

Andrzej CHARCIAREK

Katowice

Národní korpusy a polsko-česká/česko-polská překladová lexikografie

Keywords: lexicography, translation dictionary, corpus linguistics, Polish language, Czech language

Klíčová slova: lexikografie, překladový slovník, korpusová lingvistika, polština, čeština

Abstract

The article describes the role which national corpora can and should play in creating the great Polish-Czech/ Czech-Polish dictionary. The article analyses functions of the language corpora in preparing single and multi-word dictionary entries in an e-dictionary. Concrete examples have been used for the purposes of the detailed analysis of polysemy and homonymy in the process of creating dictionary entries. Also, the methodology of monolingual lexicography (e.g. identifiers and definitions) is presented in the description of dictionary entries (words or phrasemes) in the dictionary. Particular focus is put on creating dictionary entry structure and using the national corpora. Some concrete examples have been used to illustrate selectivity and incompleteness in the description of lexical unit meanings found in dictionaries not based on the texts from the language corpora. The article emphasises the significance of language corpora in creating entries as they are a useful source of frequency data in creating dictionary entries. The article is based on the corpus data from the National Corpus of Polish (NKJP) and the Czech National Corpus (ČNK).

V článku jsou popsány funkce, které poskytují jazykové korpusy pro zpracování samostatných hesel jak jednoslovných, tak víceslovných jednotek v elektronických překladových slovnících. Na konkrétních příkladech se podrobně analyzuje jev polysemie a homonymie v souvislosti se zpracováním jednotlivých hesel. Autor analyzuje využití metodologie jednojazyčné lexikografie ve zpracování hesel v překladovém slovníku. Zvláštní pozornost je věnována způsobům sestavení hesláře překladového slovníku a národním korpusům, které by měly v tomto procesu hrát významnou úlohu. Na vybraných příkladech je ukázáno, jak nepřesný a neúplný je popis významů lexikálních jednotek ve slovnících zpracovaných bez využití korpusových dat. Mnoho místa je věnováno otázce hledání a určení překladových ekvivalentů. Je

zdůrazněn význam jazykových korpusů jako zdroje frekvenčních a statistických údajů. V článku byla využita korpusová data z *Národního korpusu* polského jazyka (NKJP) a *Českého národního korpusu* (ČNK).

Je zajímavé, že v překladové polsko-české lexikografii se dosud používají ve větší míře slovníky v papírové podobě, které byly vytvořeny bez využití národních polských a českých korpusů. Dokonce i autoři nejnovějšího elektronického *Velkého polsko-českého [a] česko-polského slovníku* (LINGEA Lexicon 5) neinformují o využití korpusů. Potěšující však je alespoň skutečnost, že první takový elektronický slovník už existuje.

Je přece samozřejmé, že přišel čas převést slovníky z papírové podoby do podoby elektronické. Tato operace je přirozená, protože korpusy, jejichž využití je základem pro budování slovníků, jako takové mají datovou podobu. Navíc současný uživatel slovníků potřebuje co nejbohatší zdroj jazykových dat a příslušné nástroje k jejich využívání. Elektronická podoba slovníků umožňuje rychlé a operativní prohledávání jednotlivých hesel, která mohou mít netradiční podobu. Používání elektronického slovníku přináší v porovnání s tradiční papírovou formou slovníku celou řadu výhod, poskytuje například bohaté možnosti rozšíření hesla o multimediální prvky, animace apod. Využití hypertextu umožňuje také přejít na další text a získat tak podrobnější informace. Neexistují již žádná omezení rozsahu, která u slovníků v papírové podobě byla limitujícím faktorem.

V 21. století by se národní korpusy jako významný zdroj lingvistických dat měly stát základními zdroji při sestavování překladových slovníků. Jedná se o všechny druhy polských a českých národních korpusů, nejen o jednojazyčné synchronní korpusy psané.

Sestavování slovníku na základě korpusů není nový nápad. Na základě Národního korpusu polského jazyka (NKJP) by měl vzniknout a postupně vzniká slovník současné polštiny. Bude to první slovník polštiny tohoto druhu založený na korpusových datech.

Zatím se nemluví o tom, že by bylo plánováno vydání nového synchronního překladového slovníku polsko-českého a česko-polského založeného na korpusech. Jedním z důvodů jistě je okolnost, že dříve

uvedený elektronický slovník do jisté míry uspokojuje poptávku uživatelů, především českých polonistů, polských bohemistů, tlumočnicků a překladatelů.

Otázkou nicméně zůstává, jakým způsobem sestavovat synchronní překladový slovník a jaké korpusy by měly být použity v procesu jeho tvorby?

Je samozřejmé, že tradiční překladový slovník v papírové podobě, na rozdíl od výkladového, neobsahoval výklad hesla, kdežto uváděl jeho cizojazyčný ekvivalent resp. ekvivalenty. Obvykle takový slovník přinášel také informace gramatické a uváděl správné použití jednotlivých slov, frází a někdy vět v kontextu. Elektronická podoba slovníku umožňuje uživateli podle jeho potřeb zvolit různé režimy využití slovníku. V podstatě neexistuje již omezení na objem slovníku, díky čemuž překladový slovník může obsahovat ty informace, které dosud neobsahoval.

Pokud se týče výkladu jednotlivých hesel, překladový slovník je neuváděl. Uváděl jen identifikátory významu. Je přece těžké si představit, jak by měl slovník uvadět polysémní hesla a homonyma bez zpřesnění významů. Základní otázkou je, jakou koncepci polysémie a zároveň jejího lexikografického popisu zvolit. Podle našeho názoru je velmi zajímavá koncepce M. Grochowského, která je označována jako opoziční vůči dosavadním polským lexikografickým koncepcím. Tato koncepce spočívá v tom, že při zpracování a výkladu jednotlivých významů polysémických slov je klíčovým prvkem kontext (Grochowski, 1982, s. 83). Právě kontext slouží k určení a rozlišení významů polysémních jednotek, to znamená jejich sémantické struktury. M. Grochowski pokládá za základní popis syntaktických vlastností lexikálních jednotek hlavně valenci a slovnědruhovou spojitelnost. Je samozřejmé, že v této souvislosti jsou nejlepší pro výzkum jazykové korpusy jako zdroj objektivních empirických dat. Korpusy ukazují většinu kontextů, ve kterých se výraz nebo frazém vyskytují v textech. Vždyť korpusy umožňují výzkum významu často založený na kontextu. Je to velmi důležité nejen pro jednojazyčnou lexikografii, ale i pro lexikografii překladovou. V kontextu především je možné najít vhodný překladový ekvivalent. Z toho vyplývá, že korpus by měl

být použit při nalezení správného ekvivalentu překládaného hesla stejně jako při zpracování jednotlivých hesel.

Jak jsme již uvedli, zpracování polysémních hesel (mnohoznačných slov je v jazyce více než slov jednoznačných) je hlavní výzvou pro lexikografy. Heslo by mělo obsahovat výklady jednotlivých významů lexikální jednotky nebo frazému. Vzhledem k tomu, že se jedná o překladový slovník, neméně důležitá je volba druhu definice významu. V práci *Zadania z semantyki językoznawczej* A. Bednarek i M. Grochowski uvádějí tři druhy definice:

- sémantickou,
- gramatickou a
- pragmatickou (Bednarek, Grochowski 1993, s. 33–40).

Bez ohledu na to, že autoři berou v úvahu především jednojazyčné slovníky, v překladovém slovníku by mohly být uvedeny zvlášť definice sémantické a pragmatické (samozřejmě ve zkrácené podobě). Je však nutné si uvědomit, že ani takové definice někdy nestačí ke zpřesnění významu tzv. funkčních slov: částic, předložek, spojek apod.

Velmi zajímavou koncepci zpracování definice předložili autoři *Jiného slovníku polského jazyka* (2000). Je to jednojazyčný slovník, ale koncepci zpracování definice je možné úspěšně použít i v překladovém slovníku. Tato definice má čtyři vlastnosti:

- 1) kontextovost,
- 2) analytičnost,
- 3) neencyklopedičnost a
- 4) jasnost a přirozenost (Bańko 2001, s. 78–91).

Kontextovost znamená, že význam lexikálních jednotek se popisuje a vykládá v typických kontextech jejich použití. To vše zároveň umožňuje uživateli získávat dodatečné informace k lexikální a syntaktické spojitelnosti. Analytičnost se projevuje v podobě definice, která by měla být sestavena z významově jednodušších a vysoce frekventovaných slov. Neencyklopedičnost je vnímána jako odraz nevědeckého jazykového vědomí rodilých mluvčích. Výklad významu by měl odrážet nevědecký přístup ke kategorizaci světa a také podstatné ja-

zykové vlastnosti. Jasnost a přirozenost spočívá v tom, že každá definice by měla být jednoduchá, precizní a jasná. Navíc by se měla skládat jen z jednoduchých vět bez závorek, zkratk, vsuvek apod.

Je zajímavé, že definice zpravidla v překladovém slovníku chybí, dokonce i tehdy, když se popisují jednotlivé významy u hesel polysémních slov. Je samozřejmé, že se jedná o zkrácenou verzi definice, často jedno, dvě nebo tři slova. Je to určitě mnohem lepší způsob než číslování jednotlivých významů jazykové jednotky bez bližšího vysvětlení, co může předejít možným chybám. V tomto případě je možné použít metodu, kterou navrhli autoři sestavovaného *Velkého slovníku polského jazyka* (*Wielki słownik języka polskiego* – Instytut Języka Polskiego PAN w Krakowie). Důležité je, že tento slovník vzniká a bude vznikat celá léta na základě korpusu NKJP. Každé heslo polysémního slova obsahuje číselný identifikátor významu, který umožňuje počáteční identifikaci a rozlišování významů, např.: **pociąg** – 1. pojazd 2. popęd (WSJP).

V překladovém slovníku by měl výše uvedený identifikátor uživateli stačit. Pokud ne, lze ho rozšiřovat o vhodnou klasickou slovníkovou definici: **pociąg** – 1. pojazd – ‘zestaw wagonów z lokomotywą lub wagonem silnikowym’, 2. popęd – ‘upodobanie, które trudno jest opanować’ (WSJP)¹.

Otázkou, která je v poslední době aktuální a často diskutovaná, je otázka zachycování víceslovných lexikálních jednotek, čili frazémů. V odborné české a polské literatuře lze najít různé definice frazému více či méně shodné. Nejvhodnější jsou pro nás poněkud rozdílné názory a zároveň definice F. Čermáka a W. Chlebdy. Podle F. Čermáka frazém² je „každá ustálená anomální kombinace dvou či více prvků

¹ *Velký polsko-český [a] česko-polský slovník* (LINGEA Lexicon 5, 2010) a *Polско-český slovník* pod redakci Karla Olivy registrují jen jeden význam slova *pociąg* ‘vlak’ (PČS 2000, d. II, s. 56). Všechna spojení, která jsou v obou slovnících uvedena, se týkají jen tohoto významu.

² F. Čermák uvádí, že pojmy *frazém* a *idiom* mají prakticky překrývající se příznaky. Jediný rozdíl je v tom, že *frazém* je označením formálních aspektů jevu, a *idiom* aspektů sémantických (Čermák 2007a, s. 189).

(morfémů, slov, popř. kolokací), z nichž aspoň jeden je členem extrémně omezeného a uzavřeného paradigmatu (třídy)” (Čermák 2007a, s. 189). W. Chlebda frazém chápe velmi široce a definuje ho jako ustálené slovní spojení, které představuje název nějakého pojmu, jenž mluvčí v dané situaci reprodukuje v podobě stálého symbolu. Základní vlastností frazémů je reprodukovatelnost (Chlebda 2003, s. 177–254). Podle něho je frazémem každá jednotka, která je víceslovná, referenční (ustálená), zapamatovatelná a reprodukovatelná v určité situaci. Oba badatelé se ale shodují, že sémantické kritérium pro vymezení frazémů není vyčerpávající. F. Čermák zdůrazňuje, že tradiční rozšířené vymezení frazému jako kombinace, jejíž význam se nerovná sumě významu komponentů, však jíž nepostačuje (Čermák 2007a, s. 189). W. Chlebda nepovažuje vlastnosti frazémů (badatel prakticky nepoužívá termín frazeologismus), jako jsou obraznost, expresivnost a stylistická osobitost za podstatná kritéria pro jejich vymezení.

V klasickém překladovém slovníku frazémy netvoří samostatné položky³. Tradičně byly a jsou uváděny na konci heslových statí jednoslovných hesel. Podle našeho názoru by frazémy měly být zpracovány jako samostatná hesla stejně jako jednoslovné jednotky. V heslech frazémů jako samostatných formálně-významových jednotek by měly být také uvedeny, pokud je to nutné, příslušné identifikátory nebo definice, např.: *pięta achillesowa* ‘slaba strona’, *Achilova pata* ‘zranitelné místo, slabina’.

³ F. Čermák podrobně popisuje podobu frazémové heslové statí, která by měla odrážet komplexní poměry syntagmaticko-paradigmatické v systému a obsahovat následující obligatorní složky: A – paradigmatickou – a) forma (strukturní funkce, slovnědruhová platnost), b) sémantika lemmatu (nominativní funkce), c) gramatika (strukturní funkce), d) styl (rovinná příslušnost); B – syntagmatickou – e) forma (strukturní funkce: valence), f) sémantika kontextu (strukturní nominativní funkce), g) styl (expresivita), h) exemplifikace (Čermák 2007b, s. 627). I když autor bere v úvahu především zpracování jednojazyčného (českého) frazeologického slovníku, je zřejmé, že toto schéma stavby hesla může být použito také v překladovém slovníku.

Tento způsob umožňuje vysvětlit význam frazémů, kterým nemusí přesně rozumět někdy ani rodilí mluvčí. Navíc samostatná hesla frazémů jednoznačně určují jejich místo ve slovníku. Díky tomu se frazémy již neuvádějí několikrát pod všemi jejich plnovýznamovými komponenty.

Je třeba si uvědomit, že frazémy mají velký význam v procesu překladu. Pokud budeme vycházet z předpokladu, že víceslovných lexikálních jednotek je v každém jazyce více než jednoslovných, ve slovníku by měly frazémy mít dokonce přednost před slovy. Jestliže slova je možné jednoduše spočítat, výpočet frazémů už tak snadný není. Každopádně využití korpusu při zpracování frazémů stejně jako při zpracování celého hesláře je nezbytně nutné. Zdá se, že v překladové lexikografii se větší pozornost někdy věnovala a věnuje otázkám, jako je ekvivalence a její typy, ekvivalenty a jejich hledání, zpracování podoby samostatných hesel apod. Je samozřejmé, že všechny tyto otázky jsou podstatné pro budování překladového slovníku. Neméně důležité je však i důsledné budování hesláře. Pokud jde o velký slovník polsko-český, jeho heslář by měl v podstatě zahrnovat celou slovní zásobu současné polštiny. Je známo, že v lexikografické praxi existují tři možnosti sestavení takového hesláře s využitím:

- 1) vlastní excerpce,
- 2) existujících překladových a výkladových slovníků, anebo
- 3) kombinace obou přístupů.

Využití dosavadních překladových a výkladových slovníků má řadu výhod, ale i nevýhod. Jistě k prvním patří využití heslářů obsahujících dobře zpracovaná jednotlivá hesla. Nicméně však každý heslář předchozích slovníků musí být revidován, doplňován a korigován.

Pokud jde o využití vlastní excerpce ve stavbě hesláře velkého překladového slovníku hlavní roli by měl hrát korpus jako základní zdroj jazykového materiálu.

Dobrym příkladem tohoto přístupu je výkladový *Velký slovník polského jazyka*, vznikající na základě korpusu. Již ta hesla, která jsou v něm zpracována, jednoznačně ukazují na rozdílný přístup jeho autorů a autorů předchozích překladových slovníků při výkladu vý-

znamu jednotlivých jednotek. Uvedeme a porovnáme zpracování hesel z tohoto zdroje a dvou překladových polsko-českých slovníků:

galeria

- 1.a sztuki (instytucja zajmująca się udostępnianiem publiczności dzieł sztuki)
- 1.b budynek (budynek, w którym mieści się galeria sztuki)
- 1.c ludzie (osoby prowadzące galerię sztuki)
2. w Internecie (zbiór zdjęć prezentowany w sieci internetowej)
- 3.a balkon (długi balkon biegnący wzdłuż ściany wewnątrz lub na zewnątrz budynku)
- 3.b ludzie (ludzie zgromadzeni na galerii – balkonie)
4. handlowa (duża hala mieszcząca wiele sklepów)
5. osób (grupa osób, zwierząt lub przedmiotów wyróżniających się czymś szczególnym) (*Wielki słownik języka polskiego*)

galeria

galeria dla dziennikarzy

Obraz został wystawiony w Galerii Narodowej.

Kolekcja obrazów w Galerii Starych Mistrzów. (LINGEA Lexicon 5, 2010)

galeria

1. druh balkónu galerie
2. pavlač
3. (obrazů) galerie
4. (v divadle) galérie
5. (v pevnostech, dolech) podzemní chodba (*Polsko-český slovník* 2000)

Ve srovnání s oběma velkými překladovými slovníky jsou hesla ve *Velkém slovníku polského jazyka* zpracována precizněji a detailněji. Je to výsledek pozorného sledování množství kontextů, do kterých je konkrétní slovo *galeria* zasazeno. Zajímavé je, že *VSPJ* identifikuje 5 významů (s významovými odstíny až 8) slova *galeria*, zatímco nejnovější *LINGEA Lexicon 5* v podstatě uvádí jen 3 příklady jeho použití ve slovních a větných spojeních. Navíc 2 příklady (*Obraz został wystawiony w Galerii Narodowej a Kolekcja obrazów w Galerii Starych Mistrzów*) odrážejí stejný význam ('budynek, w którym mieści

się galeria sztuki'). Je poněkud překvapující, že tento slovník nezachycuje význam *galeria* (*handlowa*), který patří jak v polštině, tak i v češtině (*nákupní galerie*) mezi frekventované výrazy. V případném zpracování polysémního hesla *galeria* na základě NKJP by k takové situaci nedošlo, protože tento korpus registruje velký počet výskytů frazému *galeria handlowa*. NKJP také potvrzuje centrální význam polysémního slova *galeria*, který souvisí s uměním.

Polsko-český slovník uvádí také, stejně jako jednojazyčný *Velký slovník polského jazyka*, 5 významů slova *galeria*. Nicméně se však liší způsob zpracování jednotlivých významů uvedených v obou slovnících. V *Polsko-českém slovníku* se význam *galeria* (*handlowa*) neregistruje, uvádí se zastaralý význam *galeria* (*w twierdzach, fortecach, kopalniach*) a na prvním místě jako základní je uveden význam 'druh balkónu', což již neodpovídá jazykové realitě.

Podobné pochybnosti vyvolává také druhý význam, podle něhož je ekvivalentem polského slova *galeria* český lexém *pavlač*⁴.

Není náhodou, že *VSPJ* ve srovnání s dvěma uvedenými slovníky nabízí nejpodrobněji zpracované heslo. Tato skutečnost úzce souvisí s vyhodnocováním významů jazykových jednotek na základě korpusových dat, která umožňují identifikovat další existující či nové významy dosud neregistrované v předchozích slovnících. Kromě toho využití korpusů do značné míry omezuje případné chyby týkající se stylistické, lexikální a sémantické asymetrie překladových párů.

Jak již bylo uvedeno výše, velký překladový slovník by měl obsahovat frazémy jako samostatná hesla, což je v jistém rozporu s lexikografickou tradicí.

Heslo frazému se bude lišit od tradičního hesla výrazu tím, že nebude uvádět gramatické charakteristiky jednotlivých lexémů, které frazém obsahuje. V případě frazémů v překladovém slovníku je nejdůležitější explikace jeho významu a ekvivalent. Výklad významu je

⁴ Ekvivalentem českého slova *pavlač* je v polštině *krużganek*. Ekvivalence *galeria – pavlač* neexistuje.

nutný, protože ani rodilí mluvčí někdy nerozumějí významu jednotlivých frazémů, např.:

pies ogrodnika – ‘osoba, która nie pozwała innym skorzystać z czegoś, mimo że sama z tego nie korzysta’ (*WSJP*)

bulka z masłem – ‘coś bardzo łatwego do zrobienia’ (*WSJP*)

Dalším problémem je otázka variantnosti frazémů. Varianty jednotlivých frazémů by měly být v překladovém slovníku registrovány příp. s pomocí lomítka⁵: *ktoś strzelił sobie w stopę/kolano*.

Transland by měl takovou konečnou podobu: *ktoś strzelił sobie w stopę/kolano* ‘ktoś, działając nierozważnie, sam sobie zaszkodził’.

Lze předpokládat, že pečlivé zpracování hesel frazémů (translandu) má vliv na kvalitu překladového slovníku. To, co je v něm nejdůležitější, čili adekvátnost a přesnost ekvivalentů (translatů), začíná již na úrovni stavby hesláře.

O kvalitě překladového slovníku ovšem svědčí lexikografické zpracování ekvivalentů. Roman Lewicki ekvivalenty vymezuje takto:

Uznaję za translaty takie wyrażenia języka docelowego, które – niezależnie od ich statusu w systemie języka – są zdolne do przekazania informacji zawartej w jednostce tłumaczenia. Za jednostkę tłumaczenia zaś uważam każdy odcinek tłumaczonego tekstu, dla którego tłumacz poszukuje odpowiednika przy tworzeniu przekładu (Lewicki 2011, s. 160).

Z toho vyplývá, že mezi translandem a translatem neexistuje vztah souznačnosti, ale vztah informační a funkční rovnocennosti. Toto zjištění platí jak pro jednoslovné, tak i víceslovné jednotky. Je třeba si uvědomit, že překladový slovník by měl sloužit jako podpora při překladu textů. Takový přístup ke slovníku již ukazuje, že by měl být založen na textech a proces jeho sestavování by měl vypadat následovně:

⁵ Tím způsobem se registrují varianty frazémů v *Polsko-ruském slovníku překladových párů* (*Polsko-rosyjski słownik par przekładowych* 2013). V tomto slovníku se registruje množství frazémů, které ve většině překladových slovníků nejsou uvedeny.

V případě polsko-českého překladového slovníku je možné tento proces ještě více upřesnit:

polské a české texty → **překladový slovník /překladové páry**
polsko-české → **české texty**

Je samozřejmé, že základním cílem je nashromáždění velkého množství textů a pozorné sledování kontextů, ve kterých se jednoslovné i víceslovné jednotky vyskytují. Jazykové korpusy jako útvary reflektující a popisující jazykovou realitu by měly být hlavním zdrojem při rozhodování, které frazémy do velkého polsko-českého slovníku zařadit. Získané frekvenční a statistické údaje vycházející z materiálu obou korpusů mohou úspěšně pomoci objevit frazémy (často také jejich varianty) a sestavovat překladový pár, např.: *strzelić sobie w stopę / kolano* – *dát si/vstřelit gól do vlastní branky* X *dát vlastní gól* X *naběhnout na vlastní vidle*.

Výše uvedený pár ekvivalentů objevuje různé možnosti překladu v závislosti na druhu textu. Zdá se, že překladový slovník často uvádí jen jednu variantu překladu bez ohledu na to, že v různých textech můžeme najít jiné vhodné ekvivalenty. Použití korpusu umožňuje výrazné obohacení slovníku o možné překladové ekvivalenty a sestavení jejich souhrnu. Je možné předpokládat, že to bude velmi přínosné pro praktické využití slovníku.

Velkou budoucnost proto mají současné paralelní korpusy, z nichž je možné získávat překladové ekvivalenty. Na rozdíl od tradičních překladových slovníků obsahují paralelní korpusy textové ekvivalenty, které tvoří základní jednotky v procesu translace. Je však třeba konstatovat, že tyto korpusy obsahují převážně beletristické texty. Zároveň zde vznikají pochybnosti týkající se kvality překladu. Nezřídka se stává, že překlad je neprofesionální a nepřesný. Nejčastějším důvo-

⁶ Takový přístup k sestavování překladového slovníku má Wojciech Chlebda (Chlebda 2011, s. 25).

dem tohoto faktu je překladatelova nedostatečná znalost zdrojového nebo cílového jazyka, a proto některé přeložené fragmenty neodpovídají originálu, nebo úplně absentují. Navíc je třeba si uvědomit, že jazyk v beletristických textech se může významně lišit od jazyka průměrného uživatele jazyka. Autorský jazyk může být zcela odlišný od jazyka většiny populace, může být nekonvenční, plný netytických kolokací, neobvyklých idiomatických spojení a autorských metafor. Slovník by však měl uvádět jazyk příznačný pro celonárodní komunitu, jazyk standardní. Slabinou paralelních korpusů také stále zůstává přenos mluveného jazyka, neboť mluvené texty se převážně nepřekládají. Nelze však pochybovat o tom, že paralelní korpusy mohou být do budoucna užitečným zdrojem pro tvorbu překladového slovníku. Metody automatického získávání překladových slovníků z paralelních korpusů jsou již známé a použitelné v praxi. Problémem je ale relativně malý rozsah paralelních korpusů, které jsou nesrovnatelné s velkými jednojazyčnými korpusy. Posledně jmenované zřejmě zůstávají největšími z toho důvodu, že ne všechny možné kategorie textů budou přeloženy do cizího jazyka.

Závěrem je třeba konstatovat, že jazykové korpusy jsou neocenitelnými zdroji informací pro lingvisty a lexikografy při budování velkých překladových slovníků v elektronické podobě. Díky korpusům se úplně změnil pohled na současnou překladovou lexikografii a především na obsah a podobu slovníku.

Otevřenou otázkou zůstává, kdo a pro koho bude tvořit nové, kvalitativně lepší velké elektronické slovníky. Lze předpokládat, že slovníky budoucnosti budou sestaveny lingvisty-lexikografy-informatiky (nezbytné interdisciplinární vzdělání) pro všechny uživatele – lingvisty, překladatele a studenty jazyka.

Literatura

- Bańko M., 2001, *Z pogranicza leksykografii i językoznawstwa. Studia o słowniku jednojęzycznym*. Warszawa: Wydział Polonistyki Uniwersytetu Warszawskiego.
- Bednarek A., Grochowski M., 1993, *Zadania z semantyki językoznawczej*. Toruń: Wydawnictwo UMK.

- Čermák F., 2007, *Jazyk a jazykověda : Přehled a slovníky*. Praha: Karolinum.
- Čermák F., 2008, *Frazeologie a idiomatika česká a obecná. Czech and General Phraseology*. Praha: Karolinum.
- Český národní korpus. Praha: Ústav Českého národního korpusu FF UK. On-line: <http://www.korpus.cz>.
- Chlebda W., 2003, *Elementy frazematyki: wprowadzenie do frazeologii nadawcy*. Łask: Oficyna Wydawnicza LEKSEM.
- Chlebda W., 2011, *Ekwiwalencja i ekwiwalenty: między słownikiem a tekstami*. In: W. Chlebda (red.), *Na tropach tłumaczeń. W poszukiwaniu odpowiedników przekładowych*. Opole: Wydawnictwo Uniwersytetu Opolskiego, s. 21–43.
- Grochowski M., 1982, *Zarys leksykologii i leksykografii*. Toruń: Wydawnictwo UMK.
- Inny słownik języka polskiego*, 2000, red. M. Bańko, t. 1–2. Warszawa: Wydawnictwo Naukowe PWN.
- Lewicki R., 2011, *Translacje w słowniku przekładowym – poszukiwanie i ustalanie*. In: W. Chlebda (red.), *Na tropach tłumaczeń. W poszukiwaniu odpowiedników przekładowych*. Opole: Wydawnictwo Uniwersytetu Opolskiego, s. 159–165.
- Narodowy Korpus Języka Polskiego. On-line: <http://www.nkjp.pl>.
- Oliwa K. a kol., 1999, *Polsko-český slovník*, T. 1–2. Praha: Academia.
- Polsko-rosyjski słownik par przekładowych*, 2013, red. W. Chlebda. Opole: Wydawnictwo Uniwersytetu Opolskiego.
- Velký slovník polsko-český a česko-polský (elektronický slovník)*, 2010, Lexicon 5. Brno: LINGEA.
- Wielki słownik języka polskiego*. On-line: <http://www.wsjp.pl>.