

POKLADNICE ČESKÉ LEXIKOGRAFICKÉ TRADICE

Boris Lehečka; lehecka@mzk.cz; (Moravská zemská knihovna v Brně)

Článek popisuje projekt na vytvoření úložiště moderních i historických českých slovníků, které bude obsahovat (budou-li k dispozici) bibliografické údaje, metadata, digitalizáty a plné texty ve standardizované podobě, tak aby mohly sloužit širokému spektru zájemců o český jazyk a jeho slovní zásobu. Přibližuje postupy, kterými lze údaje a data automatizovaně získané z digitálních knihoven dále zpracovávat (čistit, sjednocovat a transformovat), aby se hodily k další analýze a sdílení.

<http://doi.org/10.52036/1335793X.2023.SC2.47-51>

ÚVOD

Slovníky představují významný doklad národních a kulturních dějin. Ve své podstatě reflektují jeden z prvků, který charakterizuje moderní národ. Zároveň od autorů vyžaduje výraznou reflexi mateřského, případně druhého jazyka (u překladových slovníků). Za předchůdce slovníků (Večerka, Hladká, 2017) se považují překladové glosy, které překládaly nebo vysvětlovaly neznámý výraz: v národních prostředích se často vyskytují glosy v latinských textech, ale můžeme se setkat i se zachycením češtiny pomocí znaků hebrejské abecedy. Později se začaly objevovat tematické slovníčky připojené k určitému textu, z nich se vyvinuly na textu nezávislé slovníky, nejdříve překladové, později výkladové, které k výkladu významu používají stejný jazyk jako ten, který popisují. Pro češtinu se za předchůdce posledně jmenovaného typu považuje Jungmannův Slovník německo-český (1835–1839; <https://vokabular.ujc.cas.cz/moduly/slovníky/>), sestávající z výkladů v češtině, doplněných o německé, latinské aj. ekvivalenty a citace z pramenů, případně z úzu. Za první český výkladový a dokladový slovník se považuje Příruční slovník jazyka českého (1935 až 1957; <https://psjc.ujc.cas.cz>), za poslední dosud nedokončený Akademický slovník současné češtiny (2017–; <https://slovníkcestiny.cz>). V moderní době existují slovníky v elektronické podobě, vyznačují se nejrozličnějšími funkcemi uživatelského prostředí (kombinované hledání, virtuální klávesnice apod.), některé dokonce umožňují automatizovaný přístup pomocí programového rozhraní API.

Slovníky plnily a plní různé funkce: ulehčují porozumění textu (glosy), usnadňují komunikaci mezi národy (překladové slovníky), nabízejí výběr doporučených výrazů (tzv. jazykové brusy z konce 19. století) nebo realisticky popisují užívanou slovní zásobu (tzv. deskriptivní popis moderního či historického jazyka). Podrobnější rozdělení slovníků srov. např. Novotná (2017).

Doposud byla řeč o slovnících jazykových, které se snaží popsat slova z pohledu jazyka a jeho chápání uživateli. Naproti tomu slovníky encyklopedické zprostředkovávají vědecké poznatky o světě (realitu označovanou pojmem). Ty lze členit např. na obecné (Ottův slovník naučný, 1888 až 1909, https://cs.wikisource.org/wiki/Ottův_slovník_naučný; Wikipedie, <https://cs.wikipedia.org>) nebo specializované, např. Sovák, Edelsberger (2000).

POKLADNICE ČESKÉ LEXIKOGRAFICKÉ TRADICE

Jednotlivým slovníkům, případně jejich předchůdcům, nebo úzce či široce vymezené skupině lexikografických děl se věnuje rozsáhlá odborná literatura. Přesto chybí projekt, který by se pokusil podchytit co nejširší lexikografickou produkci vázanou na český jazyk, případně české území, přičemž z historické perspektivy je vhodnější uvažovat o oblasti někdejších zemí Koruny české. Tento článek je prvním pokusem o představení takového projektu, který nese pracovní pojmenování Pokladnice české lexikografické tradice. Jeho hlavním cílem bude shromáždit údaje bibliografické, obrazové i textové a s pomocí moderních metod analýzy obsahu

nabídnout tyto materiály k dalšímu bádání a využití široké paletě uživatelů.

Jako inspirace může sloužit modul digitalizovaných mluvnic (Koupil, 2012), který je součástí Vokabuláře webového (<https://vokabular.ujc.cas.cz>) a na jednom místě zprostředkovává „digitalizované verze historických mluvnic a podobných příruček z období 16. až 19. století“ (Modul digitalizovaných mluvnic, 2006 až 2023). Každá publikace je opatřena základním popisem a charakteristikou, která ji zařazuje do dobového kontextu, a dále digitální kopii s rejstříkem, tj. seznamem lingvistických termínů pro každou jednotlivou stránku.

Při zpracování materiálů lexikografických se jako ideální jeví vycházet z dosavadní katalogizace a digitalizace, které nabízejí velké množství primárních dat i metadat. Zároveň je potřeba přistupovat k nim kriticky a vyvarovat se zbytečných duplikátů či nekvalitních metadat. To vyžaduje na jedné straně vhodně zvolenou strategii a metody vytěžování existujících úložišť (samozřejmě s ohledem na stav autorských práv, viz např. Lehečka, 2022, 14–18), na straně druhé ruční zásahy a kontrolu, případně tvorbu trénovacích sad pro automatizované zpracování. Vynaložené úsilí přinese komplexnější pohled na českou lexikografickou produkci, zvláště pokud by se podařilo najít takový způsob zpřístupnění shromážděných dat, který umožní exploataci samotných děl a ideálně i jejich srovnání, a to nejen na úrovni bibliografických údajů a metadat, ale rovněž v rámci zachycené slovní zásoby.

Před vytvořením Pokladnice je důležité si položit několik zásadních otázek a odpovědět si na ně:

1) *K čemu může soubor tematicky seskupených metadat, textů a dat lexikografických děl sloužit?*

Pomineme-li jeden z hlavních cílů, tj. nalezení významu daného slova, umožní soubor slovníků empirickou práci v různých perspektivách: diachronní (srovnání autorů s jejich následníky i předchůdci, porovnání různých vydání téhož díla), synchronní (rozsah a repertoár slovní zásoby v různých typech slovníků, vývoj pojmenování, hláskových i pravopisných podob a významů), územní (slovníková produkce českých nakladatelů, zachycení nářečních jevů) apod. Připravená data (samostatné heslové stati) poslouží jako ideální podklad pro strojové učení, neboť učící algoritmy nebudou pracovat s delším úsekem, v němž by se vyskytly výrazy, které spolu souvisejí pouze tím, že v abecedním řazení následují po sobě.

2) *Jaké bude kritérium pro zařazení díla do Pokladnice?*

Primární bude hledisko jazykové, tj. užití češtiny na straně popisované slovní zásoby, popisu významu, případně ekvivalentu. Kvůli využití heslových statí pro strojové učení je potřeba podchytit slovníky jazykové i encyklopedické. Další omezení, např. místo vzniku, národnost autora nebo minimální počet výtisků, nejsou relevantní. Z praktického hlediska, zejména velkého množství položek a objemu dat, je však rozumné (alespoň v počátečních fázích) zpracovávat do roku 1918 (vzniku samostatného československého státu) všechny slovníky včetně překladových, po tomto roce pouze jednojazyčné slovníky. Gramatiky, pravidla pravopisu a podobné jazykové příručky se do Pokladnice dostanou pouze v případě, že obsahují slovníkovou část, viz např. Pohl (1796).

3) *Jaké zdroje pro získání metadat a obrazových či textových dat využívat?*

Ideální zdroj představují digitální knihovny a existující soupisy slovníků knižní i elektronické (viz níže).

4) *Jaké typu údajů budou do Pokladnice vstupovat?*

Půjde zejména o bibliografické záznamy, odkazy na digitální kopie, samotné digitální kopie (pokud to umožní autorská práva), odkaz na existující zpřístupnění, tj. webové stránky, případně i sekundární literatura.

5) *Jak přistupovat k duplicitním údajům?*

Bibliografické záznamy téhož díla je potřeba vztáhnout k záznamu, který bude reprezentovat jedno dílo; duplicitní záznamy je nutné eliminovat, důležité je přitom zachovat informace o provenienci u starších publikací (na rozdíl od moderních děl, které existují v mnoha téměř identických exemplářích). Také u digitálních kopií je potřeba se vyvarovat duplikátů, ale zároveň podchytit digitalizáty různých exemplářů, které se mohou lišit kvalitou, barevností, čitelností apod.

Při deduplikaci je potřeba počítat s faktem, že údaje (autor, tiskárna, místo vydání apod.) nemusejí být v záznamech zpracovány jednotně a teprve po jejich sjednocení (např. převodem na autoritní záznam) může dojít k jejich duplicitě.

6) *Jaké údaje a informace se budou v Pokladnici doplňovat?*

V počátku bude žádoucí, aby se doplnilo, pokud to licence k dílu umožní, automatické rozpoznání tex-

tu (OCR, HTR) u pramenů, které jej nemají k dispozici nebo je příliš chybové. To umožní jejich další intenzivní využití. Ještě větší míru interoperability zajistí převod strojově čitelného textu do specializovaného formátu pro slovníky, např. TEI Lex-0 (viz dále). Velkou přidanou hodnotu bude mít základní popis a charakteristika lexikografického díla, které jej zařadí do dobového kontextu.

SOUČASNÝ STAV INFORMACÍ O LEXIKOGRAFICKÝCH DÍLECH

Budování ucelených sbírek lexikografických děl není moc rozšířené. Obvyklé je seskupení několika souvisejících slovníků, zejména digitalizovaných (rozpoznaných textově) pro určité časové období: viz např. Wörter-bucher (<https://www.dwds.de/d/woerterbuecher>), který propojuje digitální verze slovníků moderní němčiny (Wörterbuch der deutschen Gegenwartssprache z let 1952–1957, Große Wörterbuch der deutschen Sprache, 3. vydání z roku 1999 a některé slovníky z předchozích období, jako je Deutsches Wörterbuch Jacoba a Wilhelma Grimma). Pro český jazyk existuje např. Internetová jazyková příručka (<https://prirucka.ujc.cas.cz>), která vedle poučení o pravopisu nabízí odkazy na výskyty nalezených hesel v moderních výkladových slovnících češtiny, a dále Vokabulář webový (<https://vokabular.ujc.cas.cz/hledani.aspx>) zpřístupňující moderní slovníky historické češtiny a samostatně i vybrané dobové překladové slovníky. Sbírký digitálních kopií jsou velmi vzácné a rovněž tak soubory, které by seskupovaly publikace časově, prostorově i jazykově vzdálené.

Systém Kramerius, který využívají zejména digitální knihovny českých paměťových institucí, v jedné z posledních verzí zavedl možnost vytvářet kurátorské kolekce (v uživatelském rozhraní se označují jako „virtuální sbírky“, resp. „sbírky“ v seznamu faset), které „vznikají v rámci jednoho repozitáře... ve spolupráci s odborníky na dané téma, tj. [procházejí] návrhem, posuzováním stavu a obsahu, zařazením/vyřazením děl. Díla, která jsou pro soubor podstatná a ještě nebyla digitalizována, se zařazují do digitalizačních plánů a průběžně digitalizují, aby vznikl co nejúplnější celek pro dané téma.“ (Lehečka, 2022, s. 43) Žádný z rozsáhlých českých repozitářů (Národní knihovny, Moravské zemské knihovny a Knihovny Akademie věd, viz <https://www.digitalniknihovna.cz>) virtuální sbírku lexikografických děl zatím neobsahuje. Pokud by však vznikl přehled takových děl, mohly by po domluvě s kurátory takové kolekce vzniknout, byť

by byly vázané pouze na konkrétní repozitář (a jeho publikace) a přístup ke všem dílům z jednoho místa by se tím nevyřešil.

Publikace v českých digitálních knihovnách mívají přiřazena klíčová slova (např. cizí slova, oborová terminologie) a žánr (např. jazykové slovníky, dialektologické slovníky; uvádí je pouze MZK), podle kterých lze díla vyhledávat, resp. filtrovat. Nelze se však na ně stoprocentně spolehnout, jak o tom svědčí např. nejednotné označení (jazykové slovníky × slovníky × explanatory dictionaries), případně publikace, které nemají tyto údaje přiřazeny. Dále je možné se orientovat podle slov v názvu díla, ale ani tato cesta nemusí být stoprocentně spolehlivá, viz např. „Příručka cikánštiny“ (Lípa, 1963).

Naštěstí je možné se obrátit na existující oborové bibliografie, např. Bibliografii české lingvistiky (<https://bibliografie.ujc.cas.cz>), a existují i tematické soupisy slovníků, gramatik a obdobných příruček: pro český jazyk jsou relevantní zejména Glück (2002) a Stankiewicz (1984). Tyto seznamy mohou posloužit jako základ pro získání metadat a dat z obvyklých zdrojů, jako jsou Manuscriptorium (<https://www.manuscriptorium.com>), České digitální knihovny (<https://ceskadigitalniknihovna.cz>), portál Knihovny.cz (<https://www.knihovny.cz>) a známé zahraniční zdroje typu Google Books, HathiTrust, digitální sbírky Bavorské zemské knihovny apod. Určitě bude existovat i mnoho případů, kdy je k dispozici pouze bibliografický záznam a publikace se budou muset digitalizovat.

STRUKTURA ZÍSKÁVANÝCH DAT, JEJICH ZÍSKÁVÁNÍ A ZPŘÍSTUPNĚNÍ

ZDROJE DAT

Primárním nositelem informace o lexikografickém díle bude bibliografický záznam, který je možné získat z existujících bibliografických soupisů (Manuscriptorium, Knihopis (<http://www.knihopis.cz>), Bibliografie české lingvistiky) nebo záznamů v digitálních knihovnách. U duplicitních záznamů je potřeba zachovat položky, které se vztahují k různým exemplářům díla, a vztáhnout je k jednomu záznamu. Další podstatnou informací budou podrobné záznamy o existujících digitálních kopiích (jak ve formě digitálních obrázků, tak převodu do strojově čitelné podoby, pokud existuje a byla zpřístupněna); i v tomto případě je podstatné zachovat odkazy na odlišné digitální exempláře, které se můžou lišit kvalitou originálu, digitální

kopie, ale i ne/přítomností rozpoznávaného textu. Je tedy potřeba opatrně přistupovat k agregátorům dat, jako je Manuscriptorium, Europeana, Internet Archive, Česká digitální knihovna nebo Knihovny.cz.

ZÍSKÁVÁNÍ DAT

Při vytěžování agregátorů, digitálních knihoven i bibliografických databází bude vhodné tuto činnost automatizovat, a to vzhledem k velkému množství zdrojů, kde lze relevantní údaje hledat, k velkému počtu údajů, které je potřeba uchovávat a deduplikovat, a také s ohledem na případné aktualizace na straně knihoven nebo vzhledem k možnosti využít sadu nástrojů pro jiný typ dokumentů (např. slovníků jiného jazyka). Inspirací v tomto ohledu může být aplikace VideLibri (Van der Zander, [bez data]). Pro komunikaci s programovým rozhraním postaveným na REST API lze použít např. nástroj Kiota (Microsoft, 2023), který generuje na základě specifikace OpenAPI zdrojový kód pro komunikaci s rozhraním REST API v několika programových jazycích (C#, Go, Java, PHP, Python, Ruby, JavaScript apod.), nebo je možné použít nástroje a jazyky, které mají dobrou podporu zpracování formátů JSON a XML, např. XProc 3.0 a XQuery 3.1, resp. XPath 3.1.

ZPRACOVÁNÍ DAT

Pokud budou k dispozici data zachycující konkrétní slovník (digitální obrázky, text ve strojově čitelném formátu, ideálně standardizovaném, např. ALTO, TEI, JSON, ABBYY FineReader XML apod.) k dispozici pod vhodnou licencí, která umožní jejich sdílení nebo další využití, bude pro potřeby projektu i dalších badatelů vhodné tyto zdroje uložit, zpracovat a dále zpřístupnit. V opačném případě alespoň odkázat na originální úložiště.

Jedním z prvků, který přispěje k jednoduššímu a pokročilejšímu využití slovníků, se může stát Textbite, který vzniká v projektu NAKI III SemANT – sémantický průzkumník textového kulturního dědictví (řešitelé Vysoké učení technické v Brně, Moravská zemská knihovna v Brně, Masarykova univerzita, 2023–2027). Jedná se o obecný nástroj pro segmentaci dokumentů založený na existujících metodách, které budou upraveny pro potřeby projektu SemANT a později i Pokladnice. Textbite bude využívat základních schopností současného zpracování rozložení stránek v balíčku pero-ocr (detekce řádků a odhad hranic odstavců). Podle předběžného záměru budou detekované řádky tvořit uzly grafové neuronové sítě, která bude získávat příznaky popisující uzly jak z po-

zičního embeddingu a z příznaků konvoluční neuronové sítě zpracovávající obraz, tak z embeddingu automaticky rozpoznávaného textu. Příznaky hran grafu budou získávány také zakódováním vzájemné pozice řádků a vizuálními příznaky získanými z aktivací konvoluční sítě, ale budou doplněny i o ohodnocení neuronovou sítí, jak je pravděpodobné, že dané řádky na sebe navazují. Díky tomu bude možné získat rozdělení na řádky, přičemž dělení na části zajistí shlukovací algoritmus.

Ve struktuře slovníků, které vyšly tiskem, hraje důležitou úlohu typografická stránka díla, která pomocí jednotného použití odlišných typů, velikostí nebo řezů písma definuje mikrostrukturu hesla, tj. rozdělení informací o lexikální jednotce na specifické části, jako jsou např. heslo, slovní druh, koncovka opěrného tvaru, definice významu, text dokladu a označení jeho pramene atp. Tyto údaje výstupy poskytované aplikací PERO OCR (bez data) naneštěstí nezachycují. Vzhledem k provázanosti zmíněného nástroje Textbite s balíčkem pero-ocr by bylo vhodné, kdyby tento nástroj uměl detekovat i uvedené kvality textu. V opačném případě bude potřeba najít jiný způsob, jak je extrahovat. Existuje např. aplikace GROBID-Dictionaries (<https://github.com/MedKhem/grobid-dictionaries>), která pomocí strojového učení slovníky rozdělí na jednotlivé heslové stati i jejich dílčí části. Pro každý slovník je ovšem potřeba vytvořit trénovací data v rozsahu cca 50 stran.

Probíhají však i další výzkumy, např. v rámci nedávno ustanoveného projektu COLaF (Corpus and Tools for the Languages of France, viz <https://almanach.inria.fr/projects-fr.html>), jehož cílem je na základě vizuální podoby slovníku (např. ve formátu HTML) generovat zdrojová data ve formátu TEI Lex-0 (<https://bit.ly/tei-lex-0>). Ten vychází z obecného standardu TEI (Text Encoding Initiative, <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>) pro zachycení širokého spektra tiskopisných a rukopisných materiálů pomocí XML. TEI Lex-0 se zaměřuje pouze na metadata a text slovníků, přičemž se snaží původně široký repertoár prvků pro označení stejného jevu omezit na přijatelné minimum, které však stále dokáže postihnout různorodost lexikografické produkce. Hlavním cílem nového lexikografického standardu je najít způsob, jak v počítačově čitelné podobě zachytit různé slovníky, umožnit jejich sdílení, vzájemné propojení a srovnání. Kromě standardu samotného se postupně objevují nástroje, které k tomuto cíli směřují, např. TEI Lex-0 Publisher (<https://bit.ly/tei>

-lex-0-publisher), sloužící pro publikování slovníků, resp. generování webových aplikací pro slovníky. TEI Lex-0 představuje z tohoto pohledu jeden z vhodných formátů, který by Pokladnice české lexikografické tradice mohla využívat a nabízet.

VYUŽITÍ POKLADNICE ČESKÉ LEXIKOGRAFICKÉ TRADICE

Přepokládanými zájemci o materiály budou lingvisté zaměřeni na široké spektrum problematiky (lexikografie, lexikologie, historická mluvnice), historici, literární vědci, studenti, ale i laická veřejnost, která zde najede poučení o konkrétních lexikálních jednotkách a obecný vhled do specifické oblasti lidského poznání a bádání.

K tomu je potřeba poskytnout jim vhodné nástroje: hledání podle zvolených kritérií (období, jazyk, autor, titul, dostupnost originálu nebo kopie apod.), zobrazení digitální kopie (s využitím IIIF, odkazem na originál aj.), případně vyhledávání a zobrazení extrahovaných položek typu heslové stati, chronologické seřazení uváděných dokladů. Počítáme také s pokročilejšími funkcemi, které dokáží např. odpovědět na otázku, jakým způsobem se označoval určitý pojem v různých obdobích nebo na různých územích; kdy je poprvé zachycen a lexikograficky zpracován určitý výraz, případně jaký objem zpracované slovní zásoby je shodný pro zvolené slovníky.

ZÁVĚR

Pokladnice české lexikografické tradice představuje návrh dlouhodobého projektu na vybudování digitálního repozitáře českých slovníků, který bude obsahovat bibliografické údaje, metadata, digitalizáty a plné texty ve standardizované podobě, tak aby mohly sloužit širokému spektru zájemců o český jazyk, od laiků až po odborníky na počítačové zpracování přirozeného jazyka. Bude se ve velké míře opírat o dostupné digitální knihovny a bibliografické zdroje, čistit a sjednocovat dostupná data a dále je přetvářet na formáty určené pro slovníky a jejich sdílení.

DEDIKACE

Publikace vznikla v rámci Institucionální podpory na dlouhodobý koncepční rozvoj výzkumné organizace – Moravská zemská knihovna v Brně.

LITERATURA

GLÜCK, Helmut, 2002. *Deutsche Sprachbücher in Böhmen und Mähren vom 15. Jahrhundert bis 1918: eine teilkommentierte Bibliographie*. Berlin: de Gruyter. Geschichte des Deutschen als Fre-

mdsprache. ISBN 3110175045.

KOUPIL, Ondřej, 2012. Modul digitalizovaných mluvnic: co, k čemu, pro koho. In: Krušínský, R, Vintrová, T., eds. *Bibliotheca antiqua 2012. Sborník z 21. konference*. Olomouc: Vědecká knihovna v Olomouci, s. 123–126. ISBN 978-80-7053-298-0. Dostupné z: <http://www.vkol.cz/data/soubory/import/konf21/Bibliotheca%20Antiqua%202012-Koupil.pdf>.

LEHEČKA, Boris, David NOVÁK, Filip KERSCH, et al., 2022. *Metodika přípravy dat z digitálních knihoven pro využití v digitálních humanitních vědách*. Knihovna AV ČR. Dostupné také z: http://invenio.nusl.cz/record/511549/files/Metodika_DL4DH.pdf.

LÍPA, Jiří, 1963. *Příručka cikánštiny*. Praha: Státní pedagogické nakladatelství.

MICROSOFT, 2023. *Kiota* [online]. [cit. 2023-11-13]. Dostupné z: <https://github.com/microsoft/kiota>

Modul digitalizovaných mluvnic: Základní informace, 2006–2023. *Vokabulář webový* [online]. 20. listopadu 2020 [cit. 2023-11-14]. Dostupné z: <https://vokabular.ujc.cas.cz/moduly/mluvnice/uvodni-stranka>

NOVOTNÁ, Renata, 2017: SLOVNÍKY – TYPOLOGIE. In: Karlík, Petr, Nekula, Marek, Pleskalová, Jana (eds.), *CzechEncy - Nový encyklopedický slovník češtiny*. [cit. 2023-11-14]. Dostupné z: <https://www.czechency.org/slovník/SLOVNÍKY – TYPOLOGIE>

PERO OCR: *demonstration application* [online], bez data. Brno: Vysoké učení technické v Brně [cit. 2023-11-04]. Dostupné z: <https://pero-ocr.fit.vutbr.cz>

POHL, Jan Václav, 1756. *GRAMMATICA LINGUAE BOHEMICAE, Oder Die Böhmische Sprach-Kunst: Bestehend in vier Theilen, Benanntlich: I. Der Orthographi, oder Schreib-Kunst. II. Der Ethymologi, oder Wortforschung. III. Der Syntaxi, oder Wörterfügung. IV. Der Prosodi, oder Aussprach-Kunst. Allen Liebhabern dieser Sprache, sowohl Lehrenden als Lernenden, zu einem erforderlichen Werkzeug gefertigt*. Wien: Gedruckt und verlegt bey Johann Thomass Trattnern, kaiserl. königl. Hofbuchdr. und Buchhändlern. [8], 340, [3] s.

STANKIEWICZ, Edward, 1984. *Grammars and Dictionaries of the Slavic Languages from the Middle Ages up to 1850*. Berlin: de Gruyter. ISBN 3110097788.

SOVÁK, Miloš a Ludvík EDELSBERGER, 2000. Defektologický slovník. 3. upr. vyd. Jinočany: H & H. ISBN 80-86022-76-5. Dostupné také z: <https://www.digitalniknihovna.cz/mzk/uuid/uuid:643d4d30-96a8-11e6-9325-005056827e52>.

VAN DER ZANDER, Benito, [bez data]. *VideLibri* [online]. [cit. 2023-11-13]. Dostupné z: <https://www.videlibri.de>

VEČERKA, Radoslav a Zdeňka HLADKÁ, 2017: SLOVNÍKY ČEŠTINY. In: Karlík, Petr, Nekula, Marek, Pleskalová, Jana (eds.), *CzechEncy - Nový encyklopedický slovník češtiny*. [cit. 2023-11-14]. Dostupné z: <https://www.czechency.org/slovník/SLOVNÍKY ČEŠTINY>

■ Článek bol recenzovaný