# Raèunalni¹ki korpusi

---

## Kaj je raèunalni¹ki korpus?

**kórpus** -a m (ọ̑) **1.** *vojaška enota iz divizij in posebnih enot:* formirati korpus; poveljnik korpusa / ekspedicijski korpus **2.** publ., navadno s prilastkom *skupina, zbor:* nastopili so najboljši korpusi; baletni, godalni korpus / diplomatski korpus *diplomatski zbor*; korpus miru zlasti v ameriškem okolju *organizacija iz prostovoljcev, ki skrbi zlasti za gospodarski razvoj in pomoč nerazvitim deželam* **3.** knjiž., redko *zbirka:* korpus srednjeveških fresk ● ekspr. ta ženska je pa korpus *je zelo debela*

**korpus delikti** gl. corpus delicti

**corpus** ('kɔːpəs) *n., pl.* **-pora** (-pərə). **1.** a collection or body of writings, esp. by a single author or on a specific topic: *the corpus of Dickens' works.* **2.** the main body, section, or substance of something. **3.** *Anatomy.* **a.** any distinct mass or body. **b.** the main part of an organ or structure. **4.** *Linguistics.* a body of data, esp. the finite collection of grammatical sentences of a language that a linguistic theory seeks to describe by means of an algorithm. **5.** a capital or principal sum, as contrasted with a derived income. **6.** *Obsolete.* a human or animal body, esp. a dead one. [C14: from Latin: body]

---

**(raèunalni¹ki) korpus**: *urejena zbirka besedil, hranjena in obdelana na raèunalniku.*

---

## Tipologija

**Kriteriji**:

kolièina; kakovost (avtentiènost); enostavnost (ASCII); dokumentiranost.

**Razredi**:

- *referenèni*
  (osnova: jezikovni standard)
  Brown, BNC, Bank of English (100M)
- *govorni & govorjen*

    (avtentiènost)
    Map-Task (HCRC): 18 ur, dialog
  - *podjezikovni*
    (omejeni in specializirani)
    ATC0 (Texas Instruments): 70 ur, stolp-zrak
  - *vzorèni*
    (zgodovinski / pravni razlogi)
    Brown
  - *spremljevalni*
    (jezik v spreminjanju)
    periodika (Internet)
  - *primerljiv*
    (prevajanje)
    ECI/MCI: 27 jezikov, 92M
  - *vzporedni*
    (prevajanje)
    Hansard: kanadski parlament

# Rast

- Referenèni: Brown (1M) ’64 ... LOB (1M) ’74 ... BNC, BoE (100M) ’95
- Drugi ’angle¹ki’: Sussane, Hansard, ...
- Jeziki EU: nem¹ki, francoski, ¹panski, ¹vedski, ...
- Vzhodno-evropski jeziki: estonski, romunski, èe¹ki, ruski, ...
- Iniciative EU: ECI/MCI, Eagles, Relator, ELRA ElsNet, Copernicus
- Slovenija:
  Jo¾e Topori¹iè, urednik: Besedila slovenskega jezika. Ljubljana : Filozofska fakulteta, 1975.

# MULTEXT-East Korpus

| | **Orwell** | | | **Leposlovje** | | | **Èasopisi** | | |
|---|---|---|---|---|---|---|---|---|---|
| **En:** | D | H | C | | | | | | |
| **Bg:** | D | H | C | D | H | C | D | H | C |
| **Cs:** | D | H | C | D | H | C | D | H | C |
| **Et:** | D | H | C | D | H | C | D | H | C |
| **Hu:** | D | H | C | D | H | C | D | H | C |
| **Ro:** | D | H | C | D | H | C | D | H | C |
| **Sl:** | D | H | C | D | H | C | D | H | C |

[nazaj]

---