

**Univerzita Karlova v Praze**

**Filozofická fakulta**

*Ústav Českého národního korpusu*

*Matematická lingvistika*

## **Teze disertační práce**

*Srovnávací aspekty lotyšského a českého lexikonu:*

*Materiály k sestavení lotyšsko-českého slovníku*

*Comparative aspects of Latvian and Czech lexicons:*

*Materials for assembling a Latvian-Czech dictionary*

vedoucí práce: Prof. PhDr. František Čermák, DrSc.

konzultant: Mgr. Pavel Štoll, Ph.D.

2016

*Michal Škrabal*

Cílem této práce bylo z lexikografické perspektivy analyzovat a následně popsat přiměřeně velkou část lotyšského lexikonu pro potřeby vznikajícího lotyšsko-českého překladového slovníku. K analýze jsme účelně vybírali vesměs takové jevy, které zaujímají v jazykovém systému lotyštiny centrální pozici, a dá se tudíž očekávat i jejich hojné zastoupení v chystaném slovníku, úkazům periferním jsme se věnovali jen okrajově. Vycházeli jsme primárně z jejich stávajícího popisu v lotyšských mluvnicích a lexikografických příručkách, konfrontovali jsme jej však s korpusovými daty jakožto s reprezentativním vzorkem soudobé lotyštiny. Ukázalo se, že mnohdy si tyto dva prameny více či méně protičeří, jinými slovy: že teoretický popis jazyka není v souladu s jeho reálnou podobou, tj. aktuálním živým územ, v některých případech za ním dosti zaostává. Nejde nicméně o poznatek nikterak převratný: zastarávání slovníků je všeobecně známý fenomén, a svým způsobem i nevyhnutelný. Dnešní lexikografie však má nově k dispozici jazykové korpusy spolu s robustními nástroji pro jejich vytěžování, díky nimž se radikálně zkrátila přípravná fáze tvorby slovníku, časově vůbec nejnáročnější. Ta spočívala ve sběru jazykového materiálu (převážně formou excerpce textů) a nezřídka se protáhla na mnoho let, ba i pár dekád. Výsledný slovník tak byl zastaralý de facto již v okamžiku svého vydání. Korpusy tento handicap odstraňují: disponuje-li jazyk lexikografova zájmu vyváženým a dostatečně velkým reprezentativním korpusem, odpadá problém s akvizicí materiálu, a o to blíže se slovníkář ocitá vlastní lexikografické, tedy syntetické práci.

Co však i nadále zůstává v přípravné etapě na lexikografických bedrech, je vlastní koncepce slovníku. Tato práce má posloužit jako nezbytný teoretický fundament pro zamýšlené lexikografické dílo a zajistit tak mimo jiné profesionální, nediletantský přístup k slovníkářské práci, (nejen) v dnešní době hojně rozšířený. Chystaný lotyšsko-český slovník je vůbec první – a nejspíš nadlouho i poslední – příručkou svého druhu, a měl by tudíž být co možná nejvšestrannějším nástrojem, pomocí něhož lze přeložit širokou škálu textů: od

klasického literárního díla přes aktuální publicistiku až po odborný text. Cílem není zaplnit stávající mezeru narychlo stvořeným, okleštěným kompilátem existujících slovníků, vycházejícím vstříc pouze nejzákladnějším potřebám uživatelů. Toto vědomí odpovědnosti klade na autora slovníku nemalé nároky: k práci je třeba přistupovat maximálně zodpovědně, nejen s dobrou praktickou znalostí výchozího jazyka, ale i s nezbytnou teoretickou přípravou a v neposlední řadě i s akribií.

Struktura studie je dána chronologickým sledem jednotlivých fází tvorby slovníku. Nejprve jsme analyzovali výchozí situaci, v níž se tento projekt nachází: sledován byl aktuální stav lotyšské lexikografie, kvantitativní a korpusové lingvistiky a také výsledky na poli vzájemné, tj. česko-baltské lexikografie. Dále bylo zapotřebí definovat adresáta slovníku, od nějž se následně odvíjí další postup. Chystaný slovník byl vymezen z hlediska typologického, byly stanoveny jeho cíle. Následně jsme shromáždili různé typy pramenů – jazykové korpusy, slovníky a jazykovědné příručky, doplňkově též internetové zdroje a excerpta – a vybudovali z nich materiálovou základnu.

Důležitou roli hraje v dnešní lexikografii i technická stránka zpracování jazykového materiálu, dnes již výhradně pomocí počítačových nástrojů v podobě specializovaného lexikografického softwaru (v našem případě program TshwaneLex, opírající se o programovací jazyk XML). Klíčová je příprava tzv. definice typu dokumentu (DTD), jež pomocí sady pravidel (hovoří se občas i o „gramatice slovníku“) generuje univerzální hierarchickou, stromovou strukturu slovníkového hesla. Snad největší výhodou je – kromě její flexibility – možnost dílčích úprav této struktury i v průběhu vlastní tvorby slovníku (např. v důsledku dodatečných zjištění plynoucích z analýzy materiálu), aniž by to mělo fatální dopad na vývoj celého počínu. Obecně lze tvrdit, že lexikografický software přispívá ke konzistenci a koherenci výsledného díla a potlačuje negativní vliv lidského faktoru.

Technické stránce projektu věnujeme zvláštní pozornost záměrně: je pro zdar naší práce svým způsobem zásadní, ba přesahuje ji. Většina údajů zaznamenávaných do databáze, která poslouží za základ k vytvoření lotyšsko-českého slovníku, je veřejné povahy – měly by se tudíž do výsledného díla promítnout. Kromě toho jsou tu však také informace neveřejné, sloužící interním potřebám zpracovatele, jež lze i nadále rozšiřovat jednoduchou úpravou DTD struktury. Databáze se tak potenciálně stává v širším smyslu i znalostní databází o lotyštině vůbec – a jako takovou i pramenem k dalším projektům rázu lingvistického, translatologického či pedagogického, může být využita též pro účely strojového zpracování přirozeného jazyka. V možnosti oddělení obsahu lexikálního popisu od jeho formy (která je plně přizpůsobitelná individuálním potřebám jak lexikografa, tak cílového adresáta) vidíme mezník lexikografické teorie i praxe.

Třetí část, tvořící jádro celé studie, obsahově pokrývá jednotlivé etapy vlastní lexikografické práce: od sestavení hesláře přes tvorbu jednotlivých heslových statí a doprovodného informačního aparátu až po finální fázi editace a kontrolu rukopisu slovníku. Má teoreticko-praktickou charakter: jednotlivé problémy jsou teoreticky nazírány s oporou v autoritativních jazykovědných příručkách a konfrontovány jak s korpusovými daty, tak s popisem ve slovnících, načež jsou popsány rozhodovací strategie vedoucí k předkládaným lexikografickým řešením. Ta jsou aktuálními výstupy z lexikální databáze vznikající v rámci programu TshwaneLex, mají prozatím pracovní povahu a mohou ve finální verzi ještě doznat některých parciálních změn. Nemálo problémů se již však podařilo zdárně vyřešit a najít (alespoň z naší perspektivy) optimální formát pro lexikografický popis. Lexikograf by však neměl zapomínat, že každé slovníkové dílo zůstane vždy jen pouhou aproximací jazyka, stejně tak je třeba „vyrovnat se se svou omezenou znalostí. [...] **Slovník je uměním možného.**“ (Čermák 1995d: 231).

Konečně čtvrtá část ukazuje na třech případových studiích komplexní charakter slovníkářské práce. Lexikograf má na začátku pouze lemma, které musí doplnit o zbytek heslové stati; od analýzy, vycházející z korpusových dat a popisu lexému ve slovníkových příručkách, gramatikách a jiných zdrojích, se dobírá k syntéze. Obvykle přitom řeší vícero lexikografických problémů současně. Shrňeme-li výsledky této kazuistiky, vyjevuje se, že výsledná podoba ukázkových hesel se liší od řešení uvedených jak v jiných překladových, tak ve výkladových slovnících. Přičítáme to tomu, že jako v jediném případě bylo soustavně využito korpusových dat, respektive výchozí popis byl s nimi porovnán a náležitě zohledněn.

V prvním případě jsme odmítli výkladovými slovníky nabízenou sémantickou segmentaci lemmatu *biedrs*: 1. druh, přítel, kolega; 2. člen; 3. soudruh. Na základě ruční analýzy korpusových dat jsme objevili přehlížený význam, mimochodem frekventovanější než dobově příznakový třetí význam; podle frekvence jsme upravili rovněž pořadí na prvních dvou místech. Výslednou kostru (1. člen; 2. druh, přítel, kolega; 3. zástupce; 4. soudruh) jsme následně obalili živou tkání v podobě četných kolokací; náš slovník jich má ze všech překladových slovníků bezpochyby nejvíce.

**biedr|s** m<sub>1</sub> (Vsg ~i!) **1.** (politické strany, organizace, klubu, komise ap.) *člen* • sociāldemokrātu partijas b. *člen sociální demokracie* • arodbiedrības b. *člen odborů, odborář* • kultūras komisijas b. *člen kulturní komise* • ~a karte *členská legitimace* • ~a nauda *členský poplatek* • ~u sapulce *členská schůze* • kļūt par goda ~u *stát se čestným členem* • biržas b. <ekon> *člen burzy* (cenných papírů) || **2.** *kolega, druh, společník, kamarád* • darba b. *spolupracovník, kolega z práce* • skolas / klases / studiju b. *spolužák (kolega) ze školy / z třídy / ze studií* • dzīves b. *životní partner, druh* • komandas b. *spoluhrač, týmový kolega* • viņš ir lielisks sarunu b. *skvěle se s ním povídá, je to skvělý společník* • cīņu b. *spolubojovník, kamarád ve zbrani* • domu b. *ideový soupevník* • ceļa b. *spolucestující, soupevník* • istabas b. *spolubydlící* || **3.** <polit> *zástupce, náměstek* • ministru prezidenta b. *místopředseda vlády, vicepremiér* • Saeimas priekšsēdētāja b. *místopředseda parlamentu* || **4.** <sov> *soudruh* • b. Kalniņš *soudruh Kalniņš* • cienījamie ~i *vážení soudruzi* • ~u tiesa *soudružský soud*

V druhé studii jsme, s ohledem na dostatečnou frekvenci v korpusu, přidali k lemmatu spojky *turpretī/turpretim* její běžně neuváděnou variantu *turpretīm*, rovněž v korpusu jsme hledali rozdíly ve srovnání se spojkou *kurpretī/kurpretim*, běžně pojímanou jako synonymní. Specifický úzus konjunkcí jsme se pokusili vystihnout dostatečně širokým kontextem.

**turpretī; turpretim; turpretīm** <hov> [vše -'pre-] konj *zatímco, kdežto, naopak, naproti tomu, na druhou stranu, zato* • *zelta cena pieaugš, t. naftas cena kritīsies cena zlata vzroste, zatímco cena ropy spadne* • *man t. šķiet, ka.. mnē se naopak zdā, že...* • *Izlikties par muļķi dažreiz ir ļoti gudri. T. izlikties par gudru vienmēr ir muļķīgi. Vydávat se za hlupáka je často velmi chytré. Zato vydávat se za chytráka je vždy hloupé.*

**kurpretī; kurpretim** [-'pre-] konj *zatímco, kdežto* • *vīrieši ir galvenokārt pārstāvēti augstos amatos, k. sievietes pārsvarā ieņem zemus amatus muži jsou zastoupeni především na vysokých pozicích, zatímco ženy zastávají pozice převážně nízké*

Podřadící časová spojka *iekams* má dokonce čtyři kodifikované varianty. Popis v gramatikách či slovnících však postrádá jakoukoliv zmínku o frekvenci jednotlivých forem, informaci je potřeba hledat v korpusech. Při analýze dat vycházíme ze sémantické odlišnosti dvou základních českých ekvivalentů (*než* × *dokud* [ne–]), které konstituují samostatné významy (srov. totéž vydělení v LP, LL i LR), k nim hledáme další možné české protějšky. Kratší varianty bez koncového –s hodnotíme pro jejich nízkou četnost jako druhotné a kvůli možné kolizi v abecedním řazení s heslem *iekampt* je zpracováváme jako samostatné, odkazové heslo. Příslušný fragment v LČ pak vypadá takto:

**ieka|lt** v<sub>1</sub> [...]

**iekam; iekām** konj ↓iekams

**iekam|pt** v<sub>1</sub> [...]

**iekams; iekāms** konj **1.** *než, dříve než, předtím než* • *es ilgi domāju, ie. piekritu dluho jsem přemýšlel, než jsem souhlasil* • *ie. es beigšu, es gribētu pateikt.. dříve než skončím, rád bych řekl...* || **2.** (se záporným tvarem slovesa) *dokud* • *ie. nav par vēlu a) dokud není pozdě b) [dříve] než bude pozdě* • *ikviens uzskatāms par nevainīgu, ie. viņa vaina nav atzīta saskaņā ar likumu každý je považován za nevinného, dokud není jeho vina uznána v souladu se zákonem* • *finansējuma problēmas jautājumā nekas nemainīsies, ie. uzņēmējiem nebūs pieejami lēti kredīti v otázce financování se nic nezmění, dokud podnikatelé nebudou mít k dispozici levné úvěry*

**iekāpšan|a** f<sub>4</sub> [...]

Ve třetím případě bylo navrženo progresivní lexikografické řešení, opírající se primárně o korpusová data (zde už přecházíme od přístupu *corpus-based* k pojetí *corpus-driven*), zatímco tradiční zpracování akcentuje systémovou stránku lexému.

**ār|s**<sup>2</sup> [â] m<sub>1</sub> sg: • uz ~u • logi (durvis) veras uz ~u • izlauzties uz ~u • izvirzīt uz ~u • kažoks ar vilnu uz ~u || • staigāt / dzīvot pa ~u • bērni rotaļājas pa ~u || • no ~a • ienākt no ~a • no ~a nekas nav redzams || • ~a durvis / logs • ~a gaisa temperatūra • ~a aploks • ~a darbi || • ~a bērzs <bot> • ~a kultūras <zem>

↓

**āra** G<sub>sg</sub> *venkovní, vnější* [...]

**ārā** adv **1.** (kam) *ven* [...] || **2.** (kde) *venku* [...]

...

**ārs**<sup>1</sup> [ā] m<sub>1</sub> *ar* [...]

**ārs**<sup>2</sup> [â] ↑āra ↓no āra, pa āru, uz āru

...

**no āra** adv *zvenči, zvenku* [...]

...

**pa āru** adv *venku* [...]

...

**uz āru** adv *vně, ven, směrem ven* [...]

## Bibliografie

### a) Slovníky

- Andronov, A. V. (2002). *Matérialy dlja latyšsko-russkogo slovarja. Latviešu-krievu vārdnīcas materiāli*. Sankt-Petěrburg: Filologičeskij fakultēt SPbGU
- Baldunčiks, J. (2007). *Svešvārdu vārdnīca*. Rīga: Jumava.
- Bankava, B. (2007). *Eponīmu vārdnīca*. Rīga: Valsts valodas aģentūra.
- Bankavs, A. (2003). *Latviešu valodas saīsinājumu vārdnīca*. Rīga: Avots.
- BLJuT (2004). Vjaterē, L. – Vjaterē, I.: *Bieži lietoti jēdzieni un termini*. Rīga: Avots.
- ČL (1988). *Česko-lotyšský slovník. Čehu-latviešu vārdnīca*. Rīga: Avots.
- ČLit (2004). Piročkinas, A.: *Česko-litevský litevsko-český slovník. Čekų-lietuvių lietuvių-čekų kalbų žodynas*. Vilnius: Žodynas.
- ČLN (2006). Nikuļceva, S.: *Česko-lotyšský slovník. Čehu-latviešu vārdnīca*. Praha: Leda.
- Entner, H. (1928). *Slovníček česko-lotyšský*. Praha: Československo-Lotyšská společnost.
- Hartmann, R. R. K. – James, G. (1998). *Dictionary of Lexicography*. London: Routledge.
- ISV (2005). *Ilustrētā svešvārdu vārdnīca*. Rīga: Avots.
- Kavaleris, A. (2002). *Tā runā zonā. Latvijas argo – kriminālvides žargona vārdnīca*. Rīga: Preses nams.
- Klégr, A. (2007). *Tezaurus jazyka českého*. Praha: Nakladatelství Lidové noviny.
- Kursīte, J. (2007). *Neakadēmiskā latviešu valodas vārdnīca jeb novadu vārdene*. Rīga: Madris.
- Kursīte, J. (2009). *Tautlietu vārdene*. Rīga: Nemateriālā kultūras mantojuma valsts aģentūra.
- Kuzina, V. (1998a). *3000 latviešu sarunvalodas biežāk lietotie vārdi ar tulkojumu krievu, vācu un angļu valodā*. Rīga: Valsts valodas centrs.
- LE (2015). *Lāti-Eesti sõnaraamat*. Tallinn: Eesti Keele Sihtasutus.
- LFV (2000). Laua, A. – Ezeriņa, A. – Veinberga, S.: *Latviešu frazeoloģijas vārdnīca*. Rīga: Avots.
- LKIV (2010). Zuicena, I. – Žitareva, M. – Burkevica, O.: *Latviešu-krievu ilustrētā vārdnīca. Latyšsko-russkij illjustrirovannyj slovar*. Rīga: Zvaigzne ABC.
- LL (2003). Butkus, A.: *Latvių-lietuvių kalbų žodynas. Latviešu-lietuvišu vārdnīca*. Kaunas: Aesti.
- LLVV (1972–1996). *Latviešu literārās valodas vārdnīca*. Rīga: Zinātne. Dostupné též z: <http://www.tezaurs.lv/llvv/>
- LP (2011). *Latviešu-poļu vārdnīca. Słownik lotewsko-polski*. Rīga: Zvaigzne ABC.
- LPA (2008). *Lielsais pasaules atlants*. Rīga: Karšu izdevniecība Jāņa sēta.
- LR (2006). *Latviešu-krievu vārdnīca. Latyšsko-russkij slovar*. Rīga: Univers – Avots.
- LVBV (1973). Jakubaite, T. et al.: *Latviešu valodas biežuma vārdnīca. [Apvienotais (1.–3.) sējums]*. Rīga: Zinātne.
- LVSiv (2002). Grīnberga, E. et al.: *Latviešu valodas sinonīmu vārdnīca*. Rīga: Avots.
- LVSv (2006). Bušs, O. – Ernstsone, V.: *Latviešu valodas slenga vārdnīca*. Rīga: Norden AB.
- LVV (2006). *Latviešu valodas vārdnīca*. Rīga: Avots. Dostupné též z: <http://www.tezaurs.lv/lvv/>



- LVPPV (1995). *Latviešu valodas pareizrakstības un pareizrūnas vārdnīca*. Rīga: Avots.
- LVŽV (<sup>3</sup>2005). Mirovics, I. – Dubaus, A. et al.: *Latviešu valodas žargona vārdnīca*. Rīga: Avots.
- MLVV: *Mūsdienu latviešu valodas vārdnīca* [online]. Dostupné z:  
<http://www.tezaurs.lv/mlvv/>
- NASCS (2005). Kraus, J. et al.: *Nový akademický slovník cizích slov A–Ž*. Praha: Academia.
- Pala, K. – Všíanský, J. (2000). *Slovník českých synonym*. Praha: Nakladatelství Lidové noviny.
- Rozenvalde, I (2004). *Preses lasītāja svešvārdu vārdnīca*. Rīga: Nordik.
- Ruberte I. (2004). *Senvārdu vārdnīca*. Rīga: Zvaigzne ABC.
- SČFI (<sup>2</sup>2009). Čermák, F. et al.: *Slovník české frazeologie a idiomatiky* [sv. 1–4]. Praha: Leda.
- Siliņš, K. (1990). *Latviešu personvārdu vārdnīca*. Rīga: Zinātne.
- SSJČ (1957–1971). Havránek, B. et al.: *Slovník spisovného jazyka českého* [sv. 1–4]. Praha: Academia.
- SSČ (<sup>2</sup>1994). *Slovník spisovné češtiny pro školu a veřejnost*. Praha: Academia.
- SV: Spektors, A.: *Skaidrojošā vārdnīca* [online]. Dostupné z:  
<http://www.tezaurs.lv/sv/>
- Veisbergs, A. (2005). *Jaunā latviešu-angļu vārdnīca*. Rīga: Zvaigzne ABC.
- VPSV (2007). Skujiņa, V. et al.: *Valodniecības pamatterminu skaidrojošā vārdnīca*. Rīga: LU Latviešu valodas institūts.

## **b) Odborné publikace**

- Adamska-Sałaciak, A. (2006). *Meaning and the Bilingual Dictionary*. Frankfurt am Main: Peter Lang, 2006.
- Adamska-Sałaciak, A. (2010). Examining Equivalence. *International Journal of Lexicography*, Vol. 23, No. 4, 2010, s. 387–409.
- Andronova, E. – Andronovs, A. (2011). Latviešu valodas korpuss un tā izmantošana. In O. Bušs (ed.). *Valodas prakse: vērojumi un ieteikumi. Populārzinātnisku rakstu krājums* 6. Rīga: Latviešu valodas aģentūra, s. 41–57.
- Andronovs, A. (1998). Par vārda gramatisko raksturojumu vārdnīcā. *Linguistica Lettica* 3. Rīga: LU Latviešu valodas institūts, s. 141–151.
- Atkins, B. T. S. (1996). Bilingual dictionaries – past, present and future. In M. Gellerstam et al. (eds.). *EURALEX '96 Proceedings I–II: Papers submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*. Göteborg: University of Göteborg, Department of Swedish, s. 515–546.
- Atkins, B. T. S. (2002). Then and Now: Competence and Performance in 35 Years of Lexicography. In A. Braasch – C. Povlsen (eds.). *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, Copenhagen, Denmark, August 13–17, 2002*. Copenhagen: Center for Sprogteknologi, s. 1–28.
- Atkins, B. T. S. – Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. New York: Oxford University Press Inc.
- Auziņa, I. (2010). Towards Spoken Latvian Corpus: Current Situation, Methodology and Development. In I. Skaldiņa – A. Vasiljevs (eds.). *Human Language Technologies –*

- The Baltic Perspective. Proceedings of the Fourth International Conference HLT 2010.* Amsterdam: IOS Press, s. 39–44.
- Balázsová, L. (2008). Editácia – kľúčová etapa práce na slovníku. In A. Jarošová – A. Rangelova – J. Světlá (eds.). *Lexikografie v kontextu informační společnosti.* Praha: Ústav pro jazyk český, v. v. i., s. 133–142.
- Baldunčiks, J. (ed.) (2012). *Vārdnīcu izstrāde Latvijā 1991–2010.* Rīga: Latviešu valodas aģentūra.
- Baldunčiks, J. (2015). Vārdu un vārdu nozīmju [ne]iekļaušana vārdnīcā: vērojumi un secinājumi. In J. Baldunčiks – A. Veisbergs (eds.). *Valsts valodas komisija raksti. 7. sējums. Vārdnīcas un valoda.* Rīga: Zinātne.
- Béjoint, H. (2000). *Modern Lexicography. An Introduction.* Oxford: Oxford University Press.
- Bergenholtz, H. – Tarp, S. (eds.) (1995). *Manual of Specialised Lexicography.* Amsterdam – Philadelphia: John Benjamins Publishing Company.
- [Běřáková, M. E.] (2012). *Lotyština. Konverzace.* Lingea: Brno.
- Bezděk, J. (1974). Ekvivalent slova v jazyce cílovém. In *Otázky lexikografie a vydávání slovníků. Z materiálů VIII. mezinárodní porady představitelů socialistických zemí vydávajících slovníky. Praha 8.–9. září 1972.* Praha: SPN, s. 139–213.
- Blinkena, A. (2009). *Latviešu interpunkcija.* Rīga: Zvaigzne ABC.
- Bosák, J. (2008). Odras dynamiky jazyka v sústave funkčných kvalifikátorov. In A. Jarošová – A. Rangelova – J. Světlá (eds.). *Lexikografie v kontextu informační společnosti.* Praha: Ústav pro jazyk český, v. v. i., s. 123–131.
- Bušs, O. (2005). Pēc 25 gadiem jeb Barbarismi-2005. In S. Lagzdīņa (ed.). *Valodas prakse: vērojumi un ieteikumi 1.* Rīga: LU Akadēmiskais apgāds, s. 39–46.
- Bušs, O. (2015). Sõnaraamat ja/või eksperiment. *Keel ja kirjandus* 10, 2015, s. 744–746.
- Clear, J. (1996). Technical Implications of Multilingual Corpus Lexicography. *International Journal of Lexicography* 9 (3), s. 265–273.
- Corde, A. et al. (1998). Automatic Reversal of a Bilingual Dictionary: Implications for Lexicographic Work. In T. Fontanelle et al. (eds.). *Actes EURALEX '98: Communications soumises à EURALEX'98 (Huitième Congrès International de Lexicographie) à Liège, Belgique.* Liège: Université de Liège, Département d'anglais et de néerlandais, s. 433–443.
- Cvrček, V. – Kovářiková, D. (2011). Možnosti a meze korpusové lingvistiky. *Naše řeč* 94, 2011, s. 113–133.
- Cvrček, V. et al. (2015). *Mluvnice současné češtiny. 1/ Jak se píše a jak se mluví.* Praha: Karolinum.
- Čermák, F. (1995a). Frazieografie. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie.* Jinočany: H&H, s. 116–136.
- Čermák, F. (1995b). Komputační lexikografie. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie.* Jinočany: H&H, s. 50–71.
- Čermák, F. (1995c). Paradigmatika a syntagmatika slovníku: možnosti a výhledy. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie.* Jinočany: H&H, s. 90–115.
- Čermák, F. (1995d). Překladová lexikografie. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie.* Jinočany: H&H, s. 230–248.

- Čermák, F. (2007). Some of Current Problems of Corpus and Computational Linguistics or Fifteen Commandments and General Truths. In F. Čermák et al. (eds.). *The Third Baltic Conference on Human Language Technologies. Proceedings, October 4–5, 2007, Kaunas Lithuania*. Kaunas: Vytauto Didžiojo Universitetas, Lietuvių Kalbos Institutas, s. 61–69.
- Čermák, F. (2010a). *Lexikon a sémantika*. Nakladatelství Lidové noviny: Praha.
- Čermák, F. (2010b). Notes on Compiling a Corpus-Based Dictionary. In J. C. M. D. du Plessis (ed.). *Lexikos 20 (AFRILEX-reeks/series 20: 2010)*. Buro van die WAT: Stellenbosch, s. 559–579.
- Čermák, F. – Blatná, R. (eds.) (1995). *Manuál lexikografie*. Jinočany: H&H.
- Čermák, F. – Rosen, A. (2012). The case of InterCorp, a multilingual parallel corpus. *International Journal of Corpus Linguistics* Vol. 13, no. 3, p. 411–427.
- Ďurovič, Ľ (1953). I diskusie. In Š. Peciar (ed.). *Lexikografický sborník. Materiály z I. celoštátnej konferencie československých lexikografov. Bratislava 1952*. Bratislava: Vydavateľstvo Slovenskej akadémie ved, s. 127–128.
- Endzelīns, J. (1951). *Latviešu valodas gramatika*. Rīga: LVI.
- Erhart, A. (1984). *Baltské jazyky*. Praha: Státní pedagogické nakladatelství.
- Erjavec, T. (2004). MULTTEXT-East Version 3: Multilingual Morphosyntactic Specifications, Lexicons and Corpora. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC '2004)*. Paris: European Language Resources Association, s. 1535–1538.
- Filipec, J. (1985). Problematika konfrontace v lexikální zásobě. *Slovo a slovesnost* 46 (1985), s. 201–213.
- Filipec, J. (1995). Teorie a praxe jednojazyčného slovníku výkladového. In F. Čermák – R. Blatná (eds.). *Manuál lexikografie*. Jinočany: H&H, s. 14–49.
- Greule, A. (2008). Möglichkeiten und Grenzen einer Aufnahme von Namen in das Große deutsch-tschechische Wörterbuch. In M. Vachková (ed.). *Beiträge zur bilingualen Lexikographie*. Praha: FF UK, s. 61–69.
- Grīnberga, I. et al. (1997). *Latviešu valodas prasmes līmenis*. Strasbourg: Council of Europe.
- Grīsle, R. (2008). *Heterotonu vārdnīca un heterotonijas pētījumi*. Rīga: Zinātne.
- Hanks, P. (2000). Do Word Meanings Exist? *Computers and the Humanities* 34, s. 205–215.
- Hartmann, R. R. K. (2001). *Teaching and Researching Lexicography*. Harlow: Pearson Education Limited.
- Hausmann, F. J. (1989). Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht. In F. J. Hausmann et al. (eds.). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie, 1. Teilband*. Berlin - New York: Walter de Gruyter, s. 649–656.
- Hausmann, F. J. – Wiegand, H. E. (1989). Component parts and structures of general monolingual dictionaries: A survey. In F. J. Hausmann et al. (eds.). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie, 2. Teilband*. Berlin - New York: Walter de Gruyter, s. 328–360.

- Hausmann, F. J. – Reichmann, O. – Wiegand, H. E. – Zgusta, L. (eds.). (1989–1991). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie* [sv. 1–3]. Berlin - New York: Walter de Gruyter
- Honselaar, W. (2003). Examples of design and production criteria for bilingual dictionaries. In Piet van Sterkenburg (ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins Publishing Company, s. 323–332.
- Jakubiček, M. et al. (2013). The TenTen Corpus Family. In *7th International Corpus Linguistics Conference CL 2013*. Lancaster: UCREL, s. 125–127.
- Jansone, I. – Veisbergs, A. (2013). Vārdnīcas. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 435–452.
- Jarošová, A. (2000). Problems of Semantic Subdivision in Bilingual Dictionary Entries. *International Journal of Lexicography*, 13, 2000, s. 12–28.
- Jarošová, A. (2008a). Aspekty ekvivalencie: prekladový ekvivalent a prekladový význam. In S. Ondrejovič (ed.). *Iugi observatione... Jubilejný zborník na počesť Ľubomíra Ďuroviča*. Bratislava: Veda, s. 105–109.
- Jarošová, A. (2008b). Dimensions of Equivalence. In M. Vachková (ed.). *Beiträge zur bilingualen Lexikographie*. Praha: FF UK, s. 71–81.
- Jelínek, J. – Bečka, J. V. – Těšitelová, M. (1961). *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Praha: SPN.
- Jērane, S. (2013). *Darbības vārdu semantikas skaidrojums latviešu literārās valodas vārdnīcās* [diplomová práce]. Rīga: LU Humanitāro zinātņu fakultāte.
- Jirásek, K. (2011). Využití paralelního korpusu InterCorp k získávání ekvivalentů pro chorvatsko-český slovník. In F. Čermák (ed.). *Korpusová lingvistika Praha 2011: 1. InterCorp*. Praha: Nakladatelství Lidové noviny, s. 45–55.
- Kalme, G. – Smiltnece, G. (2001). *Latviešu literārās valodas vārd darināšana un morfoloģija. Lokāmās vārdšķiras*. Liepāja: Liepājas Pedagoģijas akadēmija.
- Kalnača, A. (2004). Darbības vārda veida kategorijas realizācija latviešu valodā. *Linguistica Lettica 13*. Rīga: LU Latviešu valodas institūts, s. 5–34.
- Kalnača, A. (2013). Morfoloģija. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 45–108.
- Kilgarriff, A. (2013). Using corpora [and the web] as data sources for dictionaries. In H. Jackson (ed.) *The Bloomsbury Companion to Lexicography*. London: Bloomsbury, s. 77–96.
- Kopeckij, L. V. (1974). Teoretické předpoklady dvojjazyčné slovanské lexikografie. In *Otázky lexikografie a vydávání slovníků. Z materiálů VIII. mezinárodní porady představitelů socialistických zemí vydávajících slovníky. Praha 8.–9. září 1972*. Praha: SPN, s. 21–96.
- Křen, M. (2006). Kolokační míry a čeština: srovnání na datech Českého národního korpusu. In F. Čermák – M. Šulc (eds.). *Kolokace*. Praha: Nakladatelství Lidové noviny, s. 223–248.
- Kušķis, J. (2004). *Ŗ, r*. S. 1.: Eraksti.
- Kušķis, J. et al. (1998). *Latviešu valoda 10.–12. klasei*. Rīga: Zvaigzne.
- Kuzina, V. (1998b). Leksisko minimumu izstrāde Latvijā. *Linguistica Lettica 2*. Rīga: LU Latviešu valodas institūts, s. 174–179.

- Laizāns, M. (2015). *Latviešu valodas korpusa izveide no emuāru tekstiem* [bakalārskā práce]. Rīga: LU Datorikas fakultāte.
- Landau, S. I. (1991). *Dictionaries. The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.
- Laua, A. (1969). *Latviešu leksikoloģija*. Rīga: Zvaigzne.
- Laua, A. (1997). *Latviešu literārās valodas fonētika*. Rīga: Zvaigzne ABC.
- Levāne, K. (2001). *Paula Bankovska romāna "Plāns ledus" pirmās nodaļas morfoloģiskā anotēšana un statistiskā analīze* [bakalārskā práce]. Rīga: LU Filoloģijas fakultāte.
- Levāne-Petrova, K. (2012). Līdzsvarots mūsdienu latviešu valodas tekstu korpus un tā tekstu atlasē kritēriji. In *Baltistica VIII priedas*, Vilnius: Vilniaus Universitetas, s. 89–98.
- Levāne, K. – Spektors, A. (2000). Morphemic Analysis and Morphological Tagging of Latvian Corpus. In *Proceedings of the Second International Conference on Language Resources and Evaluation*, vol. 2, pages 1095–1098.
- Lew, R. (2009). Towards variable function-dependent sense ordering in future dictionaries. In H. Bergenholtz – S. Nielsen – S. Tarp (eds.). *Lexicography at a crossroads: Dictionaries and encyclopedias today, lexicographical tools tomorrow*. (Linguistic insights – studies in language and communication, vol. 90.) Bern: Peter Lang, s. 237–264.
- Lokmane, I. (1988). Veida un laika nozīmju mijiedarbe latviešu valodas darbības vārda sistēmā. *Latviešu valodas kultūras jautājumi* 24. Rīga: Avots, s. 109–119.
- LVG (2015). *Latviešu valodas gramatika*. Rīga: LU Latviešu valodas institūts.
- LVGK (2008). *Latviešu valodas gramatika: koncepcija, prospekts, atsevišķu nodaļu pirmvarianti, diskusijas materiāli*. Rīga: LU Latviešu valodas institūts.
- LVMPK (2009). *Latviešu valodas morfoloģisko pazīmju kopa [beta versija, 2009-12-08]* [online]. Dostupné z: [http://www.semti-kamols.lv/doc\\_upl/TagSet.html](http://www.semti-kamols.lv/doc_upl/TagSet.html)
- Malý, I. (2013). Lotyšský sen Hanuše Entnera a Pavla Finka. In P. Štoll et al.: *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 145–155.
- Markus, D. – Auziņa, I. (2013). Fonētika. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 17–33.
- Marvans, J. (1967). Par verbu izteiksmes kategoriju mūsdienu latviešu valodā. *Latviešu valodas apcerējumi*, 9.A sēj. *Zinātniskie raksti*, 60. sēj. Rīga: Zvaigzne, s. 127–133.
- Mathiassen, T. (1996). *Tense, Mood and Aspect in Lithuanian and Latvian. Meddelelser av Slavisk-baltisk avdeling, Universitetet i Oslo, No. 75*. Oslo: Universitetet i Oslo.
- McEnery, T. – Hardie, A. (2012). *Corpus linguistics. Method, Theory and Practise*. New York: Cambridge University Press.
- Miškovský, J. (2013). Česká turistická a cestopisná literatura o Lotyšsku. In P. Štoll et al. *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 191–211.
- MLLVG I (1959). *Mūsdienu latviešu literārās valodas gramatika I. Fonētika un morfoloģija*. Rīga: LPSR Zinātņu Akadēmijas Izdevniecība.
- MLLVG II (1962). *Mūsdienu latviešu literārās valodas gramatika II. Sintakse*. Rīga: LPSR Zinātņu Akadēmijas Izdevniecība.

- MTL (2007). Joffe, D. – de Schryver G.-M.: *TshwaneLex. Professional Lexicography Software. User Guide. Version 3.0.* s. d.: TshwaneDJe HLT, Human Language Technology. Dostupné z <http://tshwanedje.com>.
- Murphy, M. L. (2003). *Semantic Relations and the Lexicon*. Cambridge: Cambridge University Press.
- Nešpore, G. (2009). Latviešu valodas pārvietošanās verbi teliskuma aspektā. In A. Kalnača (ed.). *Latvijas Universitātes Raksti*, 746. sējums. Rīga: Latvijas Universitāte, s. 64–73.
- Nikuļceva, S. (2013). Ekvivalence lexémů v Česko-lotyšském slovníku: přístup a zpracování. In P. Štoll et al. *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 213–225.
- Nītiņa, D. (1994). Ģenitīveņi (Ģenitīvs kā nelokāma vārdforma latviešu valodā). *Latvijas Zinātņu akadēmijas vēstis* 3 (560), 1994, s. 21–25.
- Nītiņa, D. (2001). *Latviešu valoda morfoloģija (konspektīvs lokāmo vārdšķiru apskats)*. Rīga: Rīgas Tehniskā universitāte.
- Och, F. J. – Ney, H. (2003). A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29 (1), s. 19–51.
- Paegle, D. (2003). *Latviešu literārās valodas morfoloģija*. Rīga: Zinātne.
- Paegle, D. (2006). Piedēkļverbu konjugācijas maiņa. *Baltu filoloģija*, Nr. XV (1/2). Rīga: LU Akadēmiskais apgāds, s. 57–79.
- Paikens, P. – Rituma, L. – Pretkalniņa, L. (2013). Morphological analysis with limited resources: Latvian example. In *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013)*, s. 267–278.
- Pořízka, P. (2014). *Tvorba korpusů a vytěžování jazykových dat: metody, modely a nástroje*. Vydavatelství FF UP: Olomouc.
- Ridings, D. (2003). Lexicographic workbench: A case history. In Piet van Sterkenburg (ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins Publishing Company, s. 204–214.
- Rosen, A. – Adamová, M. – Vavřín, M. (2014). Extrakce lexikálních ekvivalentů z paralelního korpusu. *Korpusová lingvistika Praha 2014. 20 let mapování češtiny*. Abstrakty. Praha: Ústav Českého národního korpusu, s. 177–179.
- Sekaninová, E. (1993). *Dvojjazyčná lexikografia v teórii a praxi*. Bratislava: Veda.
- Schmidt, W. P. (1991). Die Lexikographie der baltischen Sprachen II: Lettisch. In F. J. Hausmann et al. (eds.). *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie, 3. Teilband*. Berlin - New York: Walter de Gruyter, s. 2354–2361.
- Sikogukira, M. (1994). Measuring synonymy as an intra-linguistic and cross-linguistic sense relation. In A. Davies – B. Parkinson (eds.). *Edinburgh Working Papers in Applied Linguistics 5*. Edinburgh: University of Edinburgh, s. 109–118.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Skadiņa, I. – Vasiļevs, A. (2013). Valodas tehnoloģijas. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 453–475.

- Skoumalová, H. (2010). Extrakce česko-litevského slovníku z korpusu. In F. Čermák – J. Koček (eds.). *Mnohojazyčný korpus InterCorp: Možnosti studia*. Praha: Nakladatelství Lidové noviny, s. 241–247.
- Stafecka, A. (2013). Latviešu valoda: dialekti un izloksnes. In A. Veisbergs (ed.). *Latviešu valoda*. Rīga: LU Akadēmiskais apgāds, s. 253–266.
- van Sterkenburg, P. (ed.) (2003). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins Publishing Company.
- \*Suchomel, V. – Pomikálek, J. (2012). Efficient Web Crawling for Large Text Corpora. In A. Kilgarriff – S. Sharoff (eds.). *Proceedings of the seventh Web as Corpus Workshop (WAC7)*. Lyon, s. 39–43.
- Svensén, B. (2009). *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.
- Šeleryová, D. et al. (2013). *Česko-lotyšsko-litevská konverzace*. Praha: Česko-lotyšský klub.
- Škrabal, M. (2011). Slovesný vid v lotyštině z lexikografovy perspektivy. In F. Čermák (ed.). *Korpusová lingvistika 2011. 1. InterCorp*. Praha: Nakladatelství Lidové noviny, s. 262–269.
- Škrabal, M. (2012). Vēlreiz par verba veidu no leksikogrāfa viedokļa. In O. Bušs (ed.). *Valodas prakse: vērojumi un ieteikumi 7*. Rīga: Latviešu valodas aģentūra, s. 131–141.
- Škrabal, M. (2013). Moderně, či tradičně? Chystaný lotyšsko-český slovník v konfrontaci s dosavadními plody česko-baltské lexikografie. In P. Štoll et al.: *Zkušenosti a vztahy. Lotyšská a česká společnost ve 20. století*. Praha: FF UK, s. 227–233.
- Štícha, F. et al. (2013). *Akademická gramatika spisovné češtiny*. Praha: Academia.
- Štoll, P. (2013). *Lotyšská kultura a Jednota bratrská. České kontexty lotyšských kulturních tradic v 17.–20. století*. Praha: Karolinum.
- Šustere, G (2009). *Latvija skaitļos un faktos*. Rīga: Zvaigzne ABC.
- Teubert, W. (2002). The Role of Parallel Corpora in Translation and Multilingual Lexicography. In B. Altenberg – S. Granger (eds.). *Lexis in Contrast. Corpus-based Approaches*. Philadelphia: John Benjamins Publishing Company, s. 189–214.
- Vachková, M. (2007). *Kapitoly k německo-české metalexikografii I*. Praha: FF UK.
- Vachková, M. (2011). *Das große akademische Wörterbuch Deutsch-Tschechisch. Ein erster Werkstattbericht*. Frankfurt am Main: Peter Lang.
- Vachková, M. (2014). Z lexikografické dílny Velké německo-české lexikální databáze: Slovníkový příklad. *Jazykovědné aktuality* 51, 2014, č. 3–4, s. 58–70.
- Varantola, K. (2003). Linguistic corpora (databases) and the compilation of dictionaries. In P. van Sterkenburg (ed.). *A Practical Guide to Lexicography*. Amsterdam–Philadelphia: John Benjamins, s. 228–239.
- Veisbergs, A. (1998). Dažas problēmas divvalodu vārdnīcu veidošanā. *Linguistica Lettica* 2. Rīga: LU Latviešu valodas institūts, s. 153–159.
- Veisbergs, A. (2003). Elektroniskās vārdnīcas un datorizētie vārdnīcu resursi. *Linguistica Lettica* 12. Rīga: LU Latviešu valodas institūts, s. 163–175.
- Vondříčka, P. (2011). *Formalized contrastive lexical description: a framework for bilingual dictionaries. Formalizovaný kontrastivní popis lexikálních jednotek: deskriptivní rámec pro dvojjazyčné slovníky* [disertační práce]. Praha: FF UK.

- Vossen, P. (1992). The automatic construction of a knowledge base from dictionaries: A combination of techniques. In H. Tommola et al. (eds.). *EURALEX '92 Proceedings I-II: Papers submitted to the 5<sup>th</sup> EURALEX International Congress on Lexicography in of the 5th Euralex International Congress on Lexicography in Tampere, Finland*. Tampere: Tampereen Yliopisto, s. 311–326.
- Zgusta, L. (1971). *Manual of Lexicography*. Praha: Academia.
- Zgusta, L. (1984). Translational Equivalence in the Bilingual Dictionary. In R. R. K. Hartmann (ed.). *LEXeter '83 Proceedings*. Tübingen: Max Niemeyer Verlag, s. 147–154.
- Zuicena, I. (1998). Mūsdienu latviešu valodas vārdnīca: projekts un tā realizācija. *Linguistica Lettica 3*. Rīga: LU Latviešu valodas institūts, s. 121–125.
- Zuicena, I. (2012). Latviešu valodas attīstības procesu atspoguļojums „Mūsdienu latviešu valodas vārdnīcā”. In *Apvienotais Pasaules latviešu zinātnieku III kongress un Letonikas IV kongress „Zinātne, sabiedrība un nacionālā identitāte”*. Valodniecības raksti. Rīga: LU Latviešu valodas institūts, s. 236–246.
- Zuicena, I. – Migla, I. (2014). The Dictionary of Contemporary Latvian Language and its Lexicographical Process. *Workflow of Corpus-based Lexicography. COST ENeL WG3 meeting*. Bolzano, 19 July 2014. Dostupné z: <http://www.elexicography.eu/workinggroups/working-group-3/wg3-meetings/wg3-bolzano-meeting/>

### c) korpusy

Český národní korpus – InterCorp. Ústav Českého národního korpusu FF UK, Praha.

Dostupné z: <http://www.korpus.cz>

Český národní korpus – SYN. Ústav Českého národního korpusu FF UK, Praha. Cit.

Dostupné z: <http://www.korpus.cz>

Český národní korpus – SYN2015. Ústav Českého národního korpusu FF UK, Praha.

Dostupné z: <http://www.korpus.cz>

LVTK: *Latviešu valodas tīmekļa korpus* [ver. 1.0]. LU Matemātikas un informātikas institūts, Rīga. Dostupné z: <http://www.korpuss.lv>

LVK: *Līdzsvarots mūsdienu latviešu valodas tekstu korpus 2013*. LU Matemātikas un informātikas institūts, Rīga 2013. Dostupné z: <http://www.korpuss.lv>

### d) internetové zdroje

AkadTerm: <http://termini.lza.lv/term.php>

Biolib.cz <http://www.biolib.cz/>

EuroTermBank: <http://www.eurotermbank.com/>

Latvijas daba: <http://www.latvijasdaba.lv/>

Letonika.lv: <http://letonika.lv/>

Linguee: <http://www.linguee.com>

News.lv: <http://news.lv/>



**Opus:** <http://opus.lingfil.uu.se/>

**TermNet:** <http://termnet.lv/>

**Wiki ČNK:** <https://wiki.korpus.cz/>

**Wikipedia:** <https://www.wikipedia.org/>