

- Český národní korpus – ORAL2008. Ústav Českého národního korpusu FF UK, Praha 2008. Dostupný z WWW: <<http://www.korpus.cz>>.
- Český národní korpus – PMK. Ústav Českého národního korpusu FF UK, Praha 2001. Dostupný z WWW: <<http://www.korpus.cz>>.
- Waclawicová M., 2011, Jazykovězeměpisné aspekty mluvených korpusů Českého národního korpusu. In *Sborník z konference Dialektologie a geolingvistika v současné střední Evropě*, Opava. V tisku.
- Waclawicová M., 2008, Regionální mluva a nivelizační tendence (se zaměřením na oblasti pohraničí a Prahu). In *Čeština v mluveném korpusu*, eds. M. Kopřivová, M. Waclawicová. NLN a ÚČNK, Praha, 262-273.
- Waclawicová M., 2009, Regionální mluva v korpusu mluvené češtiny ORAL2006 se zaměřením na situaci v českém pohraničí. *Naše řeč*, 92, č. 2, 72-86.

Příspěvek vznikl za finanční podpory grantu MSM0021620823.

Výukové materiály založené na korpusu

Pavína Vališová

Ústav českého jazyka, Filozofická fakulta, Masarykova univerzita

Abstrakt

The aim of this paper is to describe means of exploiting the Czech National Corpus (CNC) in the teaching of Czech as a foreign language. Two approaches exist: 1) Data Driven Learning (DDL) – students work directly with corpora. CNC can become a complement to a dictionary, because it displays lemma, gender or case of words. 2) Corpus-based teaching materials – the teachers use corpora as a source of authentic contemporary language for creating their own exercises, both in KWIC or traditional format. It has been found out that Czech learners have difficulties with hands-on tasks (formulating the query, understanding large quantities of data). This research focuses on corpus-based exercises in KWIC format which have been used in the classroom. These types of tasks are complicated for the students because of limited context, unfinished sentences and unfamiliar vocabulary. Although KWIC attracts their attention and teaches them to recognize the grammatical/lexical patterns, these tasks were challenging for advanced students only. Working with wordlists of collocations was considered to be the most beneficial activity.

1. Teoretická východiska

V době, kdy se češtinu již neučí pouze zahraniční bohemisté, ale skupiny cizinců s rozličnými mateřskými jazyky i potřebami (od zahraničních manažerů po žadatele o mezinárodní ochranu), je užitečné otevřít obor čeština jako cizí jazyk různým metodickým přístupům. V posledních letech vzniká mnoho nových materiálů založených na komunikační metodě, ne všechny však využívají autentické věty nebo texty. Proto nás zajímalo, jak bychom v této oblasti mohli uplatnit Český národní korpus (ČNK) jako zdroj současného jazykového úzu. Po vzoru výzkumníků zabývajících se výukou angličtiny, jsme aplikovali metodu DDL na češtinu a vytvořili jsme cvičení z korpusových dat jak pro práci přímo s vyhledávačem Bonito, tak i pro klasickou výuku, tj. na papíře. Místo aby vysoká flektivnost češtiny byla překázkou ve využívání autentického jazyka ve výuce, pokusili jsme se z ní udělat přednost, a naše aktivity se zaměřují především na procvičování morfologie (na rozdíl od angličtiny, kde je důraz na lexikální vzory). Budování nové slovní zásoby však šlo ruku v ruce s hlavním tématem aktivit. Přestože se ukázalo, že tyto typy cvičení nejsou pro všechny typy studentů (ne všem vyhovuje induktivní přístup) ani pro všechny pokročilosti (pro nižší pokročilosti je zřejmě nutné věty z korpusu upravit), část studentů práce s konkordancemi zajímala a vnímali ji jako zajímavé zpestření výuky.

■ 1.1. Data-Driven Learning (DDL)

První, kdo se zabýval využitím korpusu v jazykové výuce, byl Tim Johns, který začal vytvářet cvičení podle Bristkého národního korpusu a používal je ve svých lekcích angličtiny pro zahraniční studenty na univerzitě v Birminghamu. Johns (1991) zavádí pojem Data Driven Learning (DDL) pro svůj přístup, jenž nezáleží pouze v použití korpusových dat, ale ve změně procesu učení, stejně jako s sebou přináší odlišné role studentů a učitele. Změna od deduktivního přístupu k induktivnímu znamená, že učitel již není znalec jazyka, ale pomáhá či ukazuje, jak se učít, a student objevuje pravidla jazyka vlastními silami. Tým proces lze sledovat u rodilého mluvčího, který svůj rodný jazyk zná, ale jeho pravidla či vzory si neuvědomuje, musí je objevit. Bernardiniová (2004) nazývá tento přístup Discovery Learning, neboť v centru se vyskytuje student, který je výzkumníkem a objevitelem. Učí se díky cvičením, která zahrnují pozorování jazykových vzorů a jejich následné zobecnění. Učitel již není autorita, která zná správnou odpověď, ale je spíše organizátorem. Johnsův proces učení tak postupuje od jazykových dat směrem k pravidlům a dá se shrnout do tří fází: 1. identifikovat, 2. klasifikovat, 3. generalizovat (Bernardiniová 2004, 16–17). Jedná se tedy o opačný postup než v případě tradiční gramaticko-překladové metody ve smyslu předvedení pravidla a následného procvičování. Tento neautorativní přístup zvyšuje studentovu samostatnost a motivuje ho.

Přestože se v posledních letech mnoho výzkumníků experimentuje s korpusovými daty ve výuce (např. Bernardiniová 2004, Boulton 2009, Braunová 2007, Granathová 2009, Estling Vannestal et al. 2007), každodenní používání korpusu v jazykové výuce je však daleko od reality, což platí i pro angličtinu. Důvodem bývá nejen nedostatečná technická vybavenost škol, ale i malé povědomí učitelů o tom, jak lze korpus ve výuce efektivně využívat. Römerová (2009, 84) vyšla z toho, že nejdříve je nutné ověřit, zda vůbec korpus může být pomocníkem a vyřešit každodenní problémy učitelů. Provedla výzkum mezi učiteli angličtiny na německých středních školách a zjistila, že jen minimum z nich se dívá v případě nejasnosti do korpusu. Nejvíce z nich dává přednost konzultaci s rodilým mluvčím. Ten však není vždy nablízku, proto mnoho z nich také využívá slovníky. Učitelé si však stěžovali, že vyhledávání ve slovnících je pomalé a často tam chybí informace o kolokacích. Korpus tak může nahradit konzultaci s rodilým mluvčím a zároveň i on-line slovník. Podle Römerové (2009, 95) by v tomto směru měli více spolupracovat korpusoví odborníci s učiteli, a to jak již při učitelském studiu, tak v celoživotním vzdělávání.

Na malou rozšířenost korpusu ve výuce má také vliv fakt, že většina experimentů se odehrává v akademickém prostředí (kde je na experimenty dostatek prostoru), aktivity jsou určeny pro vysoce pokročilé studenty a učitelé vytvářejí výukové materiály založené na korpusu jsou zároveň výzkumníci v oblasti korpusové lingvistiky (Boulton 2009, 39–40). Podobné je to s jazyky jinými než angličtina. Velmi málo výzkumů bylo provedeno pro jiné jazyky, přitom např. při výuce fletivních jazyků lze korpus využít k procvičování různých gramatických forem, ať již standardních nebo substandardních. Jsou-li cílovou skupinou výuky

slovanského jazyka mluvčí jiného slovanského jazyka, mohli bychom navíc těžit z vysokého stupně pasivního porozumění, které napomáhá při orientaci v korpusových datech.

■ 1.2. Empirický výzkum

Otázkou však zůstává, zda je DDL přístup k výuce vhodný pro všechny a zda je zrealizovatelný v rámci kurikula. Granath (2009, 59, 62) poukazuje na to, že pro některé studenty je těžké vyhledávat v korpusu i interpretovat korpusová data. Nejsou zvyklí postupovat induktivně a neumí se vyznat v množství dat bez učitelova vedení. Na podobné problémy jsme narazili při práci studentů češtiny přímo s korpusovým vyhledávačem Bonito (Vališová 2011). Zajímavé však bylo zjištění, že i prosté vyhledání slova bylo pro studenty velmi užitečné. Studenti vyhledávali slova v pádech, které se ještě neučili nebo slova, u kterých dochází k alternaci konsonantů, a je tedy obtížné odvodit nominativ. Pro zobrazení lemmatu stačí třikrát kliknout a tento postup si studenti osvojili již po prvním vyhledávání a poté byli schopni samostatně odpovědět na analogická zadání. Stejně probíhalo zobrazení morfologické značky u slov, kde měli za úkol rozeznat rod nebo pád. Zjistili jsme, že korpus je výborným doplňkem slovníku, kde bývá uveden pouze nominativ nebo infinitiv, popř. rod.

Složitější aktivity, jako např. vyhledávání kolokací, jež vyžadovalo bližší znalost manažeru Bonito a také toho, jak interpretovat nalezená data, byli studenti schopni zvládnout pouze za učitelovy neustálé asistence. Granathová (2009, 59) shodně poznamenává, že studenti potřebovali neustálou pomoc učitele jak při formulování dotazu, tak i při čtení konkordancí. Studenti, kteří při vyhledávání čelili největším problémům, byli ti, jejichž jazykové znalosti byly na nižší úrovni než u ostatních. Boulton (2009, 49–50) však ve svém výzkumu dokládá, že DDL je vhodné i pro méně pokročilé studenty. Vyzkoušel cvičení s celými větami i ve formátu KWIC a studenti měli lepší výsledky u druhého typu cvičení. Zdůrazňuje, že studenti s nižší pokročilostí jazyka se kvůli malé slovní zásobě soustředí pouze na nejbližší kontext slova, a proto snadněji odhalí vzor, který se opakuje.

Braunová (2005, 2007) naopak navrhuje vytvořit pedagogicky relevantní korpus, který bude sloužit pro vzdělávací účely. Poukazuje na to, že velké množství slov je výhodné pro lexikografy, ale pro výuku jazyků stačí menší specializovaný korpus, např. o velikosti 20 000 až 200 000 slov. Přednostmi takového korpusu není jenom velikost, ale hlavně obsah, který může zahrnovat texty nebo dokonce i videa různých zdrojů a na různá témata. Takto se však vyhneme neobvyklé slovní zásobě z beletrie či odborných textů. Důležitý je také design vyhledávače, jenž by měl být uživatelsky přátelský a snadno ovladatelný. Braunová (2007) též představuje svůj projekt korpusu pro pedagogické účely, jež obsahuje rozhovory rodilých mluvčích na rozličná témata, přičemž student může nejen číst přepsaný text, ale i shlédnout video.

Vidíme tedy, že výhody korpusu, jako je velké množství dat, reálný úzus rodilých mluvčích a rychlost při vyhledávání, nemusí být pro DDL vždy výhodami.

Poděkování

Děkuji dr. Paulu Johnsonovi za konzultace a cenné rady při tvorbě korpusových cvičení, magistřím Jiřce Lukášové a Petře Stražovské za jejich otestování ve výuce a studentům češtiny, kteří se do testování zapojili.

Literatura

- Atkins S., Rundell M., 2008, *The Oxford Guide of Practical Lexicography*. Oxford University Press, Oxford. Kapitola 10.8, Examples, 452–464.
- Boulton A., 2009, Testing the limits of data-driven learning: language proficiency and training. *ReCALL*, 21, č. 1, 37–54.
- Bernardini S., 2004, Corpora in the classroom. An overview and some reflections on future developments. In Sinclair J., *How to Use Corpora in Language Teaching*. John Benjamins, Amsterdam, 17–36.
- Bischofová B., Hrdlička M., 2005, *Čeština pro cizince B1 (Učebnice)*. Soze, Brno.
- Bořilová P., Holá L., 2009, *Česky krok za krokem 2*. Akropolis, Praha.
- Braun S., 2005, From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, 17, č. 1, 47–64.
- Braun S., 2007, Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19, č. 3, 307–328.
- Cook G., 2001, 'The philosopher pulled the lower jaw of the hen' Ludicrous Invented Sentences in Language Teaching. *Applied Linguistics*, 22, č. 3, 366–387.
- Český národní korpus – SYN2000, 2000. Ústav českého národního korpusu FF UK, Praha 2000. Dostupný z WWW: <<http://ucnk.ff.cuni.cz>>
- Český národní korpus – SYN2005, 2005. Ústav českého národního korpusu FF UK, Praha 2000. Dostupný z WWW: <<http://ucnk.ff.cuni.cz>>
- Estling Vannestal M., Lindquist H., 2007, Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, 19, č. 3, s. 329–350.
- Granath S., 2009, Who benefits from learning how to use corpora? In Aijmer, K. (ed.) *Corpora and Language Teaching*. Amsterdam, John Benjamins, s. 47–65.
- Hrdlička M., 2002, *Cizí jazyk čeština*. Praha, ISV.
- Hrdlička M., 2009, *Gramatika a výuka češtiny jako cizího jazyka. K prezentaci gramatiky českého jazyka v učebnicích češtiny pro cizince*. Praha, Karolinum.
- Hrdlička M., 2005, K otázce komunikační metody a komunikativnosti. In *Sborník Asociace učitelů češtiny jako cizího jazyka (AUČCJ) 2003-2005*, eds. Čemusová J., Holá L., Rynďová J., Praha, Akropolis, s. 17–20.
- Johns T., 1991, Should you be persuaded. Two samples of data-driven learning materials. In *Classroom Concordancing*, eds. Johns T., King P., Birmingham, English Language Research Journal 4, University of Birmingham, s. 1–13.
- Lafer B., 2008, Corpus-based versus Lexicographer Examples in Comprehension and Production of New Words. In Fontanille T., *Practical Lexicography (A Reader)*. Oxford, Oxford University Press, s. 215–218.
- Römer U., 2009, What help do teachers need and what can we offer? In *Corpora and Language Teaching*, ed. Aijmer K., Amsterdam, John Benjamins, s. 83–96.

- Rychlý P., 2008, Bonito – grafické uživatelské rozhraní systému Manatee, Verze 1.49., 1998–2003. Dostupné z WWW: <<http://ucnk.ff.cuni.cz/bonito/index.php>>
- Svensén B., 2009, *A Handbook of Lexicography, The Theory and Practice of Dictionary-Making*. Cambridge, Cambridge University Press, Kapitola 15, Examples, s. 281–288.
- Vališová P., 2011, Využití koprusu při výuce češtiny jako cizího jazyka. In *20 let didaktiky cizích jazyků*, Liberec, Technická univerzita (v tisku).

Tento článek byl podpořen děkanským grantem FF MU: MUNI/21/VAL/2011