

CJBB75 6

Je možné kombinovat empirický výzkum a introspekci

Máme kousek porušeného textu (dvě slova, z nichž známe buď poslední písmeno, nebo poslední dvě písmena). Které z následujících dvojic mohou/nemohou být českými větami (jeden z členů dvojice je verbum finitum). Dokažte to.

a)

ý	ůl
---	----

b)

š	ě
---	---

c)

í	m
---	---

d)

u	č
---	---

e)

t	r
---	---

f)

é	ou
---	----

g)

p	s
---	---

Důkaz proveďte 1) empiricky na základě analýzy dat z korpusu a 2) analyticky na základě znalostí české gramatiky.

Řešení:

Máme kousek porušeného textu (dvě slova, z nichž známe buď poslední písmeno, nebo poslední dvě písmena). Které z následujících dvojic **mohou/nemohou** být českými větami (jeden z členů dvojice je verbum finitum). Dokažte to.

Např.

s	á
---	---

<i>pes</i>	<i>štěká</i>
------------	--------------

a)

y	ůl
---	----

b)

s	ě
---	---

Důkaz proveďte 1) empiricky na základě analýzy dat z korpusu a 2) analyticky na základě znalostí české gramatiky.

- a) V korpusu vyhledáme všechny sekvence tvaru (Implicitní atribut - word) . *ý . *ůl . Získáme tak dvojice jako **zlatý důl, kulatý stůl, okřídlený vůl**. P-filtrem budeme hledat dvojice takové, že jeden z členů bude (Implicitní atribut – pos) sloveso. Buď . V, nebo V . . V obou případech nenalezneme ani jeden doklad. Empiricky jsme dokázali, že v korpusu není ani jeden doklad posloupnosti vyhledaných slovních tvarů taková, že by jeden z členů dvojice byl slovesný tvar (tudíž nemůže být ani verbum finitum).
- b) Český slovesný tvar určitý (ani neurčitý) nemůže končit na příslušné řetězce. Určité tvary mohou končit pouze na vokál –a, -á, -e, -ě, -i, -í, -y, -o, -u (-ou) a na téměř všechny konsonanty (?x, ?q, ?w). Pokud končí na –l, pak se může jednat buď o l-ové přídělníky (ve 3. os. je tvarem verb. fin.), pak ale mohou před –l předcházet z vokálů pouze [ááěěiyu], nebo o imperativ, pak ovšem před –l může předcházet pouze krátký vokál (*val, vel, šil, vol, bul, nemyl se*).

Domácí úkol:

Stejným způsobem zpracujte případy c) – g).