

Úvod

Jedním z hlavních způsobů, jakými se jazyk vyvíjí, je jeho frekvence. Frekvence slova je počet, s jakým se slovo v určitém jazyce vyskytuje. Frekvence slova je důležitým faktorem při studiu jazyka a při jeho využití. Frekvence slova je také důležitým faktorem při studiu jazyka a při jeho využití. Frekvence slova je také důležitým faktorem při studiu jazyka a při jeho využití.

Tyto skutečnosti si lidé uvědomili už dávno, především snaze poznat jedno jazyka, které se také dle frekvence rozdělují, a využití výsledků této snahy (v prvních aplikacích, zejména v oblasti tvorby učebnic pro cizince, kde se po takovém společném základu už slovně volalo). V novodobé historii přišel s poznáním frekvence jako první porovnatelný Keelingův slovník pro němčinu už z r. 1897, svou výchozí materiálovou základnou 10 milionů slov byl a dodnes je opodlný. Hlavní frekvenční slovníky tohoto typu však začaly vycházet až před 2. světovou válkou a téměř po ní. Šlo o desítky různě psaných slovníků pro němčinu, angličtinu, francouzštinu a další jazyky (Thornhille 1921 a zvl. 1944, West 1953, Juillard 1964, 1970 apod.). Největší materiálovou základnou, vycházející z psaných pramenů, byl slovník *Prüfungswörterbuch*, který bylo manuální, byla omezena na 100 000 slov nebo miliónu slov vybraných textů (třetinu až 0,5–1,5 miliónu). Výsledky z analýzy takto dosažené se užívaly hlavně už za 2. světové války, při tvorbě *Basic English*, 1970 apod. Následně se rychle naučitelé angličtiny pro zahraniční kore, kteří si se později stala modelem i pro jiné jazyky aj.

Neuvěřitelným předpokladem a motivem pro vytvoření slovníku byla ovšem vědomí, že toto jedno jazyka a tedy i jeho základní slova se dají zachytit určitým výběrem jejich mateřského množství textů zvolených z excerptů. Takto posbírané texty, ačkoliv stále jen odhad těchto textů podle jejich výskytu vzhledem ovšem zůstával problém i do pozdější doby.

Takto pojatý slovník vyšel pro češtinu v. 1961 také (Jedlička, Bečka, Těšitelová, tj. *Frekwence slov, slovník úhrubý a barevně oděšen jazyce*, dále PSSDČ), jeho testový základ má zhruba 1,5 miliónu slov; tento slovník, čerpající ze 75 knihovních pramenů a obsahující i cenat morfologické mluvy, představuje své frekvenční výsledky ve dvou podobách, v *Průběh slov podle frekvence* (obsahuje 10 000 nejčastějších slov v vzestupném pořadí) a v *Abecedním seznamu slov* (až 26 000 slov). O několik let později pak vychází i podobný menší slovník pro slovenštinu (Majstrik 1969), dále pro ruštinu, polštinu a další jazyky, přičemž šlo z těchto slovníků už měla lze použít jako je či jejich přehled podává M. Těšitelová (mj. 1977) spoluautorka prvního, více studijního slovníku češtiny, která se jako hlavní autorka k frekvenci slov věnovala ještě v označením záměru i rozsahu posudek (1985, *Frekwence slov češtiny včasněho období*), v 70–80. letech minulého století má konečně nějaký frekvenční slovník větší většina evropských jazyků a ty se také poměrně snadno a s různými úpravami využívají. Z dosavadní literatury přitom má výbornou bázi pro excerpty ovšem nutně vedle z excerptů i toho typu, že se mezi nejzákladnější a nejčastější slova dostávají například slova ohrožená a specifická (takto řadí PSSDČ) mezi nejčastějších deset tisíc slov a slova jako *hůl*, *ovnat* a *calice*, které tam vnést jen jejich opakovaný výskyt výskyt v jednom z omezeného počtu excerptovaných pramenů).

Koncem minulého století však nastupují i logické už pro-

cerptní základny, mu (František Čermák)

USPOŘÁDÁNÍ SLOVNÍKU A VÝZNAM JEDNOTLIVÝCH ÚDAJŮ / 11
(Michal Křen)

PRŮMĚRNÁ REDUKOVANÁ FREKVENCE / 15
(Jaroslava Hlaváčková)

FREKVENČNÍ SLOVNÍK APELATIV / 17
(pořadí podle abecedy)

FREKVENČNÍ SLOVNÍK APELATIV / 373
(pořadí podle absolutní frekvence)

FREKVENČNÍ SLOVNÍK APELATIV / 469
(pořadí podle průměrné redukované frekvence)

FREKVENČNÍ SLOVNÍK PROPRIÍ / 565
(pořadí podle průměrné redukované frekvence)

FREKVENČNÍ SLOVNÍK ZKRATEK / 581
(pořadí podle průměrné redukované frekvence)

DODATKY / 591

Frekvenční seznam interpunkčních znamének / 593

Frekvenční seznam grafémů / 594

Lexikální pokrytí textů / 595

I když frekvence grafémů (písmen) do určité míry i určitěmu frekvence kovákní však nepří, dosud k dispozici není a nově se intenzivně usiluje i o frekvenci víceřádkových jednotek, kolokací a frazémů (pro češtinu dosud údaje neexistují), kde však se pozornost soustřeďuje především na jejich typické páry.

Cíl Frekvenčního slovníku češtiny

Hlavním cílem a důvodem vzniku nového frekvenčního slovníku je jeho mnohokrátelná potřeba. Odpovědí na tuto potřebu byl sice již náležitý vznik nového frekvenčního slovníku pro český jazyk, nazvaného *Frekwence slov češtiny* (dále jen BSC), ve skutečnosti jde o odlišný koncept, resp. frekvenční seznam. V současném mírovském společenství a věku korpusů již není nutné i jeho takto odlišnost příliš zdůrazňovat. Tento slovník se máž přiblížit současnou panou češtinu na mnohokrátelně větším a objektivnějším korpusovém základě ČNK, a to zároveň i v mnohem větší podobě (kompaktnější vzhled) i *Frekwence slovů mluvných češtiny*, viz Čermák et al.).