

Slovnědruhá a morfoložická homonymie, homofonie a homografie v současné češtině¹

Vladimír Petkevič (Praha)

PART-OF-SPEECH AND MORPHOLOGICAL AMBIGUITY, HOMOPHONY AND HOMOGRAPHY IN CONTEMPORARY CZECH

The paper presents a classification of the types of morphological ambiguity and the types of homophony and homography in contemporary Czech occurring in the material of the SYN and SYN2013PUB corpora of the Czech National Corpus. The classification of homonymy and homography constitutes a data base for the rule-based automatic morphological disambiguation of written Czech performed in the Institute of Theoretical and Computational Linguistics at the Faculty of Arts, Charles University. As for homophony, the types presented in the paper and mainly the sets of word forms associated with these types, can be used for the disambiguation of spoken texts.

KEYWORDS

part-of-speech and morphological ambiguity, homophony, homography, contemporary Czech, language corpora

KLÍČOVÁ SLOVA

slovnědruhá a morfoložická homonymie, homofonie, homografie, současná čeština, jazykové korpusy

1. ÚVOD

V tomto příspěvku se zaměříme na morfoložickou homonymii, homofonii a homografii v současné češtině reprezentované korpusy současné češtiny SYN a SYN2013PUB (srov. Křen et al., 2013 a 2014). Studium homonymie včetně homonymie morfoložické a polysémie ve světové lingvistice (srov. např. Lyons, 1975; Deane, 1988; Langacker, 1987, 1991a, 1991b; Tuggy, 1993; Chierchia — McConnell-Ginet, 1993) i lingvistice bohemistické (srov. Těšitelová, 1966; Filipec, 1970; Filipec — Čermák, 1985; Horalík, 1967, 1968; Horálek, 2005; Čermák, 2010), poměrně zanedbávané v jazykovědě před nástupem matematické a obzvláště korpusové lingvistiky, se ukazuje jako nesmírně důležitá v souvislosti s počítačovým zpracováním češtiny, konkrétně s automatickou gramatickou anotací korpusových textů. Aby se korpusy daly kvalitně morfoložicky, syntakticky i sémanticky anotovat, musí se řešit řada problémů spjatých s homonymií slovních tvarů, syntaktických struktur a lexémů. Problematika automatické morfoložické anotace, která se řeší stochasticky a/nebo pomocí lingvistických pravidel, již byla popsána ve velkém množství publikací ve světě i u nás (Church, 1988; DeRose, 1988; Charniak, 1997; Garside et al., 1997; Brill, 1992; Hajič, 2004; Petkevič,

¹ Tento příspěvek byl podpořen z grantu GAČR Treebank češtiny na základě gramatiky, reg. č. 13-27184S.



2006). Obecně platí, že morfologii využívá syntax a syntax zase slouží sémantice. Při počítačovém zpracování jazyka slouží správná morfologická analýza a disambiguace (zjednoznačnění) automatické syntaktické analýze a ta zase případné analýze sémantické. Přitom jedna fáze zpracování jazyka zásadně ovlivňuje další: je-li vstupní věta nesprávně morfologicky analyzována nebo disambiguována, je nanejvýš pravděpodobné, že následná syntaktická analýza bude chybná, a pokud bude syntaktická analýza chybná, negativně to ovlivní analýzu sémantickou. Aby se věty obsahující homonymní, homofonní² či homografní tvary analyzovaly správně zvláště v systému využívajícím lingvistických disambiguačních pravidel, je nutno tyto tvary nejprve náležitě klasifikovat z hlediska gramatického systému češtiny, přičemž taková klasifikace nemá zdaleka jen praktický význam pro korpusovou anotaci, ale přispívá i k teoretickému obohacení poznatků o studovaném a zpracovávaném jazyce, v našem případě češtiny.

Příspěvek dále tvoří čtyři části. Ve druhé části nejprve definujeme základní pojmy: homonymii a polysémii, dále pak rozlišíme jednak pravá a nepravá homonyma, jednak plně homonymní a částečně homonymní lexémy. V hlavní, třetí části zavedeme pojem morfologické homonymie, uvedeme jej do vztahu k předchozím pojmům a popíšeme druhy morfologické homonymie.³ Ve čtvrté části pak pojednáme o druzích homofonie a homografie; v páté, závěrečné části celý příspěvek shrneme.

2. DEFINICE HOMONYMIE A POLYSÉMIE

V této části budeme definovat základní pojmy.

Definice 1. Homonymum „je lexém, který je synchronně zcela formálně podobný jinému a tato naprosto náhodná plná formální shoda dvou, popř. více lexémů nemá sémanticky ani funkčně v sobě nic příčinného, nemá žádný společný základ, který by oba výrazy sdílely“ (Čermák, 2010, s. 261).

Jako příklad můžeme uvést lexém *travička* (žena, jež někoho otrávil, vs. mladá tráva; tamtéž).

Definice 2. Polysémie „[je] vedle homonymie případ víceznačnosti izolované jazykové, především slovní formy. [...] Od lexikální homonymie se liší tím, že jednotlivá významová užití (lexie) polysémního lexému (hyperlexému) spolu sémanticky i geneticky souvisejí“ (Karlík et al., 2002, s. 320–321, autor ZH — Zdeňka Hladká).

Jako příklad se uvádí lexém *jazyk*: 1. svalnatý orgán v dutině ústní, 2. co se podobá jazyku živočichů, např. součást boty, 3. soustava vyjadřovacích prostředků (tamtéž).

Obě definice jsou lexikologické: týkají se lexémů, nikoli slovních tvarů, kterými se v tomto příspěvku zabýváme především.

2 Homofonii je nutno studovat mj. s ohledem na automatické zpracování mluveného jazyka.

3 Nebudeme se tu zabývat homonymií lexikální, jež je důležitá zvláště při automatické disambiguaci lexikálních významů.

Z hlediska protikladu psaného a mluveného jazyka lze homonyma dále dělit na **pravá a nepravá**.

Definice 3a. Pravá homonyma jsou lexémy, které mají různý význam a přitom stejnou zvukovou i grafickou podobu, tj. vyslovují se a píšou stejně.

Definice 3b. Nepravá homonyma jsou lexémy, které mají různý význam a

- buď mají stejnou zvukovou podobu a odlišnou podobu grafickou — taková nepravá homonyma se nazývají **homofona** (srov. heslo **homonymie** v ESČ, Karlík et al., 2002, s. 169–170, autor rovněž ZH — Zdeňka Hladká)
- nebo mají stejnou grafickou podobu a odlišnou podobu zvukovou — taková nepravá homonyma se nazývají **homografa** (tamtéž).

Příklady pravých homonym jsou lexémy *travička* či *role*. Mezi nepravá homonyma lze zařadit homofonní tvar [plot], který lze psát jako *plot* nebo *plod*.

Definice 3a a 3b lze upravit tak, aby se vztahovaly na slovní tvary: slovo *lexémy* v nich nahradíme výrazem *slovní tvary*. V tomto případě můžeme jako příklady pravých homonym uvést slovní tvary *okouním*, *smýčce*, *pracovně*, *již*.⁴ Mezi nepravá homonyma lze zařadit homofonní tvar [objetí], který lze psát jako *obětí* a *objetí*,⁵ nebo homofonní tvar [stesky], jež lze psát jako *stezky* či *stesky*. Mezi nepočtená homografa v češtině patří třeba tvar *divizně* (vyslovuje se [d'ivizně]), je-li tvarem substantivního lexému *divizna*, a [dyvizně], je-li to příslovce odvozené od adjektiva *divizní*.

Homonyma se dále rozlišují na základě rozsahu formálních shod homonym.

Definice 4a. Plně homonymní lexémy se shodují v celých svých paradigmatech.

Definice 4b. Částečně homonymní lexémy se neshodují v celých svých paradigmatech, jenom v jejich částech (srov. Karlík et al., 2002, s. 170).

Plně homonymními lexémy jsou např. *travička*, *kynout*, *cenit*, *raketa*, *topit*, *zahýbat*. Částečně homonymní lexémy, tj. **homonyma neúplná/částečná**, se v některých tvarech svých paradigmat liší. Existují tedy slovní tvary, které vykazují totožnost formy v rámci paradigmat alespoň dvou lexémů (tvarových systémů). Příkladem lexémů částečně homonymních je dvojice tvořená femininem *noc* a neživotným maskulinem *nok*: v průniku jejich paradigmat, tj. tvarových systémů, je homonymní tvar Lpl *no-cích*. Jiným z mnoha příkladů budiž tvar *vezme*, jež patří jak do tvarového systému slovesa *vézt* jakožto 1. os. pl. imperativu slovesa *vézt*, tak do tvarového systému slovesa *vzít* jako 3. os. sg. přítomnosti.

4 Tvar *okouním* je tvarem adjektiva *okouní* i slovesa *okounět*, tvar *smýčce* je tvarem substantivního lexému *smýčka* i *smýčec*, tvar *pracovně* je tvarem substantivního lexému *pracovna* i příslovečného lexému *pracovně*, tvar *již* je příslovce/částice i tvar vztažného osobního zájmena *jenž*, přičemž v obou interpretacích se čte jako [jiš], je to tedy pravé homonymum.

5 Tvar *objetí* je sám homonymní: je tvarem slovesného substantiva odvozeného od slovesa *objet* i *obejmout*.

3. MORFOLOGICKÁ HOMONYMIE V ČEŠTINĚ A JEJÍ TYPY

Morfologickou homonymií definujeme stejně jako Těšitelová (1966, s. 8).

Definice 5. Morfologická homonymie je jakákoli totožnost formy při funkčním rozrůznění v rámci: (a) jednoho tvarového systému, nebo (b) více tvarových systémů. Se zřetelem k této okolnosti můžeme rozlišit homonymií vlastní a nevlastní.

Definice 5a. (Morfologická) homonymie vlastní je totožnost formy při funkčním rozrůznění v rámci tvarového systému jednoho lexému.

Definice 5b. (Morfologická) homonymie nevlastní je totožnost formy v rámci alespoň dvou tvarových systémů:

- (a) buď jednoho lexému, který má tvary rozlišené vzhledem ke jmennému rodu,⁶
- (b) nebo více lexémů patřících buď k jednomu slovnímu druhu, nebo k více slovním druhům.

Morfologickou homonymií tedy chápeme jako homonymií celých tvarů, nikoli menších jednotek, např. morfů. Definice 5 přitom zahrnuje jak (a) tvary, jejichž odlišné interpretace spolu vůbec nesouvisí, a jde tedy o náhodnou tvarovou identitu, tedy homonymií podle definice 1 vztažené ovšem na tvary, tak (b) tvary, jejichž odlišné interpretace spolu naopak úzce souvisí, a vyhovují tedy vymezení polysémie v definici 2. Při takovémto chápání je vlastní homonymie příkladem polysémie tvarů, zato nevlastní homonymie je buď případem homonymie ve smyslu definice 1, nebo polysémie ve smyslu definice 2, nebo tvar není víceznačný významově, ale pouze formálně, jak nyní ukážeme.

Obecně je v systému jmen (= substantiva, adjektiva, zájmena, číslovky) klíčový pro rozlišení vlastní a nevlastní homonymie slovní druh, slovní poddruh (např. různé druhy zájmen nebo číslovek), jmenný rod a lexém, jejichž odlišnost v interpretacích slovního tvaru zakládá u jmen homonymií **nevlastní**. Pokud nějaký tvar má v paradigmatu jednoho lexému různé morfemické interpretace jakožto pouze kombinace pádu a čísla, nikoli jmenného rodu, jde o homonymií **vlastní** (srov. Těšitelová, 1966, s. 8). Například tvar *soudce* vyjadřuje Nsg, Gsg, Asg, Vsg, Npl, Apl a Vpl lexému *soudce*; tvar *hradu* vyjadřuje Gsg, Dsg a Lsg lexému *hrad* — jsou to příklady homonymie vlastní: tvar *hradu* má tři různé interpretace, ale v každé z nich patří k paradigmatu lexému *hrad*, takže je to z hlediska definice 2 tvar polysémní. V obdobném smyslu chápeme homonymií vlastní i u adjektiv, zájmen a číslovek. Ostatní typy homonymie, v nichž figuruje tvar jména, jsou typy homonymie **nevlastní**. Jsou to obecně tyto typy:

- **forma** jména patří k více tvarovým systémům téhož slovního druhu, ale různého jmenného rodu: např. substantivní tvar *prestiž* je tvarem Nsg a Asg jak maskulina neživotného, tak feminina, význam je však stejný a víceznačnost je tak jen for-

⁶ Rozlišujeme čtyři hodnoty jmenného rodu: maskulinum životné a neživotné, femininum a neutrum.

málně morfologická; substantivní tvar *model* je tvarem Nsg lexému *model* (mask. živ.) a zároveň i tvarem Nsg, Asg lexému *model* (mask. neživ.) a také tvarem Gpl lexému *modla* (fem.) — jasná homonymie druhého a třetího lexému ve smyslu definice 1 a polysémie prvního a druhého lexému; substantivní tvar *soudcích* je tvarem Lpl lexému *soudce* (mask. živ.) i *soudek* (mask. neživ.) — jasná homonymie; tvar *snížení* je tvarem adjektiva *snížený* i tvarem substantiva *snížení* — polysémie ve smyslu definice 2; adjektivní tvar *neschopnými* vyjadřuje Ipl všech čtyř jmenných rodů — víceznačnost je tu jen formálně morfologická; tvar *zavřete* vyjadřuje imperativ i indikativ přítomnosti 2. osoby plurálu téhož lexému *zavřít* — příklad polysémie (dále už nebudeme uvádět vztahy k definici 1 a 2);

- **forma** jména patří k tvarovým systémům lexémů různých slovních druhů: např. tvar *smlouvám* je jak Dpl substantiva *smlouva*, tak 1. os. sg. přítomnosti slovesa *smlouvat*; tvar *interně* je jak tvarem pozitivu adverbia, tak Dsg a Lsg substantiva *interna*; tvar *žeň* je Nsg a Asg substantivního lexému *žeň* i imperativ sg. dokonce dvou lexémů slovesných: *ženit* i *hnát*; tvar *je* je Asg neut. zájmena *ono*, Apl zájmen *on*, *ona*, *ono* i 3. os. sg. přítomnosti slovesa *být*;
- **forma** jména patří k více tvarovým systémům téhož slovního druhu, téhož jmenného rodu, ale odlišného lexému: např. *tancích* je tvarem Lpl jak substantivního lexému *tank*, tak *tanec*.

Pádový synkretismus jmen je tedy případem vlastní i nevlastní homonymie, vyskytuje se ve všech jmenných paradigmatech, tj. v každém paradigmatu se nachází aspoň jeden tvar mající více morfologických interpretací.

Příkladem vlastní homonymie v systému slovesném je homonymie tvarů 3. os. sg. a pl. přítomnosti sloves patřících ke 4. konjugační třídě a časovaných podle vzoru *prosí* a *trpí* a dnes už i *sází*, zvláště v mluvě. Ostatní typy slovesné homonymie představují homonymii nevlastní, přičemž tu rozlišujeme čtyři typy. **Forma** slovesa:

- patří k různým tvarovým slovesným systémům téhož lexému: např. několik sloves má tvar infinitivu shodný s pasivním přičestím životného i neživotného maskulina singuláru, např.

(1a) *Ten dokáže na ošetřené cesty opět **navát*** (Inf) *mnoho sněhu*. vs.

(1b) *Prach tedy nemůže být **navát*** (PassPart) *zvenčít*;

kde tvar *navát* patří k různým tvarovým systémům (infinitiv vs. pasivní participium) v rámci tvarů téhož lexému. Za tvarové slovesné (pod)systémy totiž považujeme u sloves systém přítomnosti, syntetického futura, tvarů kondicionálu, imperativu, infinitivu, systém l-ových a pasivních přičestí, systém přítomného a minulého přechodníku a infinitiv;

- patří k těmto tvarovým slovesným systémům odlišných slovesných lexémů: tvar *tají* je tvarem 3. os. přítomnosti sg. i pl. slovesa *tajit* a zároveň tvarem 3. os. přítomnosti pl. slovesa *tát*;
- patří k odlišným tvarovým slovesným systémům různých slovesných lexémů: např. tvar *bil* je jak tvarem imperativu sg. lexému *bílit*, tak tvarem l-ového pří-

čestí životného i neživotného maskulina sg. lexému *bít*; podobně tvar *vzdal* je jak tvarem imperativu sg. lexému *vzdálit*, tak tvarem l-ového přičestí životného i neživotného maskulina sg. lexému *vzdát*;

- patří k tvarovým systémům odlišných slovních druhů: např. tvar *znám* je tvarem 1. os. sg. prézentu slovesa *znát* a zároveň jmenným tvarem adjektiva *známý* v životném i neživotném maskulinu sg.; tvar *děl* je tvarem l-ového přičestí životného i neživotného maskulina sg. slovesných lexémů *dít se* i *dít*, dále tvarem imperativu sg. lexému *dělit*, ale i tvarem Gpl substantivních lexémů *dělo* a *dílo*!

4. KLASIFIKOVANÉ TYPY MORFOLOGICKÉ HOMONYMIE

Autor na základě morfoloogické analýzy korpusu SYN2013PUB roztrídil typy vlastní a nevlastní morfoloogické homonymie, přičemž každému typu je přiřazena skupina příslušných homonymních tvarů.⁷ Ukazuje se, že v češtině existují stovky tříd nevlastní morfoloogické homonymie dané kombinací:

- hodnot morfoloogických kategorií vč. slovního druhu a poddruhu a
- lexému reprezentovaného lemmatem.

Příkladem takové třídy je homonymie mezi tvary 1. os. sg. prézentu slovesa a zároveň Dpl substantiva ve femininu; patří sem mimo jiné tvary *párám*, *smlouvám*, *tížím*, *zdržím*, srov.:

- (2a) Už se s ní **párám** (sloveso) *patnáct let*. (SYN)
 (2b) Kvůli *nebezpečným* **párám** (substantivum) *museli evakuovat obyvatele nejbližších domů*. (SYN)

Nejvíce typů homonymie je:

- mezi substantivy navzájem
- mezi substantivy a dalšími slovními druhy, hlavně slovesy
- mezi slovesy a dalšími slovními druhy.

Skupina homonymních tvarů náležící jistému typu může obsahovat jediný tvar (srov. pozn. 9 níže), nebo tvarů několik, nebo několik desítek či stovek tvarů, ba dokonce tisíce tvarů. Skupinu obsahující několik desítek tvarů představuje typ homonymie mezi Dsg a Lsg substantiva ve femininu a pozitivu adverbia: *herně*, *interně*, *ploše*, *ředitelně*, *služebně*, *úhlopříčně* a další. Velmi početnou skupinu s tisíci tvary představuje typ homonymie mezi slovesnými substantivy v Nsg, Gsg, Dsg, Asg, Vsg, Lsg, Npl, Gpl, Apl, Vpl neutra a slovesnými adjektivy v Npl a Vpl mask. živ. odvozenými od pasivního přičestí: *ohrožení*, *políbení*, *snížení*...

⁷ Použil přitom však i původní manuální klasifikaci homonymních slovních tvarů, kterou vytvořil K. Oliva.

Může se stát, že hned několik tvarů v paradigmatu určitého lexému vykazuje nevlastní homonymii různých typů, tyto tvary tedy patří k různým třídám homonymie. K lexémům, jejichž několik tvarů vykazuje nevlastní homonymii a homonymní jsou i některé jejich deriváty, patří např. slovesné lexémy *jíst* a *sníst*:

- *jíst* má tyto homonymní tvary: *jí, jím, jedli, jeden, jed, jez*;
- *sníst* má tyto homonymní tvary: *sním, sníš, sní, sníme, sníte, snědí*.

Tvary *jí* a *jím* jsou homonymní se zájmennými tvary osobního zájmena *ona*; *jedle* je tvarem adverbia i tvarem substantiva *jedle*; i tvar *jedli* je homonymní s jedním z tvarů substantivního lexému *jedle*; tvar *jeden* je homonymní se základní číslovkou; *jed* je jak tvarem substantivním, tak l-ovým přičestím maskulina sg. v obecné češtině (*jed'*, tj. *jedl*)⁸ a konečně *jez* je tvarem imperativním i substantivním. Tvary *sním, sníš, sní, sníme, sníte* jsou homonymní se svými protějšky v paradigmatu slovesného lexému *snít*, tvar *snědí* je jak tvarem 3. os. pl. slovesa *sníst*, tak Npl i Vpl adjektiva *snědý*.⁹

Je typické, že jeden homonymní tvar náleží paradigmatům lexémů dvou slovních druhů, a naopak výjimečné, náleží-li paradigmatům lexémů tří, popř. více slovních druhů. Jedním z nich je například tvar *vesel*: Gpl substantivního lexému *veslo*, jmenný tvar adjektiva *veselý* v maskulinu sg. i imperativ sg. slovesného lexému *veselit se*.

5. DRUHY HOMOFONIE A HOMOGRAFIE

Vedle homonym vyznačujících se totožností formy v mluveném i psaném jazyce jsou v češtině také homofona a homografa (viz definici 3b). **Homofona** představují problém pro analýzu mluveného jazyka a lze je rozdělit na tyto hlavní typy:

- (a) progresivní a regresivní asimilace různých druhů: např. *pře[s]káče*¹⁰ (tento tvar se vlivem regresivní asimilace znělosti *zk* → [*sk*] čte jednoznačně jako *pře[s]-káče*, píše se však různě: jako *přeskáče* (3. os. sg. přítomnosti slovesného lexému *přeskákat*), nebo jako *přezkáče* (Gsg, Npl, Apl a Vpl substantivního lexému *přezkác*). Dalšími příklady jsou třeba *jich / jih, světci / svědci, schod / shod*;
- (b) neutralizace znělosti různého druhu na konci slov: *pokut / pokud, stech / steh, střet / střed, touš / touž, vás / váz*;
- (c) opozice *ě / je* po některých obojetných konsonantech: *obětí / objetí, vět / vjet*;
- (d) opozice *mě / mně*;
- (e) opozice *i / y, í / ý: mil / myl, bili / bily / byli / byly, vír / výr, vískat / výskat*;
- (f) opozice *ú / ů: kúra / kůra*;
- (g) gemináty (*nn, šš* a další) vyslovované jako jedna hláska: *cenou / cennou, trůní / trůnní, tiší / tišší*;

8 Tvar *jed* a *jed'* je navíc homofonní s tvarem *jet*.

9 Typ homonymie, který vykazuje tvar *snědí*, je zastoupen jen tímto jediným tvarem.

10 Příklad K. Olivy.

- (h) jednou souhláskou vyslovované konsonantické shluky: *praští / pražští, řidší / řiči*;
- (i) nevyslovované *j* před *s* v některých slovech: *jsem / sem, jsi / si*;
- (j) různá cizí slova, jejichž výslovnost je shodná s nějakým českým slovem nebo s jiným cizím: *gól / Gaille*.

Dále mohou existovat kombinace různých typů v témže tvaru: *vížka / výška, věd / vět / vjet*.

Homograf je v češtině velmi málo a dají se přiřadit ke třem hlavním typům:

- (i) slova cizího původu, která se vyslovují jinak než jejich české, stejně psané protějšky: jsou to slova obsahující slabiku *di, ti, nebo ni*. Příkladem **homografa** tohoto typu je slovní tvar *unikáte* (Vsg substantivního lexému *un[y]kát* a zároveň 2. os. pl. přítentu slovesného lexému *u[ň]ikat*). Podobně tvar *panice* (Gsg, Asg a Apl substantivního lexému *panic* a zároveň Dsg a Lsg substantivního lexému *panika*), či *postila* (pasivní přičestí feminina sg. či neutra pl. slovesa *postit se* a zároveň Nsg substantivního lexému *postila*, čteme *post[y]la*);
- (ii) slova obsahující *s* mezi vokály a čtená jako [s] nebo [z]: *nasála* (Nsg substantivního lexému *nasála* (nosovka), vyslovujeme *na[z]ála*, a zároveň tvar l-ového přičestí feminina sg. / neutra pl. slovesného lexému *nasát*, vyslovujeme *na[s]ála*);
- (iii) cizí slova stejně psaná, ale vyslovovaná s odlišnou kvantitou: *balkon: balk[o]n / balk[ó]n*;
- (iv) odlišná morfeematická segmentace slova, srov. *proudit* ([*proud-it*] / [*pro-udit*]), *neurologický* ([*neu-rologický*] / [*ne-urologický*]).

6. ZÁVĚR

V příspěvku jsme se zaměřili zejména na klasifikaci typů morfologické homonymie a dále pak homofonie a homografie v současné češtině. Klasifikace jednotlivých druhů morfologické homonymie a také homografie je vedle předpokládaného přínosu k teoretickému popisu češtiny i východiskem pro zkvalitnění automatické morfologické disambiguace českých textů v korpusech řady SYN Českého národního korpusu. Třídy homofonie se zase dají využít k řešení problémů disambiguace mluveného jazyka.

POUŽITÉ KORPUSY

KŘEN, M. — ČERMÁK, F. — HLAVÁČOVÁ, J. — HNÁTKOVÁ, M. — JELÍNEK, T. — KOCEK, J. — KOPŘIVOVÁ, M. — NOVOTNÁ, R. — PETKEVIČ, V. — PROCHÁZKA, P. — SCHMIEDTOVÁ, V. — SKOUMALOVÁ, H. — ŠULC, M. (2014): *Korpus SYN, verze 3 z 27. 1. 2014*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z WWW: <<http://www.korpus.cz>>

KŘEN, M. — HNÁTKOVÁ, M. — JELÍNEK, T. — PETKEVIČ, V. — PROCHÁZKA, P. — SKOUMALOVÁ, H. (2013): *SYN2013PUB: korpus psané publicistiky*. Praha: Ústav Českého národního korpusu FF UK. Dostupný z: <<http://www.korpus.cz>>

LITERATURA

- BRILL E. (1992): A Simple Rule-Based Part-of-Speech Tagger. In: *Proceedings of the Third Conference on Applied Natural Language Processing*. Trento.
- ČERMÁK, F. (2010): *Lexikon a sémantika*. Praha: Nakladatelství Lidové noviny.
- DEANE, P. (1988): Polysemy and Cognition. *Lingua*, 75.
- DEROSE, S. J. (1988): Grammatical category disambiguation by statistical optimization. *Computational Linguistics*, 14, s. 31–39.
- FILÍPEK, J. (1970): K otázce invariantu a variant v lexikální sémantice. *Slavica Slovaca*, 5, s. 272–280.
- FILÍPEK, J. — ČERMÁK, F. (1985): *Česká lexikologie*. Praha: Academia.
- GARSDALE, R. — LEECH, G. — MCENERY, A. (1997): *Corpus Annotation. Linguistic Information from Computer Text Corpora*. London — New York: Longman.
- HAIJČ, J. (2004): *Disambiguation of Rich Inflection (Computational Morphology of Czech)*. Praha: Univerzita Karlova v Praze.
- HORÁLEK, J. (2005): Přejchod homonymie v polysémii a pragmatický význam. In: P. NEJEDLÝ — M. VAJDLOVÁ — B. LEHEČKA (eds.), *Verba et historia: Igoru Němcovi k 80. narozeninám*. Praha: Ústav pro jazyk český AV ČR, s. 137–140.
- HORALÍK, L. (1967): O univerbizaci a homonymii z hlediska principu ekonomie. *Československá rusistika*, 12, s. 140–143.
- HORALÍK, L. (1968): *K problematice homonymie*. Olomouc.
- CHARNIAK, E. (1997): Statistical Techniques for Natural Language Parsing. *AI Magazine*, 18, s. 33–44.
- CHIERCHIA, G. — MCCONNELL-GINET, S. (1993): *Meaning and Grammar: An Introduction to Semantics*. Cambridge, Massachusetts: MIT Press.
- CHURCH, K. W. (1988). A stochastic parts program and noun phrase parser for unrestricted text. In: *ANLC 1988: Proceedings of the second conference on Applied natural language processing*. Association for Computational Linguistics Stroudsburg, PA.
- KARLÍK, P. — NEKULA, M. — PLESKALOVÁ, J. (eds.) (2002): *Encyklopedický slovník češtiny*. Praha: Nakladatelství Lidové noviny.
- LANGACKER, R. W. (1987): *Foundations of Cognitive Grammar*. Vol. I. *Theoretical Prerequisites*. Stanford, California: Stanford University Press.
- LANGACKER, R. W. (1991a): *Foundations of Cognitive Grammar*. Vol. II. *Descriptive Application*. Stanford, California: Stanford University Press.
- LANGACKER, R. W. (1991b): *Concept, Image, and Symbol: The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.
- LYONS, J. (1975): *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- PANEVOVÁ, J. (1980): *Formy a funkce ve stavbě české věty*. Praha: Academia.
- PETKEVIČ, V. (2006): Reliable Morphological Disambiguation of Czech: Rule-Based Approach is Necessary. In: M. Šimková (ed.), *Insight into the Slovak and Czech Corpus Linguistics*. Bratislava: Veda, s. 26–44.
- TĚŠITELOVÁ, M. (1966): *O morfologické homonymii v češtině*. Praha: Academia.
- TUGGY, D. (1993): Ambiguity, polysemy, and vagueness. *Cognitive Linguistics*, 4–3.