**Gramatika a korpus II: PLIN032 JS 2019**

**Středa: 12.00-13.30 G13**

27. 2. Dú: Varianty od kmene prézentního

6. 3. Vlastnosti adjektiv zakončených na -cí

# ***Lhoucí* nebo *lžoucí* ? *Slovotvorné dublety procesuálních adjektiv na -oucí/-ící v českých korpusech***

# *Motto*

## ***… V truchlivšti dlouhé noci tvé***

## ***v  zemi přibité v té nepřejícné zemi***

## ***sníš kníže rodné řeči žárlivé***

## ***žal <lhoucí> pod růžemi ….***

## ***(František Halas: K. H. M.)***

# Konkordance variant adjektiva *lhoucí/lžoucí* z korpusu czTenTen12

# **Adjektiva na *–oucí/-ící***

## Paradigmatické tvoření

## Opěrný tvar 3. pl. ind. préz. akt. / přechodník přítomný

## Tvarová varianta/dubleta je předpokladem slovotvorné varianty/dublety

# **KDE ?**

# Morfio

# Výpis seznamu vyhledaných dvojic

# *spící*

# *spějící*

# *chybící*

# *svářející/svářící*

# [word=".\*[kč]oucí.\*"]

# internet

## *V osm hodin začínám školení osob konajících veřejně prospěšné práce, osamocení zaměstnanci <****sečoucí****> trávu křovinořezem nebo motorovou sekačkou a uklízející obecní plochy ve dne provádějící úklid chodníků a silnic v obci!!!.*

## *My musíme poslouchat firmy, tedy ty velké, ne nějaké pekaře <****pečoucí>*** *chleba, řezníky vyrábějící paštiku, rolníky pěstující pšenici, kováře na kovadlině kovající mříže, elektrikáře spravující vedení, krejčí šijící oděvy, spisovatele píšící knihy, herce hrající roli.*

# ***tečící vlečící***

# pečící / pečicí

# Nerozlišování adjektiv na *-icí × -ící*

## (a) V označeních role/úřadu/funkce osob se účelová adjektiva nahrazují dějovými: *světící biskup, velící důstojník*.

## (b) *Chladicí věže* a podobná zařízení, pojmenovaná podle účelu, jsou mnohdy ve stálém provozu, takže volba *-ící* není významově scestná.

## (c) Rozdíl ve výslovnosti[i] *×* [í]nebývá zřetelný. (d) V cizích jazycích mívají oba typy mnohdy jediný ekvivalent: např. *Sleeping Beauty* ,Spící kráska, Šípková Růženka‘i *sleeping car/bag/pill* ,spací pytel/vagon/pilulka‘.

# [word="[dmptvz][rř]oucí.\*"]

# [word="[dmptvz]řící.\*"]

# czTenTen12

# *-oucí/-ající*

# *-oucí/-ající*

# *-oucí/-ící/-ající*

# *-oucí/-ící/-ající*

# III. třída / IV. TŘída

# Závěr

## Jak hledat v korpusu doklady (slovesa uzavřených tříd)

## Homonymie a přegenerovávání – problém nejen NLP

## Korpus jako zdroj informací o užití variant (kontext – frekvence – zdroj)

# **Literatura**

## Bauer, J., Lamprecht, A., Šlosar, D. (1986): *Historická mluvnice češtiny*. Praha: SPN.

## Cvrček, V. (2013): Co je v ČNK nového II. *KGA 2013*, 07, s. 95–97.

## Cvrček, V., Vondřička, P. (2012): *Morfio.* Dostupný z WWW: <http://morfio.korpus.cz/>.

## Dokulil, M. a kol. (1986): *Mluvnice češtiny 1.* Praha: Academia.

## Hajič J. (2004): *Disambiguation of Rich Inflection (Computational Morphology of Czech)*. Praha: Karolinum Charles University Press.

##

## Jelínek, T. (2008): Nové značkování v Českém národním korpusu. *Naše řeč*, 91, 1, 13–20.

## Komárek, M. a kol. (1987): *Mluvnice češtiny 2.* Praha: Academia.

## Osolsobě, K. (1996): *Algoritmický popis české formální morfologie a strojový slovník češtiny*. Disertační práce. Brno : MU.

## Osolsobě, K. (2009): Kajícný a nevěřícný – adjektiva na *-cí/-cný*: slovníky, gramatiky, korpusy. In: Hlaváčková, D. – Horák, A. – Osolsobě, K. – Rychlý, P. (eds.), *After Half a Century of Slavonic Natural Language Processing*, Brno : Masarykova univerzita, s. 173–183.

## Osolsobě, K. (2011): *Morfologie českého slovesa a tvoření deverbativ jako problém strojové analýzy češtiny*. Brno : MU.

## Osolsobě, K. (2013*):* Korpusy a internet jako zdroje dat pro výzkum produktivity periferního slovotvorného typu: adjektiva typu hrůzoucí (hrůza) v korpusech a na internetu*.* *Gramatika a korpus 2012*. Hradec Králové: Gaudeamus.

## Osolsobě, K. (2014): Podklady pro úpravy slovníku automatického morfologického analyzátoru. rkp.

## Petkevič, V. (2006): Reliable Morphological Disambiguation of Czech: Rule-Based Approach is Necessary. In: *Insight into the Slovak and Czech* Corpus Linguistics (Šimková M. ed.). Bratislava: Veda, 26–44.

## Rychlý, P. (2010): *Bonito2 - web-based graphical user interface to the Manatee system.*

## Spoustová D., Hajič, J., Votrubec, J., Krbec, P., Květoň, P. (2007): The Best of Two Worlds: Cooperation of Statistical and Rule-Based Taggers for Czech. In: *Proceedings of the Workshop on Balto-Slavonic Natural Language Processing*. Praha: ACL, 67–74.

## Šimandl a kol. (2016): *Slovník afixů užívaných v češtině.* Praha: Karolinum. (<http://www.slovnikafixu.cz/>)

## Šmilauer, V. (1972): *Nauka o českém* *jazyku*. Praha: SPN.

## Kolektiv autorů ÚJČ (1993): *Pravidla českého pravopisu*. Praha: Academia.

## **SYN**  *Český národní korpus - SYN.* Ústav Českého národního korpusu FF UK, Praha. Cit. 4. 3. 2014, dostupný z WWW: <http://www.korpus.cz>.

## **czTenTen12** FI MU – czTenTen12. Centrum zpracování přirozeného jazyka FI MU, Brno. Cit. 4. 13. 2014, dostupný z WWW: <http://ske.fi.muni.cz/bonito>.

## Internetová jazyková příručka. Dostupná z WWW: <http://prirucka.ujc.cas.cz/>

13. 3. Dú:

Jak vyhledáte účelová adjektiva typu ***mycí***houba*,* ***žehlicí*** prkno, ***sázecí*** systém, ***hrací***karty?

Popište postup. Tato adjektiva rovněž končí na *cí*. Je zde problém s přegenerováním dotazu?

Existují ještě nějaká adjektiva končící na *cí*, která nepatří ani do jedné ze jmenovaných skupin?

Popište je.

Konzultujte příslušné heslo/hesla ve slovníku Šimandl a kol. (2016). <http://www.slovnikafixu.cz/>