

Projekt: Lidé – Média – Komunikace

Metodické zásady

A. Obecné zásady:

1. jednotky tvořící heslář slovníku musí pocházet z původních textů publikovaných v jazycích projektu (A, B, C);
2. ekvivalenty vstupních jednotek v cílových jazycích (B, C) pochází z původních textů publikovaných v těchto jazycích (nemohou být vytvořeny lexikografy ad hoc).

Hlavním principem je „přechod od textů publikovaných v jazyce A přes slovník k novým textům v jazycích B a C”.

B. Podrobné principy:

1. **Vyhledávání v textech:** soubor textů určených k extrakci slovníkových jednotek do slovníku bude stanoven dle následujících zásad:
 - a) texty se budou vztahovat k předmětu uvedenému v názvu slovníku;
 - b) textový korpus budou tvořit různorodé texty (publicistické texty, vědecké stati, populárně-vědecké publikace, uživatelské příručky, internetové blogy, počítačová rozhraní atd.) spolu s národními a paralelními korpusy;
 - c) texty budou na střední úrovni specializace dostupné uživatelům se SŠ a VŠ humanitním vzděláním.

Celkový objem textů by měl dosáhnout stropu cca **3 milionů** grafických slov.

2. **Excerpce hesel:** soubor hesel slovníku bude obsahovat ustálené víceslovné jazykové jednotky v podobě jmenných frází, slovesných frází a jiných víceslovných frází typu:
 - a) jmenné fráze: **přibližovací čtečka, mobilní aplikace, digitální platforma, navigační lišta;**
 - b) slovesné fráze: **kdo archivuje co na disku, kdo dobil mobil, co funguje v jakém systému, kdo skenoval dokument;**
 - c) fráze: **Přidat účet, Přidat k oblíbeným, Ochrana proti sledování v soukromých oknech, Varovat před zavřením více karet.**

Excerpce jednotek z tradičních (tištěných) textů bude prováděna manuální metodou, z digitálních textů – pomocí automatizované metody vyvinuté v monografii: D. Borysowski, *Teksty zespolone a problemy leksykografii przekładowej*, Opole 2020. Počet excerpt bude asi 1000 víceslovných jednotek pro každou část slovníku (**celkem asi 5 000 jednotek**). Hranice jednotek a jejich

struktura budou ověřeny jejich zadáním jako fráze v uvozovkách do národních internetových vyhledávačů a do národních korpusů všech tří jazyků.

3. Makro- a mikrostruktura slovníku:

- a) **makrostruktura:** slovník bude překladový s tematickým a abecedním pořádkem (abecední pořadí bude samostatné v každé z tematických částí slovníku). Slovník bude mít následující části:

Úvod (v češtině)

Seznam zkratk a symbolů (v češtině)

Pět tematických sekcí s česko-polsko-ruskou strukturou

Úvod (v češtině a v polštině)

Seznam zkratk a symbolů (v češtině a v polštině)

Pět tematických sekcí s polsko-česko-ruskou strukturou

Bibliografie

- b) **mikrostruktura:** metoda prezentace jazykového materiálu slovníku je již v lexikografii ověřena, je to metoda tří sloupcové tabulky, ve které každý sloupec (svisle) představuje jeden jazyk, a každá výstupní jednotka se svými dvěma ekvivalenty zabírá jeden řetězec tří polí ve vodorovném uspořádání tabulky podle vzorce:

cesta: vesnická cesta	droga wiejska	деревенская доро́га
cesta: vesnická cesta	droga wiejska	сельская доро́га
cesta: vnitřní cesta	droga wewnętrzna	внутренняя доро́га
cestovat: kdo cestuje autostopem	kto podróżuje autostopem	кто путешествует автосто́пом

Tento způsob prezentace materiálu je čitelný, jednoznačný a pro uživatele komfortní. Slovník má překladový, a ne definičně-překladový charakter, proto budou mikro-definice u hesel uváděny pouze pro mnohoznačné jednotky.

4. **Ekvivalentizace:** v souladu s principem „přechod od textů publikovaných v jazyce A přes slovník do nových textů v jazycích B a C“ by měl slovník poskytnout uživateli takové ekvivalenty výstupních jednotek, které umožní vytváření regulérních a akceptovatelných textů v jazycích B a C. Při výběru ekvivalentů se autoři slovníku budou řídit pravidlem, že překladový ekvivalent jednotky A je v jazyce B to, co by uživatel jazyka B řekl automaticky v odpovídající řečnické situaci; v překladové lexikografii se tento princip nazývá „situační kritérium pro výběr ekvivalentů“. Ekvivalenty jazykových jednotek A v jazycích B a C navržené autory slovníku budou ověřeny v korpusu textů shromážděných

autory, v národních počítačových korpusech a v otevřeném internetovém prostoru pomocí vyhledávačů jazyků B a C.

5. **Použité lingvistické metody:**

- a) **konfrontační a kontrastivní lingvistika** (pro srovnání subsystémů tří slovanských jazyků a určení rozsahu symetrie a asymetrie mezi nimi);
- b) **metalexikografie a překladová lexikografie** (pro vypracování principů budování slovníku, jeho makrostruktury a mikrostruktury, vypracování pravidel pro výběr ekvivalentů překladu a meta-jazyka pro popis jednotek);
- c) **teorie sémantických polí** (pro vypracování pravidel pro rozdělení jazykových subsystémů na menší sémantická pole odpovídajících tematickým částem slovníku);
- d) **koncepce jazykových jednotek a jazykových produktů** (pro vypracování pravidel pro excerpci z textů skupin slov majících status skutečných jednotek daného jazyka, určování jejich hranic a rozsahu variability, pro vypracování pravidel jejich jednoznačného a konzistentního grafického záznamu);
- e) **počítačová a korpusová lingvistika** (pro vypracování pravidel automatického vyhledávání digitálních textů pro excerpci víceslovných jednotek, pravidel jejich počítačového ověřování, pravidel používání různých typů korpusů pro zpracování překladatelského slovníku a pro vypracování pravidel digitalizace slovníku s vyhledávačem).