**Gramatika a korpus II plin032**

**Středa: 7.30-9.00 G13**

23. 3. Rozbor dú.

Efektivní vyhledávání homonym slovesných tvarů pomocí počítačových nástrojů II. (I když pořád spěj, pořád někam spěj …)

**Dluž a splácej!** [[1]](#footnote-1)

V korpusech (jakkoliv jsou ty, které dnes máme k dispozici, rozsáhlé) není zachycení tvarů imperativu takové, jaké bychom si přáli. Důvody vidíme jednak v omezení výskytu tvarů imperativu, které plyne přirozeně z významů jednotlivých sloves (rozkazovat k činnosti pojmenované příslušným slovesem je mnohdy podivné až nesmyslné), jednak v omezeních pragmalingvistických (imperativ je vnímán jako tvar, kterému se v řadě komunikačních situací snažíme spíše vyhnout). Problematické je ovšem pro automatickou morfologickou analýzu i to, že v řadě případů je tvar imperativu (2. os. sg.) homonymní s nějakým často zcela významově naprosto vzdáleným (krátkým) slovem, či zkratkou (*kop, oř, kos, jez, nos, kup, div, děj, lež, boj, pas, per, tel, …*), které bývá frekventovanější, než potenciální imperativní tvar. S tím souvisí, že morfologické analyzátory s potenciálními tvary imperativu pracují mnohdy selektivně (technické omezení rozgenerování tvarů, které opakovaně způsobují chyby v lemmatizaci – srov. např. výskyt lemmatu *telit* v korpusu SYN2000). Tyto skutečnosti je třeba mít na zřeteli, chceme-li se věnovat konzistentnímu zachycení variantních tvarů imperativu.

V tomto semináří se pokusíme najít pomocí nástroje Morfio podklady pro uvedený typ homonym. Všimneme si, že nemusí vždy jít jen o homonyma a podíváme se, jak vypadá disambiguace.

***Dluž* a *splácej*!**

Jaké dvojice obecně představují slovesa podobná těm, která vidíme v uvedeném případě?

1. slovesa s kořenným ou, jejichž imperativ má alternaci (ou/u: dl**ou**žit/dl**u**ž) a jsou tedy homonymní se slovesy s kořenným u (dl**u**žit).
2. Slovesa IV. třídy (sázet: *splác****e****t*) a slovesa V. třídy (dělat: *splác****a****t*) mohou mít homonymní tvary imperativu (sáz**ej**/děl**ej**: *splác****ej***).

Použijeme nástroj Morfio: (Odkaz na toto zadání: <http://morfio.korpus.cz/0d5waQ4X>)





31 dvojic. Všímejte si, že některé dvojice by bylo možné klasifikovat jako synonyma. Které?





Celkem nalezeno 192 dvojic

Problémy ručního třídění a pozorování vlastností kandidátů na uvedené případy homonymie:

1. sloveso na [eě]t nemusí mít imperativ na [eě]j (*vrtět/vrtat*)
2. sloveso na at nemají alternace (jotaci) labiálních finál v kontextu imperativního –ej (supět, supat, zasípět/zasípat)
3. sloveso na –at není slovesem V. třídy (nepravidelné *–spat: prospět/prospat*)

Vyber slovesa, kterých se možné homonymie týkají a sleduj desambiguaci potenciálně homonymních tvarů v korpusu SYN.

Problémy: Více větších korpusů.

Vyzkoušejme řadu Araneum, CzTenTen a internet.

DÚ:

1. Existují i jiné případy krácení kořenného vokálu v imperativu, které mohou zapříčinit vznik tvarové homonymie?
2. Kdy se kořenný vokál krátí? Lze popsat pravidlem, nebo je krácení podmíněno lexikálně?
3. Homonymní čtení má i slogan: ***Nevaž se, odvaž se!*** (Chcete hubnout, nebo trojčit?). Podívejte se, jak je v případě sloves, která mají homonymní tvary, vyřešena disambiguace v korpusu SYN.
1. Hezkým dokladem homonymních imperativů od různých sloves je lichvářský slogan, na který upozornila J. Hlaváčová: *Dluž a splácej.* [↑](#footnote-ref-1)