

Retrográdní slovníky = inverzní

- Přinášejí slovní zásobu (nebo její část) určitého jazyka v obráceném abecedním řazení (tj. od konce slova, a tergo)
- Jsou především zdrojem poznání **slovotvorných typů a jejich produktivnosti v daném jazyce** (ukazují např., které sufixy mohou být produktivní ve tvorbě slovní zásoby, srov.: -ak/-ák X -an)
- Mohou vycházet z psaných textů i mluvených projevů
- Hlavní úlohu ve slovníku mají informace nesené koncovkami substantiv a sloves
- Využití: didaktické účely (tvorba učebnic); zkoumání slovo tvorby, afixů, tvarů slov, odvozenin, pravopisu; pro automatickou morfologickou analýzu při strojovém zpracování; slovník rýmů
- Aktuální retrográdní slovník na trhu chybí, používají se proto především dva níže rozebrané

Retrográdní slovník současné češtiny, M. Těšitelová, J. Petr, J. Králík, 1986

- Odlišuje se od předešlých retrográdních tím, že 1) u všech jednotek uvádí jejich **absolutní frekvenci** zjištěnou z korpusu → je tedy de facto **frekvenční retrográdní**
- 2) Kromě lexémů všech slovních druhů uvádí i uspořádání a absolutní frekvenci všech doložených **tvarů** základních slovních druhů – **substantiv a sloves**
→ informuje tedy o frekvenci všech morfologických kategorií substantiv a sloves v češtině (Př.: Vedle sebe tak stojí třeba: *filozofa* (ak. sg. m.), *socha* (nom. sg. f.), *měřítka* (gen. sg. n.)
- Je založen na kvantitativní analýze jazyka **věcného stylu současné češtiny**
- Vychází ze statistických dat zjištěných z korpusu čítajícího **540 000** slov (v letech 1972-75)
 - korpus obsahuje celkem 180 textů z publicistiky, odborné češtiny a administrativy
- Vychází ze všestranné analýzy současných psaných textů, ale do popředí se dostávají i rysy **mluvených projevů** (třetina materiálu)
- Rozdělen do tří částí, uvádí celkem 99 661 různých jednotek

1. ČÁST: SLOVNÍK LEXÉMŮ

- Celá slovní zásoba, kterou nalezneme v korpusu
- Lexikální jednotky uvedeny v základní podobě a každou doplňuje dvoumístní kód, jehož:
 - 1. číslice/sloupec označuje **slovní druh** (tradiční dělení do 10 tříd)
 - 2. číslice/sloupec blíže specifikuje slovo vzhledem k **původu** (slovotvornému) či **druhu**→ užitím kódů se předejde záměně homonym (např. heslo *ráno* je uvedeno jako ADV i SUB)
 - vpravo je uvedena absolutní frekvence doložená v korpusu (nejvyšší má spojka *a* - 19 017)

poslankyně	11	5	určitě	66	26
velvyslankyně	11	1	spojitě	66	1
zaměstnankyně	11	1	nespojitě	66	1
zástupkyně	11	2	dvojitě	64	1
Purkyně	10	2	svědomitě	66	2
Svrkyně	10	1	rozmanitě	66	1
jeskyně	11	3	znamenitě	66	3

- Př.: **dvojitě 64 1** → 1. sloupec = slovní druh: 6 = adverbium; 2. sloupec = bližší určení původu/druhu: 4 = číslovkové; frekvence slova = 1
- Př.: **poslankyně 11 5** → 1. sloupec = slovní druh: 1 = substantivum; 2. sloupec = bližší určení původu/druhu: 1 = bez určení; frekvence slova = 5
- Př.: substantivizovaná adjektiva se označují 12, zájmenná substantiva 13, vlastní jména 10; u číslovek druhý sloupec vypovídá o druhu: 1 = základní, 2 = druhová,..., u sloves značí vid: 1 = dok., 2 = nedok.

2. ČÁST: SLOVNÍK VŠECH TVARŮ SUBSTANTIV

- Všechny doložené tvary substantiv, jsou charakterizovány příslušnými morfologickými kategoriemi, které jsou zaznačeny kódem
- Př.: **drak 110211 3** → 1. sloupec = slovní druh: 1 = substantivum; 2. sloupec = bližší určení původu/druhu: 1 = bez určení; 3. sloupec = spojení (s předložkou/bez): 0 = bez; 4. sloupec = rod (u mask. i ž./než.): 2 = mask. nž.; 5. sloupec = číslo: 1 = jednotné; 6. sloupec = pád: 1 = první; 7. sloupec = nesklonnost: mezera = x, 8. sloupec = nespisovnost: mezera = x; frekvence slova = 3
- Kódy umožňují rozlišit tvarově homonymní slova, podávají také informaci o tom, jak často jsou homonymní tvary užívány v jednotlivých pádech (bez/předložkové)

drak	110211	3
mrak	110211	2
prak	110214	1
zrak	110211	1

3. ČÁST: SLOVNÍK SLOVES

- Seznam tvarů sloves, doplněny kódy vyjadřující morfologické kategorie
- Různost jejich tvarů je nižší než u substantiv či adjektiv, zato vyšší frekvence
- Př.: **trpěna 5260 24 1** → 1. sloupec = slovní druh; 2. sloupec = vid: 2 = nedok.; 3. sloupec = osoba, číslo: 6 = 3. pl.; 4. sloupec = způsob, čas, slovesný rod: 0 = ind. fut. pas.; 5. sloupec = neurčitý tvar: mezera = x; 6. sloupec = složenost, zvrátlost: 2 = slož. nezvr.; 7. sloupec = jmenný rod: 5 = n.; 8. sloupec = nespisovnost: mezera = x; frekvence = 1

Retrogradní morfemický slovník češtiny, Eleonora Slavičková, 1975

- Podstatou je uvést slova morfemicky rozčleněná na morfémy kořenové, předponové a příponové, vzniká tak jejich inventář
- Výhoda oproti dřívějším slovníkům, kde jsou pouze seřazena slova jako celky
- Nabízí téměř **64 000** slov, jsou doplněna základní klasifikační charakteristikou
- Vychází z *Česko-německého slovníku J. Volného*, konfrontován se *SSJČ*
- Využití: podkladem pro slovníky kořenových, předponových a příponových morfémů
- Popisuje vztah slova k základu slovní zásoby

1. ČÁST: RETROGRÁDNÍ SLOVNÍK MORFEMATICKY ČLENĚNÝCH SLOV

- Slova uvedena v základním tvaru, opatřena indexem 0-9 (u jednoho slova mohou být i 2 nebo 3, v tom případě první index = hlavní charakteristika, další indexy = doplňková char.)
 - Indexy jsou podrobně vysvětleny v úvodu slovníku:
0 – autosémantika domácí; **1** – synsémantika; **2** – slovo složené;
3 – autosémantika cizího původu; **4** – vlastní jména; **5** – slovo expresivní; **6** – homonymie morfémů a slov; **7** – eponyma (odvozená z vlastních jmen); **8** – slovo archaické; **9** – slova nejasného původu/ struktury
- | | |
|-----|-----------------|
| 5 | krav-in-a |
| 0,6 | trav-in-a |
| 0,6 | po-trav-in-a |
| 0 | ledv-in-a |
| 0,6 | ne-vin-a |
| 0 | dřev-in-a |
| 2 | kos-o-dřev-in-a |
| 3 | škod-l-iv-in-a |
- Př.: **0, 6 po-trav-in-a** → domácí autosémantikum (hl. char.), informace o homonymii morfémů a slov (doplňková char.)

2. ČÁST: INVENTÁRNÍ SLOVNÍK ČESKÝCH MORFÉMŮ

- kapitoly popisující **kořenové**, **prefixiální** a **sufixiální** morfémy
- obsahují také **hláskové varianty**, vycházejí z **neutrální slovní zásoby**
- u kořenových je uveden i seznam homonymních morfémů, u prefixiálních i dvojice či trojice předpon

Frekvenční slovníky

- Uvádějí frekvenci užívání slov (slovních tvarů apod.) v různých typech textů
- Sestavovány na základě statistických metod
- Slova **řazena podle četnosti užití** (popřípadě abecedně a doplněna indexy frekvence)
- Dnes slouží jako podklad pro frekvenční slovníky současné mluvené češtiny rozsáhlé počítačové korpusy mluvených textů
- Využití: **poznání jádra jazyka** → k výuce cizího jazyka (tvorba učebnic, slovníků); lingvistická

bádání, lexikografie; teorie komunikace; informatika

- *Kaedingův slovník pro němčinu* (1897) – první frekvenční slovník, základna 10 milionů slov
- Frekvenční slovníky začaly vycházet až před 2. světovou válkou a po ní → slovníky pro němčinu, angličtinu, francouzštinu a další jazyky. Rozsah jejich materiálu byl podmínkami omezen na milion.

Frekvence slov, slovních druhů a tvarů v českém jazyce

J. Jelínek, J. V. Bečka, M. Těšitelová, 1961

- Zaznamenává slovní zásobu, ale také **soustavu ohebných tvarů a slovních druhů**
- Materiál získaný ze 75 děl z let 1926 – 50 (skupiny A-H: beletrie, poezie, odborná literatura, vědecká, žurnalistika, politické projevy,..), excerpcí získáno celkem 1 623 527 výpisků
- Přináší **rozbor kvantitativních poměrů v jazyce v jistém časovém okamžiku**, i když v té době prošla společnost a s ní i jazyk sociálními, politickými a kulturními změnami
- Slova určena svou **formou a mluvnickými významy**, nikoli lexikálními → takže se nerozlišují homonyma (jen pokud mají jiný SD)
- Obsahuje cenné morfologické tabulky a 2 seznamy slov:
 1. 10 000 **nejčastějších slov**, lexikální jednotky řazené podle absolutní frekvence
 2. 26 000 **abecedně řazených slov** s připojenými údaji o absolutní frekvenci každého z nich
 - u každého hesla v abecedním seznamu se kromě celkové frekvence uvádí i počet skupin (A–H) a děl, v nichž se slovo vyskytlo

Frekvenční slovník češtiny, František Čermák, 2004

- Zachycuje jazyk **posledního desetiletí 20. st.**, vychází z reálné, **autentické**, tedy často **nekodifikované formy slov**, jak je přináší korpus
- Zdrojem je Český národní korpus – SYN2000 (nedostatečné zastoupení mluveného jazyka)

ČÁSTI:

FREKVENČNÍ SLOVNÍK APELATIV: 50 000 slov, řazení podle **abecedy** i informace o frekvenci

FREKVENČNÍ SLOVNÍK APELATIV: 20 000 slov, řazeno podle **absolutní frekvence**

FREKVENČNÍ SLOVNÍK APELATIV: 20 000 slov řazeno podle **průměrné redukované frekvence**

FREKVENČNÍ SLOVNÍK PROPRIÍ: 2 000 slov řazeno podle **průměrné redukované frekvence**

FREKVENČNÍ SLOVNÍK ZKRATEK: 1 000 řazeno podle **průměrné redukované frekvence**

- Kromě **absolutní frekvence (FRQ)** řeší i **průměrně redukovanou frekvenci (ARF)**
- **Pořadí podle průměrné redukované frekvence (rank ARF):** redukovaná frekvence je počet

segmentů, ve kterých se nachází alespoň jeden výskyt daného slova. Záleží tedy na rozložení výskytu slova v celém korpusu – zohledňuje se počet pramenů, v nichž se slovo vyskytlo. Slova, která jsou „smrštěna“ do jednoho pramene, mají redukovanou frekvenci nižší.

- Příklad: **píst, ventil** – vyskytují se často, ale v drtivé většině v pramenech jednoho žánru. Díky průměrně redukované frekvenci se nedostanou do popředí na úkor jiných slov.
- Hodnota průměrné redukované frekvence (ARF) je nižší nebo rovna frekvenci absolutní (FRQ) a je zaokrouhlena na celá čísla
- **Pořadí podle absolutní frekvence (rank FRQ):** hodnota absolutní frekvence (FRQ) je počet výskytů všech tvarů daného slova v korpusu.
- Ve slovníku je zachycena absolutní frekvence slov v jednotlivých žánrech: **beletrie, odborná literatura, publicistika** → vyjádřeno v procentech

	rank ARF	ARF	rank FRQ	FRQ	bel. (%)	odb. (%)	pub. (%)
duna	16686	112	15090	329	60	23	18
dunajský	15024	135	14020	373	16	18	66
dunění	20696	75	21377	176	83	7	10

- Zachycení kombinatoriky slov: + znamená vázanou součást frazémů, ustálených výrazů
 - Příklad: *holičkách+, bycha+* → jsou to slova především používána ve frazémích
- Nejfrekventovanější apelativa: *a, v/ve, se, být/bejt, na, ten, s/se, že, z/ze, který/kterej*
- Nejfrekventovanější propria: *Praha, Jan, Evropa, Jiří, Petr, Václav, Josef, Brno, Německo, Pavel*
- Nejfrekventovanější zkratky: *ČR, tzv, USA*

Frekvenční slovník mluvené češtiny, 2007

- Vychází z Pražského mluveného korpusu (PMK) – sestaven na základě více než 300 magnetofonových nahrávek z terénu, které byly pořízeny v letech 1988-1992 a 1992-1995
 - **neformální rozhovor** – zápis konverzace bez vědomí účastníků
 - **formální rozhovor** – iniciování rozhovoru pomocí dotazníku
- Mapuje **autentickou mluvenou češtinu a ukazuje její jádro**
- Obsahuje 780 322 slov a dalších znaků

ČÁSTI:

ABECEDNÍ SLOVNÍK MLUVENÉ ČEŠTINY

ABECEDNÍ SLOVNÍK MÁLO FREKVENTOVANÝCH SLOV (frekvence = 1)

- nejčastější slova: ten, bejt, a, že, já, mít, no, v, na, jako

Použitá literatura

SLAVÍČKOVÁ, Eleonora. *Retrográdní morfematický slovník češtiny: s připojenými inventárními slovníky českých morfémů kořenových, prefixálních a sufixálních*. 1. vyd. Praha: Academia, 1975. 645 s.

TĚŠITELOVÁ, Marie – PETR, Jan – KRÁLÍK, Jan. *Retrográdní slovník současné češtiny*. 1. vyd. Praha: Academia, 1986. 528 s.

JELÍNEK, Jaroslav – BEČKA, V. Josef – TĚŠITELOVÁ, Marie. *Frekvence slov, slovních druhů a tvarů v českém jazyce*. Praha: Státní pedagogické nakladatelství, 1961. 585 s.

ČERMÁK, František a kol. *Frekvenční slovník mluvené češtiny*. 1. vyd. Praha: Karolinum, 2007. 510 s.

ČERMÁK, František – KŘEN, Michal. *Frekvenční slovník češtiny*. Praha: Nakladatelství Lidové noviny, 2004. 595 s.

GREPL, Miroslav - KARLÍK, Petr - NEKULA, Marek - PLESKALOVÁ, Jana (Eds.). *Příruční mluvnice češtiny*. 2. vyd. Praha: Lidové noviny, 2012. 799 s.

CONFORTIOVÁ, Helena. Český retrográdní slovník. *Naše řeč*. 1989, roč. 72, č. 2, s. 81-84.