

Obsah

Články / Articles

O jazyce a stylu korespondence Jiřího Voskovce a Jana Wericha
Language and style of the correspondence between J. Voskovec and J. Werich /3
Jana Hoffmannová, František Štícha

Dva pohledy na vývoj českého poválečného syntaktického myšlení
Two views on the Development of the Czech post-war Syntactic Research /34
Petr Karlík, Jarmila Panevová

Syntaktické konstrukce typu *Včely se hemží na zahradě* – *Zahrada se hemží včelami*
The Syntactic Constructions *Bees are swarming in the garden* – *The garden is swarming with bees* /54
Václava Kettnerová

Nejen obecná čeština. Od *Obyčejných věcí* k *Slezskému románu*
Not Only Common Czech. From *Common Things* to *The Silesian Novel* /72
Petr Mareš

Recenze / Reviews

František Čermák, Vladimír Petkevič, Alexandr Rosen (eds.):
Korpusová lingvistika Praha 2011, 1–3. /84
Kateřina Najbrtová

Klára Osolobě:
Morfologie českého slovesa a tvoření deverbativ jako problém strojové analýzy češtiny /90
Karel Pala

Obsah

Zprávy / News

Korespondenční texty a jejich lingvistické využití /96

Klára Osolsobě

Pokyny pro autory /99

Instructions for authors /101

Objednávka /103

Order form /104

O jazyce a stylu korespondence Jiřího Voskovce a Jana Wericha

Jana Hoffmannová, František Štícha

Ústav pro jazyk český AV ČR, v. v. i.

hoffmannova@ujc.cas.cz, sticha@ujc.cas.cz

Language and style of the correspondence between Jiří Voskovec and Jan Werich

ABSTRACT: The aim of this article is to present a relatively complex linguistic analysis of the three-volume correspondence between Jiří Voskovec and Jan Werich (V+W). Focusing on the dialogic structure of these letters, its remarkable coherence and stylistic uniformity, we examine the research possibilities such text types may offer in the field of dialogue stylistics. This highly private correspondence draws attention to the transitions between written and spoken language, fluctuations between using literary and common Czech, as well as a frequent use of vulgarisms, which we interpret as a natural part of the authors' shared idiolect. Another building block in the stylistic uniformity of the entire correspondence that is worth highlighting is the linguistic playfulness and language-based humor, including V+W's fondness for using, manipulating, and commenting on specific set phrases, in which they also intertwine Czech and English. Although English holds a prominent position both in terms of frequency of use and the variety of expressions, the letters actually display features of multi-lingual texts; this quality comes out also in the form of stating the authors' opinions about Czech (or other languages). The multi-lingual flavor is not far from issues of intertextuality, which manifests itself in the form of a variety of quotations, paraphrases, references, or ad-hoc created fake proverbs. Another part of our analysis concerns V+W's creativity in word-formation (especially on the part of Voskovec), as well as various grammatical phenomena. In morphology, the texts show a conspicuously high frequency of participles and the archaic short-form adjectives; in syntax, we observe the stylistic function of certain passive constructions, causativization of non-causative verbs, special kinds of ellipsis, and idiosyncratic patterning in marking information structure.

KEY WORDS: correspondence; dialogue stylistics; linguistic playfulness; intertextuality; word-formation; participles; short-form adjectives; passive constructions

KLÍČOVÁ SLOVA: korespondence; stylistika dialogu; hra s jazykem; intertextualita; tvoření slov; přechodníky; jmenné tvary adjektiv; pasivní konstrukce

Dva pohledy na vývoj českého poválečného syntaktického myšlení

Petr Karlík

Filozofická fakulta Masarykovy Univerzity, Brno

pkarlik@phil.muni.cz

Jarmila Panevová

Matematicko-fyzikální fakulta Univerzity Karlovy, Praha

panevova@ufal.mff.cuni.cz

Two views on the Development of the Czech post-war Syntactic Research

ABSTRACT: The authors present their respective views on the development of the Czech post-war syntactic studies. Their approach is influenced by the fact that they were educated by the different syntactic schools: thus the paper is a combination of Prague's and Brno's views. V. Šmilauer's *Novočeská skladba* (*Syntax of Modern Czech, 1947*) is understood as a source of the contemporary research of the Czech syntax. The paper describes the results reached by individual investigators as well as the results of the research teams. According to the authors' opinion, Two-Level Valency Syntax (represented by F. Daneš and his close collaborators and reflected in the Czech Academic Grammar) and Functional Generative Grammar (developed by P. Sgall and his colleagues) form the main paradigms of the Czech syntax since 1960. Both theories incorporate the results of the classical Praguian functional approach as well as results of the generative paradigm. The authors conclude that the Prague's and Brno's views on the development of Czech syntactic studies are not incompatible but rather complementary and that the methods of formal and corpus linguistics are attractive and useful for the young researchers.

KEY WORDS: syntax, linguistic theory, structuralism, generative grammar, two-level valency syntax, functional generative description, grammar, lexicon

KLÍČOVÁ SLOVA: syntax, teorie jazyka, strukturalismus, generativní gramatika, dvourovňová valenční syntax, funkční generativní popis, gramatika, slovník.

Syntax, linguistic theory, structuralism, generative grammar, two-level valency syntax, functional generative description, grammar, lexicon

Syntaktické konstrukce typu *Včely se hemží na zahradě – Zahrada se hemží včelami*¹

Václava Kettnerová

Matematicko-fyzikální fakulta Univerzity Karlovy, Praha

kettnerova@ufal.mff.cuni.cz

The Syntactic Constructions *Bees are swarming in the garden – The garden is swarming with bees*

ABSTRACT: In this paper, we focus on the relation between Czech syntactic constructions *Bees are swarming in the garden – The garden is swarming with bees*. We refer to this relation as a lexical-semantic conversion. This lexical-semantic conversion is associated with changes in the correspondence between two situational participants – Bearer of action and Location – and valency complementations. These changes result in a permutation of these participants which affects the prominent surface syntactic position of subject. We demonstrate that whereas the first syntactic variant refers to a simple action, the latter expresses rather a state arising from this simple action. In the latter case, the participant Location is usually holistically interpreted. We assume that the holistic effect of this participant follows from the fact that a state is generally attributed to each participant expressed in the subject position as a whole. Furthermore, syntactic variants with formal subject *to* and subjectless constructions (usually conceived as formal paraphrases of the pairs of the observed constructions) are discussed. Finally, a lexicographic representation of these syntactic structures is presented.

KEY WORDS: lexical-semantic conversion, process and processual state, holistic effect, lexicographic representation

KLÍČOVÁ SLOVA: lexikálně-sémantická konverze, proces a procesuálně pojatý stav, holistický efekt, lexikografická reprezentace

¹ Za pečlivé přečtení a cenné připomínky děkuji především prof. Jarmile Panevové a doc. Markétě Lopatkové.

Nejen obecná čeština

Od *Obyčejných věcí* k *Slezskému románu*

Petr Mareš

Filozofická fakulta Univerzity Karlovy, Praha

petr.mares@ff.cuni.cz.

Not Only Common Czech. From *Common Things* to *The Silesian Novel*

ABSTRACT: The non-standard language has become, in the course of the twentieth century, an important and legitimate component of Czech literary texts, but the most widely used non-standard variety, the so-called common Czech, prevails. The Moravian (or Silesian) varieties appear much less. This article analyses the language of two prosaic works which operate to a significant extent with Silesian dialects, *Obyčejné věci* (Common Things) by Jan Vrak (1998) and *Slezský román* (The Silesian Novel) by Petr Čichoň (2011). The use of dialects is, in both pieces of prose, supported by the localization of some parts of the storyline to the territory of Silesia. The intricately composed experimental prose by Jan Vrak confronts the typically Silesian elements with the elements of other varieties and presents the emphasized language diversity as a part of the main character's search for his personal and family identity. *Slezský román* works with language in an easier way and highlights primarily the language specificity of the Hlučínsko region, affected also by the influence of German.

KEY WORDS: non-standard language, language varieties, Common Czech, Moravian dialects, Silesian dialects, literary text

KLÍČOVÁ SLOVA: nespisovný jazyk, jazykové variety, obecná čeština, moravské dialekty, slezské dialekty, literární text

Recenze / Reviews

F. Čermák, V. Petkevič, A. Rosen (eds.): Korpusová lingvistika Praha 2011, 1–3.

František Čermák (ed.): Korpusová lingvistika Praha 2011 – 1 Inter-Corp. Praha: Nakladatelství Lidové noviny, 2011. 372 s.

František Čermák (ed.): Korpusová lingvistika Praha 2011 – 2 Výzkum a výstavba korpusů. Praha: Nakladatelství Lidové noviny, 2011. 323 s.

Vladimír Petkevič – Alexandr Rosen (eds.): Korpusová lingvistika Praha 2011 – 3 Gramatika a značkování korpusů. Praha: Nakladatelství Lidové noviny, 2011. 225 s.

Kateřina Najbrtová

Filozofická fakulta Masarykovy univerzity, Brno

145034@mail.muni.cz

Výsledkem konference KorPLing 2011, pořádané Ústavem Českého národního korpusu FF UK v září téhož roku, je rozšíření edice Studie z korpusové lingvistiky o další tři svazky, jež představují sborník konferenčních příspěvků.

Ve svazcích je obsaženo přes šest desítek česky či anglicky psaných článků vycházejících z příspěvků přednesených v průběhu konference lingvisty nejen z celé České republiky, ale také Itálie, Maďarska, Německa, Polska, Rakouska, Slovenska, Spojených států a Španělska.

První část sborníku nese název *InterCorp* a je věnována srovnávacím pohledům skrze více než 15 odlišných jazyků na jejich různé vrstvy – slootovorbou, lexikografií, frazeologií, gramatiku.

Publikace *Výzkum a výstavba korpusů* obsahuje kromě tří plenárních přednášek články zabývající se nejen výstavbou korpusů, ale i možnostmi jejich vyěžování v nejrůznějších oblastech – fonologii, morfologii, aplikaci ve výuce atd.

Posledním svazkem je *Gramatika a značkování korpusů*. Je zaměřen na výzkum gramatiky (převažují synchronní korpusy) a korpusový výzkum syntaxe, druhá půle svazku je zasvěcena spíše technicky orientovaným článkům o aspektech značkování korpusů.

Jen stěží lze v omezeném rozsahu recenze věnovat patřičnou dávku pozornosti všem příspěvkům demonstrujícími rozsáhlé a takřka neomezené využívání korpusů jak synchronních, tak diachronních v nejrůznějších relacích vůči jazyku (jazykům).

Články svazku *InterCorp* lze dělit do více kategorií. Někteří z autorů rozšiřují výsledky svého bádání publikované u příležitosti mezinárodní konference *InterCorp* v roce 2009, popř. na ně navazují.

A. Klégr, M. Malá a P. Šaldová se ve svém příspěvku věnují anglickým předložkovým ekvivalentům českých předložkových frází s předložkami *v/ve* a *s/se*. V této studii, představující součást projektu, jenž zkoumá anglické překladové ekvivalenty nejméně frekventovaných českých předložek, autoři četné příklady podrobně kategorizují dle způsobu, kterým jsou české předložkové fráze realizovány. V důkladném popisu této problematiky spatřují podmínku pro tvorbu kvalitních lexikografických hesel.

Možnosti využití *InterCorpu* při vytváření slovníku studuje i K. Jirásek, tentokrát ve vztahu českého a chorvatského jazyka a užití sémanticky správných ekvivalentů výrazů obou zmíněných jazyků. K. Jirásek upozorňuje na to, že strojově zpracovaný materiál paralelního česko-chorvatského korpusu není dosud postačujícím zdrojem pro tvorbu velkého slovníku (což ostatně zdůrazňuje i ve svých předchozích studiích založených na paralelním korpusu). Jak poukazuje, i v případě dobudování dostatečně rozsáhlého korpusu je v korpusové lexikografii nezbytnou nutností počítat s prací zkušeného lexikografa, schopného překlenout případné mezery způsobené nedostatkem rozsahu i obsahu korpusových dat.

Otázkami víceméně překladatelskými či lexikografickými se dále zabývají např. M. Vachková, M. Šemelík a V. Marková (vybraná německá adjektiva a jejich české protějšky ve vztahu k syntakticko-sémantickému kontinuu) či M. Nábělková (slovenské podvojně a trojně předložky a jejich české ekvivalenty). Dlouhodobě se otázkou správnosti překladů mezi češtinou a ruštinou zabývá L. Stěpanova – ve svém příspěvku o problematice překladu ruských neologismů navazuje na svou dřívější studii o typologii chyb při překladech frazémů.

Možnou podobnost funkce českých demonstrativních zájmen *ten, ta, to* a anglického určitého členu zkoumá M. Souček. Dle našeho očekávání demonstrativa neplní funkce anglického členu určitého – a musíme si nutně položit otázku, zda lze učinit relevantní závěr, aniž by autor sledoval mluvený jazyk.

To, že paralelní korpus poskytuje i možnost studia slovtvorné oblasti jazyka, popř. ověření sémantické správnosti překladů, dokazují mj. příspěvky P. Štichauera a P. Čermáka nebo P. Štichauera a L. Bratánkové.

T. Káňa navazuje ve svém příspěvku na výsledky výzkumu zaměřeného na srovnání německých a anglických překladových ekvivalentů českých deminutiv.

Deminutivní význam může být v závislosti na jazyku vyjádřen různými soubory prostředků, tedy nikoli nezbytně deminutivem samotným. Tímto zjištěním rozšiřuje poznatky, k nimž dospěl v předchozím výzkumu: analytické jazyky tendují k vyjadřování deminutivního významu analytickým způsobem, zatímco pro češtinu jsou typická syntetická deminutiva. Autor ukazuje, že ne všechna slova s deminutivními sufisy představují opravdová sémantická deminutiva. Svou pozornost obrací k rozdílu mezi sémantickými a formálními deminutivy z hlediska překladatelské praxe.

Zda je výhodné hledat v paralelních korpusech odpovědi na překladatelské problémy, způsobené existencí rozdílných morfologických a gramatických kategorií různých jazyků, řeší ve svých příspěvcích např. J. Jindrová, V. Mazaná či L. Chlumská.

Tematicky podobnou studii předkládá P. Nádeníček. Přechylování v češtině srovnává s angličtinou, němčinou a ruštinou, docházejí k zajímavému a podnětnému výsledku: míra přechylování ve zmíněných jazycích nesouvisí s jejich genetickou ne/příbuzností, roli hraje (dnes stále akcentovanější) politická korektnost.

Využití možnosti přímého srovnání mezi dvěma či více jazyky pro bádání v oblasti syntaxe demonstrují články H. Pelouškové (popisuje struktury s korelátém *es* v němčině a víceméně neúspěšně se pokouší nalézt české paralely) a K. Mojžíšové (ověřuje schopnost vytýkácích konstrukcí plnit v češtině obdobné funkce jako v angličtině).

Svazek *Výzkum a výstavba korpusů* je členěn do dvou tematických celků: první se zabývá aspekty výstavby korpusů, druhý pak prezentuje různorodé výsledky výzkumu korpusových dat. Oběma celkům předcházejí tři plenární přednášky.

F. Čermák (*Korpusy včera, dnes a zítra*) shrnuje historii vývoje korpusů a korpusové lingvistiky, podává zprávu o jejich současné situaci a zároveň nabízí výhled do budoucnosti. Od dřevních dob prvních korpusů ze 60. let 20. stol. se dostává až k dnešnímu stavu, k otázkám spolehlivé anotace, jež jediná zajistí relevantní výstupy, a k problémům, které je třeba vyřešit pro „zítra“: vztahu korpusu a jazykové reality (důraz na mluvený jazyk), rozpoznání metaforičnosti jazyka (potencialita úzu jazykových jednotek) a získávání zdrojových textů.

K. Pala a P. Rychlý (*Do We Need Very Large Web Corpora?*) objasňují potřebu velkých korpusů získávaných z webu a popisují techniky a nástroje užívané pro budování takto rozsáhlých korpusů. Takovéto korpusy hrají nezastupitelnou roli: jsou referenční a umožňují vyhledávání pomocí různých typů anotací, čímž se liší od klasických internetových vyhledávačů, které uvedené výhody nemají, byť zahrnují větší rozsah dat. Zároveň autoři neopomíjejí ani možnost využití totožného postupu pro tvorbu menších, avšak specializovaných korpusů

(např. terminologických, a to především v dynamicky se rozvíjejících oborech, v nichž terminologie není ještě zcela ustálená).

Přiblížit langue, jehož plný rozsah je ukryt v mozcích uživatelů jazyka, se pokouší F. Štícha (*O skrytosti langue v korpusovém parole aneb k definici kvantové lingvistiky*), přičemž využívá dvou příkladů slovosledných a dvou slovtvorných. Nelze než souhlasit s upozorněním autora, že teprve prací s kvanty studovaných jevů získaných z velkého množství korpusových dat za pomoci korpusového manažeru (tedy z parole) lze poodhalit skutečný rozsah langue, které může kromě entit lexikonu (v autorově terminologii materiální a strukturní entity) obsahovat i entity jiného charakteru.

O budování korpusu rozhlasových nahrávek informuje V. Štěpánová s tím, že tento korpus bude primárně napomáhat výzkumu vybraných výslovnostních jevů.

J. Hoffmannová a M. Mikulová věnují pozornost důležitému aspektu dnes stále populárnějších korpusů mluveného jazyka: syntaktické anotaci. Jakkoli je tento proces pro korpusy psaného jazyka víceméně zvládnut (viz PDT), při anotování mluvených korpusů je v tradiční podobě de facto nevyužitelný pro získání spolehlivých výstupů. Řešením je tzv. standardizace mluveného jazyka, tj. jeho zjednoduštění, odstranění rysů typických pro mluvený jazyk (opakování, elipsy, redundance, vycpávková slova apod.) a vytvoření korektních vstupů pro anotaci. Tento novátorský a unikátní postup umožní, aby standardizovaný jazyk byl posléze anotován klasickým postupem typickým pro jazyk psaný a zároveň zachovával významy původně zaznamenaných segmentů.

Průkopnický postup anotování korpusů na syntaktické rovině bez účasti vyškolených pracovníků představují D. Hlaváčková, M. Grác a P. Sedlářová. I přes zdánlivě nedostatečnou kvalifikaci anotátorů (studenti prvních ročníků VŠ, obor český jazyk) vykazuje proces vysokou míru spolehlivosti a otevírá cestu relativně nízkonákladové anotaci.

Oblasti lexikální pragmatiky je věnován objevný příspěvek F. Esvana, jenž na základě korpusového výzkumu dramaticky rozšiřuje možné významy sloves *prodávat* a *kupovat*. Správnost významového rozčlenění podkládá odlišným fungováním vidových opozic pro ten který význam.

Využití korpusových dat při výuce češtiny jako cizího jazyka zkoumají K. Schoene (frekvence jednotlivých pádů s ohledem na pořadí jejich výuky) a P. Vališová. P. Vališová představuje našemu prostředí dosud neznámou inovativní metodu Data Driven Learning spočívající v přímém využití korpusových dat při výuce. Autorka navrhuje postupy, jak vytvářet na základě korpusu vhodná cvičení, a zároveň evaluuje jejich úspěšnost – metoda je sama o sobě v případě flektivní češtiny využitelná jen pro velmi pokročilé studenty (studenti

pracují s autentickým jazykem). V neposlední řadě podává autorka návrhy na možné úpravy formátu cvičení.

Morfologická rovina je tématem příspěvků M. Vojtekové a D. Kovářkové. M. Vojteková zkoumá variantnost koncovek neživotných maskulin v češtině a slovenštině a dochází – dle našeho mínění vcelku předvídatelně – k názoru, že zatímco distribuce některých koncovek se shoduje díky příbuznosti obou jazyků, jiné se v důsledku rozdílného vývoje liší. D. Kovářková předkládá na základě korpusového výzkumu argumenty pro interpretaci negace jako gramatické kategorie.

Webovou aplikaci SyD sloužící pro studium variability v korpusech češtiny představují V. Cvrček a P. Vondříčka. Aplikace umožňuje v uživatelsky příjemném prostředí sledovat poměry zadaných variantních slov v průběhu doby či frekvenci a stylistické rozložení těchto variant.

Tabuizovanou oblast nadávek zkoumá V. Schmiedtová v korpusech Karla Čapka a Bohumila Hrabala a srovnává je s vulgarismy obsaženými v mluveném korpusu ORAL2008. Její zjištění průkazně a jednoznačně potvrzují nelichotivý předpoklad: užití nadávek v průběhu doby narůstá, avšak jejich emocionální síla slábne.

Třetí a poslední díl *Gramatika a značkování korpusů* je taktéž rozdělen do dvou tematických částí: úvodní je věnována výzkumu gramatiky v korpusech a v podstatě tak navazuje na předcházející svazek, následující část se zaměřuje na technické aspekty budování korpusů.

Korpusy jako vhodné zdroje dat pro studium slovtvorby představuje K. Osolobě na materiálu deverbativních substantiv zakončených řetězcem [áí]rna. K výzkumu výhodně využívá i webového rozhraní Deriv, umožňujícího získávat podklady ze strojového slovníku češtiny.

Slovtvorbou se zabývá i I. Kolářová, která obrátila pozornost k derivaci sloves od substantivních názvů osob.

L. Válková podává detailní popis doposud dosti opomíjeného jevu, a to distribuce příklonného -s. Výchozím bodem zkoumání jsou korpusy ORAL2006 a ORAL2008, na jejichž základě se jí daří formulovat pravidla ovlivňující výskyt příklonného -s ve spontánně mluvených projevech. Vhodným doplněním tohoto popisu by mohlo být srovnání s Korpusem soukromé korespondence – dialogický charakter textů KSK je důvodem relativně vysokého výskytu tvarů druhé osoby, a tudíž i tvarů s volným morfémem -s.

Předmětem zájmu několika příspěvků jsou různé aspekty syntaxe, např. členská negace a její vyjadřování (K. Veselovská), slovosled konstrukcí s obligatorními participanty (K. Rysová) či pozice infinitivu v podmínkových větách (K. Mílová).

V. Baisa zkoumá ve slovníku valenčních rámců VerbaLex možnost nahrazení příkladů vytvořených lexikografy za reálná data z korpusů. Typizuje obtíže, jež mohou při extrakci dat nastat, a faktory ovlivňující ne/kvalitu výsledků.

Oddíl technických příspěvků orientovaných na syntaktickou anotaci otevírá T. Jelínek analýzou chyb statistického závislostního parseru. Zároveň představuje algoritmus pro automatické odstranění chyb a v závěru předkládá přesvědčivé příklady využití syntakticky označovaného korpusu.

Týž autor, tentokrát s V. Petkevičem, popisuje základní rysy systému anotace psaných korpusů, přičemž akcentuje jeho klíčové fáze: lemmatizaci a morfologickou desambiguaci. Dlouhodobou snahou je utvořit spolehlivý kognitivně plausibilní systém. Ideální způsob morfologické desambiguace češtiny představuje hybridní model, tj. postup založený na kombinaci statistické (stochastické) metody s metodou pravidel (buď automaticky vyvozených z textů, či formulovaných lingvisty).

H. Skoumalová navazuje na tato zjištění a srovnává úspěšnost tagování korpusů odlišnými metodami měřením úspěšnosti těchto metod při tagování testovacího korpusu. Získané výsledky přehledně evaluuje a navrhuje opatření, jež by mohla zlepšit úspěšnost taggerů.

Popis problémů a nesnází při anotaci žákovských korpusů (tj. korpusů textů utvořených nerodilými mluvčími), anotačního schématu a jeho chyb z pera S. Škodové, B. Štindlové, J. Hany a A. Rosena tuto tematickou oblast uzavírá.

Všichni autoři článků dokazují, že korpusy a strojově zpracovaný jazyk obecně umožňují široké spektrum pohledů na různé jazykové aspekty. Badatelé ve svých závěrech prezentují mnohdy překvapivá fakta o fungování jazyka, která se nepochybně mohou stát velkou inspirací pro další lingvistický výzkum. Zamyslíme-li se hlouběji nad tematickou pestrostí příspěvků a otázkami, jež přímo či nepřímo kladou, je zřejmé, že využívání korpusů pro studium jazyka ještě zdaleka neodhalilo všechny své možnosti.

LITERATURA A DALŠÍ ZDROJE

ČERMÁK, F. – KOCEK, J. eds. (2010): *Mnohojazyčný korpus InterCorp: Možnosti studia*. Praha: Nakladatelství Lidové noviny.

ČERMÁK, F. – CORNESS, P. – KLÉGR, A. eds. (2010): *InterCorp: Exploring a Multilingual Corpus*. Praha: Nakladatelství Lidové noviny.

Pražský závislostní korpus 2.0 [online]. 2006 [cit. 22. března 2012]. Dostupné z: <<http://ufal.mff.cuni.cz/pdt2.0/index-cz.html>>.

Klára Osolobě: Morfologie českého slovesa a tvoření deverbativ jako problém strojové analýzy češtiny

Osolobě, Klára: Morfologie českého slovesa a tvoření deverbativ jako problém strojové analýzy češtiny. Opera Universitatis Masarykianae Brunensis FF, č. 401, Brno: Muni Press 2011, 220 s.

Karel Pala

Fakulta informatiky Masarykovy univerzity, Brno

pala@fi.muni.cz

Kniha seznamuje čtenáře s výsledky, k nimž autorka dospěla v oblasti propojení české formální morfologie a tvoření slov. Jádrem práce je formální popis vybrané oblasti české slovtvorby, konkrétně deverbativních substantiv tvořených od slovesného tvaru, od slovesného kmene a od slovesného základu (kořene). Popis má podobu formálních pravidel a může posloužit k testování pravidelných slovtvorných procesů, a to jak na úrovni elektronických morfologických databází, tak i na úrovni elektronicky přístupných korpusů.

Úvodní kapitola obsahuje informace o automatické analýze přirozeného jazyka realizované v poslední době na ÚJČ FF MU ve spolupráci s Centrem zpracování přirozeného jazyka Fakulty informatiky Masarykovy univerzity (CZPJ FI MU). Východiskem tu je dlouholetý výzkum v oblasti formálního popisu české morfologie, který započal na půdě Kabinetu počítačové lingvistiky Ústavu českého jazyka FF MU (dříve UJEP) koncem 80. let minulého století. Výsledky jedné jeho etapy jsou shrnuty v autorčině disertační práci *Algoritmický popis české formální morfologie a strojový slovník češtiny* (Osolobě, 1996). Ten se posléze stal lingvistickou bází některých aplikací v oblasti počítačového zpracování při-

rozeného jazyka, z nichž nejvýznamnější je nepochybně morfologický analyzátor *ajka* (Sedláček, 2004), používaný mj. k anotacím korpusů budovaných na FI MU i FF MU. Jeho novou verzí pak je analyzátor *majka*, jehož autorem je pracovník Centra ZPJ FI MU Pavel Šmerk (Šmerk, 2010).

Ve druhé kapitole najde čtenář informace o automatické morfologické analýze a počítačovém slovníku českých kmenů jako východiscích pro zkoumání vztahů mezi formální morfologií a tvořením slov. Počítačový slovník českých kmenů čítá v současnosti cca 400 000 položek a představuje tak dostatečně úplnou a spolehlivou datovou základnu pro formální popis českého tvoření slov. Každému kmenu ve slovníku kmenů je přiřazen morfologický vzor, pomocí něhož se generují uspořádané trojice: základní tvar (lemma) – generovaný tvar (slovní tvar) – slovní druh a morfologická značka (tag) obsahující informace o příslušných gramatických kategoriích.

Jedna verze analyzátoru *ajka* zahrnuje pravidla tvoření tvarů slov podle jednotlivých tvarotvorných vzorů a takto i definice některých pravidelných derivací. Tak například součástí definice tvarotvorných vzorů substantiv pojmenovávajících osoby jsou pravidla odvozování adjektiv na *-ův* (maskulina životná) a na *-in* (feminina označující převážně živé osoby). Součástí definic tvarotvorných vzorů adjektiv je propojení s derivačními vzory definujícími např. derivace: a) tvarů komparativu a superlativu příslušných adverbí, b) adverbí paradigmaticky tvořených od adjektiv, a další.

Zde je třeba zdůraznit komplexnost tvarotvorných a slovotvorných pravidel (vzorů) ve vztahu k definicím vzorů sloves. Na základě pravidel přiřazených jednotlivým kmenům tak lze generovat jak jednoduché tvary určité (tvary indikativu přítomného/futura aktiva a imperativu), tak neurčité (tvary participia *l*-ového, participia pasivního, přechodníků přítomného i minulého a infinitivu). Autorka přesvědčivě ukazuje, že algoritmičtý popis formální morfologie poskytuje dobře navazující a spolehlivá kritéria pro formulaci formálních pravidel tvoření slov.

Ve třetí kapitole autorka mluví o pravidelnosti derivačních procesů z hlediska počítačového zpracování přirozeného jazyka, které dovolují formulování spolehlivých derivačních pravidel, na druhé straně však také ukazuje, jaké jsou meze „slovotvorného stroje“. Věnuje pozornost popisům českého tvoření slov

v českých klasických gramatikách a ukazuje, že jsou pro potřeby algoritmického popisu v mnoha aspektech neúplné, i když dobře posloužily jako výchozí. Proto bylo nejdříve třeba vypracovat metodu formálního popisu pravidelných derivací v češtině. Jeho základem je vzájemný vztah (formy a významu) slova základového a slova odvozeného. Na rovině formy autorka vychází z grafické podoby slova / slovního tvaru. Jak slovo základové, tak slovo odvozené lze chápat jako řetězec grafémů (písmen). Na rovině významu se vychází z obecného významu slovního druhu a gramatických kategorií, které jsou zachyceny v morfologické značce u každého kmene v morfologickém slovníku kmenů.

Navržený popis má formu substitučních pravidel formulujících substituce na úrovni formy za určitých podmínek. Pravidla zachycují slovotvorné procesy jako operace nad řetězci grafických znaků / písmen (kmenů uložených v počítačovém slovníku morfologického analyzátoru *ajka*), jejichž podmínkou jsou definovatelné vlastnosti zadaných řetězců (gramatické informace obsažené v gramatických značkách). Na základě lingvisticky stanovené hypotézy, tedy souboru substitučních pravidel obsahujících definované podmínky, lze z morfologického počítačového slovníku kmenů automaticky extrahovat n-tice jednotek, které se nalézají ve slovníku a splňují danou hypotézu.

V kapitole čtvrté se probírají vzájemné vztahy formální a derivační morfologie poskytující východiska pro automatickou slovotvornou analýzu, která se opírá o pravidelnosti formálních změn, k nimž dochází při derivaci. Slova základová, od nichž potenciálně mohou být derivována slova odvozená, lze definovat na rovině gramatické abstrakce zakódované v morfologických značkách. Tyto vztahy jsou do různé hloubky popisovány v klasických popisech české slovotvorby. Autorka zde nově formuluje formální derivační pravidla, testuje je a doplňuje stávající popisy o poznatky získané při testování. V této souvislosti se dále věnuje segmentaci slovesných tvarů a deverbativ a pochopitelně i problematice morfonologických alternací, konkrétně těch, které doprovázejí tvoření slovesných tvarů i některých pravidelných derivací od slovesného kmene (adjektivum slovesné na *-ný/-tý*, substantivum slovesné na *-ní/-tí*). Při segmentaci

slovesných tvarů a interpretaci segmentů se autorka přidrží systému použitého v *Mluvnici češtiny 2* (MČ 2, 1986).

V kapitole páté autorka podrobně probírá morfonologické alternace, k nimž dochází při tvoření slovesných tvarů a některých derivací od slovesných základů. K nim formuluje pravidla, která se pak dále využívají v jednotlivých kapitolách věnovaných formálním pravidlům popisujícím tvoření deverbativ, a to v tabulkových přehledech alternací u příslušného derivačního typu podle jednotlivých typů morfů. Alternace lze rozčlenit na kombinační (syntagmatické), v nichž jsou hláskové proměny vázány na hláskové okolí základu, a volné (paradigmatické), které na hláskovém okolí základu nezávisí a představují mnohonásobně početnější skupinu. Popis a rozbor alternací v této kapitole obsahuje nově a komplexně formulované poznatky o jednotlivých typech alternací, které velmi dobře poslouží i v dalším výzkumu.

V šesté kapitole najde čtenář informace o softwarovém nástroji *Deriv* (Šmerk et al., 2009) umožňujícím automatickou extrakci dat z počítačového slovníku českých kmenů, a tedy i vlastní automatickou slovotvornou analýzu.

Nástroj má podobu interaktivního webového rozhraní a disponuje dvěma základními sadami funkcí. První sada obsahuje funkce vyhledávací – umožňují extrakci dat ze slovníku českých kmenů (cca 400 000 jednotek) začleněných nyní do morfologického analyzátoru Majka (Šmerk, 2010). Data jsou vyhledávána a extrahována na základě zadavatelných parametrů. Tyto parametry jsou: a) formálně definovatelné vlastnosti hledaných jednotek, b) interpretace definovaných jednotek zachycené v podobě morfologických značek.

Druhá sada funkcí umožňuje automaticky extrahovaná data prohlížet, třídit, editovat a ukládat na různé úrovni abstrakce. Požadované jednotky lze definovat jednak na základě jejich formálních vlastností (jako řetězce znaků s použitím regulárních výrazů), jednak na základě jejich gramatických vlastností vyjádřených odpovídajícími morfologickými značkami.

V rámci vyhledávacích funkcí lze definovat i nejrůznější vztahy mezi formálními a gramatickými vlastnostmi jednotek. K tomuto účelu slouží funkce

pro vyhledávání dvojic (popř. n-tic), jež lze definovat jako substituční pravidla. Jimi lze postihnout vlastnosti dvojic základové slovo – odvozené slovo, případně i vhodných n-tic. Takto automaticky získaná data lze navíc interaktivně sledovat. Automatický nástroj *Deriv* je propojen jednak s internetovým prohlížečem *DebDict* (<chrome://debdict/content/debdict.xul>), který umožňuje zobrazovat hesla odpovídající automaticky generovaným jednotkám v řadě výkladových slovníků (např. SSJČ, SSČ, PSJČ, SCS), jednak s korpusem SYN2000 (ev. i dalšími korpusy), takže se lze jednoduše přepnout do korpusu na vhodný seznam konkordančních řádků.

Výsledky automatické analýzy lze dále testovat pomocí nástroje *Deriv* z hlediska možných chyb v pokrytí pravidel. To lze udělat porovnáním počtu potenciaálních derivátů a počtu správných derivací pokrytých navrženými pravidly a získat tak i míru přegenerování. Nástroj *Deriv* umožňuje pomocí funkce *práce s adresář* navrhnout dostatečně strukturovaný systém podadresářů, což usnadňuje orientaci a systematické ukládání získaných hromadných dat.

Je potřeba ocenit, že tato kapitola může posloužit i jako stručný manuál pro práci s nástrojem *Deriv*. Lze jej najít na adrese <http://deb.fi.muni.cz/deriv>, pro přístup k němu je však potřeba spojit se s autorkou.

Poslední tři kapitoly (VII, VIII. a IX.) jsou jádrem autorčiny práce (i rozsahem – s. 55–200) a představují hlavní výsledky, jichž autorka dosáhla. Nebudeme je probírat v detailech a jen konstatujeme, že v nich čtenář najde:

a) podrobná formální pravidla odvozování deverbativních substantiv tvořených od slovesného tvaru, od slovesného kmene a od slovesného základu (kořene). Jinými slovy, pravidla představují dynamický formální popis derivace deverbativ, jehož fungování je otestováno na hromadných datech zmíněných výše (slovník kmenů analyzátoru *ajka* a korpusy). V rámci jednotlivých typů jsou zkoumány skupiny odvozených slov tvořených od základů jednoho slovního druhu (sloves) jedním derivačním prostředkem (sufixace, konverze, doprovodné morfonologické alternace).

b) podrobnou kvantitativní analýzu automaticky generovaných dat (deverbativ), která poskytuje nový pohled na popisované deverbativní slovo-tvorné procesy.

c) derivační slovník deverbativ analyzovaných typů (pochopitelně v elektronické podobě, v tištěné by byl příliš rozsáhlý). Fyzicky jsou získaná data uložena na serveru Centra ZPJ FI MU na adrese <http://deb.fi.muni.cz/deriv> (v příslušných adresářích). Slovník obsahuje automaticky generovaná i ručně zpracovaná data ve formě dvojic fundující sloveso – odvozené deverbativum (97 951 dvojic). Ke každé dvojici lze získat frekvenci v korpusu SYN2000 a odkaz na odpovídající heslo v různých českých slovnících díky vazbě na korpusový manažer Manatee/Bonito a slovníkový systém DebDict. Případným zájemcům o zmíněná data doporučuji obrátit se přímo na autorku (osolobe@phil.muni.cz), volně přístupná zatím nejsou.

Závěrem bych rád konstatoval, že práci K. Osolobě pokládám za průkopnickou, a to v několika ohledech:

- *metodologicky* – autorka pracuje s novou, v dané oblasti prakticky dosud nepoužitou metodologií: navrhla formalizovaný (a tedy algoritmitizovatelný) popis vybraných slovotvorných procesů v češtině a použila programového nástroje *Deriv* umožňujícího hromadné zpracování výchozích dat. Na rozdíl od technik používaných v klasických českých gramatikách a založených na exemplifikaci zkoumaných jevů doložených nejvýše desítkami příkladů umožňuje nová metodologie díky formálnímu popisu a počítačovému zpracování získané výsledky otestovat a ověřit jejich úplnost i přesnost,

- *dosaženými novými výsledky*, které nebyly dosud pro češtinu k dispozici. Byla použita kvalitní a rozsáhlá výchozí data – počítačový slovník kmenů rozsahem pokrývá prakticky celou češtinu – téměř 400 tis. českých kmenů (pro porovnání SSJČ čítá cca 192 tis. hesel). Z toho slovesných kmenů je cca 42 tis., substantivních – cca 131 tis., adjektivních – cca 170 tis., adverbialních – 41 tis.

- v současnosti jde o *nejúplnější popis* tvoření deverbativ v češtině, celkem bylo detailně zpracováno 29 sufixů, a to 16 substantivních a 13 adjektivních. Je potřeba zdůraznit, že manuální zpracování zmíněných dat by nebylo v rozumném čase prakticky možné. Autorka jasně prokazuje, že nové techniky využívající formálního popisu a softwarových nástrojů umožňují získat kvantitativně i kvalitativně nové výsledky a obecně vedou i k novému pohledu na jazyk jako systém i na jeho fungování.

- autorka věnovala také pozornost *homonymii* způsobené překryvy vycházejícími z různých možností formální segmentace analyzovaných jednotek. Druhým typem homonymie potom je homonymie na úrovni jednotlivých typů morfů (lexikálních i derivačních). Konkrétní příklady homonymie jsou bohatě dokumentovány v poznámkách k jednotlivým substitučním pravidlům. K tomu je potřeba poznamenat, že získaná data naznačují možná řešení. U některých

pravidel dochází díky homonymii k tzv. přegenerování, tj. dostáváme dvojice, jež v jazyce nemají místo. Tyto případy přispívají k poznání slabých míst, která se mnohdy opakují, a jejich analýza vede k optimalizaci formálního popisu a stává se východiskem pro přesnější formulaci pravidel a také pro jejich manuální kontrolu.

LITERATURA

MČ 2: KOMÁREK, M. – KOŘENSKÝ, J. – PETR, J. – VESELKOVÁ, J. (eds.) (1986): *Mluvnice češtiny, 2: Tvarosloví*. Praha: Academia.

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

SEDLÁČEK, R. (2010): *Morphematic analyser for Czech*. Disertační práce. Brno: Fakulta informatiky MU.

ŠMERK, P. et al. (2009): Exploring Derivation Relations in Czech with the Deriv Tool. Slovko Conference, Smolenice/Bratislava

ŠMERK, P. (2010): *K počítačové morfologické analýze češtiny*. Disertační práce. Brno: Fakulta informatiky MU.

Korespondenční texty a jejich lingvistické využití

Klára Osolsobě

Filozofická fakulta Masarykovy univerzity, Brno

osolsobe@phil.muni.cz

24. listopadu 2011 se na Ústavu českého jazyka Filozofické fakulty Masarykovy univerzity v Brně konalo kolokvium na téma Korespondenční texty a jejich lingvistické využití. Uspořádala je (spolu se svými doktorandkami) doc. Zdeňka Hladká jako součást projektu Soukromá korespondence 20. století. Rozšíření korpusových zdrojů, nové možnosti lingvistického využití (GA ČR P 406/10/0823). Setkání se účastnili nejen renomovaní badatelé z domova i zahraničí, ale i začínající lingvisté (studenti a čerství absolventi – magistři).

Kolokvium zahájila Zdeňka Hladká (Brno) příspěvkem na téma *Zpracování soukromé korespondence v Ústavu českého jazyka FF MU*, v němž shrnula dosažené výsledky a další perspektivy budování a využívání specializovaných žánrově vymezených korpusů soukromé korespondence. Obci korpusových lingvistů není třeba připomínat, že doc. Hladká je vedoucí týmu, který připravil Korpus soukromé korespondence (ksk – dopisy) přístupný všem uživatelům Českého národního korpusu (<http://ucnk.ff.cuni.cz/dopisy.php>), ale také širší elektronickou publikaci *Čeština v současné soukromé korespondenci*. Zkušenosti získané při budování bohatě využívaných korpusů i ověřené metodologické postupy jsou nyní zúročovány při přípravě diachronně pojatého korpusu soukromé korespondence, který má zmapovat celé 20. století, ale také při vedení četných studentských prací zaměřených na vytěžování již vybudovaného korpusu soukromé korespondence.

Sociolingvistickými aspekty žánru korespondenčních textů se ve svém příspěvku *Historická sociolingvistika a korespondence* zabýval první zvaný řečník Marek Nekula (Regensburg) a otevřel tak blok příspěvků badatelů zkoumajících různé jazykové aspekty korespondenčních textů dob minulých. Na jazyk dopisů adresovaných B. Němcové se zaměřil jeden z editorů čtyřsvazkového vydání korespondence B. Němcové Robert Adam (Praha).

Čeština v dopisech patriotické inteligence v období národního obrození byla tématem příspěvku *Literalita, gender a spisovnost v sociálně neúplném společenství* Jiřiny van Leeuwen (Jena). Některé jazykové aspekty odeslané i přijaté česky psané korespondence Bedřicha Smetany analyzovala na bázi korpusově zpracovaného materiálu Lucie Rychnovská (Brno). Jazyk a styl černínské korespondence 17. století (Dopisy Humprechta Jana Černína z Chudenic a jeho matky Zuzany Černínové z Harasova) představila Hana Borovská (Brno). Stylistické aspekty soukromé korespondence období humanismu z hlediska vyššího a nižšího stylu rozebírala Anna Zitová (Praha).

Přednáškou druhého zvaného řečníka Jany Hoffmannové (Praha) *Od dopisů významných osobností přes úřední písemnosti k mailům a SMS: ne/formálnost, ne/zdvořilost, ne/psanost (mluvenost?)* byl zahájen blok věnovaný dnes téměř ohroženému druhu soukromého dopisu s ohledem na jeho nejružnější podoby, moderní, ba nejmodernější nevyjímaje. Od sociolingvistiky se úhel pohledu přesunul ke stylistice a pragmatice.

Jiří Zeman (Hradec Králové) svým příspěvkem *Časopisecké dopisy jako specifický způsob komunikace* vyvolal diskusi o autentičnosti textů v rubrikách dopisů čtenářů z hlediska předpokládaných redakčních zásahů.

Referáty podvečerního bloku poukázaly na metodologické problémy práce s korpusovými daty a na otevřené možnosti využití korpusů korespondenčních textů.

Mária Šimková (Bratislava) porovnávala ve svém příspěvku *Manuální a korpusová analýza korespondencie (úvahy na báze korespondencie I. Kadlečička s L. Vaculíkom)* statistické a ruční zpracování korpusových dat. Úpravám nástrojů pro automatickou morfologickou analýzu textů s vyšším počtem nestandardních variant byl věnován příspěvek Dany Hlaváčkové (Brno). Zkušenosti získané jak při úpravě automatických nástrojů primárně určených pro zpracování standardní podoby psaného jazyka, tak při ruční desambiguaci automaticky značkových dat mohou být inspirací a poučením pro řešení analogických problémů při přípravě dalších korpusů.

Hypokoristika z křesťanských jmen v Korpusu soukromé korespondence bylo téma příspěvku Jany Machalové (Brno). Zaměřila se v něm na rozdíly mezi daty získanými dotazníkovou metodou a daty získanými z korpusů. Propria zastoupená v korpusu ksk analyzovala z hlediska lemmatizace a z hlediska slovtvorby.

Kolokvium bylo zakončeno v přátelském duchu ukázkami žertovných a komických pasáží z korespondenčních textů. Prezentaci několika zajímavých dopisů i dalšího obrazového materiálu k životu a dílu Bedřicha Smetany zakončila Lucie Rychnovská (Brno) několika hudebními ukázkami a podtrhla tak širší

kulturně historický rámec zkoumání jazyka, jímž si lidé psali, píší a snad ani v budoucnu psát nepřestanou.

Hlavní přínos kolokvia *Korespondenční texty a jejich lingvistické využití* spatřujeme v tom, že byl dán prostor jak lingvistům, kteří svůj zájem dlouhodobě zaměřují na nejrůznější aspekty jazyka korespondence, tak začínajícím badatelům, pro které je práce jejich starších kolegů a učitelů inspirací a vzorem. Během jediného dne se organizátorům podařilo otevřít prostor pro seznámení s tím, co a jak se již podařilo udělat. Řadu témat (lexikografické využití korpusů korespondenčních textů, budování a způsoby využívání korpusů korespondence, pohled na stylistické, sociolingvistické a pragmalingvistické zvláštnosti jazyka písemných dialogů) zmíněných na brněnském kolokviu lze nadále rozvíjet. Na výsledky práce badatelů, pro které dopisy v koši končit nemají, se můžeme jen těšit.

Pokyny pro autory

Redakce přijímá pouze původní, dosud nepublikované práce v češtině, slovenštině, angličtině nebo němčině. Pokud je anglický nebo německý příspěvek napsán nerodilým mluvčím, redakce autora vyzývá, aby svůj text nechal korigovat rodilým mluvčím-lingvistou.

Redakce žádá autory, aby své příspěvky ve formátu *.doc zaslali elektronicky na adresu **kga@ujc.cas.cz**. Zároveň je třeba na adresu redakce (**Korpus – gramatika – axiologie, Univerzita Hradec Králové, KČJL PdF, Rokitanského 62, 500 03 Hradec Králové**) zaslat jeden anonymizovaný vytištěný exemplář.

Povinnou součástí příspěvku psaného v jakémkoli jazyce je anglický abstrakt v rozsahu 100 až 150 slov a klíčová slova v jazyce příspěvku a angličtině. U anglicky nebo německy psaných příspěvků o češtině žádáme též název, abstrakt a klíčová slova v češtině.

Základní struktura příspěvku:

- název příspěvku;
- jméno a příjmení autora;
- pracoviště, e-mailová adresa;
- anglický abstrakt, klíčová slova;
- text;
- reference

Úprava příspěvku:

- zejm. pokud chcete do textu vkládat tabulky či grafy, doporučujeme nastavit už pro rukopis velikost stránky A5, popř. užít písmo Cambria vel. 9 pro základní text, vel. 8 pro ostatní texty;
- příspěvek by měl být členěn na kapitoly (a event. subkapitoly); ty se označují čísla desetinné soustavy do maximálně čtyř číslic (1; 1.1; 1.1.2; 1.1.2.1); poté, je-li to nutné, lze užívat římských číslic (I, II) nebo písmen (A, a);
- doklady analyzovaných jevů je vhodné průběžně číslovat a číslo z obou stran uzávorkovat: (1); jednotlivé doklady se píší zkraje strany vždy v samostatném odstavci; pokud se jako doklad daného jevu uvádí několik vět (kvůli žánru, stylu, kontextu atp.), není nutno je číslovat, ale píší se v jednom bloku, shora i zdola vyčleněném z textu, jiným typem písma;
- doklady jsou na konci opatřeny citací zdroje; je-li doklad převzat z korpusu, stačí citovat tento korpus; má-li zavedenou zkratku, pak tuto zkratku, např. SYN2006pub;
- případná morfologická anotace příkladových vět by se měla dít v duchu pravidel Lipské konvence (Leipzig Glossing Rules).
- tabulky je třeba průběžně číslovat, opatřit názvem a umístit na stránce vlevo;
- poznámky jsou umístěny pod čarou;
- citace v textu: (Daneš, 1969, s. 86) / ...jak uvádí Daneš (1969, s. 86);
- úpravu referencí provádějte podle následujícího vzoru:

DUŠKOVÁ, L. (1984): Some contrastive notes on interrogative sentences in Czech and in English. *Philologica Pragensia*, 72–82.

BIRJULIN, N. A. – CHRAKOVSKIJ, V. S. (2001): Imperative sentences: theoretical problems. In: V. S. Chrakovskij (ed.), *Typology of Imperative Constructions*. München: Lincom Europa, 3–50.

ZIFONUN, G. – HOFFMANN, L. – STRECKER, B., et al. (1997): *Grammatik der deutschen Sprache*, Bd. 2. Berlin – New York: Walter de Gruyter.

Redakce si vyhrazuje právo příspěvek odmítnout nebo vyžádat jeho úpravu.

Příspěvky procházejí posouzením nejméně dvou nezávislých oponentů, kteří nejsou členy redakce.

Instructions for authors

The journal solicits only original, previously unpublished work. Articles can be written in Czech, Slovak, English, or German. Please be advised that if the article is written in a language that is not the author's native language, authors are asked to have their manuscript carefully checked by a native speaker.

Manuscripts should be submitted both electronically and in one hard copy. Please email the article as an attached file to the address **kg@ujc.cas.cz** and send one hard copy, without showing the author's name, to the postal address of the journal: **Korpus – Gramatika – Axiologie, Univerzita Hradec Králové, KČJL PdF, Rokitanského 62, 500 03 Hradec Králové**. The printed copy should be written in Microsoft Word, font size 12 pts., line spacing 1.5, and with sufficient margins.

Each submission must be accompanied by an abstract of 100–150 words, written in English.

The file with the electronic submission should contain the following:

- submission title;
- author's first and last name;
- affiliation and the author's email address;
- abstract in English;
- text of the article;
- references.

Formatting guidelines

The paper should be divided into **sections** (and, possibly, **subsections**). These are numbered in a decimal system, up to four levels (e.g. 1; 1.1; 1.1.2; 1.1.2.1). If additional subsections are needed, use Roman numerals (I, II) or letters (A; a).

Attested examples should be numbered sequentially, by using numbers enclosed in parentheses: (1). Examples are not indented but each example forms a separate paragraph. If an example consists of multiple sentences (for stylistic, contextual, etc. purposes), they are written in one contiguous sequence, not as separate (and numbered) paragraphs. The examples should be separated from the main text by a single line and typed in a different font, preferably in Verdana, size 9 pts, single spaced.

Each example is accompanied by reference to its source. If the example is taken from an electronic corpus, it is sufficient to name the corpus; if the corpus has a standard abbreviation, use the abbreviation, e.g. ČNK.

Tables must be numbered and each table must be accompanied by a caption. Tables are left-justified.

Footnotes are numbered sequentially and placed at the bottom of the page.

Literature cited in the text: (Daneš, 1969, 86) / ... Daneš (1969, 86)

Bibliographic references are formatted as follows:

DUŠKOVÁ, L. (1984): Some contrastive notes on interrogative sentences in Czech and in English. *Philologica Pragensia*, 72–82.

BIRJULIN, N. A. – CHRAKOVSKIJ, V. S. (2001): Imperative sentences: theoretical problems. In: V. S. Chrakovskij (ed.), *Typology of Imperative Constructions*. München: Lincom Europa, 3–50.

ZIFONUN, G. – HOFFMANN, L. – STRECKER, B., et al. (1997): *Grammatik der deutschen Sprache*, Bd. 2. Berlin – New York: Walter de Gruyter.

The editors reserve the right to reject unsuitable submissions and to request revisions. Each submission will be reviewed anonymously by two reviewers, not by the journal's editors.

Objednávka

Objednávka časopisu Korpus - gramatika - axiologie				
Odběratel	Název instituce:			
	Dodací adresa:			
	Fakturační adresa:			
	IČO / DIČ / NIP / VAT number:			
	IBAN:			
BIC (SWIFT):				
Objednávám	- zasílání kusu/kusů počínaje číslem			
	- jednorázové zaslání čísla v počtu kusu/kusů			
	s úhradou v měně:	Kč	€	USD
Další informace pro dodavatele:				
Dodavatel	Název instituce: Univerzita Hradec Králové, Pedagogická fakulta			
	Adresa: Rokitanského 62, 500 03 Hradec Králové, ČR			
	IČO: 62690094, DIČ: CZ62690094			
	Bankovní spojení: Česká spořitelna, a.s., číslo účtu: 2733582/0800			
	IBAN: CZ65 0800 0000 0000 0273 3582			
BIC (SWIFT): GIBACZPX				
Datum:	Za odběratele:			
<p>Objednávku odešlete na adresu redakce: Korpus – gramatika – axiologie, Univerzita Hradec Králové, Pedagogická fakulta, KČJL, Rokitanského 62, 500 03 Hradec Králové, Česká republika, nebo elektronickou formou na adresu: kga@ujc.cas.cz.</p> <p>Cena bude objednateli fakturována ve zvolené měně: 60 Kč/ks + poštovné a balné (vč. DPH), nebo 11 €/ks včetně poštovného, balného a bankovního poplatku (bez DPH), nebo 15 USD/ks včetně poštovného, balného a bankovního poplatku (bez DPH).</p>				

Order form

ORDER FORM for <i>Korpus – gramatika – axiologie</i>				
Subscriber	Institution:			
	Delivery address:			
	Billing address:			
	IČO / DIČ / NIP / VAT number:			
	IBAN:			
BIC (SWIFT):				
Ordering information	- regular delivery of copy/copies, starting with issue No.			
	- single issue No., number of copies			
	currency of payment:	Kč	€	USD
Additional information for distributor:				
Distributor	Institution: Univerzita Hradec Králové, Pedagogická fakulta			
	Address: Rokitanského 62, 500 03 Hradec Králové, Česká republika			
	IČO: 62690094, DIČ: CZ62690094			
	Bank: Česká spořitelna, a.s.; Account no.: 2733582/0800			
	IBAN: CZ65 0800 0000 0000 0273 3582			
BIC (SWIFT): GIBACZPX				
Date:	Subscriber signature:			
<p>Order form should be mailed to: Korpus – gramatika – axiologie, Univerzita Hradec Králové, Pedagogická fakulta, KČJL, Rokitanského 62, 500 03 Hradec Králové, Česká republika, or electronically to the following address: kga@ujc.cas.cz.</p> <p>The price on the invoice will be quoted in the currency selected by the subscriber: 60 Kč/copy + shipping and handling (incl. DPH), or 11 €/copy, including shipping, handling, and a bank transfer fee (without DPH), or 15 USD/copy, including shipping, handling, and a bank transfer fee (without DPH).</p>				