek, zejména má jinou strukturu exemplifikace (viz níže).

Pro vnitřní třídění bylo uplatněno formální hledisko, zvolená typologie reflektuje především utvářenost zkratky (její formální stránku). Lišíme tedy abreviační útvary:

grafické (textové)

- jednosložkové – prosté (= 1 slovo v pozadí): *p., sl., kpt.* (tyto rysy mají i značky *A, At*)
 - složené (= víceslovný výraz nebo složené slovo v pozadí): *apod., tzv.* (tyto rysy mají i některé značky, např. *Ah, kP, mm*)
- vícesložkové: *a. s., arm. gen., př. n. l.*

graficko-fónické

- iniciálové – čtou se dle písmen: *AGB, IQ, SMS / sms, VŠ* (nelexikální u Hrbáčka 1979)
 - čtou se spojitě jako slovo: *SAV, ÚLUV* (akronyma)
- smíšené: *MěstNV*
(oba druhy mohou být součástí vícesložkové zkratky, např. *FF UK, AV ČR*)

fónické – (u Hrbáčka 1979 lexikální zkratky)

- zkratková slova
 - původní: *alobal, Čedok, Spofa*
 - přehodnocená slovnědruhová platnost: *ičo, dik, aids* (substantivní i adjektivní), u nichž proběhlo zeslovnění dle Dolníka 2003
 - modifikovaná: *jipka / JIPka, esemeska / smska / SMSka* (označují tu samou skutečnost, stejný denotát, jsou motivovaná zkratkou, ale přípona *-ka* je jen gramatikalizační)
 - odvozená: *ičovník, socdemácký* (jsou motivována zkratkou, ale s jiným denotátem)
- zkrácená slova: *krimi, limo, deci*

Zatímco pro dosavadní slovníkové zpracování byly relevantní dvě základní distinkce: zkratky vs. značky a zkratkové slovo vs. zkratkové jméno, pro LDB Praled

13 Viz zde stať B. Procházkové a J. Světlé.

byl uplatněn model, který umožňuje adekvátně zpracovat kterýkoliv prvek patřící do široké skupiny abreviačních útvarů – reflektujeme tedy více strukturních a funkčních rysů, aniž bychom aspirovali na plnou reflexi lexikologického zpracování tohoto typu lexikálních jednotek při jejich lexikografické prezentaci.

Řešení pro typ Zkratka/značka uplatněné v LDB Praled je obdobné jako pro typ Podslovní částí, kde po výběru Typ hesla > **podsl.** se aktivuje následně okno **Typ podsl.** U zkratek/značek je postup následující:

Typ hesla > jednosl. / vícesl. / podsl. / **zkratka/značka**

V novém okénku lze vybrat z nabídky **Typ zkratky/značky** následující možnosti:

zkratka (jednoduchá/jednosložková): *apod., pí, doc., RNDr., kpt*

zkratka vícesložková: *a. s., AV ČR*

značka: *Au*

značka složená: *mA* (miliampér), *kg* (kilogram)

první část složené značky: *m-* (*mili-*, např. v *mm* = milimetr), *dk-* (*deka-*, např. v *dkg* = dekagram)

druhá část složené značky (zatím nebyla zaznamenána)

Slovní druh/typ

zkratka v platnosti podst. jm.

zkratka v platnosti příd. jm.

zkratka v platnosti přísl. (*pp* =pianissimo)

zkratka v platnosti slovesa (*cf. / cfr., v. / viz*)

Nástroje pro zpracování a vnitřní třídění zkratek/značek se postupně doladují. Cílem je dosáhnout takové třídění, při kterém by bylo možné rozšířit základní vyhledávání dle typu jednotky (Zkratka/značka) o vyhledávání dle Typ zkratky/značky a výhledově dle jednotlivého slovnědruhového zařazení abreviačních útvarů.

5. Realizační stránka – typologie abreviačních útvarů a jejich zpracování v LDB Pralex

Zpracování abreviačních útvarů v LDB Pralex začalo vytvořením prvotního seznamu jednotek a jejich vložení do programu Praled. Zdrojem pro heslář se stal *Frekvenční slovník češtiny* (1000 zkratek) a SSČ (cca 650 zkratek), k dílčímu doplňování hesláře je využíván průběžně SZ ED.

Jak již bylo zdůrazněno, určujícím (identifikačním) rysem zkratkového útvaru je jeho plné znění. Z tohoto důvodu bylo zvoleno zpracování různých významů odpovídajících rozdílnému plnému znění jako homonyma (na rozdíl od SSJČ a SSČ). Přihlédneme-li k původu jednotky, další distinkce relevantní pro zpracování zkratkových útvarů je rozlišení domácí vs. cizí jednotky – ty potřebují též dešifrování plného znění ve zdrojovém jazyce a informaci o původu. Z tohoto hlediska se při zpracování zkratek/značek vyskytují následující typy zpracovatelských situací:

- a) jednotka nepotřebuje další/jiný výklad než dešifrování plného znění (*ap., atd., pí, fa, dr., př. n. l., a. s., ČR, AV, dj*, též zkrácená slova (*limo, deci*);
- b) jednotka potřebuje dešifrování i výklad: *IČO / ičo, IQ, IT*, zejména pak ty zkratky, které se používají v platnosti adjektiva;

- c) jednotka je cizího původu, potřebuje tudíž dešifrování plného znění ve zdrojovém jazyce: *ecu / ECU, P. F. / pf, PR*;
- d) heslo potřebuje jen výklad, dešifrování je podpůrné (původní zkratková slova) – *alobal, radar, laser*.

Pro zpracování typu jednotek Zkratka/značka byl navržen formulář reflektující specifika zkratk a značek v užším smyslu (tzv. neslov) jako jádra skupiny abreviačních útvarů. Je určen pro zpracování zkratk a značek v užším smyslu (mimo speciální případy¹⁴), a to jak apelativních, tak propriálních (srov. též Knappová 2000). Za pomoci tohoto formuláře jsou zpracovávány též jednotky přehodnocované, tj. pouze zeslovněné – *IČO / ičo, AIDS / aids* atp. Zkratková slova všech ostatních typu bývají zařazována a zpracovávána jako příslušný slovní druh, příp. jiná užší kategorie, např. substantiva (*radar, laser, umprum, esemeska* atd.) a zkratková jména (viz dále).

Vedle povinných položek určených zásadami zpracování hesel obecně¹⁵, při zpracování jednotky typu Zkratka/značka se v nástroji ZÁHLAVÍ povinně uvádí plné znění jednotky, které může být i v cizím jazyce (azbuku transliterujeme). U těchto zkratk je nutné rozlišovat plné znění v původním jazyce a jeho český překlad – do pole Plné znění zapisujeme pouze plné znění jednotky, nikoliv překlad, např. **ASEAN**, Plné znění: The Association of Southeast Asian Nations.

Spojení zkratky/značky a jejího plného znění je realizováno též programově: v případě, že je plné znění v LDB Pralex zpracováno, pak jsou obě jednotky propojeny aktivním proklikem, který je funkční v obou směrech. Tomuto účelu slouží nástroj ZKR./EKVIVALENTY umožňující vyhledání a proznačení zkratky či plného znění v LDB Pralex. V náhledu se pak aktivní proklik zobrazuje s barevným odlišením. Dochází tak k jistému zdvojení či vrstvení uvedené informace, které je pro databázové zpracování účelné, i když pro slovníkové zpracování se může jevit jako redundantní, např. **Bk**, plné znění: berkelium; výklad: Původ definice: zkratky-Kos (v chemii) *berkelium*. Na seznamu ZKR./EKVIVALENTY je umístěn výraz *berkelium* v červené barvě.

Specifické řešení bylo zvoleno též pro poznámku k výkladu. Sleduje se zde zpracování jednotky nejen ve výkladových slovnících češtiny, ale také v slovnících speciálních (Kos 1983, SZ ED) např. **AGB**, Pozn. k výkladu: • ssc 0, ssc 0, scs-DEB 0, zkratky-Kos 0, • zkratky-ED: AGB, Agrobanka.

Ve srovnání s podstatnými jmény je nástroj Exemplifikace u zkratk/značek redukován – obsahuje jen tři bloky, a to E1 Upravené doklady, E2 Větné doklady k užití a E3 Větné doklady k definici. Toto rozvržení se ukázalo jako výhodné zejména kvůli faktu, že mnohé jednotky pocházejí z různých méně známých věcných okruhů a v textech se velice často vyskytuje jejich plné znění (pro zaručení správné interpretace)¹⁶. Uvedením dostupných kontextů, na jejichž základě lze zkratku definovat, je připraven mate-

14 Jde o jistou zbytkovou množinu zahrnující specifické skupiny jednotek, které zatím do LDB Pralex nezařazujeme: jednotky obsahující písmeno z jiné abecedy, (např. Ω = ohm, π = Ludolfovo číslo; \vee = vektor, M = množina), modifikovaná písmena s úpravou (\$, £), číslicové značky (m^2 , H_2O), ideogramy (§, %) a zkratky obsahující číslici v iniciální pozici, např. 3D. Nezařazujeme ani emotikony, i když mají k této skupině velmi blízko, viz Niševa 2003.

15 Viz zde stat J. Světlé *Struktura a zpracování hesel v lexikální databázi Pralex*.

16 Tendence k ekonomičnosti vyjadřování se tak vyrovnává tendencí k explicitnosti sdělení.

riál pro budoucí slovníkové zpracování, který lze zužitkovat jak pro výkladový slovník, tak i pro případné slovníky speciální (např. slovník zkratek, slovník proprií apod.).

Je třeba podotknout, že rozlišení podle pojmenovací funkce apelativa vs. propria jde napříč formálním tříděním: velká část iniciálových zkratek představuje zkratky jmen institucí – *KB* (= Komerční banka, a. s.), *AV ČR* (= Akademie věd České republiky) atp. Propriální platnost mají zejména zkratky, ale i některé značky (např. kódy států). Tato platnost je u typu Zkratka/značka signalizována v poznámce k celému heslu, např. **AGB**, plné znění Agrobanka, výklad: *bývalá Agrobanka, a. s.*; Poznámka celkově • propr.

Propriální ráz je častý i u fónických zkratek, kde rozlišujeme zkratková slova apelativní (tradiční zkratková slova a zkrácená slova) oproti zkratkovým slovům propriálními (tradičně nazývána zkratková jména).

Apelativa jsou řazena a zpracovávána jako běžná substantiva¹⁷ s informací o vzniku slova, která je umístěna do poznámky k slovo tvorbě (Celk. pozn. ke slovo tvorbě), např. *radar*, Typ hesla: jednosl. podstatné jméno; výklad: Původ definice: SŠJČ, *radiolokátor založený na odrazu rádiových elektromagnetických vln*; exemplifikační bloky (E1 – E14)¹⁸. Celk. pozn. ke slovo tvorbě: Zkratkové slovo z angl. výrazu radio detection and ranging.

Zkratková jména, tím že jsou zařazena mezi propria, se zpracovávají za pomoci speciálního formuláře pro propriální jednotky, který má zjednodušenou strukturu exemplifikace (obsahuje dva exemplifikační bloky – E1 Větné doklady a E2 Upravené doklady). Jako typ hesla je zaznamenáno „jednosl. vlastní jméno“ a informace o vzniku slova je též v poznámce k slovo tvorbě, např. *Četka*, definice (upravená autorem): *Česká tisková kancelář (ČTK)*; Celk. pozn. ke slovo tvorbě: Zkratkové jméno z ČTK (Československá / Česká tisková kancelář).

Závěr

S ohledem na zvolené téma a také vzhledem k omezenému rozsahu nebylo zde možné podrobně popsat zásady zpracování zkratek/značek jako zvláštního typu databázové jednotky. Jak již bylo řečeno, nástroje pro zpracování abreviačních útvarů v LDB Praled představují jistý model, který umožňuje adekvátně zachytit kterýkoliv prvek patřící do široké skupiny abreviačních útvarů. Při budování LDB Pralex v rámci stávajícího výzkumného záměru neaspirujeme na plné zpracování tohoto typu lexikálních jednotek, neboť je databáze zamýšlena jako příprava a materiálový podklad pro slovníkové zpracování, další práce budou tedy určeny typem a rozsahem budoucího lexikografického výstupu.

Literatura

Akademický slovník cizích slov. Red. V. Petráčková, J. Kraus a kol. Praha: Academia 1998. 834 s. (ASCS)

¹⁷ Viz zde stať B. Procházkové a J. Světlé.

¹⁸ Viz podrobněji stať M. Koutové.

- CEJP, Ladislav: Poznámky o místě zkratk v jazyce. In: Slovo a slovesnost 1940, roč. 6, s. 37–42.
- DANEŠ, František: K výslovnosti tzv. iniciálových zkratk. In: Naše řeč 1989, roč. 72, č. 3, s. 170–172.
- DOLNÍK, Juraj: Lexikologia. Bratislava: Univerzita Komenského 2003. 236 s.
- Encyklopedický slovník češtiny. Praha: Nakladatelství Lidové noviny 2002. 604 s. (ESČ)
- FILIPEC, Josef: Zkratky a značky jako organická část výkladového slovníku (osnova pro teorii lexikografie). Pracovní materiál, s. 1–6 (bez vnočení).
- Frekvenční slovník češtiny. Praha: Nakladatelství Lidové noviny 2004. 596 s. (FSČ)
- FURDÍK, Juraj: Slovenská slovtvorba. (Teoria, opis, cvičenia.). Ed. M. Ološtiak. Prešov: Nauka 2004. 200 s.
- HELCL, Miloš: Zkratková slova. In: Naše řeč 1949, roč. 33, č. 9–10, s. 161–170.
- HAUSER, Přemysl: Nauka o slovní zásobě. Praha: SPN 1980, 195 s.
- HAVRÁNEK, Bohuslav – JEDLIČKA, Alois: Česká mluvnice. Praha: SPN 1981. 568 s.
- HORECKÝ, Jan: Niečo o akronymách. In: Kultúra slova 1980, roč. 14, č. 5, s. 149–152.
- HRBÁČEK, Josef: Typy zkratkových pojmenování v ruštině a češtině. In: Československá rusistika 1967, roč. 12, č. 3, s. 156–162.
- HRBÁČEK, Josef: Zkratky a lexikální systém. In: Slavica Pragensia, 1968, roč. 10, s. 275–283.
- HRBÁČEK, Josef: Jazykové zkratky v češtině. AUC – Philologica, Monographia 78, Praha: Univerzita Karlova 1979. 125 s.
- JADACKA, Hanna: Skrótwce. In: Wielki słownik poprawnej polszczyzny PWN. Red. A. Markowski. Warszawa: Wydaw. Naukowe PWN 2006, s. 1666–1668.
- KNAPPOVÁ, Miloslava: Obchodní jméno jako fenomén jazykový a sociologický. In: Onomastyka polska a nowe kierunki językoznawcze: materiały z XI Ogólnopolskiej Konferencji Onomastycznej 15–17 czerwca 1998 Bydgoszcz-Pieczyska. Eds. M. Czachorowska – Ł. M. Szewczyk. Bydgoszcz: WSP 2000, s. 123–131.
- KOCHOVÁ, Pavla: Internacionální iniciálové zkratky. In: Internacionalizmy v nové slovní zásobě. Sborník příspěvků z konference, Praha 16.–18. června 2003. Usp. Z. Tichá – A. Rangelova. Praha: ÚJČ AV ČR 2003, s. 157–165.
- KOCHOVÁ, Pavla: Nové iniciálové zkratky. In: Neologizmy v dnešní češtině. Praha: Ústav pro jazyk český AV ČR 2005, s. 242–246.
- KOS, Zdeněk: Zkratky, značky, akronymy. Praha: Horizont 1983. 289 s.
- MARKOWSKI, Andrzej: Skróty. In: Wielki słownik poprawnej polszczyzny PWN. Red. A. Markowski. Warszawa: Wydaw. Naukowe PWN 2006, s. 1664–1666.
- MICHAJLIŠIN, Bogdan Pavlovič: Проблеми й перспективи абрєвіального термінотворення. In: Słowotwórstwo, semantyka i składnia języków słowiańskich. T. 1. Eds. M. Blicharski – H. Fontański. Katowice: Wydaw. Uniwersytetu Śląskiego 1999, s. 53–61.
- Mluvnice češtiny 1. Academia, Praha 1986. 566 s. (MČ 1)
- MÜLDNER-NIECKOWSKI, Piotr: Wielki słownik skrótów i skrótowców. Wrocław: Wyd. Europa 2007.
- NIŠEVA, Božana: K jednomu zvláštnímu typu internacionalizmů – tzv. emotikony a akronymy. In: Internacionalizmy v nové slovní zásobě. Usp. Z. Tichá – A. Rangelova. Praha: ÚJČ AV ČR 2003, s. 220–228.

- Nová slova v češtině 1. Slovník neologizmů. (O. Martinová a kol.) Praha: Academia 1998. 356 s. (SN 1)
- Nová slova v češtině 2. Slovník neologizmů. (O. Martinová a kol.) Praha: Academia 2004. 568 s. (SN 2)
- OLOŠTIAK, Martin: Abreviačná a slovotvorná motivácia. In: Slovo o slove. Sborník Katedry komunikačnej a literárnej výchovy Pedagogickej fakulty Prešovskej univerzity 11. Prešov: Prešovská univerzita 2005, s. 39–50.
- POHNERT, Bohuslav: Ke zkratkám a značkám v technické literatuře. In: Naše řeč 1973, roč. 56, č. 3, s. 134–137.
- Příruční mluvnice češtiny. Eds. P. Karlík – M. Nekula – Z. Rusínová. Praha: Nakladatelství Lidové noviny 1995. 799 s. (PMČ)
- Příruční slovník jazyka českého. Praha: SPN 1935–1957.
- Pravidla českého pravopisu. Praha: Academia 2003. 391 s. (PČP)
- ROMPORTL, Milan: Výslovnost zkratkových slov. In: Naše řeč 1951, roč. 35, č. 2, s. 46–50.
- RUDNIK-KARWATOWA, Zofia: Złożeniowce we współczesnej polszczyźnie? In: Reverendissimæ Halinæ Satkiewicz cum magna æstimatione. Pod red. G. Dąbkowskiego. – Warszawa: Plejada 2008, s. 307–315.
- SAWICKA, Grażyna: Struktura współczesnych skrótowców. In: Język w mieście: problemy kultury i poprawności. Ed. K. Michalewski, Łódź: Wydaw. UŁ 1998, s. 153–165.
- Slovník spisovné češtiny pro školu a veřejnost. Red. J. Filipec – F. Daneš – J. Machač (1. vyd.) – V. Mejstřík (2. a 3. vyd.). 3., opravené vyd. Praha: Academia 2003. 647 s. (SSČ)
- Slovník spisovného jazyka českého. 8 sv. Red. B. Havránek a kol. 2., nezměněné vyd. Praha: Academia 1989. 569 s. (SSJČ)
- Slovník zkratek. Praha: Encyklopedický dům 1994. 231 s. (SZ ED)
- STANĚK, Vladimír: Poznámka k psaní skloňovaných zkratek typu *dikové* / *DIKové*. In: Naše řeč 1995, roč. 78, č. 5, s. 271–272.
- VLKOVÁ, Věra: Pojmenování vzniklá z písmenných zkratek. In: Naše řeč 1984, roč. 67, č. 3, s. 165–166.

Jednopísmenné zkratky a značky ve slovnících a v lexikální databázi Pralex¹

Zdeňka Tichá

Ústav pro jazyk český AV ČR, v. v. i.

One-Letter Abbreviations and Acronyms in Dictionaries and the Pralex Lexical Database

This paper deals with a group of entries in the Pralex lexical database, namely with the one letter abbreviations and acronyms, which form a specific and very diverse group. Firstly, the author shortly describes the concept of chosen abbreviations and acronyms; secondly, she focuses on the methodology of processing of this group of entries in the lexical database.

1. Úvod

Zkratky a značky jsou nedílnou součástí lexikální databáze Pralex – je to jeden ze čtyř základních typů zpracovávaných jednotek (jednoslovné, víceslovné, podslavní části a zkratky/značky). Při tomto základním rozřídování (pod Typ hesla) je stanovena společná kategorie zkratka/značka, teprve na další úrovni rozlišujeme, zda jde o zkratku či značku, aby bylo možné vyhledat buď celý soubor najednou, nebo jen jeden dílčí typ. V následujícím textu budeme používat slovo „zkratka“ jako zastřešující označení jak pro zkratku, tak pro značku (nebude-li uvedeno jinak). Soustředíme se zde především na zpracování zkratk a značek **jednopísmenných**, protože v LDB Pralex tvoří specifickou, velmi různorodou skupinu. V jejím rámci můžeme vydělit – dle grafického hlediska – skupinu zkratk (psány s tečkou) a skupinu značek (psány bez tečky). Nejdříve nastíníme zvolené pojetí zkratk a značek, poté se zaměříme na metodické postupy při zpracování této dílčí skupiny abreviačních útvarů².

2. Pojetí a vymezení zkratk a značek jako objekt lingvistického zájmu

Zkracování obecně je považováno za okrajový způsob tvoření slov (Hauser 1980, s. 155); tvoření zkratk stojí na okraji slovo tvorby a zasahuje do ní vytvářením zkratkových slov (*Mluvnice češtiny 1*, 1986, s. 517, dále MČ 1)³.

Při vymezení zkratk je v MČ 1 uplatněn kombinovaný procesuálně-rezultativní přístup; uvádí se, že při tvoření zkratk nastává tzv. krácení mechanické (na rozdíl od krácení sémantického, mezi které je řazena univerbizace, elipsa); redukce „slova či víceslovných výrazů na některé hlásky (písmena) nebo skupiny hlásek (písmen), které svým nápadným složením nebo pravopisem a současně vžitou podobou signalizují v povědomí adresáta projevu význam nezkráceného výrazu [...] Zkratka a plná podoba mají v principu stejný význam...“ (MČ 1, s. 517).

1 Příspěvek vznikl v rámci výzkumného záměru ÚJČ AV ČR, v. v. i., *Vytvoření databáze lexikální zásoby českého jazyka počátku 21. století* (AV0Z90610521), jehož vedoucím je doc. RNDr. K. Oliva, Dr.

2 K typologii v rámci celé skupiny viz zde stať A. Rangelove.

3 Část F. Tvoření zkratk a zkratkových slov, autor M. Churavý, s. 517–526.

MČ 1 podává rovněž podrobnou typologii zkratek, a to podle jejich realizace a užívání v textu, s podrobnějším dalším členěním podle strukturních vlastností:

- čistě grafické (textové podle Filipce, pracovní materiál): *apod., tzv., fi*; patří sem rovněž značky, např. *A, km*;
- graficko-fónické (iniciálové), např. *SRPŠ* („sdužení rodičů a přátel školy“), *TOS* („továrna obráběcích strojů“);
- fónické (zkratková slova), např. *kilo, metro*.

Vymezení **značky** nalezneme v *České mluvnici* (Havránek – Jedlička 1981, s. 74): „... zkrácením slov nebo slovních spojení (domácích i cizích) vznikly ustálené značky měř, vah, měn, fyzikálních veličin a jednotek, chemických prvků, matematických pojmů, pojmů hudebních apod. Za značkami tečku nepíšeme.“; srov. též *Pravidla českého pravopisu* (1998). Značky využívají rovněž písmena jiných abeced i nepísmenné znaky, jak upozorňuje MČ 1 (s. 518).

Při formování našeho přístupu k abreviačním útvarům jsme měli k dispozici starší pracovní materiál *Zkratky a značky jako organická část výkladového slovníku*, jehož autorem je pravděpodobně J. Filipce. Autor hovoří o typech zkracování, tvoření zkratk a značek, a jako jeden ze způsobů zmiňuje „vypsání prvního písmene slova nebo slov daného spojení anebo charakteristické počáteční skupiny písmen, slabiky nebo i více slabik“. Autor dále zmiňuje rovněž psaní zkratky s tečkou (*p., p*), s malým nebo velkým písmenem (např. *a, A*) a psaní písmenných značek (*a – ,ar‘, A – ,ampér‘*).⁴

3. Utvářenost zkratk/značek

Jak již bylo řečeno na začátku, zkratky/značky vznikají redukcí formální stránky nějakého pojmenování; u jednopísmenných zkratk je tato redukce maximální – z původní formy zbývá jeden znak. Zvláštními případy jsou **kombinace jednopísmenné zkratky a číslice**. Ze starších zkratk sem patří např. *V1, V2* s výslovností [fau ajnc, fau cvaj] ve významu ‚zbraň odplaty (kódové označení německých bezpilotních raketových střel vyvinutých 1942)‘ – z *německého Vergeltungswaffe*; z novějších např. označení formátů papíru (*A0 – A11*), státní poznávací značky letadel (např. *A6 – Spojené arabské emiráty; A7 – Katar; V8 – Brunej*); mezinárodní označování grafických technik (*C2 – mědirytina; X3 – linoryt*); označení ve sportovní oblasti (*K2 – kajak pro dvojice; K4 – kajak pro čtveřice*)⁵.

Zaznamenali jsme i případy **variant jednopísmenné a vícepísmenné zkratky**. Vznik takovýchto variant, zejména té vícepísmenné, je motivován s největší pravděpodobností snahou o větší srozumitelnost zkratky. Slovník *Zkratky, značky, akronymy* (dále Kos 1983) uvádí tyto varianty: *v. / viz, b. / baln.* – (lat.) ‚lázeň‘; *i. / int.* – (lat.) ‚introductio (úvodní část skladby), předehra‘; *n. / nom.* – (lat.) ‚nomen (jméno, název skladby)‘; *p. / pag.* – (lat.) ‚pagina (stránka /knihy/)‘; *p. / pí – ,paní‘; s. / seg.* – (it.)

4 V LDB Pralex neuvádíme značky obsahující písmeno z jiné abecedy (*Ω – ,ohm‘, Π – ,Ludolfovo číslo‘; v – ,vektor‘, M – ,množina‘*), modifikována písmena s úpravou (*\$, £*), číslicové značky (*m², H₂O*) ani ideogramy (*\$, %*).

5 Všechny příklady jsou ze *Slovníku zkratk* (1994).

,segno (znamení, od něhož se má skladba opakovat)'; *s.* / *sin.* – (it.) ,sinistra (levá ruka – v hudbě)'. Tento typ variantnosti se projevuje též u zkratk domáciho původu. Při práci na lexikální databázi jsme zaznamenali případy jako např. *č.* / *čís.* – ,číslo', *č.* / *čes.* – ,český' (a také *Č.* / *Čes.* – ,Český' ve vlastních názvech), *s.* / *sl.* – ,slečna', *s.* / *str.* – ,strana, stránka'; mezi variantní zkratky patří dále např. *f.* / *fr.* / *franc.* – ,francouzština, francouzský'; *n.* / *něm.* – ,němčina, německý'; iniciály křestních jmen, např. *F.* / *Fr.* – František, *A.* / *Al.* – Alois.

4. Původ zkratk

Za **zkratky české** považujeme takové zkratky, které vznikly na pozadí českých pojmenování, v rámci českého jazyka, např. *b* ve významu ,(u nákladních vozů ČSD) označení pro vůz řady Z s čelní plošinou s čelními dveřmi' (Kos 1983); *h.* ve významu ,hora, horský'; *j.* ve významu ,jezero, jezerní' (*Slovník zkratk* 1994, dále SZ ED).

Zkratky, které označujeme jako **cizí**, můžeme rozdělit na dvě skupiny. Do první skupiny bychom zařadili zkratky **internacionální** – jsou rozšířeny a užívají se ve více jazycích, rys cizosti se u nich postupně stírá. Již Hrbáček (1979, s. 46) upozorňuje na to, že internacionální povaha je u zkratk častým rysem. U zkratk vzniklých v nedávné minulosti či vznikajících v této době můžeme vysledovat jazyk původu (nejčastěji to bývá angličtina), u zkratk, se kterými se setkáváme delší dobu, je však jejich původ mnohdy obtížné zjistit.

Skupina zkratk/značek internacionálních se dělí na jednotlivé podskupiny dle věcného okruhu, ke kterému patří, např.: značky fyzikálních veličin (*A* – ,ampér', *p* – ,pond', *k* – ,koňská síla'); chemické značky (*O* – ,kyslík', *H* – ,vodík', *I* – ,jód', *B* – ,bor'); matematické značky (*R* – ,pravý úhel'; *B* ve významu ,byte (jednotka množství informace)'); značky vztahující se k terminologii z oblasti hudby (*i.*, *lat.* ,introductio, úvodní část skladby, předehra'; *K* – ,Köchel' (číslování hudebních skladeb podle Köchelova seznamu); lingvistiky (*m.* – ,maskulinum') i dalších oborů; mezinárodní poznávací značky motorových vozidel (*M* – ,Malta', *U* – ,Uruguay'), označení římských číslic (*I* – ,římská číslice 1'; *V* – ,římská číslice 5') a další. Mnoho z nich vzniklo na základě latinských pojmenování.

Do druhé skupiny řadíme zkratky, které vznikly **v cizím jazyce** a jsou omezeny na užívání v něm (v cizím jazykovém prostředí), např. v angličtině *B* ve významu ,bachelor, bakalář'; *M* ve významu ,Monday, pondělí'; v němčině *B* ve významu ,Bundesstraße (silnice 1. třídy)'; *r.* ve významu ,rechts, vpravo, napravo, doprava'⁶.

5. Zpracování zkratk a značek v referenčních zdrojích

5.1. Vzhledem ke specifice zkratk a značek bylo při jejich zpracování v LDB Pralex přihlédnuto ke **slovníkům speciálním**, a to ke SZ ED a k publikaci Kos 1983. Na rozdíl od všeobecných slovníků výkladových, v nichž je uvádění zkratk výběro-

6 Všechny příklady jsou ze SZ ED.

vé, ve slovnících zkratk je zachycen rozmanitější materiál, co se oborů týče, a liší se rovněž způsob zpracování. Jednotlivé významy zkratk jsou v těchto speciálních slovnících uváděny zvlášť, bez číslování, resp. bez označení významů homonymních. To je však pro potřeby rozsáhlejšího slovníku (resp. databáze) nevýhodné, neboť počítáme se vzájemným proodkazováním jak mezi jednotlivými zkratkami, tak mezi zkratkou a plnovýznamovým slovem.

Podobně zpracovává jednotlivé zkratky slovník zkratk uveřejněný na internetu (*Zkratky.cz*), s tím, že zde není uváděn původ jednotlivých zkratk a zpracování se omezuje na dešifrování plného znění, příp. na překlad cizojazyčného plného znění. Nalezneme zde např. zkratku *A* v 9 různých významech, jeden z nich je zpracován následujícím způsobem⁷:

A

Obor:	Ekonomie
Význam:	Account
Popis:	Překlad: účet

Jednopísmenné **zkratky**, tj. takové jednotky, které jsou v textu psány převážně s tečkou, jsou ve skupině jednopísmenných zkratk a značek v menšině.

SZ ED u písmene malého *b* a velkého *B* uvádí celkem 41 zkratk a značek, 6 bychom zařadili mezi zkratky, a to s upozorněním, že nejde o zkratky české, ale o zkratky užívané s největší pravděpodobností pouze v cizích jazycích. Jsou to zkratky typu: *b., fr.* ‚billet, směnka‘; *b., lat.* ‚balneum, lázeň‘; *b., lat.* ‚bis, dvakrát‘; *b., něm.* ‚bei, u‘; *B., fr.* ‚billet, směnka‘; *B., it.* ‚basso, nízký, hluboký; potichu‘. Podobně i ve slovníku Kos 1983 je uvedeno celkem 17 zkratk a značek, z nichž bychom mezi zkratky zařadili pouze dvě, opět cizí, a to: *b.* nebo *baln.* (lat.) ‚balneum – lázeň‘; *B.* (it.) ‚basso – nízký, hluboký; též potichu /v hudbě/‘.

Jednopísmenné **značky** jsou ve skupině jednopísmenných zkratk a značek zastoupeny častěji. Vrátime-li se k výše uvedenému příkladu (malé *b* a velké *B*), pak v rámci SZ ED je mezi 41 jednopísmennými zkratkami a značkami 17 jednopísmenných značek⁸, z nichž většinu bychom zařadili ke značkám internacionálním. Jsou to následující jednotky:

b ‚bes (jednotka hmotnosti)‘; *b* ‚bit (jednotka množství informace)‘; *b* ‚konstanta Wienova zákona‘; *b* ‚pohyblivost (mobilita) nabitých částic‘; *b* ‚ryv (časová změna vektoru zrychlení přímočarého pohybu)‘; *b* ‚točivost, moment hybnosti‘; *B* ‚baryonové číslo‘; *B* ‚bas (v partituru)‘; *B* ‚Belgie (mezinárodní poznávací značka motorových vozidel)‘; *B* ‚Besselova epocha‘; *B* ‚byte (jednotka množství informace)‘; *B* ‚Čína (státní poznávací značka letadel)‘; *B lat.* ‚borum, bor (chemická značka)‘; *B* ‚magnetická indukce, hustota magnetického toku‘; *B* ‚modrobílé hvězdy s povrchovou teplotou 11 000 – 18 000 K‘; *B* ‚modrý obor spektra nebo magnituda v něm určená‘; *B* ‚třída požáru (hoření kapalných látek hořících plamenem)‘.

7 <http://www.zkratky.cz/A/654>.

8 Stranou ponecháváme zkratky či značky, které mají v tomto slovníku označen anglický či německý původ, např. *b, angl.* ‚bachelor, bakalář‘; *b, angl.* ‚bill, účet, faktura; návrh zákona‘; *B, angl.* ‚barge, barže (vlečný nákladní člun)‘; *B, angl.* ‚born, narozený‘; *B, něm.* ‚Bundesstraße (silnice 1. třídy)‘ atp. – je jich celkem 18. Dodejme, že psaní zkratk a značek se může v jednotlivých jazycích odlišovat.

Podobně ve slovníku Kos 1983 je mezi 17 jednopísmennými zkratkami a značkami jednopísmenných značek 15, opět většinou internacionální povahy:

b (angl.) ,barn – jednotka plošného obsahu / 10^{-28} m²/'; *b* ,bes – jednotka hmotnosti používaná v Itálii /l kg/'; *b* ,u nákladních vozů ČSD označení pro vůz řady Z s čelní plošinou s čelními dveřmi'; *b* ,konstanta Wienova zákona'; *b* ,pohyblivost /mobilita/ nabitých částic'; *b* ,součinitel vazby /nukleonů v atomovém jádru/'; *b* ,ryv /časová změna vektoru zrychlení přímočarého pohybu/'; *b* ,točivost; moment hybnosti'; *B* ,bas /v partituru/'; *B* ,chemická značka bóru'; *B* ,mezinárodní poznávací značka motorových vozidel Belgie'; *B* ,státní poznávací značka tchajwanských letadel'; *B* ,u osobních vozů ČSD označení pro vůz druhé třídy'; *B* ,hmotnostní schodek; hmotnostní úbytek'; *B* ,průzvučnost'.

5.2. Ve výkladových slovnících jsou jednotlivé významy zkratky a značky zpracovávány nejčastěji v rámci jednoho polysémního hesla; pokud jde o značku, uvádí se tato charakteristika v rámci hesla explicitně (*značka*). Zaměříme se především na *Slovník spisovného jazyka českého* (dále SsJČ) a *Nový akademický slovník cizích slov* (dále NASCS), neboť v *Slovníku spisovné češtiny pro školu a veřejnost* (dále SsČ) je zpracování zkratk řešeno jako dílčí slovníková příloha (samostatná abecední řazení). SsJČ zachází s jednotkami tohoto typu různě, zpracovávají se zejména:

- 1) zkratky institucí souvisejících (převážně) se socialistickým společenským zřízením v ČSSR (*ONV, KSČ, SNB*);
- 2) zkratky názvů států, mezinárodních organizací a jiných institucí (*NDR, RVHP, UNRA*);
- 3) zkratky vědeckých a jiných titulů (*MUDr., JUDr.*);
- 4) textové a ostatní zkratky (*apod., j. č., cca, P. T., rkp.*);
- 5) zkratková slova a zkratková jména (*alobal, Svazarm, umprum*).

Některé zkratky nemající samostatné heslo, např. *confer* [kon-] (lat.) srovnej, porovnej (zkr. *cf.*); *dráha 2.* železnice a zařízení s ní související; zařízení pro dopravu vázanou na kolejnice ap. vůbec: Československé státní dráhy (zkr. *ČSD*);

Pro porovnání zpracování v SsJČ a NASCS se podívejme na heslo **a**:
SsJČ:

a

1. česká samohláska (*nízká, střední*) a písmeno ji označující
2. (vysl. a někdy též ps. á) neskl. s. *jm. samohlásky a písmene a*, á: vyslovit a; napsat velké A; [x] neřící (ani) a (á), ani b (bé) *neřící nic, nijak se nevyjádřit*; kdo řekne á, ať řekne také bé *kdo něco začne (napoví), ať to také dokončí (dopoví)*; od a (á) do zet *všechno, od začátku až do konce*
3. [á] *označení prvního členu, prvního místa v něj. řadě n. jedné ze dvou n. více věcí vůbec: A, B..., a), b)... (při označování odstavců, při výčtu ap.)*; třída II. A; krevní skupina A; mužstvo A; vitamín A (hovor. A-vitamín)
4. [á] neskl. s. hud. *jm. tónu a příslušné noty; jm. struny, kt. má tento tón*: komorní a; tónina A dur, a moll; zazpívat a; - prasklo mi a
5. a (značka) *ar*
6. A (značka) *ampér*

NASCS:

a fyz. 1. zn. pro ar ; 2. zn. pro atto-

a zkr. <f> na hodinkách označující zrychlení chodu (avance)

6. Zpracování zkratek/značek v LDB Pralex

6.1. Zdroje zkratek/značek pro heslář LDB Pralex

Výchozím zdrojem zkratek pro lexikální databázi Pralex je *Frekvenční slovník češtiny* (2004, 1000 zkratek) a SSČ (cca 650 zkratek). Tyto zkratky byly do databáze vloženy programově. Průběžně jsou při zpracovávání stávajících zkratek přidávány další (homonymní) zkratky, resp. významy zkratek. Jejich zdrojem je především korpus SYN⁹. Důvodem pro průběžné přidávání je zejména to, že v případě zařazení zkratky do hesláře LDB Pralex by měly být uvedeny všechny její doložené (homonymní) významy.

Na základě zvolených kritérií jsou cíleně přidávány vybrané zkratky ze SZ ED¹⁰. V první fázi šlo o zkratky začínající písmenem *a, A*, s tím, že jde o pokusný soubor; na základě práce s těmito zkratkami budou dále postupně zpřesňována kritéria výběru jednotek do hesláře. Doplnkově jsou do hesláře přidávány zkratky ze slovníků *Nová slova v češtině. Slovník neologizmů 1, 2* (Martincová a kol. 1998 a Martincová a kol. 2004).

6.2. Vyhledávání zkratek v korpusových textech

Jednopísmenné zkratky/značky jsou specifické z různých ohledů, včetně práce s jazykovým materiálem. Nejobecněji je třeba říci, že základní metodologické přístupy stanovené pro LDB Pralex se uplatňují i u této skupiny jednotek. Všechny zkratky, resp. jednotlivé významy zkratek, které byly vloženy do hesláře programově, ověřujeme z hlediska doloženosti v korpusu *syn* (od ledna 2011 *orig_syn*). V případě zkratek, a to především jednopísmenných, prakticky nepoužíváme vyhledávání kolokací prostřednictvím programu Word Sketch Engine. Pracujeme především s korpusovým manažerem Bonito. Vyhledávání jednopísmenných zkratek je velmi obtížné a má svá specifika. Nejsložitější je vyhledávání jednopísmenných značek, jejichž grafická podoba je bez tečky, tudíž vyhledávaná grafická sekvence je minimální (pouze 1 znak); vyhledávání zkratek iniciálových anebo textových je jednodušší, neboť bývá vyhledáván řetězec znaků.

Na základě zpracování zkratek ve speciálních slovnících definujeme v korpusovém manažeru Bonito zejména pozitivní filtry, s jejichž pomocí vyhledáváme příklady na užití zkratky v daném významu. Při vyhledávání je nutné pamatovat na velká / malá písmena, na tečky (zadávají se s mezerou) a na další znaménka (pomlčka apod.), která se mohou v okolí zkratky vyskytovat. Uvedme příklad pro vyhledávání písmene *A* samostatně a s použitím pozitivního filtru:

A – 731 516 výskytů vyhledáno bez filtrů.

A s pozitivním filtrem písmen.* / od -5 do 5 pozic – 492 výskytů, např.

opus.rokvyd=1996,opus.txtype=NOV,opus.id=pece2“ Číslo 1408 přečtu snadno , ale nedokážu rozeznat , jestli za ním následuje písmeno <A> nebo B . „

opus.rokvyd=1998,opus.txtype=NOV,opus.id=mcbainV některých zdejších domech byly byty

9 Od roku 2011 byl původní korpus *syn* přejmenován na *orig_syn*.

10 V LDB Pralex v poli Zdroj pro heslář je tento zdroj označen jako Zkratky-ED podle nakladatelství Encyklopedický dům.

v přízemí označené jen písmeny <A> , B , C atakdále , nikoli čísly .

opus.rokvyd=1999,opus.txtype=NOV,opus.id=bellona <A> písmeno R a J patrně znamenají Robert a John a ty cifry znamenají částky , které jim chtěl odkázat . „

A s pozitivním filtrem písmen.* / od -1 do 0 pozic – 225 výskytů, např.

opus.rokvyd=1996,opus.txtype=NOV,opus.id=pece2“ Číslo 1408 přečtu snadno , ale nedokážu rozeznat , jestli za ním následuje písmeno <A> nebo B . „

opus.rokvyd=1998,opus.txtype=NOV,opus.id=mcbainV některých zdejších domech byly byty v přízemí označené jen písmeny <A> , B , C atakdále , nikoli čísly .

opus.rokvyd=2004,opus.txtype=NOV,opus.id=popel Měla několik laciných stříbrných prstýnků a dvě malá neumělá tetování inkoustovou tužkou : křížek poblíž palce a přeškrtnuté písmeno <A> . Zato na zápěstí měla profesionálně vytetovaný trnový náramek .

A (značka) ve významu ,ampér‘

Pozitivní filtr: *ampér.* / od -1 do 1 pozice – 8 výskytů*, např.

opus.rokvyd=1996,opus.txtype=POP,opus.id=hlavol Voltmetr V1 ukazuje napětí (efektivní hodnotu) U 1 = 60 V a voltmetr V2 pak U 2 = 80 V . Wattmetr W ukazuje činný příkon P = 120 W a ampérmetr <A> proud (efektivní hodnotu) I = 2 A . Vlastní spotřebu měřicích přístrojů můžeme zanedbat .

a – 12 998 619 výskytů

Pozitivní filtr: *tón.* / od -1 do 1 pozice – 1549 výskytů*

Pozitivní filtr: *tón / od -1 do 1 pozice – 465 výskytů*

Pozitivní filtr: *tón / od -1 do 0 pozice – 357 výskytů*

opus.rokvyd=2001,opus.txtype=ENC,opus.id=atlashudProto byl zaveden komorní **tón <a>** 1 o frekvenci 440 Hz při teplotě [20&Cgrad ;] (na 2 . mezinárodní konferenci o ladění v roce 1939 v Londýně) .

opus.rokvyd=2001,opus.txtype=ENC,opus.id=atlashud Komorní tón <a> 1, podle něhož se ladí , určuje přitom svými 440 Hz absolutní tónovou výšku celkové soustavy .

opus.rokvyd=1996,opus.txtype=PUB,opus.id=mf961227V Bachově sonátě pak při interpretovně skvělé pohyblivosti přijde líto , že Graf utíká od konců frází a pokud setrvá , má nevyrovnané vibrato , v Maraisovi (ve 3 . variaci) má nejistý tón <a> v Sonátě Carla Ph .

a (značka) ve významu ,ar‘

Pozitivní filtr: *ar / od -10 do 10 pozic – 66 výskytů*; pozitivní filtr: *výměr.* / od -3 do 3 pozic – 477 výskytů* – ani v jednom případě není zachycena značka v daném významu. Hledání značky *a* ve významu ,ar‘ bylo bez výsledku, resp. značku *a* ve významu ,ar‘ se nepodařilo doložit. S podobnou situací jsme se setkali u dalších homonymních významů značky *a* – u všech uvádíme „*nepodařilo se doložit v syn*“.

Velmi často dochází také k tomu, že při hledání dokladu na příslušný význam zkratky nalezneme v korpusu další (homonymní) významy anebo doklady výskytu zkratky jiné, která ale může pocházet ze stejné věcné/sémantické skupiny; zde např. značky *a* ve významu ,ar‘ a *ha* ve významu ,hektar‘, které patří do skupiny jednotek plošného rozsahu.

6.3. Zpracování homonymních zkratek/značek v LDB Pralex

U zkratk se homonymie projevuje v mnohem větší míře než u jednoslovných výrazů z jednotlivých slovních druhů. Důvodem této homonymie je to, že jednotlivé

zkratky a značky jsou z různých oborů, a denotát, který označují, je zcela odlišný. Ukazuje se, že především u jednopísmenných zkratek a značek je množství homonymních jednotek značné.

Této tematiky se ve svém příspěvku dotkl i Michajlišin (1999, s. 60), který upozorňuje (ve volném překladu): „Stejně jako plným názvům, je abreviaturám vlastní homonymie. Přitom, čím menší je početní stav písmen, tím větší je různorodost homonymních zkratek.“ S tímto konstatováním musíme – na základě dosud zpracovaných zkratek v naší lexikální databázi – plně souhlasit.

Homonymní zkratky jsou zpracovávány v souladu s obecnými zásadami zpracování homonym v LDB Pralex; homonyma označujeme římskými číslicemi v pořadí podle slovních druhů a zkratky uvádíme až na posledním místě, proto je vždy třeba nejprve zkontrolovat a případně založit všechna ostatní relevantní homonyma (pokud je ještě nemáme v hesláři), u všech určit slovní druh a typ hesla (mimo jiné kvůli vyhledávání a propojování komponentů víceslovných lexikálních jednotek) – zejména u jednopísmenných hesel, srov.: *a: a I a II* – podstatné jméno; *a III* – spojka; *a IV* – částice; *a V* – citoslovce; *a VI* až *a IXX* zkratky s homonymními významy. Číslování homonym je nutné pro vzájemné hypertextové proodkazování jak mezi jednotlivými zkratkami, tak mezi zkratkou a jejím nezkráceným jednoslovným protějškem, aby se v nabídkovém seznamu u ekvivalentního hesla našlo to správné jí odpovídající homonymum a propojení bylo korektní.

V LDB Pralex zaznamenáváme např. tyto případy: *č. I / čes. I* ve významu ‚český‘; *č. II / čes. II* ve významu ‚česky‘; *č. III / čís. ve významu ‚číslo‘*. Z toho je patrné, že zpracování zkratek a značek má dosah pro zpracování dalších slovních druhů.

6.4. Zpracování variant

Obdobně jako u jiných typů lemmat, i u zkratek se všechny varianty zpracovávají zvlášť, jako samostatné položky databáze, a to v souladu se zásadami pro zpracování variant v LDB Pralex. V případě jednopísmenných zkratek jde např. o variantu *a* (v první fázi zpracování 19 homonymních lemmat, jejich počet bude pravděpodobně redukován) – *á* (podstatné jméno I, II x předložka III x citoslovce IV).

Shrnutí

Zpracování zkratek (jak jednopísmenných, tak i dalších) v LDB Pralex probíhá na základě zásad zpracování, které jsou postupně upřesňovány. Na základě práce s materiálem dochází k úpravám těchto zásad tak, aby byly v souladu s obecnými zásadami práce s hesly v LDB Pralex, ale zároveň aby postihovaly specifika zkratek a značek. Zkratky a značky zpracované v rámci LDB Pralex najdou (v přísném výběru) využití v rámci budoucího výkladového slovníku češtiny, ale mohou být využity i při přípravě slovníku speciálního – slovníku zkratek.

Literatura

- FILIPEC, Josef: Zkratky a značky jako organická část výkladového slovníku (osnova pro teorii lexikografie), pracovní materiál, s. 1–6, příruční knihovna m. 146. (bez vnoření).
- Frekvenční slovník češtiny. Praha: Nakladatelství Lidové noviny 2004. 596 s.
- HAUSER, Přemysl. Nauka o slovní zásobě. Praha: SPN 1980, 195 s.
- HAVRÁNEK, Bohuslav – JEDLIČKA, Alois: Česká mluvnice. Praha: SPN 1981. 568 s.
- HRBÁČEK, Josef: Jazykové zkratky v češtině. AUC – Philologica, Monographia 78, Praha: Univerzita Karlova 1979. 125 s.
- KOS, Zdeněk: Zkratky, značky, akronymy. Praha: Horizont 1983. 289 s.
- MICHAJLIŠIN, Bogdan Pavlovič: Проблеми й перспективи абрєвіального термінотворення. In: Słowotwórstwo, semantyka i składnia języków słowiańskich. T. 1. Eds. M. Blicharski – H. Fontański. Katowice: Wydaw. Uniwersytetu Śląskiego 1999, s. 53–61.
- Mluvnice češtiny 1. Praha: Academia 1986. 566 s. (MČ 1)
- Nová slova v češtině. Slovník neologizmů 1. (Martincová, Olga a kol.) Praha: Academia 1998. 356 s. (SN 1)
- Nová slova v češtině. Slovník neologizmů 2. (Martincová, Olga a kol.) Praha: Academia 2004. 568 s. (SN 2)
- Nový akademický slovník cizích slov. Praha 2005. 880 s. (NASCS)
- Pravidla českého pravopisu. Academia, Praha 1998. 391 s.
- Slovník zkratk. Encyklopedický dům, Praha 1994. 231 s. (SZ ED)
- Slovník spisovné češtiny pro školu a veřejnost. 3. vyd., Praha 2005. 647 s. (SSČ)
- Slovník spisovného jazyka českého. Praha 1960–1971. (SSJČ)

K aspektom tvorenia antonymných lexém s prefixom *ne-* (nominácia, lexikalizácia, pragmatické komponenty)¹

Klára Buzássyová

Jazykovedný ústav Ľudovíta Štúra SAV

Some aspects of Formation of Antonymous Lexemes with Prefix *ne-* (Nomination, Lexicalization, Pragmatic Components)

In this study, we try to provide more detailed view of particular means of expressing antonymy. Our aim is to contribute to more reasoned introduction of antonyms (opposite expressions) as facultative formal-semantic information, which, apart from synonyms, support explanation of particular lexical items (entries) formed while using a negative prefix *ne-* in an unfinished third volume of The Dictionary of the Contemporary Slovak Language M-O. Our study is mainly lexicology-lexicographical and follows function-structural as well as communicative approach, which has required certain reinterpretation of antonymy types in the existing literary resources.

1. Úvod

Lexémy so záporovým (negačným) prefixom *ne-* tvoria súčasť podstatne širšej oblasti antonymie v jazyku. V tejto štúdiu sa zameriame osobitne na lexémy so záporovým prefixom *ne-*, a to na pozadí formálne rovnakých lexikálnych jednotiek, ktoré zohrávajú v slovenčine aj inú funkciu ako funkciu pravých antonym. K téme nás inšpirovalo redigovanie hesiel tretieho zväzku Slovníka súčasného slovenského jazyka (ďalej SSSJ M – O) a potreba prispieť k spresneniu uvádzania/neuvádzania opozitných, ale aj synonymných jednotiek ako dokresľujúcich výklad heslových slov, ktorými sú lexikálne jednotky utvorené pomocou prefixu *ne-*.

Treba povedať, že autorky Slovníka súčasného slovenského jazyka (SSSJ) majú iba v pomerne malej miere oporu v uvádzaní opozít v predchádzajúcich lexikografických dielach, či už prekladových, ale najmä v predchádzajúcich výkladových slovníkoch slovenského jazyka. Informácia o opozitných jednotkách sa totiž raz pre iný žáner, stručnosť spracovania slovníka (porov. Krátky slovník slovenského jazyka, 2003; ďalej KSSJ), raz z iných dôvodov, napr. pre menšiu rozpracovanosť niektorých aspektov lexikografickej teórie v čase vypracovania diela (porov. Slovník slovenského jazyka, 1959 – 1968; ďalej SSJ), uvádzala doteraz skôr iba okrajovo, vo veľmi obmedzenej miere.

V príspevku nás budú zaujímať v rámci naznačeného zúženia rozličné typy antonymných dvojíc, ktorých jeden člen sa tvorí pomocou záporovej predpony *ne-*. Základné poučenie o antonymii, resp. o opozitných vzťahoch v lexikálnej zásobe nájdeme v lexikológiách, porov. *Česká lexikologie* (Filipec – Čermák, 1985) a *Lexikológia* (Dolník, 2003). Toto novšie dielo s ambíciou slúžiť ako vysokoškolská učebnica

1 Príspevok bol vypracovaný v rámci grantového projektu VEGA č. 2/0018/11 *Slovník súčasného slovenského jazyka – 4. etapa (koncipovanie a redigovanie slovníkových hesiel a s tým súvisiaci lexikologicko-lexikografický výskum)*.

aj ako autorské vedecké dielo, obsahuje popri základných poznatkoch o antonymii a jej typoch aj diskusiu o sporných, otvorených problémoch týkajúcich sa tohto javu (takto autor postupuje aj pri iných témach lexikálnej sémantiky). F. Čermák hovorí o opozitnosti vo všeobecnosti – „opozitnosť (tradične antonymie)“ – ako o oblasti výskytu lexém s protikladným, opozitným významom. Za základný problém oblasti výskytu lexém s protikladným významom právom pokladá určenie podstaty protikladu a miery jeho platnosti (Čermák, 1997, s. 262). Autori spomenutých komplexných lexikologických príručiek, ani autori štúdií, napr. M. Pisárčiková (1974) alebo špeciálne zameraných monografií, napr. L. A. Novikov (1973) nevenovali explicitnú pozornosť antonymii ako lexikologicko-lexikografickému problému, teda miestu, ktoré majú antonymá v lexikografických dielach, osobitne vo výkladovom slovníku. O objasnenie antonymie práve z tohto aspektu nám pôjde predovšetkým. Ukazuje sa totiž, že uvádzanie opozít k lexémam utvoreným záporovým prefixom *ne-*, ale aj synonymým vyžaduje prehliť zásady, podľa ktorých sa dopĺňa výkladová zóna hesiel v SSSJ, osobitne v pripravovanom treťom zväzku M – O. Zároveň tým v širšom pláne sledujeme cieľ prispieť k adekvátnejšiemu prezentovaniu paradigmatických a syntagmatických vzťahov skúmaných lexikálnych jednotiek v naznačenom výseku tejto nezanedbateľnej a pritom zaujímavej zložky slovenskej lexiky.

V štúdiu sa dotýkame týchto bodov:

1. všeobecné (základné) poznatky o antonymách (v logike a v jazyku); kontrárne a kontradiktorké, ako aj iné vzťahy protikladných členov dvojíc slov (a viacčlených radov jednotiek);
2. centrálné (pravé, úplné) antonymá, čiastočné antonymá; miera platnosti protikladov pri týchto typoch;
3. slovnodruhový aspekt: substantíva, adjektíva, príslovky (častice, číslovky); slovesá vo vzťahu k antonymii (opozitnosti);
4. lexikalizácia, terminologická nominácia;
5. intenzifikácia; vyjadrovanie intenzifikácie lexikálno-sémanticky;
6. eufemizácia (pragmatické komponenty významu) ako podnet na tvorenie jednotiek v prefixom *ne-*.

2. Logický a jazykový základ antonymie

Analýza prefixálnych antónym so záporovou predponou *ne-* vyžaduje uviesť základné poznatky o logickej a jazykovej stránke tohto javu.

Užitočné je ujasniť si vzťah k logickému základu antonymie, ale zároveň si uvedomiť, že jazykové (a teda aj jazykovedné, konkrétne lexikologicko-lexikografické) hľadisko, ak má adekvátne reflektovať vzťahy, v ktorých potenciálne antonymá existujú a fungujú v rámci systémových vzťahov na paradigmatickej osi, musí zohľadňovať aj fungovanie príslušných analyzovaných jednotiek na syntagmatickej osi, teda v rozmanitých kontextoch. Tie, ako je známe, lexikografi ilustrujú v slovníkoch vhodne vybranými príkladmi v dokladovej zóne heslových statí. Táto časť spracúvaných hesiel zároveň ukazuje, čím všetkým jazykové hľadisko presahuje logický základ antonymie. Podrobnejšie poznatky o logickom základe antonymie podáva J. Dolník

(2003). Odkazujeme na všeobecne orientovanú časť o antonymii aj na časť nazvanú Diskusia o antonymii, kde autor poukazuje na vývin výskumu tohto javu (Dolník, 2003, s. 98 – 104). **Antonymiu** definuje autor ako **sémantický** alebo **formálno-sémantický vzťah medzi lexiami** (v prípade mnohoznačných jednotiek sa antonymné vzťahy týkajú vždy jednotlivých významov), v dôsledku čoho významy sú **v polárne protikladnom význame** (vo vzťahu kontrastu). Kontrast je krajný protiklad. Kontrastnosť významov antonymných lexém podmieňuje to, že antonymá sa vzájomne predpokladajú a vzájomne sa vylučujú. Aj M. Pisárčiková (1974), ktorá venovala pozornosť antonymii v rozsiahlej štúdií (pričom nadviazala na súdobé výskumy najmä ruských a českých bádateľov), poukazuje na protikladnosť ako základnú vlastnosť antonymie. Zdôrazňuje, že protikladnosť antonym funguje v jazyku natoľko, nakoľko pravidelne sa v reči vybavujú ako protikladné sémantické jednotky.

Bádatelia sa zhodujú v názore, že **antonymia je spätá s negáciou**, čo je v niektorých prípadoch aj formálne vyjadrené predponami, napr. *známy – neznámy, cenný – bezcenný*; inokedy je negácia (v rozličnom stupni) implicitná. Existujú, ako to z definície vyplýva, sémantické antonymá (s rozličnými koreňmi) *starý – mladý, častý – zriedkavý*, aj formálno-sémantické antonymá ako *chybný – bezchybný, známy – neznámy, súlad – nesúlad* (opozitnosť, protikladnosť má vtedy predponu alebo afixoid ako formálny exponent, ukazovateľ tohto vzťahu).

Dôležitý je aj rozdiel medzi **protikladnými (logicky kontrárnymi) pojmami a protirečivými (logicky kontradiktorickými) pojmami**. Kým protikladné, kontrárne pojmy majú každý aj istý kladný obsah, nezahŕňajú celý rodový pojem, protirečivé, kontradiktorické pojmy sú také dvojice, v ktorých protirečivý člen iba neguje obsah druhého člena, takže nemá kladný obsah. V interpretácii kontradiktorického vzťahu sa bádatelia rozchádzajú. Napríklad autori Českej lexikologie (Filipec – Čermák, 1985) hovoria v prípadoch ako sú *čítať – nečítať, robotník – nerobotník, český – nečeský* o kontradiktórnych (slov. kontradiktorických) antonymách; J. Dolník (2003) na rozdiel od nich popiera, že by sa o kontradiktorickom vzťahu dalo hovoriť ako o súčasť antonymie. K týmto prípadom sa vrátíme nižšie, v bodoch o slovných druhoch, intenzifikácii a pragmatických komponentoch významu. Tu spomeňme len to, že J. Dolník cituje I. Němca (1969), ktorý delí lexikálny zápor na zápor bez kladného synonyma (ten vyjadruje len popieranie) a na zápor s kladným synonymom, ktorý vyjadruje kontrastný opak (čiže antonymnosť). Podľa toho *nefajčiar, nečeský, neďaleko* sú z logického hľadiska kontradiktorické pojmy, ale z jazykového hľadiska nepatria do jedinej skupiny, lebo slová *nefajčiar, nečeský* nemajú formálne kladné synonymá a vyjadrujú len popieranie, kým k slovu *neďaleko* existuje synonymum *blízko* a vyjadruje kontrastný opak.

3. Typologické hľadisko: pravá a čiastočná antonymia a predpoklady jej uplatnenia

3.1. Centrálny typ antonym (pravá antonymia) existuje medzi dvojicami slov vyjadrujúcimi protiklady, opozitnosť ako krajné hodnoty toho istého radu jednotiek. Sú to predovšetkým antonymá zastúpené kvalitatívnymi (kvalifikačnými) adjektí-

vami s rozličnými koreňmi *dobry* – *zly*, *mlady* – *stary*. Medzi nimi existuje aj stredná hodnota. Napríklad rozmer podľa členenia „príznač nezodpovedajúci norme“ (*vel'ký*, *malý*) má hodnoty „väčší ako norma“ (*vel'ký*) alebo „menší ako norma“ (*malý*); *stredný* = „zodpovedajúci norme“. Dvojice so záporovou predponou *ne-* predstavujú tento typ (podľa M. Pisárčikovej, 1974, s. 183 – 192) vtedy, keď iný prostriedok na vyjadrenie v jazyku nie je a krajným členom sa stáva slovo so záporovou predponou (M. Pisárčiková hovorí o záporovej častici): *ľudský* – *neľudský*, *kultúrny* – *nekultúrny*.

V skoncipovaných heslách SSSJ M – O je značný počet antonym so záporovou predponou *ne-*, ktoré môžeme zaradiť medzi antonymá tohto centrálného typu. Slovnodruhovo sú to predovšetkým prídavné mená a slovotvorne a svojím lexikálnym významom s nimi súvzťažné príslovky: *nekonvenčný* – op. *konvenčný*; *nekonvenčne*; *nekoordinovaný* – op. *koordinovaný*; *nezladený* – op. *zladený*. Sémantizáciami sa v slovníku určuje, na aké objekty sa vlastnosti označené adjektívami vzťahujú: **neobsadený** – ktorý nie je trvalo obsadený, zaplnený (o priestore, o mieste, o funkcii)...; op. *obsadený*: *neobsadený byt*; *neobsadená lavica*; *neobsadené územie*; *neobsadený post veľvyslanca*. (Príklady hesiel budeme uvádzať v redukovanej podobe typograficky prispôsobenej textu štúdie. Heslá neprešli úplnou redakciou a editáciou.) Pravými antonymnými adjektívami sa často pomenúvajú vlastnosti, ktoré môžu byť podľa druhu určovaného substantíva fyzikálnymi vlastnosťami aj vlastnosťami zasahujúcimi mentálnu oblasť človeka, napr.: **neobratný** – ktorý nie je pohotový (v pohyboch, vo vyjadrovaní, v správaní)...; op. *obratný*: *fyzicky*, *manuálne neobratný človek*; *neobratná slovná argumentácia*; **nekonvenčný** – **1.** ktorý nie je založený na konvenciách, na ustálených všeobecne prijímaných spôsoboch a vzoroch správania, myslenia...; *netradičný*; op. *konvenčný*: *nekonvenčný štátnik*, *politik*; *nekonvenčné hudba*, *poézia*, *próza*, *fotografia*

2. *i odb.* využívajúci inú techniku, iné postupy...; *syn.* *neštandardný*: *nekonvenčná doprava*; *nekonvenčný spaľovací motor*; *nekonvenčné zbrane*.

3.2. V rozpracovanom slovníku (SSSJ M – O) sa systematicky zachytáva **synonymný vzťah antonymných dvojíc**, ktorých člen zasiahnutý negáciou je tvorený domácou záporovou predponou *ne-* a internacionálnymi predponami, napr. *politický* – op. *nepolitický*, *apolitický*; *kompatibilita* – op. *inkompatibilita*, *nekompatibilita*.

Spracúvanie týchto lexikálnych jednotiek vyžaduje starostlivú koordináciu, lebo ide často o súvzťažňovanie presahujúce práve spracúvaný zväzok tohto viacväzkového slovníka, porov. *netvorivý* – op. *tvorivý*, *kreatívny*, resp. v konkrétnom hesle: **netrénovaný** – **1.** ktorý netrénuje, nenadobúda isté schopnosti tréningom, nacvičovaním: *nekonvenčný človek sa ťažšie vyrovnáva s chorobami*... **2.** ktorý nebol zdokonalovaný pravidelným nacvičovaním, upevňovaním; op. *cvičený*, *školený*: *spievať netrénovaným hlasom*; *Slovinčina je pre netrénované ucho takmer nezrozumiteľná* – pre človeka bez kontaktu s týmto jazykom. Medzi opozitami by sa mohlo doplniť aj pomenovanie *trénovaný*, ktoré už môžeme pokladať za adjektivizované prídavné meno, lebo funguje nielen v prísudkovej, ale aj v prívlastkovej pozícii, a to i nezávisle od lexikálno-gramatickej kategórie vidu ako slovesnej kategórie.

3.3. **Antonymá sú úzko späté so spájatelnosťou slov.** Sémantika antonymických dvojíc umožňuje podľa M. Pisárčikovej takmer úplné splynutie okruhov syntagma-

tického spájania, porov. *vysoký – nízky človek*; podobne to platí o spájateľnosti so substantívami *postava, dom, strom, stĺp, podpätok, hlas, tlak*.

Naša skúsenosť s antonymami tvorenými pomocou predpony *ne-* a ich spájateľnosťou v porovnaní s podobami bez záporovej predpony ukazuje, že situácia je tu zložitejšia. Podľa dokladových častí skúmaných adjektív v SSSJ zameniteľnosť členov potenciálnych antonymných dvojíc v rovnakých kontextoch častejšie platí v prípade minimálnych (dvojslovných) kontextov. V širších kontextových dokladoch spravidla test zameniteľnosti druhým členom antonymnej dvojice neplatí. Kritérium syntagmatickej spájateľnosti s rovnakým okruhom slov nie je splnené. V daných konštrukciách či vo vetných kontextoch sú potrebné aj ďalšie zmeny, aby dané výpovede boli zmysluplné. porov. *mravný – nemravný človek; zakázať nemravnú reklamu// *zakázať mravnú reklamu; nenáhodne* – príznačne pre to, čo je usmerňované vôľou, nie náhodou; *syn. úmyselne, zámerne; op. náhodne: nenáhodne zoradené citáty/náhodne zoradené citáty*; ale *Porovnanie tu uvádzam vonkoncom nenáhodne/*Porovnanie tu uvádzam vonkoncom náhodne*.

Na druhej strane dnešná možnosť používať početné nástroje počítačovej lexikografie (napr. Sketch Engin) umožňuje získavať zo Slovenského národného korpusu (verzia prim.-5.0.-publ.-all; ďalej SNK) pomerne spoľahlivé údaje o spájateľnosti členov skúmaných dvojíc, o frekvenčnom zastúpení adjektívno-substantívnych spojení a o ich výskyte v textoch rozličných žánrov (a komunikačných sfér). Napr. adjektívum *neudržateľný* sa s najvyššou frekvenciou spája so substantívami *stav* (50), *situácia* (63), *hľadisko* (43), o tematických okruhoch svedčia spojenia *neudržateľné pomery, neudržateľné vedecké poznatky, neudržateľné vzorce spotreby, neudržateľné modely poľnohospodárskej produkcie*. Pri protikladnom adjektíve *udržateľný* sú najfrekvenciovanejšie lexikalizované spojenia *trvalo udržateľný život* (259), (*trvalo*) *udržateľný rast* (494), *rozvoj* (175), *udržateľné poľnohospodárstvo* (135). Najbližšiu podobnosť môžeme vidieť v spojeniach *neudržateľné modely poľnohospodárskej produkcie – udržateľné poľnohospodárstvo*. V SSSJ M – O je v heslovej stati **neudržateľný** druhý význam s výkladom „ktorý sa nemôže zachovať v nezmenenom stave pre neprimeranú intenzitu, veľkosť“; *syn. neúnosný; op. prijateľný adekvátne ilustrovaný typickými spojeniami neudržateľný stav, neudržateľná situácia, neudržateľný rozvoj regiónu*. Ako prvý sa zachytáva význam „ktorý alebo ktorého nemožno zastaviť, udržať“...; *syn. nezadržateľný* dokumentovaný spojeniami zo športovej oblasti: *neudržateľný hráč*. Z analýzy stručne ilustrovanej uvedenými príkladmi môžeme odporúčať tento čiastkový záver. Pri rozhodovaní o uvádzaní/neuvádzaní opozitnej jednotky, antonyma, neodporúčame postupovať iba podľa potenciálneho antonyma, ktoré sa lexikografovi môže vo vedomí vybaviť ako slovotvorne stvárnená lexikálna jednotka. Pri lexikografickom spracovaní treba intuitívny postup nositeľa jazyka overovať a korigovať na základe skúmania syntagmatickej spájateľnosti. A keď kritérium zameniteľnosti antonym nie je vo väčšine dokladov (s frekventovaným výskytom) splnené a nie sú splnené ani iné kritériá, opozitum sa neuvádza.

3.4. Za čiastočné antonymá so záporovou predponou *ne-* pokladáme také dvojice, pri ktorých **negačná forma je stredným členom** medzi protikladnými, krajnými členmi radu, nie krajným členom. Interpretácii tejto skupiny venovali bádatelia veľa

pozornosti. *Nemladý* je z hľadiska pojmu *mladý* „starý“, ale z hľadiska pojmu *starý* je to „nie celkom starý“. Tieto pomenovania s predponou *ne-* sa uvádzajú do súvislosti so **stupňovaním adjektív** a najmä s pomerným a tzv. absolútnym stupňovaním (Poldauf, 1948; Morfológia slovenského jazyka 1966, s. 215; Buzássyová, 1979; Buzássyová, 1984; Mluvnice češtiny. 2. Tvarosloví, 1986, s. 79).

Možno poukázať na to, že ani táto skupina nie je z hľadiska významov a funkcie stredných členov skúmaných radov celkom homogénna, aj keď tu môžeme odhaliť podobnosti. Porov. *pekný – nepekný – škaredý; veľký – nevelký – malý; chutný – nechutný – odporný; chuť – nechuť – odpor*. V jednotlivých prípadoch existuje aj úplná antonymická paradigma: *veľký – nevelký – nemalý – malý*. Medzi členmi *nevelký* a *malý* je vzťah synonymie.

Z hľadiska lexikografického spracovania SSSJ sa žiada zdôrazniť, že pri jednotkách vstupujúcich do týchto radov treba venovať osobitnú pozornosť súhre vzťahov synonymie a možných antoným. Porov. heslovú stať **nemladý** – ktorý je v pokročilejšom období života, ktorý už nie je mladý; *syn.* starší, postarší: *nemladý ženích, manželský pár; nemladý, skúsený režisér*. V KSSJ je prídavné meno **nezdravý** spracované ako polysémická jednotka s tromi významami: **1.** chorý; svedčiaci o nedostatku zdravia: *nezdravý otec, nezdravá farba tváre* **2.** škodlivý zdraviu: *nezdravé prostredie* **3.** škodlivý, nepriaznivý: *nezdravé názory, pomery*. Aj v SSSJ M – O sa zachytávajú v podstate tieto tri významy. Iba pri treťom význame sa uvádza aj opozitum: **3.** odlišný od normálneho, prirodzeného, zvyčajného stavu, škodlivý, nežiaduci; *op.* zdravý: *nezdravý vývoj v spoločnosti; nezdravé pomery, praktiky*. Adjektíva s predponou *ne-* vystupujúce ako stredné členy príslušných radov môžu mať štylisticky aj eufemizujúci účinok (pomenovanie *nemladý* použité namiesto slova *starý*). Porov. aj príslovku **nemlado** – so zreteľom na pokročilejšie obdobie života, staršie, postaršie: *vyzerat nemlado; pôsobiť nemlado*. Keďže neoznačujú polárne protikladné pojmy, v zásade sa neuvádza ani nemusí pri nich uvádzať opozitum.

4. Slovnodruhovú hľadisko

Výrazy so (záporovou) predponou *ne-* sa vyskytujú v lexikálnych jednotkách **rozličných slovných druhov**. Obmedzíme sa tu na to, ako sa prejavuje negácia pri menách (substantívach a adjektívach), pri slovesách aj pri niektorých tzv. malých slovných druhoch.

4.1. V podstatných menách *kresťan – nekresťan, športovec – nešportovec*, ale aj v **adjektívach** *slovenský – neslovenský* kladné podoby vyjadrujú istý príznak, spravidla **príslušnosť k nejakej skupine**. Slová so záporovou časticou okrem toho, že tento príznak nemajú, zahrnujú celú ostatnú rovnorodú skutočnosť. Tento vzťah možno podľa M. Pisárčikovej (1974, s. 189) charakterizovať ako **vzťah medzi vyňatou časťou a celkom**. Slová typu *nekresťan, neeurópsky* neobsahujú podľa autorky protikladný význam k slovám *kresťan, európsky*. Ku *kresťan* by protikladom (kontrárnym) bolo slovo *pohan* (vyznávač mnohobožstva); *kresťan* (vyznávač jednobožstva). J. Dolník (2003, s. 100) zdôrazňuje rozdiel medzi komplementárnym formálno-obsa-

hovým binárnym členením a formálnym členením, kde záporné členy v protirečivom vzťahu iba popierajú obsah kladných členov. Substantívam, ako sú *nefajčiar*, *neplatič*, *Neslovák*, *Neslovan*, *neslovenský*, *neslovanský*, *Neróm*, *nerómsky*, *nesionistický* a i. si zaslúžia našu pozornosť aj ako jazykové prostriedky, ktoré reagujú citlivo na život daného jazykového spoločenstva, na mimojazykové podnety vzniku lexikálnych významov. V príspevku (Buzássyová 1986) sme poukázali na to, že pomenovania typu *neplatič* (*nájomného*) vznikli zo spoločenskej potreby pomenovať osobu ako takú, ktorá nevykonáva nejakú potrebnú činnosť. Kým je niečo, istá činnosť v spoločensky uznávanej norme, konkrétne, kým občania platili načas povinné poplatky za nájomné, za rozličné služby, pomenovanie nositeľa/činiteľa deja ako opačnej činnosti (vyhýbanie sa plateniu) nebolo potrebné. Situácia sa zmenila a v súčasnosti sa pomenovanie *neplatič*, *neplatiči* ešte väčšmi lexikalizovalo ako pred štvrtstoročím. V SSSJ (M – O) je heslová stat' **neplatič** bohato doložená, porov. **neplatič** – *admin. profes.* kto sústavne neplatí isté (povinné), stále poplatky, neplatiteľ: *chronickí neplatiči; neplatič nájomného bytu; vymáhať pohľadávky od neplatičov výživného*. V spracovaní sa rešpektuje, že lexikálna jednotka **neplatič** bola utvorená len formálne ako kontradiktorká jednotka. V skutočnosti má svoj vymedzený lexikálny význam, ktorý vznikol ako primárny, nie akoby v (kvázi)protiklade k nejakému *platič; takáto jednotka v slovenčine nejestvuje. Z jazykového hľadiska neutrálnymi protikladnými pomenovaniami (bez negatívneho konotatívneho, pridaného príznaku) sú pomenovania *platiteľ*, *platca*. Majú, resp. budú mať v ďalšom zväzku SSSJ samostatné heslové state. Ako všeobecnejšie širšie pomenovanie je v slovníku spracované aj heslové slovo **neplatiteľ** – 1. kto za niečo neplatí 2. kto nie je povinný platiť isté poplatky; *syn.* neplatca; *op.* platca: *neplatiteľ poistného, dane*.

Charakteristiku vzťah vyňatej časti a celku možno vhodne uplatniť pri interpretácii lexikálnych jednotiek, ktoré vznikli v slovenčine až po spoločenských a politických zmenách po novembri 1989, ako je napr. *neklieň banky*, ale aj *Neslovák*, *Nečech*, ak máme na mysli posudzovanie lexikálnej jednotky *Neslovák* vo vzťahu k názvom zahŕňajúcim v rodovom pojme národ dimenziu „príslušníci iných národov, resp. národností žijúcich v Slovenskej republike“ *Slováci a Neslováci* (t. j. Česi, Maďari, Rusíni, Ukrajinci, Poliáci, Rómovia atď.). *Neklieň banky* je výraz označujúci tých ľudí, ktorí sa uchádzajú o nejakú službu/nejaké služby danej banky, a pritom nie sú jej stálymi klientmi. Toto členenie je relevantné vzhľadom na rozdielne spoplatnenie príslušných služieb pre klientov a neklieňov. Porov. heslo s výkladom pomenovania *klient* v SSSJ H – L: **klient** – (stály) zákazník využívajúci platené služby najmä v oblasti peňažníctva, poisťovníctva, advokácie,... a pod.: *vkłady klientov banky*.

Treba rozlišovať medzi lexikalizovanými (prípadne aj terminologizovanými) jednotkami lexikálnej zásoby a jednotkami, ktoré vznikajú iba z pomenovacej potreby konkrétneho jednotlivého textu (komunikátu) a ktoré nie sú lexikalizované. Takúto funkciu má slovo *nelekári* v štúdiu M. Pisárčikovej (1974) v kontextovom použití *lekári nech tu ostanú, nelekári sa môžu rozísť*. O tomto rozdielne textových nominácií oproti lexikalizovaným nomináciám, pomenovaniam ako lexikálnym jednotkám lexikónu sme písali v štúdiu Buzássyová (2000). V mikrosystéme *vojak* – *nevojak*, *generál* – *negenerál*, *plukovník* – *neplukovník*, ktorý tvoria základné a negované podoby pomenovaní ako nositeľov príznaku, ktorým je vojenská hodnosť, iba dvojica *vojak*

– *nevojak* predstavuje v slovenčine lexikalizovaný paradigmatický vzťah, ktorý nájdeme zachytený vo väčších výkladových slovníkoch (a aj v niektorých prekladových slovníkoch). Pomenovanie *negenerál*, resp. *negeneráli* sa vyskytlo v knihe historika Ľubomíra Liptáka *Storočie dlhšie ako sto rokov: „Generál Rudolf Viest, veliteľ povstaleckej armády a Generál Ferdinand Čatloš, minister národnej obrany,... ešte ako negeneráli agitujú v irkutskom tábore na Sibíri...“* (Lipták, 1998). Nominácia *negeneráli* je typická jednotka použitá pre potrebu konkrétneho textu. Nepatrí do slovníka, do lexikónovej časti slovnej zásoby.

Treba podotknúť aj to, že slovníky jednotlivých jazykov sa odlišujú v tom, aké lexikalizované jednotky so záporovou predponou sa v nich vyskytujú, lebo táto stránka jazykov súvisí s kultúrou a dejinami daných jazykových spoločností. Vo viacerých slovníkoch slovanských jazykov nájdeme lexikálne jednotky typu *neplavec*, *nefajčiar*, *nevodič*, *nemotorista* (resp. ich ekvivalenty), ale napríklad v staropolskej lexike je zastúpená trieda substantív s predponou *nie-*, ktorými sa popiera obsah základového slova, iba veľmi skromne: *nieliterat* – ten, kto nepozná litery, písmená, *niezslachcié* – človek, ktorý nepatrí do šľachtického stavu (Kleszczowa, 1998, s. 87). V Historickom slovníku slovenského jazyka (1991 – 2008) aj v SSSJ nájdeme pomenovanie *nezeman* – kto nie je príslušníkom šľachtického stavu najnižšieho stupňa.

Poznámka k slovu *neslovenský*. V prípade 225 výskytov adjektíva *neslovenský* v SNK prevládajú kontexty svedčiace o význame prostej negácie v zmysle „časť vyňatá z nejakého celku“: *neslovenské obyvateľstvo*, *neslovenský výraz*, *priezviská neslovenského pôvodu*. Ale na internete sa najmä z publicistických textov nájdú aj doklady svedčiace o negatívnom konotatívnom (pridanom) príznanu: *neslovenský* = kto sa nespráva ako (roduverný) Slovák: *vyčítajú nám, že sme neslovenskí* (= zaujatí protislovensky). V publicistických dokladoch v SNK však pomerne frekventovane doložené kontexty adjektíva *protislovenský* (770 x) svedčia o prevažujúcom priamom používaní tohto lexikálneho prostriedku s daným významom, a to najmä z tlače v rokoch, keď sa zvyšuje napätie medzi susednými štátmi v dôsledku rozdielných prístupov najmä k problematike národnostných menšín, 1994 – 1998, 2009). Z tejto skupiny adjektív so záporovou predponou *ne-* je aj adjektívum *neeurópsky* spracované v SSSJ ako dvojjvýznamové: **neeurópsky** – **1.** nepochádzajúci z Európy; netypický pre Európanov: *ľudia z neeurópskych krajín*; *neeurópsky film*; *neeurópske kultúry, tradície* **2.** odporujúci charakteru, cieľom Európskej únie: *neeurópsky zákon*; *poslanci s neeurópskymi postojmi*. Treba však počítať aj s metaforicky posunutým hodnotiacim významom, aký predstavujú spojenia: *patríť do inej, neeurópskej sústavy hodnôt*; *poľutovaniahodný, neeurópsky krok*.

4.2. Dvojice **slovies** so záporovou predponou *ne-* a so svojimi nepredponovými členmi patria spravidla medzi dvojice s nulovou antonymiou (Pisárčiková, 1974, s. 189). Zápor tu nie je lexikalizovaný, neguje význam slova len v konkrétnom použití: *spievať* – *nespievať*, *spať* – *nespať*. Výnimku tvoria prípady *byť* či *nebyť*, kde záporné *nebyť* má svoju presnú sémantiku. J. Dolník (2003, podrobnejšie pozri s. 101) rozlišuje medzi slovesami, ktoré sú v antonymnom vzťahu a ktoré v takomto vzťahu nie sú. Na základe logického testu (*X nevidí, ale je slepý, *X nepočuje, ale je hluchý) dokazuje

že synonymné výrazy nemôžu byť v odporovacom vzťahu, a preto uvedené slovesá (zmyslového vnímania) *nevidieť*, *nepočuť* sú skutočnými antonymami, kým slovesá *nehovoriť*, *nezapamätať si* nie sú antonymami vo vzťahu k nenegovaným náprotivkom. V koncepcii SSSJ sa uplatňuje riešenie, ktoré sa prijalo pri budovaní SNK, že totiž slovesá so záporovou predponou *ne-* sa nezaraďujú medzi antonymá (lexikalizované záporu), ale pokladajú sa za flektívne morfológické tvary daného slovesa. Pokladajú sa teda za súčasť morfológických údajov a doklady na *ne* sa môžu nachádzať v heslových stadiach slovík bez záporovej predpony. Z tohto riešenia vyplýva niekoľko dôsledkov. Treba počítať s tým, že záporové slovesné tvary sa neraz vyskytujú ako súčasť výkladových perifráz vo výkladových zónach rozličných heslových statí, ale samy nemajú a ani nebudú mať „svoju“ heslovú stať. Redigovanie SSSJ M – O ukazuje, že zásady nespracúvať slovesá so záporovou predponou *ne-* ako samostatné heslové state sa kolektív autorov slovníka nemusí vzdávať. Spomenuté záporové slovesá sa vyskytujú vo výkladoch rozličných slovík, alebo rozličných deverbatívnych adjektív utvorených adjektívizáciou pasívnych alebo činných prídavných slov. Porov. časti výkladov hesiel: **odmyslieť si** – vedome *nevziať* niečo do úvahy, obísť niečo, *nedbať* na niečo; **neuznaný** – ktorý nie je prijatý, pretože *nevyhovuje* určitým kritériám: *medzinárodne*, *právne neuznaný štát*. Záporová predpona *ne-* sa môže vyskytovať v modálnych slovesách aj v odvodeninách od modálnych slovík: **neuznášaniashopnosť** *nemožnosť*, *neschopnosť* *záväzne sa uznieť*, *dohodnúť* na niečom...; *op. uznášaniaschopnosť*. O vzťahu kladu a záporu, modálnosti a podiele negácie na modálnosti sú zásadné zistenia M. Grepla, ktoré sformuloval v príspevku na medzinárodnej syntaktickej konferencii v Brne a potom aj v ďalších svojich dielach. „Do modalít v zásadě... nepatří klad a zápor. Klad a zápor jsou primárně záležitostí lexikálního obsahu věty“ (Grepl, 1973; porov. aj Grepl – Karlík 1986, s. 82 – 110; Příruční mluvnice češtiny, 1995; Pavlovič 2003, s. 38).

V KSSJ je pri slovese **zaplatiť** prihniezdovaná aj príslovka v podobe spojenia *je to na nezaplatenie*. Vo Valenčnom slovníku slovenských slovík (na korpusovom základe) (Nižníková, 2006) slovesá so záporovou predponou *ne-* (*neuznať*) takisto neboli predmetom zachytenia v základnom súbore analyzovaných slovík. Záporná podoba slovesa *uznať* sa vyskytuje v dokladoch ilustrujúcich daný negovaný význam: *Okolité krajiny nový štát neuznali. Mexiko muselo uznať nezávislosť Texasu. Rozhodca neuznal gól*. Poznnamenávame, že Valenčný slovník slovenských slovík základným záberom zaradených slovík nadväzuje na KSSJ, aj keď ho dopĺňa na základe SNK.

4.3. Časť prísloviak so záporovou predponou *ne-* svedčí o tom, že v slovenčine existujú popri príslovkových významoch formálne tých istých lexém aj **časticové významy**.

V SSJ je lexikálna jednotka *nepochybné* spracovaná iba ako príslovka prihniezdovaná k adjektívu *nepochybný*: **nepochybný** *adj.* o ktorom nemožno pochybovať; **nepochybné** *prísl.* bezpochyby, isto, istotne. (Nezabúdame, že druhý zväzok SSJ vyšiel v roku 1960, teda ešte pred publikovaním akademickej *Morfológie slovenského jazyka*, 1966, ktorá priniesla J. Mistríkom spracovanú prvú dôkladnú analýzu nového slovného druhu častíc v slovenskej jazykovede.) V KSSJ sa popri spracovaní príslovy stručne zachytáva aj časticový význam: **nepochybné** I. *prísl.* *nepochybné*

zistiť; II. **nepochybne** čast. vyjadruje veľký stupeň istoty: *nepochybne sú tu aj iné dôvody*. V SSSJ sa príslovky a častice spracúvajú ako slovnodruhové homonymá. Medzi jednotlivými druhmi častíc sa registrujú ich synonymné vzťahy, ktoré sú pri tomto slovnom druhu veľmi bohaté: **iste**², **isto**² čast. **1.** vyjadruje istotu, presvedčenie o správnosti tvrdenia; *syn.* istotne, určite, nepochybne: *zajtra isto prídem....***3.** v spojení *celkom iste* zdôrazňuje istotu presvedčenie o platnosti výroku, výpovede; *syn.* istotne, zaiste, bezpochyby.

V tejto slovnodruhovej a tematickej oblasti pociťujeme potrebu doriešiť odkazovanie (mieru odkazov na výrazy iného slovnodruhového zaradenia). Pri predložke **nedbajúc na predl. s A** sa odkazuje na sloveso **dbať** spracované v prvom zväzku slovníka. V heslovej stati substantíva **nepoznanie**, ktoré funguje najmä ako filozofický termín môže byť odkaz na príslovku **nanepoznanie** cez jej pravopisný variant **na nepoznanie**.

5. Lexikalizácia. Terminologické pomenovania

5.1. V predchádzajúcej časti štúdie sme sa už dotkli **problému lexikalizovaných jednotiek a textových jednotiek**. V príkladoch na analyzované protikladné jednotky sa už vyskytli aj lexikálne jednotky, ktoré sa používajú aj ako odborné výrazy. V nadväznosti na A. Jedličku (1964) M. Pisárčiková (1974, s. 187) predpokladá paralelnosť vyjadrovania kladných a záporných podôb dvojíc adjektív na *-tel'ný, -ný*: *riešiteľný – neriešiteľný, ovládateľný – neovládateľný, predvídateľný – nepredvídateľný*, ktorá je potrebná najmä v odbornom vyjadrovaní. Autorka pokladá tieto dvojice za pravé antonymá, aspoň potenciálne je vždy možné vyjadriť protikladnosť, pričom túto protikladnosť obsahuje slovo bez ohľadu na kontext, nezávisle od neho. Preto je pre ne charakteristický aj syntagmatický faktor, splynutie sfér spájania. Uvedené dvojice sú krajnými bodmi logického radu. stredný člen sa dá vyjadriť opisom: *riešiteľný – čiastočne riešiteľný – neriešiteľný*.

Bohatý materiál, na základe ktorého sa spracúva SSSJ (SNK má dnes už vyše 750 mil. textových slov), ukazuje, že pri spracovaní adjektív s sufixom *-tel'ný* a formulovanie ich terminologického významu nemôžeme vždy abstrahovať od kontextu. Medzi týmito pomenovaniami so záporovou predponou *ne-* je nemálo takých, kde negovaná podoba vznikla primárne ako základná, ako terminologické pomenovanie (a to v rozličných odboroch). Majú svoj špecifický terminologický význam. Pri jeho identifikácii kontext zohráva dôležitú úlohu: **neziskový** – neprinášajúci zisk; nezameraný na dosahovanie zisku; *syn.* bezziskový: *neziskový sektor; neziskový projekt; neziskové združenie občanov*. Ekonomický termín *nezisková organizácia* má význam „právnická osoba, ktorej predmet činnosti je poskytovanie všeobecne prospešných služieb (napr. v sociálnej, zdravotnej, vzdelávacej oblasti)“. Kladný člen *ziskový* má sčasti odlišnú spájateľnosť: *ziskové štátne podniky; investícia bude zisková*. Najmä termíny v úzkom zmysle slova nemožno v kontextoch zamieňať kladnými náprotivkami. Preto sa pri spracovaní adjektíva *neziskový* neregistruje pravé opozitum, nejestvuje ako pravidelne sa vyskytujúce opozitum v tých istých kontextoch alebo v ich prevažnej väčšine. Je, pravda, možný výskyt opozitného člena v špecifických kontextoch:

Svojpomocné spolky preberajú úlohy, ktoré neakceptuje ziskový sektor. Porov. aj ekonomický a finančný termín **nezdaniteľný** – nepodliehajúci daňovej povinnosti, ktorý sa nezdaňuje: *nezdaniteľná časť príjmu; Ústavní činitelia majú okrem základného platu aj rôzne formy nezdaniteľných príplatkov (tzv. paušálne náhrady).* (Profit 2001). V polysémických heslách môže byť vhodné uviesť opozitum v niektorom význame: **nenávratný** – **1.** ktorý sa už nijako nemôže vrátiť: *nenávratná minulosť...* **2.** ktorý nie je určený na opätovné použitie; *op. vratný, návratný: nenávratné obaly, téglíky; zvýšené investície na tento rok súvisia s nákupom dvoch nových liniek, jednou na návratné fľaše, jednou na nenávratné plastové fľaše.* (Sme 1996)... **3. fin.** ktorý je koncipovaný tak, aby ho nebolo potrebné vrátiť: *nenávratný príspevok na kúpu bytu; nenávratná dotácia na obnovu historickej budovy.*

Podoba adjektíva so záporovou predponou *ne-* môže mať uvedený protikladný člen dvojice, opozitum, pričom ide o zameniteľnosť obidvoch členov protikladu v istých spojeniach. Heslová stať kladného člena má často bohatšiu a rozmanitejšiu spájatelnosť, nejde o úplne symetrický vzťah a ani o symetrické spracovanie v jednotlivých zväzkoch slovníka, napríklad v slovníku SSSJ M – O zatiaľ pracovná (nezredigovaná) verzia spracovania adjektíva **neexkluzívny** má dva významy, oproti heslu **exkluzívny** v SSSJ A – G (2006) so štyrmi významami: **neexkluzívny** – **1.** ktorý nie je vyberaný, určený výlučne pre niekoho, pre niečo; *op. exkluzívny. neexkluzívny predaj; neexkluzívna dohoda, licencia...* **2.** neodlišujúci sa od zvyčajného každodenného; *syn. bežný, všedný; op. exkluzívny, mimoriadny.* Porovnajme si heslové state ďalšej dvojice: **nekomerčný** – ktorý nie je zameraný na zisk zo svojej činnosti; *neprinášajúci zisk, neziskový, neobchodný; op. komerčný: nekomerčný charakter publikácie, internetovej stránky; nekomerčná kinematografia...*; **komerčný** – **1.** súvisiaci s komerciou, obchodný; *zameraný na zisk: komerčná banka obchodná; komerčná sféra; komerčné prostredie; komerčné ciele podujatia, komerčné médiá...* práv. ekon. *komerčný právnik* právnik, ktorý poskytuje právnickú pomoc fyzickým a právnickým osobám vo veciach súvisiacich s podnikateľskou činnosťou... **2. 2. stupeň -nejší** tvorený so zreteľom na zisk, zameraný jednostranne na obchodnú stránku prinášajúcu zisk: *komerčné sklady; komerčnejšie ladené albumy; nakrúcať komerčné filmy.*

Prechodný vzťah medzi pravou antonymnosťou a zníženým stupňom antonymie, resp. neantonymnosťou sa prejavuje v tom, že nezáporný člen dvojice môže vstupovať do protikladu s viacerými jednotkami. *komerčná televízia – nekomerčná televízia, verejnoprávna televízia.*

Za špeciálny typ protikladných, ale nie antonymických dvojíc možno pokladať podľa I. A. Novikova terminologické dvojice *dokonavý – nedokonavý vid; kovy – nekovy; rovnoramenný – nerovnoramenný trojuholník* (na osobitné vyčlenenie takýchto dvojíc zhodne poukazujú aj M. Pisárčiková, 1974; J. Pavlovič, 1985). V SSSJ niektoré bežnejšie termíny majú podľa koncepcie spôsob spracovania monokolokabilných slov, ktoré majú veľmi obmedzenú spájatelnosť. Iným možným spôsobom je typ spracovania adjektíva *neeuclidovský*; porov. napr. **neeuclidovský** – nerešpektujúci Euklidove axiómy; iný ako euklidovský: *neeuclidovský model sveta; neeuclidovské hladisko*; mat. *neeuclidovská geometria* pripúšťajúca množstvo mysliteľných priestorov.

6. Intenzifikácia a intenzifikácia vyjadrená lexikálno-sémanticky

6.1. Kvalifikačné adjektíva a s nimi významovo súvzťažné príslovky so záporovou predponou *ne-* môžu mať aj lexikálny význam (alebo častejšie jeden z lexikálnych významov) **intenzifikačného charakteru**. Tento význam (a pragmatickú funkciu) chápeme a navrhujeme interpretovať ako význam vyjadrený lexikálno-sémanticky. Nemá totiž osobitnú morfému ako formálny ukazovateľ, je to význam vyjadrovaný v rámci lexikálnych jednotiek so záporovou predponou *ne-* v kooperácii s istými základovými slovami.

Prípadom, na ktoré tu poukážeme, nevenovala osobitnú pozornosť J. Hansmanová (2010, s. 401 – 448), ktorá najnovšie podrobne skúma intenzifikáciu a jej prejavy pomocou jazykových prostriedkov ako súčasti kategórií, ktoré chápe ako gramatické, resp. lexikálno-gramatické (stupňovanie, antonymia, deminutívnosť, augmentatívnosť, synonymia, frazeológia). Východiskovým, a dodajme, určujúcim uhlom autorkinho pohľadu je filozoficko-logické východisko pojmu intenzifikácie vyjadrené v tézach: intenzifikácia sa vzťahuje na kvantitu príznamu (kvality); intenzifikácia sa interpretuje ako silnejší vs slabší prejav kvality. V chápaní gramatickosti spomenutých oblastí mohla autorka nájsť oporu v tom, že veľké syntetické gramatiky jazykov (aj stručnejšie jednozväzkové gramatické príručky) často predstavujú dané jazyky v úplnosti, teda aj s poučením o slovtvorbe a slovnej zásobe, pričom také kategórie ako deminutívnosť či stupňovanie adjektív a prísloviak sú mierou jednotiek, ktorých sa týkajú, produktívnosťou tvorenia skutočne hraničnými oblasťami gramatiky a lexiky. (Výklad o stupňovaní adjektív zahŕňa informáciu o formálnej aj významovej zložke stupňovacích tvarov aj v *Morfológii slovenského jazyka*, 1966; aj v *Mluvnici češtiny*, 1986; príkladom komplexnej príručky môžeme uviesť z nedávnych rokov napríklad *Průruční mluvnice češtiny*, 1995).

J. Hansmanová skúmala jazykový priestor ako fiktívny priestor, v ktorom sa používateľ orientuje na základe svojej kompetencie. Kompetencia sa chápe ako schopnosť orientovať sa v množstve intenzifikátorov (prvkov a štruktúr) a možnostiach ich používania. V priestore intenzifikácie základ a funkciu dominanty má kategória stupňovania – vo vzťahu k nej sa registrujú ďalšie intenzifikačné kategórie.

Gramatickosť antonymickej operácie spočíva podľa J. Hansmanovej v realizácii negačnej operácie záporovými morfémiami *ne-*, *bez-*, *a-/an-*, *anti-*, *pa-*, *tiež-* *kvázi-*, *pseudo-*, napr. *nemotorný*, *bezkonkurenčný*, *asymetria*, *kváziumenie*. Účasť lexikálnych prostriedkov zabezpečujú častice *nie*, *aleže*, *aleba*, *figu*, *kdeže* a *i*. Pri sémantických antonymách je nositeľom intenzifikácie lexikalizovaný prvok (lexikálna morféma), napr. *hovorit' – mlčať*, *zdravý – chorý*, *vojna – mier*. Deminutívnosť a augmentatívnosť sú vysoko gramatické; ako intenzifikátory veľkosti fungujú všetky deminutívne formanty. Synonymná intenzifikácia sa realizuje čisto lexikálne. Intenzifikačné synonymá ako lexikalizované jednotky môžu v kooperácii s lexikálnymi jednotkami (príslovkami miery) vyjadrovať ďalšie gradácie všeobecného parametra: *milý – chutný – neskutočne chutný*, alebo môžu vyjadrovať mieru vo vzťahu k sile: *buchnúť – pleštit' – silno pleštit'*. Gramatickosť frazeologickej intenzifikácie sa prejavuje v modeloch prirovnania s operátorom *ako* (*jasné ako krištál*), v tautologických modeloch (*ako kričí, tak kričí*).

Pri realizácii intenzifikácie lexikálne jednotky kooperujú s morfológickými gramatickými prvkami. Oblasť prieniku s naším vymedzením lexikálno-sémantickej intenzifikácie a intenzifikácie vo vymedzení J. Hansmanovej predstavujú okrem prísloviak miery iné slovné druhy s významom a funkciou adverbii miery, medzi ktorými autorka uvádza napríklad *absolútne, bezhranične, bezmedzne, náramne, nedostatočne, nepomerne, nezmyselne, nevýslovne, nezvyčajne, notoricky, sakramentsky, neodvotálne, nepredstaviteľne* a i. (Hansmanová, 2010, s. 447).

6.2. Kritériá určenia rozsahu pôsobnosti (či skupiny) intenzifikátorov vyjadrovaných lexikálno-sémanticky v našom vymedzení. Ak lexikálna jednotka spĺňa v lexikálnom systéme a vo fungovaní v reči (v komunikácii) funkciu intenzifikátora, spravidla alebo aspoň často ide o dvojznačnú alebo polysémickú jednotku, ktorej základný, primárny význam je do istej (variabilnej) miery sémanticky oslabený, vyprázdnený, a to v prospech všeobecnejšieho lexikálneho významu pri adjektívach „veľký, veľmi veľký“; pri prísloviakoch „veľmi, nadmieru, väčšmi“. Druhým kritériom je spôsob prezentácie, interpretácie tohto intenzifikačného významu – vo výklade sa uplatňujú sémy (príznamy) „značný, majúci veľkú intenzitu, presahujúci normu“ a i. Tretím kritériom je sémantika lexikálnych jednotiek, s ktorými sa daný intenzifikačný význam porovnáva, pri polysémických jednotkách význam, ktorý je v skúmanej lexikálnej jednotke základný a ku ktorému sa pridáva ako ďalší v poradí intenzifikačný význam. Skúsenosť s koncipovaním a redigovaním SSSJ ukazuje tendenciu, že intenzifikačné významy možno zistiť (identifikovať) väčšinou medzi (kvalifikačnými) adjektívami so záporovou predponou *ne-* od základových slov, ktorými sú adjektíva označujúce mieru, rozmer, dimenziu, adjektíva označujúce deje a činnosti z psychickej (mentálnej, kognitívnej) oblasti človeka a adjektíva v užšom zmysle slova hodnotiace. Menej sa možno stretnúť s adjektívami a s nimi súvzťažnými prísloviakmi označujúcimi fyzické vlastnosti človeka.

Príklady: **nepomerný** – 1. nezodpovedajúci istému pomeru, istej proporcii; *syn.* disproporčný, neprímeraný, neúmerený: *nepomerná menšina; nepomerný výskyt žien v politike* 2. presahujúci normálnu mieru, prípustnú hranicu; *syn.* nadmerný: *nepomerná úspora času; nepomerné zdražovanie; vynaložiť na niečo nepomerné množstvo energie; nepomerne* – 1. príznačne pre niečo, čo nezodpovedá istému pomeru, istej proporcii, neúmerne rozdielne; *syn.* disproporčne, neprímerane: *úzka skupina finančníkov vlastní nepomerne veľkú časť svetového kapitálu* 2. presahujúci prípustnú, normálnu hranicu; oveľa väčšmi; *syn.* nadmieru, značne: *kúpiť si nepomerne účinnejší prací prášok; týndžerky sú nepomerne vyspelejšie ako ich chlapčenský rovesníci; pracuje nepomerne výkonnejšie ako jej kolegyně; neoceniteľný* – majúci takú hodnotu, ktorá sa ani nedá vyjadriť, mimoriadne cenný; *syn.* nedoceniteľný, nezaplatiteľný, vzácny: *poskytnúť niekomu neoceniteľnú službu; naberat' neoceniteľné skúsenosti; neuveriteľný* – 1. ktorý sa vyznačuje takými neobvyklými vlastnosťami, znakmi, ktoré sa nepredpokladali, neočakávali, ktoré udivujú; *syn.* prekvapujúci, neočakávaný: *neuveriteľný príbeh, zážitok; kaskadérske kúsky boli čoraz náročnejšie, neuveriteľnejšie* 2. ktorý sa vyznačuje veľmi veľkými rozmermi, veľmi veľkou intenzitou, hodnotou, obrovský: *neuveriteľný neporiadok; neuveriteľná vytrvalosť, bojovnosť; neuveriteľné množstvo ľudí;*

netušený – 1. ktorý nebol dosiaľ známy, a preto neočakávaný; *syn.* nepredvídaný: *odhaliť netušenú minulosť partnera* 2. ktorý prevyšuje zvyčajnú mieru, intenzitu; *syn.*

neuveriteľný, neobyčajný: *netušené možnosť cestovania, kríza netušených rozmerov; netušene* – 1. bez predchádzajúceho očakávania, nepredpokladajúc výskyt niečoho; *syn.* nepredvídane, nečakane: *pri splavovaní dreva sa netušene a nechtiac okúpali; netušene našla na dovolenke životnú lásku* 2. prevyšujúci zvyčajnú mieru, intenzitu; *syn.* prekvapujúco, neuveriteľne: *klásť netušene tuhý odpor; vyvolať netušene silné emócie*

Funkciu intenzifikátora, prostriedku s intenzifikačným významom môžu mať aj posunuté významy s posunom (metaforického charakteru) utvorené (a lexikalizované) od iných typov adjektív, porov. **nekresťanský** – 1. iný než kresťanský: *nekresťanské náboženstvá* 2. *expr.* priveľký, nesmierny: *zaplatiť nekresťanské peniaze; nekresťanské vydieranie nemilosrdné* (výklad podľa KSSJ).

Intenzifikačné významy lexikálnych jednotiek sa zachytávajú už v prvom zväzku Slovníka súčasného slovenského jazyka A – G (2006), porov. výklad heslových statí *bezbrehý, bezbreho, bezbrehosť: bezbrehý* – 1. vyzerajúci ako bez brehov, ktorý nemá viditeľné brehy; *syn.* rozsiahly: *bezbrehý tok, bezbrehé močariská* 2. (často o ľudských pocitoch) ktorý nemá vymedzenú šírku, mieru, trvanie, platnosť; *syn.* neobmedzený, nekonečný: *bezbrehý čas; bezbrehý smútok; bezbrehá dôvera, oddanosť* 3. (s negatívnym postojom hovoriaceho k označenému javu) prekračujúci únosnú mieru; *syn.* prehnaný: *bezbrehá bezočivosť* (L. Lenčo); *bezbrehá privatizácia vodných zdrojov na Slovensku*.

6.3. Možno uvažovať o tom, do akej miery sú intenzifikačné významy citovaných lexikálnych jednotiek **expresívne** v miere, že ich prezentácia vyžaduje uvádzanie kvalifikátora expresívnosti.

Nazdáваме sa, že sémantická odvodenosť v rámci adjektív (a prísloviek) typu *netušený, neoceniteľný, nepredstaviteľný; netušene, neoceniteľne, nepredstaviteľne* predstavuje významový posun do abstraktnejšej polohy, ktorý nie je veľmi výrazný, „ostrý“, preto nevyžaduje uvádzať kvalifikátor expresivity, expresivita sa nepocituje veľmi silno. Pri výraznejšom posune oproti primárnemu významu základového slova, ako je to v prípade lexikálnej jednotky *kresťanský – nekresťanský*, sa expresivita uvedomuje výraznejšie. Je namieste, keď sa vyznačuje v lexikografických opisoch; porov. 2. význam tohto slova aj v SSSJ M – O: **nekresťanský** – 2. *expr.* ktorý je prehnaný čo do veľkosti, intenzity: *nekresťanský úrok; zaplatiť nekresťanské peniaze; žiadať za byt nekresťanské nájomné*.

Obdobná je expresivita pri opise významu súvzťažnej príslovky *nekresťansky*, porov. v spojeniach typu *Televízny prijímač bol drahý, nie najkvalitnejší, ale stál nekresťansky veľa peňazí*. (P. Andruška).

K adjektívam s intenzifikačným významom sa antonymá (opozitá) neuvádzajú a ani nie je podľa nášho názoru náležité ich uvádzať, lebo význam a funkcia týchto adjektív v komunikácii je prasto iná ako negačná. Skôr prichádza do úvahy uvádzanie synonym.

7. Pragmatický, postojový význam

Adjektíva so záporovou predponou *ne-* môžu v slovenčine spĺňať aj **eufemizačný význam** funkciu. Hovoriaci používajú isté výrazové prostriedky na zmierenie, zjemnenie nepriaznivej, neželanej skutočnosti, o ktorej je reč. Existujú ja-

zykové prostriedky, ktoré sú plne lexikalizované, aj také, ktoré sa ešte plne nelexikalizovali. Prvú skupinu reprezentujú príklady, ktoré sa už neraz analyzovali pri skúmaní antonymie. M. Pisárčiková (1974) hovorí o prípadoch, ktoré majú inú stylistickú hodnotu alebo odtienok v porovnaní s antonymami s rozličnými koreňmi, hoci sa s nimi významovo kryjú: *vidomý – nevidomý*, *nevidiaci – slepý* (*nevidomý* je zastaraný výraz a *zjemňujúce* slovo oproti *slepý*), *nepočujúci* je *zjemnené* vyjadrenie oproti *hluchý*. V niektorých kontextoch má *zjemňujúci* odtieň aj adjektívum *nezdravý* oproti *chorý*. V posledných desaťročiach pribúda spoločensky *zjemňujúcich* pomenovaní. porov. *knižnica pre nevidiacich a slabozrakých*; oproti *invalid*, *invalidi*, kde eufemizmom je opisné vyjadrenie *človek/ludia so zníženou pracovnou schopnosťou*. Eufemizačná tendencia sa dnes uplatňuje aj ako tzv. politická korektnosť, keď sa namiesto protikladných výrazov, v ktorých sa väčšmi exponuje národnostné a rasové hľadisko dáva prednosť všeobecnejšiemu pomenovaniu *neprispôsobiví občania* (v tomto názve sa národnostný či rasový motív neuplatňuje, ako je to v pomenovaniach *Róm – rómsky*). Už sme spomenuli, že adjektívum *neslovenský* môže mať eufemizujúci účinok oproti priamejšiemu, explicitnejšiemu vyjadreniu pomenovanej vlastnosti *protislovenský*. (V päťdesiatych rokoch v začiatkoch studenej vojny podobnú funkciu zohrávalo v Spojených štátoch severoamerických pomenovanie *neamerický*, porov. *Výbor pre neamerickú činnosť*. Neamerickou činnosťou sa mysleli aktivity ľudí, ktorých bolo možné obviňovať z činnosti namierenej proti USA, teda z protiamerickej činnosti.)

Záver

1. V štúdiu sme sa usilovali o prehľadnejší pohľad na jeden výsek z výrazových prostriedkov na vyjadrovanie antonymie s cieľom prispieť k premyslenejšiemu uvádzaniu antonymných (opozitných) výrazov ako doplňujúcej formálno-sématickej informácie, dokresľujúcej popri údajoch o synonymii výklad lexikálnych jednotiek (heslových statí) utvorených pomocou negačného prefixu *ne-* v rozpracovanom treťom zväzku Slovníka súčasného slovenského jazyka M – O. Ide teda primárne o lexicologicko-lexikografickú štúdiu v rámci koncepcie slovníka zohľadňujúcej funkčno-štruktúrny aj komunikačno-pragmatický prístup. Tento prístup si vyžiadala istá reinterpretácia typov antonymie v doterajšej literatúre.

2. Pri centrálnom type (pri pravých antonymách) odporúčame dôsledne uvádzať opozitá. Treba však počítat' s tým, že intuitívny prístup nositeľa jazyka (vrátane lexikografa) je užitočné overovať a korigovať testovaním syntagmatickej spájateľnosti. Kritérium zameniteľnosti členov antonymnej dvojice v tých istých kontextoch väčšmi platí v minimálnych (dvojslovných) kontextoch ako pri širších kontextoch.

3. Treba počítať so škálovitým prechodom – neostrou hranicou medzi pravými antonymami a čiastočnými antonymami – ostrý logický protiklad kontrárny : kontradiktorkický sa v prirodzenom jazyku „rozostruje“, znejasňuje takými javmi, ako sú polysémia, metaforické prenosy, synonymia.

4. Najmä z lexikografického hľadiska je užitočné rozlišovať medzi lexikalizovanými jednotkami lexikálnej (slovnej) zásoby, lexikónu, a na druhej strane nelexikalizo-

vanými jednotkami, ktoré vznikajú z nominačnej (pomenúvacej) potreby konkrétneho jednotlivého textu.

5. Pri fungovaní lexikálnych jednotiek so záporovou predponou (negačným sufixom) *ne-* v textoch ako intenzifikačných prostriedkov je v popredí ich funkcia ako intenzifikátorov, antonymnosť sa nevedomuje, nie je prítomná (alebo je iba na formálnom pláne), a teda nie je vhodné ju registrovať. Podobná situácia je pri prostriedkoch používaných v eufemizačnej funkcii.

6. Ak je v lexikografickom diele, akým je výkladový slovník národného jazyka, v koncepcii aj v realizácii dôležité, ako autori pristupujú k synonymii, o to dôležitejšia je súhra synonymie a antonymie (opozitnosti) pri spracúvaní heslových statí, ktoré predstavujú lexikálne (a medzi nimi slovotvorné, resp. morfológické) prostriedky negácie.

Literatúra

BUZÁSSYOVÁ, Klára: Príspevok k vymedzeniu neutralizácie v kategórii stupňovania.

In: Jazykovedný časopis, 1979, roč. 30, č. 1, s. 6 – 17.

BUZÁSSYOVÁ, Klára: Špecifiká stupňovania v slovenčine. In: *Studia Academica Slovaca*. 13. Prednášky XX. letného seminára slovenského jazyka a kultúry. Red. J. Mistrík. Bratislava: Alfa 1984, s. 135 – 150.

BUZÁSSYOVÁ, Klára: Novšie názvy negatívnych vlastností a stavov. In: *Kultúra slova*, 1986, roč. 20, č. 10, s. 335 – 340.

BUZÁSSYOVÁ, Klára: Syntagmatická a paradigmatická dimenzia jednotiek nominácie. K problému textových a lexikónových jednotiek nominácie. In: *Słowotwórstwo i inne sposoby nominacji*. Materiały z 4 konferencji Komisji Słowotwórstwa przy Międzynarodowym Komitecie Slawistów. Katowice 27 – 29 wrzesnia 2000 r. Redaktorzy tomu K. Kleszczowa – L. Selimski. Katowice: Wydawnictwo Gnome 2000, s. 17 – 23.

ČERMÁK, František: Jazyk a jazykověda. Přehled a slovníky. 2. vyd. Praha: Pražská imaginace 1997. 460 s.

DOLNÍK, Juraj: Lexikológia. Bratislava: Univerzita Komenského 2003. 236 s.

DOLNÍK, Juraj: Orientácia v jazykovom priestore. In: *Jazykovedný časopis*, 2007, roč. 58, č. 1, s. 11 – 19.

FILÍPEK, Josef – ČERMÁK, František: Česká lexikologie. Praha: Academia 1985. 284 s.

GREPL, Miroslav: K podstatě modálnosti. In: *Otázky slovanské syntaxe*. III. Red. J. Burián. Brno: Universitas J. E. Purkyně 1973, s. 23 – 38.

GREPL, Miroslav – KARLÍK, Petr: *Skladba spisovné češtiny*. 1. vyd. Praha: Státní pedagogické nakladatelství 1986. 474 s.

HANSMANOVÁ, Júlia: Intenzifikácia. In: *Morfologické aspekty súčasnej slovenčiny*. Ed. J. Dolník. Bratislava: Veda 2010, s. 401 – 448.

Historický slovník slovenského jazyka. VII. Z – Ž. Dodatky. Red. M. Majtán. Bratislava: Veda 2008. 542 s.

JEDLIČKA, Alois: K adjektivům s významem možnosti (nemožnosti) v nové spisovné češtině. In: *Prace filologiczne*. 13. zv. 2. Red. H. Koneczna – S. Skorupka. Warszawa: Państwowe Wydawnictwo Naukowe 1996, s. 199 – 207.

- Krátky slovník slovenského jazyka. 4. doplnené a upravené vydanie. Red. J. Kačala – M. Pisárčiková – M. Považaj. Bratislava. Veda 2003. 996 s.
- KLESZCZOWA, Krystyna: Staropolskie kategorie słowotwórcze i ich perspektywiczna ewolucja. Rzecowniki. Katowice: Wydawnictwo Uniwersytetu Śląskiego 1998. 186 s.
- LIPTÁK, Ľubomír: Storočia dlhšie ako sto rokov. Bratislava: Kalligram 1998. 356 s.
- Mluvnice češtiny. 2. Tvarosloví. Red. svazku: M. Komárek – J. Kořenský – J. Petr – J. Veselková. Praha: Academia 1986. 536 s.
- Morfológia slovenského jazyka. Red. J. Ružička. Bratislava: Vydavateľstvo SAV 1966. 896 s.
- NĚMEC, Igor: O lexikálním záporu v češtině. In: Slovo a slovesnost, 1969, roč. 30, s. 337 – 346.
- NIŽNÍKOVÁ, Jolana: Valenčný slovník slovenských slovies. 2. zväzok (na korpusovom základe). Prešov: Prešovská univerzita v Prešove, Filozofická fakulta 2006. 166 s.
- NOVIKOV, L. A.: Antonimija v russkom jazyke (Semantičeskij analiz protivopoložnosti v leksike). Moskva. Izdatel'stvo Moskovskogo universiteta 1973. 290 s.
- PAVLOVIČ, Jozef: Prostriedky negácie v spisovnej slovenčine. In: Slovenská reč, 1985, roč. 50, č. 5, s. 265 – 274.
- PAVLOVIČ, Jozef: Negácia v jednoduchej vete. Bratislava: Slavistický kabinet SAV 2003. 138 s.
- PISÁRČIKOVÁ, Mária: Prefixálne antonymá so záporovou časticou ne- v slovenčine. In: Jazykovedné štúdie. 12. Peciarov zborník. Red. J. Ružička, J. Oravec. Bratislava: Veda 1974, s. 183 – 192.
- POLDAUF, Ivan: O podstatě stupňování přídavných jmen. In: Časopis pro moderní filologii 1948, roč. 31, s. 109 – 122.
- Průruční mluvnice češtiny. Eds. P. Karlík – M. Nekula – Z. Rusínová a kolektiv autorů Ústavu českého jazyka Filozofické fakulty Masarykovy univerzity v Brně. Praha: Nakladatelství Lidové noviny 1995. 800 s.
- Slovenský národný korpus – prim.-5.0.-public.-all. Bratislava: Jazykovedný ústav Ľ. Štúra SAV. Dostupný z WWW:<http://korpus.juls.savba.sk>.
- Slovník slovenského jazyka. 6. zv. Red. Š. Peciar. Bratislava: Vydavateľstvo SAV 1959 – 1968.
- Slovník súčasného slovenského jazyka. A – G. [1. zv.]. Red. K. Buzássyová – A. Jarošová. Bratislava: Veda 2006. 1134 s.
- Slovník súčasného slovenského jazyka. H – L. [2. zv.]. Red. A. Jarošová – K. Buzássyová. Bratislava: Veda 2011. 1088 s.

Sloveso ako lexikografická výzva¹

Jana Hašanová, Nicol Janočková
Jazykovedný ústav Ľudovíta Štúra SAV

Verb as a Lexicographic Challenge

This paper deals with partial grammatical issues of verbs, which seem to be problematic at the final integration works on the second volume of the *Dictionary of the Contemporary Slovak Language* (2011). The following topics are discussed: the question of recording imperative with processive verbs of a type *chladnúť* (= lose temperature); the participial passive of reversible and irreversible verbs (e.g. grammatical homonym *informovaný* could mean „he/she was informed by someone“ as well as „he/she informed themselves“); the variability of grammatical forms of verbs with the conjugation types *česať* – *češem* and *chytat'* – *chytám*; impersonal verbs and impersonal forms of personal verbs; the relationship between the reversible and irreversible verbs, which could be manifested only as a reversible lexeme (reflexive tantum *kamarátiť sa* = to be friends with someone) or as two independent lexemes (*halit'* – *halit' sa*; *kynožiť* – *kynožiť sa*; *hnevať* – *hnevať sa* = make sb. angry – be angry) or as a reversible form of irreversible verb (reversible passive of irreversible verb, e.g. *krátiť* – *krátili sa im dôchodky* = their pension was cut).

1. Úvod

Sloveso ako „kráľovský“ slovný druh je zaujímavé z gramatického, lexikálneho aj slovtvorného hľadiska. Na samotný opis sloviess vo výkladovom slovníku vplýva bohatosť slovesných gramatických kategórií a ich realizácií v podobe gramatických tvarov, prehodnocovanie tvarov na samostatné lexikálne jednotky či na syntaktické konštrukcie. V našom príspevku sa zameriame na niektoré problémové javy v gramatickom a lexikálnom spracovaní sloviess vo výkladovom type slovníka, konkrétne v 2. zväzku Slovníka súčasného slovenského jazyka. H – L (2011; ďalej SSSJ2).

Závery, ktoré sme pri lexikografickom spracovaní prijali, vyplynuli z editačného procesu sloviess v SSSJ2, ktorý prebiehal v letných mesiacoch v roku 2010. Editácia, pri ktorej sa vykonalo niekoľko rôznorodých analýz všetkých sloviess druhého zväzku, ukázala, ako sa slovesá prejavujú v komunikácii, v ktorých prípadoch dochádza k variantnosti gramatických tvarov, ktoré gramatické tvary sú potenciálne, ale ne-realizovateľné, ako v mnohých prípadoch splyývajú gramatické tvary s osobitnými slovesnými lexémami (v *paneláku sa ešte nekúri* – gramatický tvar slovesa *kúriť*; z *komína sa kúri dym* – zvrtné sloveso *kúriť sa*) a pod.

Budeme sa venovať týmto témam: priebehu samotnej editácie SSSJ2, spôsobu zachytenia imperatívnych gramatických tvarov, participiálnemu pasívu pri zvrtných a nezvrtných slovesách, variantnosti gramatických tvarov sloviess pri slovesnom type *česať/chytat'*, neosobnosti, zvrtným a nezvrtným slovesám.

1 Príspevok bol vypracovaný v rámci grantového projektu VEGA č. 2/0018/11 *Slovník súčasného slovenského jazyka – 4. etapa (koncipovanie a redigovanie slovníkových hesiel a s tým súvisiaci lexikologicko-lexikografický výskum)*.

2. Editácia SSSJ2

Prvou fázou editácie bolo zjednocovanie spracovaných hesiel podľa slovných druhov, ktoré vykonávali tri dvojice editoriek. Každá dvojica mala pridelený konkrétny slovný druh. Úlohou editačnej dvojice, ktorá zjednocovala slovesá, bolo kontrolovať heslár, výslovnosť, gramatický aparát, variantnosť, aspekt, väzbu, výklad, terminológiu, frazeológiu a pod. Nové gramatické otázky riešili editorky s doc. PhDr. M. Sokolovou, CSc., prostredníctvom hlavnej redaktorky Mgr. A. Jarošovej, CSc. Pri slovesách išlo predovšetkým o tieto slovesné témy – genus verbi, imperatív, aspekt, neosobnosť a reflexívnosť. Autorky ešte pri samotnom koncipovaní hesiel vychádzali z existujúcej kodifikácie (Pravidlá slovenského pravopisu, 2000, ďalej PSP; Krátky slovník slovenského jazyka, 2003, ďalej KSSJ; Pravidlá slovenskej výslovnosti, 1996, ďalej PSV), no ak sa na základe dokladového materiálu ukázali sporné prípady (napr. možná variantnosť gramatického tvaru, výslovnosti, príp. prevaha dosiaľ nekodifikovaného tvaru), tie osobitne posudzovala kodifikačná komisia.

V druhej fáze editácie sa zjednocovali heslá zoradené podľa abecedy, pričom pozornosť sa sústredila na zapracovanie pripomienok odborných konzultantov (konzultovalo sa 105 odborov so 75 špecialistami), na kontrolu citátových dokladov, zjednocovanie formy pri uvádzaní prirovnaní, frazém, parémii, lexikalizovaných spojení a dokladov z oblasti etikety, na kontrolu homonymie, variantnosti a iných čiastkových prác. Osobitne dôležité bolo posúdenie diela zo strany recenzentov a následné spracovanie ich poznámok a komentárov. SSSJ2 recenzovali: PhDr. L. Králik, CSc., prof. PhDr. J. Mlacek, CSc., prof. PhDr. S. Ondrejovič, DrSc. a doc. PhDr. M. Sokolová, CSc.

3. Gramatický aparát slovesného heslového slova

Aký gramatický aparát sloves je ideálne zachytávať? SSSJ sa snaží ukázať pomerne široké spektrum gramatických tvarov. Na porovnanie uvádzame príklad spracovania slovesa *háčkovať* v Slovníku slovenského jazyka (1959 – 1968; ďalej SSJ), PSP, KSSJ a SSSJ2:

SSJ:

háčkovať, -uje, -ujú *nedok.* (čo i *bezpredm.*)

PSP:

háčkovať -uje -ujú *nedok.*

KSSJ:

háčkovať *nedok.*

SSSJ2:

háčkovať -kuje -kujú -kuj! -koval -kujúc -kujúci -kovaný -kovanie *nedok.*

SSSJ má oproti uvedeným dielam najbohatšie deklarované gramatické tvary slovesa: 3. os. sg., 3. os. pl., imperatív (2. os. sg.), préteritum, prechodník, činné prídavné prí-

tomné, trpné prídavné, slovesné podstatné meno a vid. Préteritum v ženskej podobe sa zachytáva pri slovesách patriacich do slovesného typu *chudnúť*, pretože v ženskej podobe préterita dochádza k zánikovej vokalickej alternácii (*chudnúť* – *chudol*, *chudla*).

Zápis gramatického aparátu má jednotnú podobu – pri skracovaní sa uvádza časť koncovky s poslednou nemennou hláskou, nejde teda o gramatickú či morfológickú koncovku. V mnohých prípadoch sú uvedené tvary len potenciálnou podobou. Ak však z nejakého dôvodu sloveso netvorí vybraný tvar, neuvádzame ho, napr. perfektiwa nemajú činné prídavné prítomné (*urobiť* – **urobiaci*).

3.1. Imperatív

Uvádžanie imperatívnej formy prináša so sebou otázku, či treba tento tvar uvádzať pri procesných slovesách (typ *kamenieť*), resp. pri procesných slovesách, kde sa tvorí len v niektorom z významov (typ *chladnúť*), príp. či ho pri niektorých slovesách netreba uvádzať len v zápornej podobe (typ *hlivieť*). Ide o skupinu slov patriacich k slovesnému typu *rozumieť*

(**-ieť**): *hlivieť*, *horieť*, *hrbatieť*, *hrčkavieť*, *hrdzavieť*, *hrudkovatieť*, *hubovatieť*, *chabieť*, *chľpatieť*, *chorľavieť*, *chrastavieť*, *chúlostivieť*, *intenzívniť*, *jalovieť*, *jasniť*, *jemniť*, *kalniť*, *kamenieť*...; a k slovesnému typu *chudnúť* (**-úť**): *hladnúť*, *horknúť*, *hustnúť*, *chabnúť*, *chladnúť*...

kamenieť -nie -nejú -nej! -nel -nejúc -nejúci -nenie nedok.

0. (o ľudskom tele) stávať sa meravým, nehybným (od neočakávaných, silno pôsobiacich pocitov); syn. tuhnúť, drevenieť: *ruky, nohy mu kamenejú*; *žalúdok jej nebezpečne kamenie oťažieva*; *Cítil, ako mu kamenie tvár*; *musel vyzerat', akoby sa v ňom zastavili všetky životné pochody*. [E. Hlatký];

dok. -> **skamenieť**

chladnúť -dne -dnú -dni! -dol -dla -dnúc -dnúci -dnutie nedok.

1. stávať sa chladnejším, chladným, studeným, strácať teplo; syn. ochladzovať sa: *na stole chladne večera*; *voda, káva chladne*; *chlapcovi chladli ruky*; *vzduch čoraz väčšmi chladol*; *Priložil som na chladnúcu pahrebu raždie a pofúkal som uhlíky*. [V. Mináč]; *Smrť je meravá a studená, priateľu, telo chladne a človek tichne na večnosť*. [P. Andruška]

2. (o citoch) stávať sa menej intenzívnym; syn. slabnúť, ochabovať: *jej láska a túžba chladnú*; *naše nadšenie pomaly chladlo*; *Intuitívne vytušila, že jeho vzťah k Irene chladne*. [H. Dvořáková];

dok. k 1 -> **vychladnúť**, k 2 **ochladnúť**

hlivieť -vie -vejú (ne)hlivej! -vel -vejúc -vejúci -venie nedok.

expr.

1. tráviť čas ležaním, vylihovať; syn. vyvalovať sa: *vstaň, nehlivej tol'ko!*; *Hlivel v posteli až do pol deviatej*. [P. Jaroš]; *Keď zistil, že môže hlivieť aj celý deň, už mu na piatich minútach tak nezáležalo*. [A. Vášová]; *Mohol by si i dlhšie poležať, ale práve také hlivenie bolesti ešte pomôže*. [J. Štiavnický]

2. tráviť čas v nečinnosti, nerobiť nič užitočné, zaháľať, leňošiť: *pohodlné hlivenie pri mori*; *Len občas ho zamrzí, že hlivie v Hradci, a pritom treba urobiť tol'ké veci*. [P.

Štrelinger]; *Hriech by bolo zarábať doláre iba jedným párom rúk, keď doma hlivie ďalší pár, synove.* [H. Zelinová]

V niektorých prípadoch uvádzame negovanú formu imperatívu s morférou (*ne-*) v zátvorke, ktorá zmierňuje dojem „nabádania k zlému“, napr. (*ne*)*hlivej!*; (*ne*)*luž!*/*(ne)luhaj!*; (*ne*)*jeduj sa!*; (*ne*)*kmáši!*; (*ne*)*hrb sa!*. Negovaná forma imperatívu procesných slovík sa častejšie využíva v prenesenom význame, napr.

Prosím, láska moja, nechladni, miluj ma.

(zdroj: <http://chiarasnicket.blog.cz/1010/nezladam-to>)

Pri uvádzaní tvaru imperatívu slovík typu *horárčiť*, *hubárčiť*, *chuligánčiť*, *chatárčiť* (4. intencný typ, 2. valenčný typ, t. j. subjektovo-bezobjektová štruktúra) so zakončením tvarotvorného základu sonóra + neznelá šumová spoluhláska vychádzame z výsledkov analýzy M. Sokolovej – I. Bónovej (2008), kde autorky odporúčajú dvojtvar **-i/-o**, ale konštatujú prevahu podoby s imperatívnou morférou **-i**:

„Vzhľadom na nízku (takmer nulovú) frekvenciu konfigurácií *-jč*, *-nč*, *-nš* na konci slova je len prirodzené, že imperatív od uvedených slovík sa tvorí predovšetkým s *-i* a nulová koncovka je len jej fakultatívnym variantom“ (op. cit., s. 276).

Z dôvodu veľmi nízkej frekvencie imperatívu tohto typu slovík (*Chod' často do lesa, oddychuj, hubárči!*) sme sa rozhodli uvádzať nevariantnú podobu imperatívu s morférou **-i** aj napriek riziku homonymie s 3. os. sg. prézenta (*on roky hodinárči, chalupárči*).

3.2. Participiálne pasívum pri zvrtných a nezvrtných slovesách

Na rozdiel od slovesných podstatných mien zo zvrtných slovík, kde morféma *sa* môže v danom substantíve zostať (*kajať sa* – *kajanie* aj *kajanie sa*), pri participiálnom pasíve morféma *sa* vždy vypadáva.

Participiálne pasívum sa obyčajne dá utvoriť od:

- osobných tranzitívnych alebo osobných intranzitívnych slovík (pretože majú objekt, ktorý sa pri pasíve presúva do pozície subjektu, napr. *uverejňovať správy v regionálnych denníkoch* → *správy sú uverejňované v regionálnych denníkoch*; *Peter vychudol* → *Peter je vychudnutý*),
- od dokonavých objektových slovík (*uhasiť oheň* → *oheň je uhasený*).

Participiálne pasívum sa môže vzťahovať iba na niektorú z lexii slovesa, napr. *kandidovaný* „ktorého kandidovali na/za niekoho“, nie „*ktorý kandidoval/kandiduje“, teda tvorí sa k 2. významu slovesa *kandidovať*:

kandidovať [-d-] *-duje -dujú -dují! -doval -dujúc -dujúci -dovaný -dovanie* nedok. i dok.

1. (za koho; na čo; do čoho; 0) prejavovať, prejavíť záujem, uchádzať sa o nejaký úrad, funkciu, hodnosť, členstvo; byť kandidátom: *k. za prezidenta republiky, za poslanca, za starostu obce, za rektora; k. na post primátora; k. do akademického senátu; k. ako nezávislý, občiansky kandidát; v komunálnych voľbách nekandiduje; Celkovo aj do regionálnych parlamentov kandidovalo veľmi málo žien.* [Slo 2002]

2. (koho) navrhovať, navrhnúť niekoho na zvolenie: *k. člena strany za poslanca; k. do Európskeho parlamentu aj nezávislého poslanca; zástupca kandidovaný opozíciou*

Tvorenie participiálneho pasíva nie je viazané absolútne na tranzitívnosť slovesa (napr. tranzitívne slovesá *chopiť*, *hrnúť* netvorí participiálne pasívum). **Netvorí sa od:**

- statických slovies: *Mlieko obsahuje vodu.* – **Voda je obsahovaná v mlieku.*,
- neosobných slovies 7. a 8. intencného typu (*smáďí ma, prší*), pretože nemajú agens, ktorý by sa mohol odsunúť zo svojej pozície (**som smádený, pršaný*).

Pri súvzťažných dvojiciach nezvratné – zvrtné sloveso sa vyskytujú tri prípady tvorenia participiálneho pasíva:

a) participiálne pasívum je homonymné pri nezvratnom a zvrtnom slovese:

informovať – informovať sa

informovaný:

– „ktorého *informovali*“

– „ktorý *sa informoval*“

Iné príklady: *latinizovať – latinizovať sa, chytiť – chytiť sa*

b) participiálne pasívum je iba pri nezvratnom slovese:

hanobiť: hanobený „ktorého *hanobia/hanobili*“

hanobiť sa: 0

Iné príklady: *jatrit’ – jatrit’ sa, chváliť – chváliť sa, hubiť – hubiť sa*

c) nezvratné a zvrtné sloveso sú bez participiálneho pasíva:

hrnúť – hrnúť sa, jasniť – jasniť sa, chopiť – chopiť sa

4. Variantnosť gramatických tvarov pri slovesnom type *česať/chytať*

Pri slovesách tohto typu sa prejavuje dynamická tendencia prechodu z časovacieho typu *česať* k typu *chytať*. Sú to napr. slovesá: *hatať, hltat’, hrkotat’, hrmotat’, chichotat’ sa, chechotat’ sa/chechtat’ sa, chľastat’, chleptat’, chmátat’, chrapotat’, chrochtat’, chrupotat’, chvastat’ sa, chvátat’, jagotat’ sa, jachtat’, klapotat’, klepotat’, klokotat’, klopotat’, koktat’, kutat’, kutat’ sa, kyvotat’, kyvotat’ sa, lalotat’, lapotat’, ligotat’ sa, lulotat’*.

Podľa Morfológie slovenského jazyka (1966, s. 446; ďalej MSJ) „...koreň slovies vzoru *česať – češem*... má rozličné alternácie koncových spoluhlások. Základný konsonant je v infinitívnom kmeni. Sú to tieto alternácie:

t – c: *blikotat’ – blikoce, motat’ sa – moce sa,*

d – dz: *vládat’ – vládze,*

k – č: *skákat’ – skáče, plakat’ – plače,*

c – č: *sácat’ – sáče,*

dz – dž: *hádzat’ – hádže, sádzat’ – sádzže,*

s – š: *písať – píše, česať – češe...*“

Pre slovenčinu ako flektívny jazyk je charakteristická tendencia po stálosti, transparentnosti tvarotvorného základu (porov. slovenské *ruk-a – na ruk-e*, české *ruk-a – na ruk-e*). Alternácie na konci tvarotvorného základu pri ohýbaní (*vládat' – vládze, hádzat' – hádže, sácat' – sáče*) idú proti tejto tendencii, keďže do istej miery zastierajú hranicu tvarotvorného základu. Preto sa prechod niektorých slovies zo slovesného typu *česať* do slovesnému typu *chytat'* často vysvetľuje ako úsilie o zachovanie stálosti tvarotvorného základu, resp. ako snaha vyhnúť sa kakofonickému stretnutiu hlások (*klepotat' – klepocúc/klepotajúc*) pri slovesách, ktoré napodobňujú (väčšinou príjemne pôsobiace) prírodné zvuky. Tento prechod sa prejavuje v existencii niektorých variantných gramatických tvarov, napr. *blikotať – blikoce/blikotá, motať sa – mocú sa/motajú sa, jagotať sa – jagocúc sa/jagotajúc sa*.

Keďže variantnosť znamená veľkú tvarovú dynamiku, a tým aj nestálosť, predstavuje to určitý problém pre lexikografické spracovanie. Po preskúmaní dokladového materiálu slovies patriacich do SSSJ2 sme stanovili päť skupín, ktoré zohľadňujú mieru ich prechodu z jedného časovacieho typu k druhému.

1. typ *česať* bez variantov:

kízať *kíže kížu kíž! kízal kížuc kížuci kížanie*

2. typ *česať/chytat'* PRECHODNÍK a ČINNÉ PRÍČASTIE, ostatné *česať* (najmä zvu-komalebné):

hatať *-le -lú -li! -lal -lúc/-lajúc -lúci/-lajúci -lanie*

hahotať sa *-oce sa -ocú sa hahoc sa! -otal sa -ocúc sa/-otajúc sa -ocúci sa/-otajúci sa -otanie (sa)*

3. typ dôsledné varianty *česať/chytat'*:

luhať *luže/luhá lužú/luhajú (ne)luž/(ne)luhaj! luhal lužúc/luhajúc lužúci/luhajúci luhanie*

4. typ dôsledné varianty s opačným poradím *chytat'/česať* (okrem činného príčastia prítomného):

látať^{1,2} *láta/láce látajú/lácu látaj/lác! látal látajúc/lácuc látajúci látaný látanie*

4.a typ *chytat'/česať* PRÉZENT a IMPERATÍV, ostatné *chytat'*:

hatať *hatá/hace hatajú/hacú hataj/hac! hatal hatajúc hatajúci hataný hatanie*

5. typ prechod k vzoru *chytat'* bez variantov:

klaktať *-tá -tajú -taj! -tal -tajúc -tajúci -tanie*

5. Osobné a neosobné slovesá, resp. neosobné významy osobných slovies

Osobné slovesá majú kategóriu osoby, paradigmu určitých aj neurčitých tvarov, napr. *žijeme, žijete, žitie*. Môžu tvoriť bezpodmetový zvratný tvar typu *žije sa, žilo sa dobre*, pri ktorom gramatický subjekt nie je prítomný, no činiteľ, resp. nositeľ deja sa uvedomuje (MSJ, 1966).

Neosobné slovesá majú iba kategóriu slovesného času a spôsobu, nemajú kategóriu osoby, nevyjadrujú teda kongruenciu v osobe, čísle a mennom rode. Rozlišujeme neosobné slovesá ako osobitné lexémy, tu patria slovesá zo 7. a 8. intencného typu (*Smáďí ma., Sneží.*), a neosobné tvary osobných sloviess (*Krv mu kvapčí z nosa. – Zo strechy kvapčí.*).

Vo všeobecnosti neosobné tvary osobných sloviess majú podobu bezpodmetového zvratného tvaru – deagentizácia (*idem do kina – ide sa do kina*), nezvratného tvaru (*kvapnúť ocot do polievky – kvaplo mu za golier*) alebo môže ísť o reflexívne dispozičné konštrukcie (*zle sa nám tu býva*).

Reflexívna dispozičná konštrukcia je konštrukcia, v ktorej pôvodný agens, formálne vyjadrený datívom, je nositeľom dispoziácie (*kniha sa mi číta rýchlo, tieto šaty sa jej dobre nosia*). Konštrukcia má obligatórny kvalifikátor, ktorý vyjadruje pozitívnu alebo negatívnu dispoziáciu a je obyčajne stvárnený príslovkou (*dobre, príjemne, zle, ťažko*). Frekventované sú neosobné dispozičné konštrukcie s bezpodmetovým zvratným tvarom (*zle sa nám tu býva, dobre sa ti smeje*). Tieto konštrukcie sa nedajú pretransformovať na konštrukciu typu **dobre čítam knihu, *tieto šaty dobre nosím; *zle tu bývame, *dobre sa smeješ* (porov. M. Sokolová, 1993, s. 41 – 43).

V SSSJ2 sa neosobnosť spracúva trojakým spôsobom:

I. neosobné slovesá alebo neosobné významy (pri polysémii):

a) neosobné jednovýznamové slovesá vzhľadom na svoju neosobnosť môžu mať redukovaný gramatický aparát – 3. os. sg. a préteritum:

lúpať² -pe -palo nedok.

0. neos. (koho v čom) vyvolávať ostrú, bodavú, nepríjemnú bolesť, bolieť, pichať, lámať (napr. pri reumatizme); syn. štiepať: *lúpe ho v kostiach; lúpalo ma v kolene; lúpe ma v krížoch, asi bude pršať*

b) polysémické slovesá, ktoré majú niektorý z významov neosobný:

hrmiet' -mí -mia (ne)hrm! -mel -miac -miaci -menie nedok.

1. neos. znieť silným, ohlušujúcim zvukom pri blýskaní (počas búrky): *už zrána začalo h.; To nehrmí, to dunia kanóny za Dunajom.* [A. Chudoba]; *Od severu sa valili ťažké mraky, blýskalo sa a hrmelo.* [V. Krupa]

2. vydávať silný ohlušujúci zvuk, pripomínajúci burácanie hromu; ozývať sa takýmto zvukom; syn. dunieť, hukotať, rachotiť, rachotať: *delá, výstrely hrmeli; sála hrmela potleskom; Za riekou hrmel vlak, ale nás nemohol zbadat' nik.* [K. Jarunková]; *Dolu Oravou hrmeli vojská Jána Sobieskeho, keď tiahol na pomoc Viedni proti Turkom.* [A. Habovštiak]

3. expr. hovoriť mohutným, hromovým hlasom, obyč. v určitom citovom rozporení, v rozrušení: *Kostol bol plný a renomovaný kazateľ z Budína už hodný čas hrmel z kazateľnice.* [A. Hýkisch]; *Hned' po zápase hrmel na hráčov a na druhý deň ani s chladnejšou hlavou neboli jeho slová mäkšie.* [VNK 2002]

F [robí] *len tak hrmí* pracuje veľmi rýchlo; *ved' [ešte] nehrmí* ešte je času dost';

opak. **hrmievať** -va -vajú -val;

dok. -> **zahrmiet'**

II. neosobné tvary v rámci osobného významu slovesa:

a) bezpodmetový zvrätaný tvar

chlpatieť -tie -tejú -tej! -tel -tejúc -tejúci -tenie nedok.

0. stávať sa chlpatým, zarastať: *chlapci v puberte začínajú ch.*; *hrudník chlpatie*; *od hormónov jej chlpatajú nohy*; *nadmerné chlpatenie rúk*; neos. po týchto liekoch sa chlpatie;

dok. -> **schlpatieť, ochlpatieť**

b) reflexívna dispozičná konštrukcia

chytať -tá -tajú -taj! -tal -tajúc -tajúci -taný -tanie nedok. ...

6. hovor. (v loptových hrách) byť brankárom: *ch. v reprezentačnom mužstve*; *roky chytá za miestny futbalový klub*; *brankár chytať perfektne, nesústredene*; *Pri nedelnej prechádzke mi ukázal miesto v bývalej záhrade škôlky, kde ho chytanie v bráne začalo baviť.* [D. Dušek]; neos. doma sa brankárovi chytá lepšie ako na ihrisku súpera ...

III. osobné a neosobné sloveso ako varianty:

V SSSJ2 sa prípad variantných heslových slov, ktoré sa uplatňujú ako osobné i neosobné, vyskytol len jedenkrát:

chumeliť sa -lí sa -lilo sa, neos. i chumeliť -lí -lilo -lenie nedok.

0. (o snehu) husto padať, snežiť: *sneh sa chumelí*; *neustále chumelenie*; *neos. vonku (sa) chumelí*; *celú noc chumelilo*; *nad ránom prestalo ch.*; *vo svetle lúčov bolo vidno ako (sa) chumelí*; *Chumelí sa, chumelí, padá sypký sneh.* [Š. Žary]; *pren. Zasa sa chumelilo a padali aj nadávky.* [Milan Ferko] *strhla sa bitka*

V SSSJ2 ako neosobné neuvádzame:

- slovesné konštrukcie typu *byť* + participiálne pasívum dokonavých sloviac (*je napčené*),
- konštrukcie typu *byť* + príslovka (*bolo chladno*),
- konštrukcie *byť* + modálna príslovka + sloveso (*bolo treba hovoriť*).

6. Zvrätané a nezvrätané slovesá

Pri lexikografickom spracovaní reflexívnych sloviac vychádzame zo súčasného chápania komponentu *sa*, ktorý má v slovenčine štyri morfoloické funkcie (MSJ, 1966; M. Sokolová, 1995):

- a) prázdna lexikálna morféma
- b) zámeno v reflexívnej alebo recipročnej funkcii
- c) derivačná morféma
- d) gramatická morféma.

a) sa ako **prázdna lexikálna morféma** – vyznačujú sa ňou reflexíva tantum, t. j. slovesá, ktoré nemajú nezvrätanú podobu (*chichotať sa, kamarátiť sa, lúčiť sa*). Pri

slovesách tohto typu uvádzame v gramatickom aparáte pri slovesnom podstatnom mene komponente *sa* v zátvorke, pretože doklady na tento tvar sa vyskytujú v obidvoch podobách:

Po spovedi malo prísť kajanie a práve vtedy, keď som sa mylne nazdával, že prichádza, jej mama mi priniesla tento zahanbujúci list.

(L. Ťažký: Zjavenie Sabíny, 1997; zdroj: Slovenský národný korpus; ďalej SNK)

A priam sa až utápame a dusíme v samých úvahách o vinách, hriechoch, svedomí, zrade, kajaní sa...

(L. Ballek: Zlatý stôl, 2000; zdroj: SNK)

Napriek kritike a nepochopeniu, s ktorým sa často stretáva v katolíckych kruhoch, a dokonca aj vo svojej kúrii, znovu otvoril diskusiu o kajaní sa a ľútosti, ktorú začal ešte v júni 1994 na kardinálskom zhromaždení.

(Sme, 1999; zdroj: SNK)

SSSJ2:

kajať sa -já sa/-ja sa -jajú sa -jaj sa! -jal sa -jajúc sa -jajúci sa -janie (sa) nedok.

1. náb. prejavovať ľútosť nad spáchanými zlými činmi, konať pokánie, snažiť sa o polepšenie života: *k. sa podľa pokynov kňaza; k. sa pred smrťou; kajajte sa, lebo nastal čas pokánia; Hrnuli sa hriešnici a kajali sa zo svojich hriechov.* [J. F. Kunik]

2. (z čoho; pred kým; 0) žiadať o prepáčenie: *úprimne, verejne sa k.; k. sa nahlas, potichu; k. sa z omylov; k. sa pred otcom, pred matkou; na kajanie nie je nikdy neskoro; A ona neprišla plačky, nekajala sa.* [I. Habaj];

dok. -> **pokajať sa**

b) sa ako **zámeno v reflexívnej funkcii** (*haliť sa, krémovať sa*) alebo ako **zámeno v recipročnej funkcii** (*kynožiť sa, klat' sa*), pričom sloveso existuje zároveň aj v nezvratnej podobe (*Vlasy jej halila šatka. Halím sa do osušky.*).

ba) zvrtné slovesá so zámenom sa v reflexívnej funkcii sa spracúvajú v nezvratnej aj zvratnej podobe samostatne:

haliť -lí -lia hal'! -lil -liac -liaci -lený lenie nedok.

kniž.

1. (koho, čo (do čoho)) dávať na niečo, na niekoho prikrývku, šaty al. niečo, čo prikrývku, šaty pripomína, zakrývať, zahaľovať: *h. nevestu do závoja; kapucňa mu halila tvár; Čelo nemladé a vlasy, čo ich aj šatka halila, jednako bolo vidieť, že osiveli.* [R. Jašík]; *Bol nižšej postavy, ktorú halil mierne ufúlaný biely plášť.* [KŽ 1968] ...

haliť sa -lí sa -lia sa hal' sa! -lil sa -liac sa -liaci sa -lenie sa nedok.

kniž.

1. (do čoho) dávať na seba prikrývku, šaty a pod., obliekať sa: *h. sa do pláštá; halím sa do osušky; halila sa do plédu; Kostolné šaty, ako zriedkavo sa do nich halím.* [V. Bednár]; pren. *Hlúpy obsah má tendenciu haliť sa do emócií a silných slov.* [DF 2000]...

bb) zvrtné slovesá so zámenom sa v recipročnej funkcii vyjadrujú význam: „navzájom, vzájomne, jeden s druhým, medzi sebou“; reciprocita sa môže vyskytnúť v jednovýznamovom (*kľbiť sa*) aj vo viacvýznamovom slovese (*láskat' sa*). Vzhľadom na svoju väzbu tieto slovesné heslá môžu mať redukovaný gramatický aparát takto:

– **jednovýznamové recipročné slovesá s väzbou (s kým, s čím)** majú úplný gramatický aparát:

kľbiť sa *-bi sa -bia sa kľb sa! -bil sa -biac sa -biaci sa -benie sa* nedok.

0. kniž. (s čím) navzájom sa prestupovať, spájať sa do jedného celku: *originality a odlišnosti sa dokážu dokonale k.*; *v hokejovom tíme sa kľbi skúsenosť s mladou dravosťou*; *Je to jeden zo vzácných ľudí, u ktorých sa národná hrdosť kľbi s obrovskou mierou vzdelanosti.* [Sme 1999];

dok. -> **skľbiť sa**

– **jednovýznamové recipročné slovesá, ktoré nemajú väzbu (s kým, s čím)**, majú redukovaný gramatický aparát, t. j. uvádzajú sa **len plurálové podoby určitých tvarov** – 3. osobu plurálu, imperatív v pluráli, préteritum v pluráli, pretože bez tejto väzby sa doklady vyskytujú iba v pluráli, na rozdiel od väzby (s kým, s čím), ktorá sa spája so singulárom:

kynožiť sa *-žia sa (ne)kynožite sa! -žili sa -žiac sa -žiaci sa -ženie sa* nedok. <mad.>

0. expr. spôsobiť si navzájom smrť, usmrcovať sa, zabíjať sa; syn. kántriť sa, hľušiť sa: *národy sa často navzájom kynozia*; [Uvažoval som] *nechať zaslepených Achájcov, aby sa ďalej kynozili v nezmyselnej vojne.* [J. Lenčo];

dok. -> **vykynožiť sa**

– pri **viacvýznamových zvrtných slovesách** môže dochádzať ku kombinácii **reciprocity** (*kone sa láskali hlavami*) a **nereciprocity** (*láskat' sa s bábätkami, s mačiatkami*). V takomto prípade sa uvádza v záhlaví plný gramatický aparát:

láskat' sa *-ka sa -kajú sa -kaj sa! -kal sa -kajúc sa -kajúci sa -kanie sa* nedok.

1. (s kým, s čím) nežným, milým zaobchádzaním prejavovať al. vyjadrovať citovú náklonnosť; príjemne sa pohrávať; syn. maznať sa: *l. sa s bábätkami, s mačiatkami*; *Ona si predstavovala rodinu s deťmi, s ktorými sa bude môcť láskat'.* [Š. Janega]; *Vzniká spontánne [dielo], keď mám v rukách hlinu a môžem sa s ňou láskat' dlhší čas.* [LT 2001]; pren. *l. sa s predstavou, s myšlienkou so záľubou sa zaoberať*; *Slnko sa láskalo s lístím.* [E. Dzvonič] *odrážalo sa na lístí*

2. (0; s kým) vzájomne sa príjemne dotýkať dlaňami, rukami, ústami pri prejave lásky spojenej s erotickou príťažlivosťou; pohlavne sa vzrušovať; syn. hľadiť sa, hladkať sa, maznať sa: *l. sa v spálni*; *kone sa láskali hlavami*; *Zvádzali chlapa [vily], začali sa s ním láskat', volali ho tancovať.* [LT 2000];

dok. -> **poláskat' sa**

c) sa ako **derivačná morféma** – tvorí nové slová, má rovnakú funkciu ako prefixy a sufixy; slovesá existujú v nezvrtnej aj zvrtnej podobe (*chrlit' – chrlit' sa, konať – konať sa, hnevať – hnevať sa*) a väčšinou sa odlišujú týmito sémantickými príznakmi:

Nezvrtné sloveso <i>chlpit', helenizovať', kučeraviť'</i>	Zvrtné sloveso <i>chlpit' sa, helenizovať' sa, kučeraviť' sa</i>
akčnosť	procesnosť
tranzitívnosť	intranzitívnosť
kauzativnosť	nekauzativnosť

Pri lexikografickom spracovaní sa vyskytli aj také prípady, keď nezvrtné a zvrtné sloveso vystupujú ako varianty na úrovni lexémy (pri jednovýznamovom slovese, napr. *kutrať', kutrať' sa*) alebo na úrovni konkrétnej lexie, významu (pri viacvýznamovom slovese, napr. *karamelizovať', karamelizovať' sa*). Aký status má komponent *sa* v týchto prípadoch? Nazdávame sa, že ho môžeme klasifikovať ako formálne derivačnú morfému (keďže existuje nezvrtná aj zvrtná podoba slovesa), ale sémanticky ako prázdnu morfému (keďže nevytvára lexému s novým významom, ide o slovtvorný variant):

kutrať', kutrať' sa -re (sa) -rú (sa) -ri (sa)! -ral (sa) -rúc (sa)/-rajúc (sa) -rúci (sa)/-rajúci (sa) -ranie (sa) nedok.

0. expr. (kde) prehŕňať sa, prehrabávať sa s cieľom nájsť niečo, hľadať; syn. *kutať', kutať' sa, prekutať', dolovať'*: *k. v kabelke; k. (sa) v zásuvkách, po skriniach; po krátkom kutralí na povale našla, čo hľadala; kutrať' sa v zemi kopať'; Malému sa páči pec. Kutre v nej kutáčom.* [V. Šikulová]; pren. *Dychtivo neúnavne kutral sa v pamäti, túlal sa bludiskom spomienok.* [J. Lenčo]

karamelizovať' -zuje -zujú -zuj! -zoval -zujúc -zujúci -zovaný -zovanie nedok. i dok. potrav.

1. (čo) meniť, zmeniť pálením na karamel: *k. cukor; rozpustiť na panvici med a nechať ho k.*

2. i **karamelizovať' sa** meniť sa, zmeniť sa pálením na karamel: *cukor (sa) karamelizuje;*

dok. i -> **skaramelizovať'**, k 2 i -> **skaramelizovať' sa**

d) sa ako **gramatická morféma** je súčasťou gramatického tvaru v dvoch prípadoch:

- pri bezpodmetovom zvrtnom tvare, t. j. pri deagentizácii nezvrtných slovies (*hovoríť' – ťažko sa mi hovorilo; ísť' – ide sa na vec; klebetíť' – klebetilo sa na môj účet*);
- pri zvrtnom pasíve (*krátiť' – krátili sa im dôchodky*).

Tieto prípady sa spracúvajú pri príslušnom nezvrtnom slovese ako exemplifikácia neosobného tvaru osobného slovesa (*klebetíť'*) alebo ako príklad na zvrtné pasívum (*krátiť'*):

klebetíť' -tí -tia (ne)klebet'! -til -tiac -tiaci -tenie nedok.

1. (o kom, o čom; s vedľajšou vetou; 0) šíriť klebety, neoverené, nepresné, často zveličené informácie, obyč. zo súkromia niekoho, ohovárať: *k. o susedoch; celá dedina klebetila o svadbe; klebetili o nej, že sa nestará o domácnosť; nechcel, aby o nich klebetili; Tu ste stále na očiach, ľudia hneď klebetia.* [P. Vilikovský]; neos. klebetilo sa na môj účet

2. hovor. (s kým; o kom, o čom; 0) rozprávať sa (obyč. dlhšie) o nedôležitých veciach, debatovať, zhovárať sa: *k. o móde, o deťoch; k. so sestrou o nápadníkoch; kamarátky do rána klebetili; o čom ste klebetili?; kto ti klebetil? kto ti to povedal, prezradil;*

opak. **klebetievať** -va -vajú -val

krátiť -ti -tia krát! -til -tiac -tiaci -tený -tenie nedok.

(čo)

1. robiť kratším vzhľadom na dĺžku ako lineárny rozmer, skracovať: *k. text; k. nohavice; k. si cestu ísť kratšou cestou; nesprávne krátenie slabík*

2. robiť kratším vzhľadom na dĺžku ako časový rozsah niečoho, skracovať: *k. rozhlasové, televízne vysielanie*

3. robiť menším, zmenšovať, redukovať, obmedzovať: *k. dane, odmeny, výdavky; krátené objednávky; krátili sa im dôchodky, odmeny*

T mat. *krátiť zlomok deliť rovnakým číslom bez zvyšku čitateľa aj menovateľa*

F *krátiť si čas/dlhú chvíľu niečím robiť niečo a tak odháňať nudu;*

opak. **krátievať** -va -vajú -val

Častým lexikografickým problémom je zamieňanie si zvratného pasíva so zvratným slovesom. Riešenie však nie je také jednoduché, ako by sa mohlo na prvý pohľad zdať, a to preto, lebo bohatý dokladový materiál v mnohých prípadoch ukazuje postupnú lexikalizáciu zvratného pasíva na zvratné sloveso (vtedy sa komponent *sa* pri zvratnom pasíve prehodnocuje z gramatickej morfémy na derivačnú morfému). Ide o také prípady ako:

Obyvateľstvo v blízkosti maďarských hraníc sa postupne hungarizovalo.

Médiá sa rýchlo komercionalizujú.

Konštituuje sa nová politická strana.

Tieto spojenia, resp. vety majú charakter pôvodného zvratného pasíva. Z významového hľadiska je zrejmé, že niekto spôsobil alebo niečo spôsobilo, že obyvateľstvo v blízkosti maďarských hraníc sa hungarizovalo; že existuje určitý agens, činiteľ, ktorý médiá komercionalizuje; že politickú stranu konštituuje nejaký agens (osoba, inštitúcia a pod.). Napriek tomu sme príslušné slovesá v týchto kontextoch spracovali už ako samostatné zvratné slovesá. Lexikalizácia takéhoto zvratného pasíva na zvratné sloveso je dlhodobý proces a vytvorené zvratné slovesá majú z lexikálno-sémantického hľadiska charakter procesného slovesa.

V našom príspevku sme predstavili niektoré čiastkové problémy, s ktorými sme sa stretli v záverečných prácach na druhom zväzku *Slovníka súčasného slovenského jazyka*. Slovník sa okrem iného snaží ukázať fungovanie slovesa v komunikácii cez rôznorodosť jeho gramatických tvarov, zachytávať aj jeho dynamiku tým, že ponúka variantnosť foriem a ich prehodnocovanie na morfolologickej aj lexikálnej rovine. Niektoré gramatické javy možno interpretovať iba pomocou širšieho kontextu (napr. zvratné sloveso verzus zvratné pasívum, obojvidovosť slovesa, homonymia). To však kolидуje s požiadavkou koncízности a stručnosti, ktorú si slovníkové diela vo svojej podstate vyžadujú. Napriek tomu sa autorsko-redaktorský kolektív tohto diela usiloval nájsť vhodný spôsob prezentácie každého slova v jeho bohatom jazykovom spektre.

Literatúra

- Krátky slovník slovenského jazyka. Red. J. Kačala – M. Pisárčiková – M. Považaj. 4. doplnené a upravené vyd. Bratislava: Veda 2003. 988 s.
- Morfológia slovenského jazyka. 1. vyd. Red. J. Ružička. Bratislava: Vydavateľstvo SAV 1966. 896 s.
- Pravidlá slovenského pravopisu. 3., upravené a doplnené vydanie. Bratislava: Veda 2000. 592 s.
- KRÁL, Ábel: Pravidlá slovenskej výslovnosti. 3. vyd. Bratislava: Slovenské pedagogické nakladateľstvo 1996. 626 s.
- Slovenský národný korpus – verzia prim-4.0-juls-sane. Bratislava: Jazykovedný ústav Ľ. Štúra SAV 2009. Dostupný z WWW:<http://korpus.juls.savba.sk>.
- Slovenský národný korpus – verzia prim-5.0-juls-sane. Bratislava: Jazykovedný ústav Ľ. Štúra SAV 2011. Dostupný z WWW:<http://korpus.juls.savba.sk>.
- Slovník slovenského jazyka. Red. Š. Peciar. 1. vyd. Bratislava: Vydavateľstvo SAV 1959 – 1968. 6 zv.
- Slovník súčasného slovenského jazyka. H – L. [2. zv.]. Red. A. Jarošová – K. Buzássyová. Bratislava: Veda 2011. 1088 s.
- SOKOLOVÁ, Miloslava: Sémantika slovesa a slovesný rod. 1. vyd. Bratislava: Veda 1993. 110 s.
- SOKOLOVÁ, Miloslava: Kapitoly zo slovenskej morfológie. 1. vyd. Prešov: Slovacontact 1995. 180 s. SOKOLOVÁ, Miloslava – BÓNOVÁ, Iveta: Tvorenie imperatívu v slovenčine a lexikografická prax. In: Slovenská reč, 2008, roč. 73, č. 5, s. 271 – 280.

Ustálené prirovnania: problémy ich chápania a lexikografického zachytenia¹

Alexandra Jarošová

Jazykovedný ústav Ľudovíta Štúra SAV

Similes: Problems of Their Understanding and Lexicographic Processing

The purpose of this paper is to critically analyse a concept of “phraseologically bound meaning” or “phraseologicalized component” as an implicit or explicit theoretical threshold of similes processing (as well as other types of phraseological units) in Slovak academic dictionaries. In lexicographic practice, it was just a “phraseologicalized component itself”, which used to be interpreted, and not a simile or a phraseological unit as a whole. This did not only result in a slight change of meaning and consequently in different processing of a certain phraseological unit under the various headwords – parts of the phrase, but also in a maximally narrowed comprehension of similes as a type of phraseological unit. Considering the fact that the concept of “phraseologically bound meaning” was theoretically overthrown, the author supposes that it is necessary to eliminate its consequences in the lexicographic practice.

1. Úvod

Bezprostredným dôvodom napísania tohto príspevku je prehodnotenie spracovania ustálených prirovnaní v 2. zväzku Slovníka súčasného slovenského jazyka (2011; ďalej SSSJ) v porovnaní s ich zachytením v 1. zväzku (SSSJ, 2006).

Koncepcia z roku 1992 (časť Inštrukcie) obsahovala aj článok Základné zásady spracovania frazeológie, ktoré boli v roku 2000 prehĺbené a precizované najmä so zreteľom na spôsob výkladu a na používanie lexikografickej techniky pri opise vlastností frazém. V Inštrukciách sa píše: „V zhode s našou lexikografickou tradíciou pri intenzitných prirovnaniach s explicitným vyjadrením vlastnosti sa neuvádza, napr. *biely ako sneh*, *čerstvý ako ryba*. Pri prirovnaniach s implicitným vyjadrením vlastnosti sa výklady uvádzajú; zvyčajne sú jednoslovné. Napr. (*je tam*) *ľudí ako maku* veľmi. Výklady tohto typu sa uvádzajú aj pri slovesných prirovnaniach: *hanbiť sa ako pes* veľmi.“ (Inštrukcie, s. 71; porov. aj Krátky slovník slovenského jazyka, 1987, s. 27).

V prehĺbenom a spresnenom variante Zásad sú niektoré ustálené spojenia so spojkou *ako* uvedené v kategórii Vlastné frazémy, a nie v kategórii Prirovnania, pretože sa nedali vykladať pomocou intenzitnej príslovky *veľmi*, prípadne pomocou inej príslovky v spojení s príslovkou *veľmi*, napr. *poskladať sa/poskladaný ako cólštok* – zmenšiť sa, zmenšený na minimum, *horieť ako fakľa* – a) silno b) byť červený od hanby, od vzrušenia, *byť/ prísť/ priletieť ako fúria* – nazlostený, rozzúrený; *voziť sa/prevážať sa ako gróf* – po pansky, pohodlne; *mať sa/žiť si ako gróf* – v dostatku, v blahobyte. Vidíme, že výkladová perifrása neobsahuje časť prirovnania pred spojkou *ako*, ktorá sa zvyčajne charakterizuje ako porovnávací základ (napr. v uvedených prirovnaniach *horieť ako fakľa*, *voziť sa/prevážať sa ako gróf*). F. Čermák (2009) sa

1 Príspevok bol vypracovaný v rámci grantového projektu VEGA č. 2/0018/11 *Slovník súčasného slovenského jazyka – 4. etapa (koncipovanie a redigovanie slovníkových hesiel a s tým súvisiaci lexikologicko-lexikografický výskum)*.

vo svojej rozsiahlej recenzii 1. zväzku SSSJ vyjadril k tejto praxi kriticky: „některé frazémy mají výkládaný význam vztahující se pouze ke své části, nikoliv k celku“.

Už v priebehu redigovania 1. zväzku (autorka týchto riadkov sa stala spoluriešiteľkou grantu v roku 2000 a členkou redakčného kolektívu v roku 2002) sa ukazovali určité slabé miesta tohto spôsobu spracovania, ale situácia nedozrela na zmenu koncepcie v danom bode, pretože zásady koncipovania hesiel SSSJ sa dopracúvali a spresňovali v desiatkach iných oblastí (napr. v priebehu rokov 1999 až 2003 autor-ky a redaktori vypracovali vyše 40 dodatkov ku koncepcii).

Aj v prvom koncepte 2. zväzku SSSJ sa časť prirovnaní vykladala v súlade so Zásadami takto: *krátky ako zajačí chvost* – veľmi; *zbiť niekoho ako koňa* – veľmi; *hnať sa do niečoho ako divý* – veľmi; *chudobný ako kostolná myš* – veľmi; *kliat' ako kočiš/pohan* – veľmi; *začervenat' sa až po koreničky vlasov* – veľmi; *lačný ako vlk* – veľmi; *zvládnuť niečo na jednotku* – veľmi dobre; *letieť s vetrom o preteky* – veľmi rýchlo; *vážiť niečo na lekárnických váhach* – veľmi presne; *hovorit'/kázať ako s kanclá* – veľmi učene; *spí, akoby ho do vody hodil* – tvrdo; *držať sa niekoho, niečoho ako pes ježa* – pevne.

Opäť v súlade s pôvodnou koncepciou bola časť prirovnaní v prvom koncepte zaradená medzi frazémy, a nie medzi prirovnania, napr.: *lepiť sa na niekoho ako mu-chy* – byť veľmi príťažlivý pre opačné pohlavie; *ústa jej idú [ako namastené]* – veľa a rýchlo rozpráva; *ide mu to na úžitok ako psovi tráva* – nemá to preňho nijaký význam; *hučí to tam ako v úli* – je tam rušno, živo; *hrnúť sa do niečoho ako baran* – nerozvážne konať.

2. Teoretické uchopenie prirovnaní v slovenskej jazykovede

V jazykovednom časopise boli v roku 1970 uverejnené dve zásadné štúdie týkajúce sa chápania ustálených spojení ako špecifického typu frazém so zreteľom na ich typológiu (Mlacek, 1970) a so zreteľom na ich lexikografické spracovanie (Smieškova, 1970).

J. Mlacek na základe analýzy prác venujúcich sa prirovnaniu prišiel k záveru, že autori pod ustáleným prirovnaním chápu „mnohé navzájom od seba odlišné javy“ (op. cit., s. 172). Z hľadiska metodologickej primeranosti nepovažoval za správne vychádzať pri posudzovaní kategórie prirovnania len zo sémantiky: „Z binaristickej povahy nášho chápania jazykových jednotiek vyplýva, že aj prirovnanie musí sa vyznačovať významom i osobitnou formou, ak sa má chápať ako jazyková jednotka“ (op. cit., s. 173). Nevyhnutným lexikálno-syntaktickým komponentom prirovnania je prirovnávací spojka. V slovenčine je takýmto základným a štylisticky nepríznačným prostriedkom spojka *ako*. Popri nej sa v tejto funkcii zriedkavo uplatňujú aj iné spojky, resp. spojovacie výrazy (*akoby, ani, ako keby, ani čoby*). Týmto vymedzením ustáleného prirovnania (stálosť formy a významu) sa autorovi podarilo naznačiť hranicu medzi voľnými a ustálenými prirovnaniami. Ustálenosť formy autor nechápe ako jej petrifikovanosť. Poukazuje na existenciu morfológických, syntaktických aj lexikálnych variantov. V riešení otázky, ktorá sa v tom čase často pertraktovala v literatúre, či ustálené je celé prirovnanie alebo len prirovnávací časť, sa J. Mlacek priklonil k názoru, že platnosť frazeologickej jednotky má celé prirovnanie. Napo-

kon autor ponúka štruktúrnú klasifikáciu ustálených prirovnaní urobenú na základe konštelácie komponentov a ich slovnodruhovej charakteristiky. Táto štúdia má aj dnes určujúci význam pre chápanie prirovnaní v slovenskej frazeológii, zvlášť na pozadí alternatívnych riešení, ktoré sa napokon ukázali ako neudržateľné (túto problematiku rozoberáme nižšie). Napokon aj práce z celkom nedávneho obdobia stavajú na Mlackovom vymedzení tejto kategórie: „prirovnanie sa v rámci frazeologického fondu jazyka vyníma svojou charakteristickou syntaktickou výstavbou, v ktorej je dôležité explicitné formálne vyjadrenie prirovnania“ (Kačala, 2007, s. 201).

Prejdeme k inšpiratívnej práci E. Smieškovej (1970), ktorá bezprostredne súvisí s našou témou – spracovaním ustálených prirovnaní vo výkladovom slovníku. Autorka vystupuje proti tvrdeniam niektorých autorov, že v prirovnaniach typu *spieva ako slávik* (Mann, 1953, s. 108) a *bledý ako stena*, *bledý ako smrť* (F. Kočiš, 1961, s. 274) sa prirovnáva skutočne, nie obrazne, a preto takéto prirovnania nie sú frazeologickou jednotkou. Autorka poukazuje na to, že ide o prirovnania reprodukovateľné v jazyku ako hotové celostné jednotky, a teda nejde o individuálne prirovnania. Proti Kočišovmu zdôrazňovaniu priameho významu komponentov argumentuje tým, že základ prirovnania v spojení *bledý ako stena*, teda vlastnosť *bledosť*, „nie je v realite ani podstatným, ani stálym atribútom steny ako takej, o obraznosti a nereálnosti prirovnania nemôže byť ani pochybnosti“ (op. cit., s. 192).

E. Smiešková, podobne ako J. Mlacek, sa musela vyrovnáť s problémom, či v spojení so štruktúrou prirovnania treba za frazému považovať len „tú časť, ku ktorej sa prirovnáva, teda počnúc od prirovnávacej spojky“, alebo frazémou je komparatívne spojenie ako celok (op. cit., s. 193). V prípade, že za frazému sa považuje len komparatívny zvrat (spojka + prototyp prirovnania, teda komparatum), potom ostatné súčasti spojenia (prirovnávaná časť, teda komparandum + základ prirovnania, teda tertium comparationis) sa označujú za stály vonkajší kontext (napr. Žukov, 1967).

Pretože s týmto problémom úzko súvisí spôsob lexikografického zachytávania prirovnaní, venuje autorka najväčšiu pozornosť analýze lexikálneho obsadenia a výkladu týchto frazeologických jednotiek v slovníkoch. Z tohto hľadiska člení ustálené prirovnania na niekoľko typov.

Konštatuje, že najmenší problém predstavujú (A.) frazémy, ktoré majú štruktúru prirovnania, ale vo výklade sa frazeologický zvrat neuplatňuje samostatne, funguje len frazeologický význam celku: *chodiť ako pes bez gazdu* – prejavovať bezradnosť, *prasknúť ako mydlová bublina* – ukázať sa bezvýznamným, *stratiť význam*, *mať niečo*, *niekoho rád ako koza nôž* – nenávidieť.

Osobitne posudzuje aj (B.) skupinu ustálených prirovnávacích zvratov, ktoré sa nespájajú s ustáleným prirovnávaným objektom: *akoby uťal*, *ako keby nič*, *ako na potvoru*, *ako na hriech*, *ako naschvál*, *ako na rozkaz*, *akoby čarovným prútikom udrel*, *ako hrom z jasného neba*. Nazdávame sa, že tieto prirovnania možno opísať s poukazom na závažné sémantické okolie, teda s poukazom na sémantický okruh sloviess (uvedením niekoľkých najfrekventovanejších), ktoré sa s týmito zvratmi fakultatívne spájajú: [niečo stíchlo, utíchlo, ustalo al. niekto stíchol, zmĺkol] *akoby uťal*, [tváriť sa, správať sa, konať] *ako keby nič*. Charakteristický verbálny kontext možno naznačiť aj pomocou zástupného kategoriálneho slovesa *ako na potvoru* [sa niečo deje al. niekto niečo urobil].

V tretej skupine (C.) ide o prirovnania, v ktorých substantívum (komparandum) vystupuje „vo svojom nominatívnom význame“, ale celé prirovnanie je často „viazané rýmom alebo rytmicky, takže vystupuje ako petrifikovaný celok“: *hlavička ako makovička, detí ako smetí, chlapec ako dubec, Slovák ako repa, peňazí ako pliev, zlata ako blata* a pod. (op. cit., s. 195). Pri tomto type prirovnaní slovníkový problém lexikálneho obsadenia frazémy zdanlivo nevzniká. Tu poznamenávame, že hoci tieto prirovnania najčastejšie vystupujú v predikatívnej funkcii, napr. *je/bol to chlap ako hora*, *XY je/bol chlap ako hora*, nijako nie je zriedkavé ich použitie vo funkcii prístavku, podmetu alebo predmetu: *Odrazu prišiel jeden, volali ho Peter, chlap ako hora, uj!* (P. Jaroš), *Traja chlapi ako hory sa pred mrholením uchýlili do plechovej búdky na autobusovej zastávke* (Sme 2003), *Myká hlavou, blúdi očami po chlapoch ako hory, najradšej by mal krídla* (F. Hečko). V Slovníku českej frazeologie a idiomatiky je lexikálne obsadenie týchto prirovnaní riešené v prospech uvádzania relačného verba (*je*) a formálneho komparanda (*to*) ako najčastejších fakultatívnych komponentov (*to je*) *chlap jako hora*, (*to je*) *den jako vymalovaný* (Čermák, 1983, s. 483).

Najväčšie ťažkosti z hľadiska odlišenia komponentov frazémy od „slov obvyklého kontextu“ predstavuje štvrtá skupina frazém (D.), v ktorej si prvá časť prirovnania zachováva svoju relatívnu lexikálnu a gramatickú samostatnosť: *hluchý ako poleno, bledý ako stena, dobrý ako kus chleba, kradnúť ako straka, triasť sa ako osika*. Autorka je presvedčená, že tieto ustálené spojenia sú „celostné frazeologické jednotky, teda s ich prirovnávanou aj prirovnávacou časťou“ (op. cit., s. 195). Zastáva názor, že prirovnanie musí byť v slovníku zachytené v neskrátenej podobe, pretože prirovnávací zvrät nefunguje samostatne. Fakultatívnosť niektorého komponentu treba vyjadriť zápisom v zátvorkách (*lahký*) *ako pierko*, (*studený*) *ako lad*, (*opustený*) *ako sirota*. Takisto treba uvádzať variantné podoby odlišujúce sa derivačne alebo morfológicky usúvzťažnenými podobami *bledý ako stena – zblednúť ako stena, ustatý ako kôň – ustať ako kôň, spí ako zarezaný – spí, akoby ho zarezal*. Chápanie prirovnania ako celostnej frazeologickej jednotky znemožní, aby sa lexikograf dopúšťal takých chýb, aké sa vyskytli v Slovníku slovenského jazyka (1958 – 1968; ďalej SSJ). Napr. prirovnanie *silný ako buk* sa v hesle **silný** uvádza ako exemplifikácia heslového slova a v hesle **buk** ako frazéma.

Ťažkosti s priradovaním frazém k jednotlivým významom heslového slova, podľa toho ako si význam relevantného komponentu interpretuje lexikograf, privádzajú autorku k presvedčeniu, že všetky frazeologické jednotky je potrebné uvádzať na konci heslovej state. Dnes sa v tomto bode dosiahol všeobecný konsenzus.

Ďalší zásadný problém lexikografického zachytenia prirovnaní je ich výklad. Autorka poukazuje na to, že v SSJ sa pri prezentácii významu prirovnaní postupuje nejednotne. Výklady sa odlišujú tým, že raz interpretujú celé prirovnanie, inokedy len jeho časť, prípadne absentujú úplne. Takto rozdielne sa vykladajú aj synonymné alebo totožné prirovnania. Najčastejšie sa vykladá len prirovnávací zvrät: *opitý ako snop* – veľmi, celkom, *ožratý ako snop* – veľmi. Ale niekedy sa uvedie úplnejšia perifráza so základom prirovnania: *zdravý ako buk* – celkom, veľmi zdravý (s týmto výkladom za značkou frazeológie v hesle **zdravý**, ako voľné spojenie bez výkladu v hesle **buk**), *zem (tvrdá) ani skala* – veľmi tvrdá. E. Smiešková v tejto súvislosti hovorí, že postup uvádzať vo výklade len príslovku je protirečením vzhľadom na posudzovanie

ustáleného prirovnania ako celostnej jednotky, ale pripúšťa, že lexikografom môže vyhovovať pre svoju úspornosť.

Posledné, z hľadiska nášho príspevku relevantné, zistenie E. Smieškovej sa týka dôležitosti uvádzania sémantizačnej poznámky ako doplnenia výkladu. Tieto poznámky poukazujú na typ porovnávaného objektu a pri lexikografickom spracovaní chýbajú: napr. *čierny ako žúžol'* (najmä o vlasoch, fúzoch apod.), *čierny ako Cigán* (najmä o pleti), *tvrdý ako podošva* (o zle uvarenom mäse). E. Smiešková prihliada na rozdielne typy prirovnaní najmä so zreteľom na spôsob ich lexikografického spracovania. Je zrejmé, že do tejto typológie sa premietla aj vtedy akceptovaná teória o rozdielnej nominačnej hodnote komponentov prirovnania (komponenty „v priamom význame“, komponenty „v prenesenom význame“). Videli sme však, že vo všetkých prípadoch chápe prirovnanie ako celistvú frazeologickú jednotku, a takto ju odporúča aj vykladať.

Pristavme sa pri koncepcii E. Kučerovej, ktorá tiež riešila otázku prirovnaní z hľadiska teoretického, ale aj aplikačného – lexikografického (Kučerová, 1980; Kučerová, 1981; Kučerová, 1982; Sekaninová – Kučerová, 1982).

Autorka pracuje s vinogradovovským konceptom „frazeeologicky viazaného významu“ (Kučerová, 1974, s. 223), ktorý pomenúva „frazeeologizovaným významom“. Tento koncept jej slúži ako kritérium rozdelenia frazeologicky ustálených slovných spojení na podskupinu „spojení s jedným frazeologizovaným komponentom“ a podskupinu „vlastných čiže celkových frazeologických spojení“. V podskupine „ustálených spojení s frazeologizovaným komponentom“ rozlišuje ďalšie tri podskupiny (Kučerová, 1980). Prvá má štruktúru prídavné „meno + podstatné meno“ (***deravá pamäť, osí driek, mliečne zuby***). Druhá má štruktúru „sloveso + príslovkový výraz“ (***bežať s vetrom o závod, spustiť na plné plúca, revať z plného hrdla, robiť o dušu spasenú***). Tretia skupina sú spojenia, ktorých „genetický typ má štruktúru prirovnania“ (***sprostý ako tlk, zdravý ako ryba, čierny ako uhol', robiť ako kôň, zbiť niekoho ako psa***).

V závislosti od „stupňa frazeologizácie“ delí autorka „ustálené spojenia so štruktúrou prirovnania“ medzi kategóriu „spojení s (jedným) frazeologizovaným komponentom“ (teda spojenia, v ktorých sa podľa autorky frazeologizoval len komparatívny zvrat) a kategóriu „vlastné frazeologické spojenia“ (teda spojenia, v ktorých sú podľa autorky frazeologizované všetky komponenty). „Ustálené spojenia so štruktúrou prirovnania“ autorka ako celok nepovažuje za prirovnania, ale len za frazeologickú jednotku ako takú a pomenovanie „ustálené prirovnania“ dáva dôsledne do úvodzoviek. To, že vo frazeológii nejestvuje de facto skupina prirovnaní zdôvodňuje takto: „Spojenie si ponechalo z motivujúceho typu štruktúru, ale táto štruktúra nemá vplyv na význam spojenia ako celku. Komparácia vo frazeologizačnom procese sémanticky zanikla, preto tento typ frazémy nemožno kvalifikovať ako prirovnania“ (Kučerová, 1980, s. 240).

3. Zachytenie prirovnaní v doterajších akademických slovníkoch

Pozrime sa na zásady spracovania prirovnaní v SSJ (1958 – 1968), v Krátkom slovníku slovenského jazyka (1987; ďalej KSSJ) a vo Veľkom slovensko-ruskom slov-

níku (1979 – 1995; ďalej VSRS). Ide o slovníky, na ktoré sa prihliadalo pri tvorbe koncepcie uplatnenej v 1. zväzku SSSJ (2006).

SSJ nevychádza pri uvádzaní prirovnaní (a frazém všeobecne) explicitne z koncepcie frazeologicky viazaného významu, ale zreteľ na to, či je komponent akoby v priamom význame alebo nie, sa v tomto slovníku do istej miery uplatnil, pretože status prirovnania sa mení v závislosti od toho, či heslové slovo je totožné s komponentom „v prenesenom význame“ alebo s komponentom „v priamom význame“ (úvodzovky tu naznačujú podmienenosť týchto pomenovaní). V prvom prípade má spojenie status „ustáleného prirovnania“, frazémy a v druhom prípade status voľného spojenia. Napr. prirovnanie *bledý ako stena* je v hesle **stena** umiestnený vo frazeologickej zóne a vykladá sa ako „mŕtvolne bledý, veľmi bledý“ a v hesle **bledý** sa uvádza ako voľné spojenie bez výkladu. Podobne prirovnanie *zem (tvrdá) ani skala – veľmi tvrdá* (v hesle **skala**), *tvrdý ako kameň, ako skala, ako oceľ, ako žula* (v hesle **tvrdý** ako súčasť voľných spojení bez výkladu). Takýto prístup je teoreticky aj prakticky neudržateľný. Treba však pripomenúť, že SSJ bol v tejto oblasti „prvolezec“ a problematika prirovnaní nebola u nás teoreticky spracovaná, čo sa zmenilo až v sedemdesiatych rokoch (Smiešková 1970, 1974; Mlacek, 1970). Videli sme, že E. Smiešková kriticky hodnotila aj prax slovníka vykladať len prirovnávací zvrät: *ožratý ako snop* – veľmi. Okrem toho, tieto krátke výklady a forma prirovnaní neboli zjednotené v prípade uvádzania jednej frazémy v rozličných heslách, ktoré predstavovali niektorý z komponentov frazémy. Prirovnanie *hrešiť ako pohan* je v hesle **hrešiť** uvedené ako voľné spojenie bez výkladu, v hesle **pohan** ako voľné spojenie s výkladom „veľmi škaredo“, ale synonymické spojenie *hreší, až sa hory zelenajú* je uvedené vo frazeologickej zóne (hoci pri každom z komponentov – **hora**, **hrešiť**, **zelenáť sa** – v inej podobe a s iným výkladom).

KSSJ (1987) považuje ustálené prirovnania za špecifický typ frazeologických jednotiek, ktoré „uvádza ako frazeologizmy vtedy, keď ide o nepriamu motiváciu (zvýraznila A. J.) alebo o prirovnanie so slovesom: *hluchý ako peň, pristane mu ako psovi piata noha*“ (op. cit., s. 27). Ktoré sú to tie iné ustálené prirovnania (azda s priamou motiváciou?), a teda nie sú uvádzané ako frazeologizmy, z uvedenej formulácie jasné nie je. Z hľadiska súčasnej frazeologickej teórie sa ukazuje, že „cesta viac-menej subjektívneho určovania motivačného vzťahu medzi významom frazémy a jej zložkami je neproduktívna a pre jej vágnosť nemá väčší praktický ani teoretický úžitok (Mlacek – Ďurčo a kol., s. 108). F. Čermák (1983, s. 474) hovorí o nemožnosti doslovného, priameho chápania podobnostného vzťahu.

KSSJ trochu zjednotil rozdielnu lexikografickú prax SSJ v tom zmysle, že uvádza prirovnania vo frazeologickej zóne na konci hesla. Výklady prirovnaní sú v KSSJ jednotne zredukované na príslovku, alebo sa „pri prirovnaniach s explicitným vyjadrením vlastností“ neuvádzajú vôbec: *zdravý ako buk, orech, ako ryb(ič)ka ap.* – celkom (takto v hesle **zdravý**, pri hesle **buk** bez výkladu, teda bez príslovky „celkom“); *bledý ako stena* – veľmi (takto v hesle **stena**, v hesle **bledý** bez výkladu).

Členenie frazém vo VSRS, na ktorom koncepcne a redaktorsky participovala E. Kučerová, je urobené podľa jej kritéria úplnosti frazeologizácie ustáleného spojenia, teda frazémy sa členia na 1) spojenia s jedným frazeologizovaným komponentom a 2) vlastné frazeologické spojenia. (Sekaninová – Kučerová, 1982, s. 18 – 19)

Pri aplikovaní svojej koncepcie na VSRS ju E. Kučerová z pragmatických dôvodov modifikovala. Skupinu nazývanú v slovníku „spojenia s (jedným) frazeologizovaným komponentom“ zúžila len na spojenia so štruktúrou „adjektívum + substantívum“ (*orlí nos, osí driek, opičia láska, mliečne zuby, medvedie labky*). Tieto spojenia sú v slovníku odsadené za grafickou značkou Δ .

Druhú graficky osobitne vyčleňovanú podskupinu „tvoria ustálené prirovnania a skrátené prirovnania“. Autorka pripomína, že v jej koncepcii aj tieto prirovnania predstavujú typ spojení s jedným frazeologizovaným komponentom, ale „v súlade s tradíciou založenou na ich genetickom prototypu, vyčleňujú sa vo VSRS ako osobitná skupina za značkou \square “, napr. *červený ako mak, ústa ako malina, bledý ako mŕtvola, je ako obrázok* (Sekaninová – Kučerová, 1982, s. 19). To značí, že v koncepcii VSRS sa za prirovnania považujú iba tie spojenia s komparatívnou štruktúrou, ktoré majú frazeologizovaný len jeden komponent. Ostatné spojenia s komparatívnou štruktúrou sa za prirovnania nepovažujú. Tu treba povedať, že napriek odvolaniu sa na tradíciu, toto chápanie z tradície nevychádza, pretože ako J. Mlacek (1970), tak E. Smiešková (1970) posudzujú všetky spojenia s komparatívnou štruktúrou ako jeden celok. Súlad s tradíciou je zrejme len v tom, že pripustila pomenovanie „ustálené prirovnania“, ktoré v svojej teórii neuznáva.

Tretia teoreticky vyčleňovaná skupinu spojení s frazeologizovaným komponentom, teda „sloveso + príslovkový výraz“ (*kričať na celé, plné hrdlo* – kričať veľmi hlasno), sa už v úvode VSRS nespomína a dané jednotky sa spracúvajú v zóne „vlastných frazeologických spojení“. Vidíme, že hoci v zúženom a modifikovanom moduse, ide autorke v podstate len o vyčlenenie (za značkami Δ a \square) ustálených spojení s jedným frazeologizovaným komponentom (*osí driek, bežať s vetrom o závod, čierny ako uhol, najesť sa, čo hrdlo ráči, sprostý ako tlk, robiť ako kôň*), ktoré sa zvyknú v SSJ, v KSSJ a čiastočne aj v 1. zväzku SSSJ vykladať len adjektívom, príslovkou alebo adjektívom s príslovkou. Teda vykladá sa len „frazeologizovaný komponent“ frazémy (v príkladoch zvýraznený) a nie frazéma ako celok.

Čo potom predstavuje v komparatívnom spojení nefrazeologizovaný komponent? Podľa E. Kučerovej je to závažný kontext – „závažným kontextom môže byť nielen prídavné meno, ale aj sloveso a príslovka ...“ (Kučerová, 1981, s. 173). Ak je to tak, potom by azda bolo primeranejšie dávať tieto slová do zátvorky [‘hluchý’] ako peň – celkom, aby bolo zreteľné, že sa vykladá len časť mimo zátvorky. Pod závažným kontextom však v našej koncepcii chápeme nie slovo, ktoré sa nedá v slovnom zvrate substituovať (je jediné možné alebo limitovane variantné), ale skupinu slov jednej sémantickej triedy, ktoré sa nedajú vymenovať ako fakultatívne komponenty spojenia a teda ich substituovateľnosť je pomerne široká, porov. vyššie uvedený príklad [tváriť sa, správať sa, konať] *ako keby nič*. Okrem toho, tvrdenie, že v uvedených prirovnaniach (*bledý ako stena*) ide o realizáciu determinatívneho vzťahu, kde „frazeologizovaný komponent“ (komparatívny zvrät) plní úlohu príslovky, nie je nesporné. Tieto prirovnania vystupujú vo vete ako jeden celok a ako celok plnia úlohu predikátu. Aj v spojeniach *bledý ako stena* je frazéma súčasťou menného prísudku (spona je tam skoro vždy prítomná), takže ani tak nejde o spojenie *bledý ako stena*, ale *byť bledý ako stena* (najčastejšia realizácia je v tretej osobe: *je bledý ako stena, bol bledý ako stena*).

4. Spracovanie prirovnaní v SSSJ

SSSJ (1. zv., 2006) pri spracovaní prirovnaní uplatnil určitý hybridný model.

Na jednej strane vychádzal z E. Kučerovej, ktorá pod „prirovnaniami“ (toto označenie pripúšťa len ako pragmatický ústupok tradícii) chápe skupina spojení so štruktúrou prirovnania „s jedným frazeologizovaným komponentom“ typu *čierny ako havran, zdravý ako buk, bledý ako stena, robiť ako kôň, bežať ako bez duše*. Ostatné ustálené spojenia so štruktúrou prirovnania sa nechápu ako „prirovnania“ ani v tomto veľmi podmienenom zmysle, ale ako vlastné frazémy (*ide mu to na úžitok ako psovi tráva; chodiť [okolo niečoho] ako mačka okolo horúcej kaše*). Pretože SSSJ na rozdiel od VSRS nevyčleňoval skupinu menných frazém *deravá pamäť, osí driek, labutia šija* ap. ako osobitný typ (podľa E. Kučerovej ide tiež o spojenia s jedným frazeologizovaným komponentom), tak skupina chápaná v SSSJ ako prirovnania ostala vlastne jedinou predstaviteľkou kategórie ustálených spojení s jedným frazeologizovaným komponentom. Videli sme, že z troch pôvodných podskupín, ktoré E. Kučerová vyčlenila v skupine spojení s jedným frazeologizovaným komponentom, sa vo VSRS osobitne uplatnili dve a v SSSJ už len jedna. Zdá sa, že erózia tejto teoreticky vyčlenenej skupiny v procese jej praktickej lexikografickej aplikácie takisto svedčí o nedostatočnej nosnosti konceptu frazeologizovaného komponentu.

V skoncipovanom texte SSSJ sa táto skupina úzko chápaných prirovnaní spracúvala za značkou {K} (= komparácia) a v tlačenom texte sa vyčlenenosť tejto skupiny prejavovala, tak, že sa ocitla na začiatku frazeologického odseku a až potom nasledovali tzv. „vlastné frazémy“. Hybridnosť modelu použitého v SSSJ sa prejavila v tom, že hoci koncepcne sa model inšpiroval statusom prirovnania vo VSRS (určeným E. Kučerovou), ale pre spôsob výkladu musel hľadať inšpiráciu inde, pretože VSRS pochopiteľne nepracuje s výkladom, ale s cudzojazyčným ekvivalentom (a v teoretických prácach E. Kučerovej sa nehovorí, že vykladať by sa mal len „frazeologizovaný komponent“ prirovnania). Tradícia vykladať prirovnania jedným slovom prichádza z lexikografickej praxe SSJ a KSSJ. Vykladanie len „frazeologizovanej časti“ ustáleného spojenia sa netýka iba prirovnaní, ale aj iných typov frazém. Skupina ustálených spojení typu *orlí nos, osí driek, labutia šija* sa v SSJ spracúva medzi ilustračnými príkladmi s kvalifikátorom prenesenosti: *pren. labutí krk štíhly, tenký; labutia šija štíhla, tenká*. Podobne postupuje KSSJ: *pren. labutia šija, labutí krk dlhá (-ý) a ladne prehnutá (-ý)*. Typ *kričať na celé/plné hrdlo* sa tak v SSJ ako aj v KSSJ spracúva vo frazeologickej zóne hesla a vykladá sa len príslovkou/príslovkami. V SSJ má frazéma *kričať na celé hrdlo* heslo **kričať** výklad „silne, hlasno“ a v hesle **hrdlo** má podobu *revať, kričať z plného, celého hrdla* s výkladom „veľmi hlasno“.

V prvom koncepte 2. zväzku SSSJ boli pomocou redukovaného výkladu spracované okrem prirovnaní aj iné typy frazém: *hromží, až je zemi ťažko* – veľmi; *smrť/zubatá kosila* – zabíjala; *hádať sa do krvi* – veľmi ostro; *hovoriť z vysokého koňa* – povýšene; *hovoriť dve na tri* – nezmysly.

Zdá sa, že redukované výklady v slovníkoch súvisia s nazeraním na určité frazémy, ako na konglomerát komponentov, z ktorých niektoré sú sémanticky posunuté (transponované), a iné sú použité v priamom, doslovnom význame. Môžeme v tom vidieť pozostatok vinogradovskej tradície s jej rozlišovaním frazém podľa stupňa

autonómnosti vstupných komponentov. Súčasná frazeologická teória s takouto sémantickou klasifikáciou neráta a takisto neráta s konceptom „frazeologicky zviazaného významu“ (a s ním spojeným konceptom „frazeologizovaného komponentu“). Vidí v ňom „sophisticke a špekulatívne hľadanie nejakého samostatného sémantického statusu komponentu frazémy, čo možno považovať za neproduktívny prístup, pretože sa isté nesystematické implikácie vydávajú za regulárne vlastnosti“ (Mlacek – Ďurčo a kol., 1995, s. 50). F. Čermák (2001, s. 9 – 11) priamo odmieta delenie frazémy na frazeologizovanú, teda idiomatickú časť a nejaký neidiomatický, doslovný „zvvyšok“.

Mnohým autorkám sa nedarilo odlišovať prirovnania v úzkom chápaní (v ktorých je akoby frazeologizovaný iba komparatívny zvrat) od ostatných prirovnaní (keď je spojenie frazeologizované ako celok). Nasledujúce spojenia umiestnili autorky v prvom koncepte do zóny prirovnaní {K}, hoci evidentne nejde o prirovnania s „jedným frazeologizovaným komponentom“, ktoré by sa dali vykladať len príslovkou; dané spojenia treba vykladať ako celok a podľa pôvodnej koncepcie by patrili do zóny {F} = vlastné frazémy: *frfotať ako starý hrniec* – prejavovať svoju nespokojnosť, šomrať; *horieť ako fakľa* – červenať sa; *kvitnúť ako ruža* – byť pekný, svieži; *mať jazyk ako britva/žihadlo* – príkro sa vyjadrovať; *visí to na ňom ako halena* – má príliš voľné, široké oblečenie; *bojovať ako levica* – veľmi sa o niečo usilovať; *doskakovať ako kohút* – zapárat; *kričí/vykrikuje ako [ženy] na jarmoku* – správa sa neprístojne.

Nasledujúce spojenia zaradili autorky medzi „vlastné frazémy“, hoci vykladali len „frazeologizovaný komponent“ a podľa pôvodnej koncepcie by patrili do zóny {K}: *tváriť sa/stáť ako kopa al. kôpka nešťastia* – utrápene, sklúčene; *smiať sa ako hrdlička* – pekne; *tol'ko je toho ako hydu* – veľmi veľa; *držať sa niekoho, niečoho ako pes ježa* – pevne; *hladný ako vlk/pes* – veľmi; *hádať sa do krvi* – veľmi, ostro; *šetriť s každým halierom* – veľmi; *hanbiť sa do krvi* – veľmi; *presný ako hodiny* – veľmi; *mať sa/najesť sa ako na hodoch* – veľmi dobre; *má sa ako hrach pri ceste* – veľmi zle; *klepať sa ako osika* – veľmi; *jačí/kričí/reve akoby ho z kože drali* – veľmi.

Stávalo sa, že to isté prirovnanie bolo v rozličných heslách zaradená do rozličných kategórií: *hrklo v ňom ako v starých hodinách* – v hesle **hrknúť** v {K} a v hesle **hodiny** v {F}; *leje ako/ani z krhly* – v hesle **liať** v {K} a v hesle **krhla** v {F}; *chudobný ako kostolná myš* – v hesle **kostolný** v {F} a v hesle **chudobný** v {K}; *chodiť okolo niečoho ako mačka okolo horúcej kaše* – v hesle **horúci** v {K}, v hesle **chodiť** a **kaša** v {F}; *klepať sa ako osika* – v {F}; *chvieť sa ako osika* – v {K}. Tento stav tiež svedčí v prospech toho, aby sa ustálené prirovnania uvádzali v hesle pokope, na jednom mieste.

Inokedy autorky použili redukovaný výklad (vykladali len komparatívny zvrat, takže prvú časť frazémy brali v doslovnom význame) a v súlade s koncepciou zaradili celé spojenie do zóny {K}. Na prvý pohľad problém nevznikol. Ale bližší pohľad ukázal, že takýto redukovaný výklad skreslil skutočný význam prirovnania: *hládí ako habakuk* – zamračene (primeranejšie by bolo „je namosúrený“, pretože komponent *hládí* netreba chápať vždy doslovne; význam prirovnania opisuje aj stav mysli); *sedieť ako kvaka* – nečinne, bez záujmu (primeranejšie by bolo „nič nerobiť, o nič sa nezaujímať“, pretože *sedieť* netreba chápať doslovne); *vyzerat' ako strašiak do konopí* a) neupraveno b) nevkusne, výstredne (primeranejšie by bolo „je neupravený; b) je

nevkusne, výstredne oblečený“); *čakať ako koza na nôž* – bezmocne (primeranejšia by bola perifráza „byť zmierený s tým, čo príde, trpne očakávať nepriaznivý osud“, pretože prirovnanie opisuje model správania a komponent *čakať* netreba chápať doslovne); *dostať sa do niečoho/niekam ako Pilát do kréda* – nevinne (návrh nového výkladu: „dostať sa k niečomu významnému bez vlastného pričinenia, bez vlastnej zásluhy, nevedomky, nevinne“).

Niekedy sa pod redukovaným výkladom „skrývali“ dva významy, alebo dve prirovnania. Napr. prirovnanie *klúčkovat' ako zajac* – rýchlo a šikovne. Tu sme navrhli nové spracovanie: *klúčkovat' ako zajac* – a) rýchlo a šikovne sa vyhýbať prekážkam b) konať vyhýbavo a opatrnícky. Pozrime sa na ďalšie prirovnanie: *mať ruky/zuby ako lopaty* – veľké. Toto prirovnanie obsahuje dve frazeologické jednotky: *mať ruky ako lopaty* – mať veľké a široké dlane; *mať zuby ako lopaty* – mať široké predné zuby. Iný príklad: *vrhnúť sa na niečo ako kobyľky* – pahltnie a nenásytne. Návrh nového spracovania: *vrhnúť sa na niečo ako kobyľky* – (o väčšom množstve ľudí) a) pahltnie, nenásytne sa pustiť do jedla b) dravo, bez rozpakov sa usilovať niečo získať. Podobne prirovnanie *rovný ako lata* – vzpriamený. Tu treba rátať s dvoma významami pôvodnej frazémy s odlišnými neživotnými referentmi a ďalšou frazémou, ktorá má životný referent: *rovný ako lata* – a) (o chrbte) nápadne vzpriamený b) (o povrchu) vyrovnaný, hladký; hrub. *rovná ako lata* (o žene) chudá, obyč. s plochou hrudou

Na základe editačnej analýzy hesiel 2. zväzku SSSJ sa redaktorský a autorský kolektív sa rozhodol 1) spracúvať prirovnania ako jeden celok, teda nevyčleňovať prirovnania s jedným frazeologizovaným komponentom do osobitnej skupiny a 2) vykladať všetky prirovnania ako nedeliteľný celok, napr. *tma ako/ani v kochu* – hustá, čierna tma; *spí, akoby ho do vody hodil* – tvrdo spí; *hovorí, ako mu huba narástla* – hovorí priamo, nepretvaruje sa; *ide [si] ako kráľovná* – kráča, vykračuje si hrdo, pyšne; *hladiet' ako hrom do buka/duba/putne* – nahnevane, nevraživo zazeráť.

Záver

Pre spracovanie prirovnaní v SSJ a KSSJ je príznačné ignorovanie celistvého významu niektorých prirovnaní (ale aj iných typov frazém) a vyčleňovanie v nich komponentov s priamym významom a komponentov s posunutým, „frazeologicky viazaným významom“. Zaužívala sa prax vykladať iba tento „frazeologizovaný komponent“. V koncepcii VSRS sa za skutočné prirovnania považovali len ustálené spojenia komparatívneho typu, ktoré mali „frazeologizovanú“ prirovnávaciu časť (komparandum) a prirovnávací základ mal priamy význam.

Prvý zväzok SSSJ nadviazal na chápanie prirovnaní v koncepcii VSRS a na prax redukovaného výkladu uplatňovanú v SSJ a KSSJ.

Vzhľadom na to, že teoreticky je koncept frazeologicky viazaného významu prekonaný nazdávame sa, že je potrebné dôsledky tohto chápania odstrániť aj z lexikografickej praxe. Možno sa opýtať, či je možné zaužívaný spôsob výkladu akceptovať aspoň ako lexikografickú techniku, ktorá šetrí miesto. Aj na túto otázku odpovedáme záporne.

Redukovaný spôsob výkladu nielenže skresľuje význam prirovnania, ale skresľuje podstatu prirovnania ako celistvej idiomatickej jednotky, ktorá nevyjadruje podobnostný vzťah priamo. Od 2. zväzku sa v SSSJ všetky prirovnania (vyčlenené na základe prítomnosti formálneho znaku – spojky) spracúvajú ako jeden frazeologický typ a každé prirovnanie sa vykladá ako jeden významový celok plne idiomatickej povahy.

Literatúra

- ČERMÁK, František: Česká přirovnání. In: Slovník české frazeologie a idiomatiky. Přirovnání. 1. vyd. Praha: Academia 1983, s. 463 – 492.
- ČERMÁK, František: Substance of idioms: Perennial problems, lack of data or theory? In: International Journal of lexicography, 2001, roč. 14, č. 1, s. 2 – 19.
- ČERMÁK, František: Slovník súčasného slovenského jazyka A-G, red. K. Buzássyová a A. Jarošová, Veda 2006 Bratislava, 1132 s. In: Slavica, 2009, roč. 78, č. 1 – 2, s. 113 – 119. (Recenzia)
- KAČALA, Ján: Prirovnanie a jeho funkcie v básňach Milana Rúfusa. In: Frazeologické štúdie V. Princípy lingvistickej analýzy vo frazeológii. Ed. D. Baláková - P. Ďurčo. Ružomberok, Katolícka univerzita – Filozofická fakulta 2007, s. 201 - 209.
- KOČIŠ, František: Poznámky k spracovaniu frazeológie v Slovníku slovenského jazyka. In: Slovenská reč, 1961, roč. 26, č. 5, s. 269 – 280.
- Krátky slovník slovenského jazyka. Red. J. Kačala – M. Pisárčiková. 1. vyd. Bratislava: Veda 1987. 592 s.
- KUČEROVÁ, Eleonóra: Frazeologizované menné slovné spojenia. In: Jazykovedné štúdie. 12. Peciarov zborník. Red. J. Ružička. Bratislava: Veda 1974, s. 222 – 228.
- KUČEROVÁ, Eleonóra: Ustálené slovné spojenia s frazeologizovaným komponentom (v ruštine a slovenčine). In: Slavica Slovaca, 1980, roč. 15, č. 3, s. 237 – 248.
- KUČEROVÁ, Eleonóra: Sémantická analýza „ustálených prirovnání“. In: Jazykovedné štúdie. 16. Materiály z vedeckej konferencie o sémantike (Smolenice, 14. - 17. januára 1980). Red. J. Ružička. Bratislava: Veda 1981, s. 135 – 138.
- KUČEROVÁ, Eleonóra: Sémantická analýza adjektívnych ustálených slovných spojení s frazeologizovaným komponentom (na ruskom a slovenskom materiáli). In: Slavica Slovaca, 1982, roč. 17, č. 1, s. 18 – 31.
- MAN, Oldřich: Ustálená spojení a frazeologické jednotky, jejich podstata a hranice. In: Lexikografický sborník. Materiály z I. celoštátnej konferencie čs. lexikografov, konanej v dňoch 5. – 7. júna 1952 v Bratislave. Bratislava: Vydavateľstvo Slovenskej Akadémie vied 1953, s. 101 – 110.
- MLACEK, Jozef: Základné typy ustáleného prirovnania. In: Jazykovedný časopis, 1970, roč. 21, č. 2, s. 172 – 188.
- MLACEK, Jozef – ĎURČO, Peter a kol.: Frazeologická terminológia. Bratislava: Stimul 1995. 160 s.
- SEKANINOVÁ, Ella – KUČEROVÁ, Eleonóra: Stavba hesla vo Veľkom slovensko-ruskom slovníku II. In: Veľký slovensko-ruský slovník. 2. L - O. Red. E. Sekaninová. Bratislava: Veda 1982, s. 9 – 20.

- Slovník slovenského jazyka I. – V. diel. Red. Š. Peciar. Bratislava: Vydavateľstvo Slovenskej akadémie vied 1959 – 1965.
- Slovník súčasného slovenského jazyka. A – G. [1. zv.]. Red. K. Buzássyová – A. Jarošová. Bratislava: Veda 2006. 1134 s.
- Slovník súčasného slovenského jazyka. H – L. [2. zv.]. Red. A. Jarošová – K. Buzássyová. Bratislava: Veda 2011. 1088 s.
- SMIEŠKOVÁ, Elena: Ustálené prirovnania vo frazeológii (so zreteľom na lexikografické spracovanie). In: Jazykovedný časopis, 1970, roč. 21, č. 2, s. 189 – 201.
- SMIEŠKOVÁ, Elena: Malý frazeologický slovník. 1. vyd. Bratislava: Slovenské pedagogické nakladateľstvo 1974. 296 s.
- Velký slovensko-ruský slovník. 1. zv. Red. D. Kollár. 2. – 6. zv. Red. E. Sekaninová. Bratislava: Veda 1979—1995.
- ŽUKOV, Vlas Platonovič.: Izučenie ruskej frazeologii v otečestvennom jazykoznanii poslednich let. In: Voprosy jazykoznanija, 1967, roč. 15, č. 5, s. 25 – 38.