

Aby výsledky tejto práce boli hodnovernejšie (ak vezmeme do úvahy to, že štatistika je vraj získavaním presných čísel z nepresných údajov), rozhodli sme sa nepracovať so vzorkou jedného či dvoch písmen abecedy, ako znelo pôvodné zadanie. Na tomto mieste sa chceme poďakovať kolegom V. Benkovi a M. Šimkovej za radu a pomoc pri získaní materiálu potrebného na štatistické spracovanie z počítačovej verzie Krátkeho slovníka slovenského jazyka. Zo zoznamu heslových slov (vrátane prihniezdovaných) sa pomocou špeciálneho programu vybralo každé dvadsiate slovo so základným slovnodruhovým určením. Charakteristiku sledovanej slovnej zásoby podávame tak, ako ju uvádzajú autori KSSJ v úvodných statiach: „Cieľom KSSJ je zachytiť a lexicograficky opísať najpoužívanejšiu slovnú zásobu súčasnej spisovnej slovenčiny v rozsahu asi 50 000 slov.“ Na inom mieste čítame „KSSJ zachytáva jadro slovnej zásoby s protívahou jej okrajových oblastí, ako sú nové slová a oproti nim niektoré slová zastarávajúce, knižné slová a na druhej strane hovorové slová, ako aj frekventované nespisovné výrazy atď.“ Tieto údaje považujeme za dôležité pri hodnotení výsledkov štatistiky.

Rozsah materiálu na spracovanie bol 5 % z celkového počtu slov spracúvaných v KSSJ, dovedna 3 150 slov. Podľa toho treba poznamenať, že autori slovníka podhodnotili jeho rozsah. Ten je (spolu s prihniezdovanými) približne 63 000 slov. Výsledky štatistického spracovania uvádzame v priložených tabuľkách. Prvá z nich obsahuje kompletné údaje o sledovanej vzorke slov, čiže nielen údaje v percentách, ale i absolútny počet lexém v jednotlivých písmenách abecedy. V druhej a tretej tabuľke sú už iba percentuálne údaje o jednotlivých písmenách: na základe údajov v nich možno porovnávať navzájom prevzatia z klasických jazykov, v nasledujúcom prípade prevzatia z klasických jazykov oproti prevzatiám z iných jazykov a domácim slovám. Najskôr by sme však chceli podrobnejšie opísať postup spracovania a triedenia získaného materiálu.

Z tridsiatich troch písmen abecedy pri vyberaní každého dvadsiateho slova neboli v našej vzorke vôbec zastúpené štyri – dz, ô, q, x. Dvojslovné je zastúpené dž a jednoslovné w a y. Treba však konštatovať, že aj pri takomto zanedbateľnom zastúpení sa pomerne verne zachytávajú stav slovnej zásoby z hľadiska pôvodu slov. Aj z tohto dôvodu (a isteže aj kvôli úplnosti údajov) sme ich zaradili do našej prehľadnej tabuľky. Písmená, ktoré žiadne údaje neposkytujú, sme do súhrnnej tabuľky nezaradili.

Takto získanú vzorku slov sme triedili podľa pôvodu v opore o dva slovníky: *Slovník cudzích slov* Márie Ivanovej-Šalingovej (1983) a *Machkov Etymologický slovník* (1968). Orientácia len na *Slovník cudzích slov* (ďalej SCS) prinášala jedno úskalie – nie je v ňom zachytená istá časť lexikálnych prevzatí z klasických jazykov, ktoré sa už vnímajú ako domáce slová, teda časť, ktorú z hľadiska sledovanej problematiky považujeme za dôležitú. Napr. lexikálne prevzatia *múr, krst, kresťan, octovitý, stmelovať, škola, kachle* by sme márne hľadali v SCS, no Machek ich uvádza ako cudzie slová latinského,

prip. gréckeho pôvodu. Najviac nám chýbalo jedno dielo – etymologický slovník slovenčiny, ktorý by mohol byť spoľahlivým zdrojom informácií aj o lexikálnych prevzatiach z iných jazykov, prirodzene aj z latinčiny a gréčtiny.

Slová indoeurópskeho pôvodu, o ktorých sme nenašli jednoznačné vyjadrenie ako o prevzatých, sme medzi lexikálne prevzatia nezaradovali. V opačnom prípade by sme museli medzi prevzatia začleniť aj značnú časť pôvodom domácej slovnej zásoby, ktorá sa vyvinula z indoeurópskeho základu. Týmto slovám sú blízke slová s rovnoznačným významom i podobnou hláskovou stavbou v latinčine a gréčtine, ktoré pochádzajú z rovnakého zdroja. Týchto prípadov bolo viacero. Zvyčajne boli kvalifikované ako všeslovanské a týkali sa takých oblastí ľudského života, ktoré dané spoločenstvo potrebovalo označovať (aj v tom sme nachádzali oporu pre ich zaradenie medzi domáce slová).

Po roztriedení slov podľa pôvodu sme vypočítali percentuálne zastúpenie slov podľa pôvodu v jednotlivých písmenách abecedy a napokon sme to isté urobili pre celú vzorku sledovaných lexém (3 150 slov).

Slová v tabuľke sú roztriedené do stĺpcov nasledovne:

1. stĺpec – označený aG – udáva absolútny počet lexém gréckeho pôvodu v jednotlivých písmenách sledovanej vzorky slov.
 2. stĺpec – G % – vyjadruje percentuálne zastúpenie lexikálnych prevzatí z gréčtiny v sledovanej vzorke slov.
 3. stĺpec – aL – podobne ako prvý pre latinčinu.
 4. stĺpec – L % – rovnako ako druhý pre latinčinu.
 5. stĺpec – G+L – zahŕňa skupinu slov, ktoré sa do slovenčiny nedostali priamo kontaktom s gréčtinou a latinčinou, ale boli z odovzdávajúcich jazykov prebrané prostredníctvom iných jazykov, prešli teda už ich jazykovým filtrom. Sem sme zaradili aj tzv. hybridné slová – rôznajazyčné spojenia (zväčša grécko-latinské), slová so slovenskými koreňmi alebo predponami. Ide teda o takú skupinu slov, ktoré sú istým spôsobom viazané na gréčtinu alebo latinčinu, no nemožno ich jednoznačne zaradiť do prvých dvoch skupín. Túto skupinu slov by bolo možné ešte rozčleniť, no nazdávame sa, že pre cieľ našej práce je takéto rozdelenie dostačujúce. Každé ďalšie členenie by výsledky skôr zneprehľadnilo.
 6. stĺpec – G+L % – vyjadruje predchádzajúci údaj v percentách.
 7. stĺpec – aG,L – je súčtom prvých troch údajov, 1. 3. a 5. stĺpca.
 8. stĺpec – G,L % – percentuálne uvádza počet gréckych a latinských lexikálnych prevzatí bez ohľadu na ich cestu do slovenčiny.
 9. stĺpec – aI – zachytáva počet lexikálnych prevzatí zo všetkých ostatných cudzích jazykov.
 10. stĺpec – I % – vyjadruje predchádzajúci údaj v percentách.
 11. a 12. stĺpec aS, S % sú slová pôvodom slovenské (ide., všeslovanské, psl.).
- Posledný riadok, označený znakom +, udáva celkový počet a percentuálne vyjadrenie v celej sledovanej vzorke 3 150 slov. Získali sme tak na porovnanie zaujímavé údaje.