

Možnosti využitia SNK v pedagogickej praxi

Lucia Gianitsová

Oddelenie Slovenského národného korpusu, Jazykovedný ústav L. Štúra SAV, Bratislava
<http://korpus.juls.savba.sk>

Príspevok bol prednesený na vedeckom seminári Učebné materiály vo vyučovaní materinského jazyka, slohu a literatúry (Banská Bystrica 3. – 4. 2. 2004).

1. Pri výskume jazyka i pri jeho výučbe je potrebná materiálová základňa – *textové zdroje*. Najmä pri tvorbe slovníkov bol a stále ostáva potrebný rozsiahly textový materiál. Na jeho základe lingvisti tvoria heslá, ktoré okrem ucelených výkladov slova a jeho funkcií obsahujú reálne výskyty slov tak, ako sa nachádzajú v konkrétnych textoch. Tradične teda tvorbe slovníkov predchádzal zber citátov z textov (excerpcia napr. z umeleckých diel, publicistiky, odborných štúdií a i.), ktoré obsahovali dané slová. Aj svetovo známy *Oxford English Dictionary*, ktorý obsahuje 500 000 hesiel, bol založený na takomto súbore (či korpuse) citácií. Za 71 rokov práce (do roku 1928, keď vyšiel jeho posledný, 12. diel) vyše 2 000 spolupracovníkov zozbieralo asi 5 miliónov citácií, čo predstavuje asi 50 miliónov slov (Šulc, 1999, s. 28). Podobne aj pri príprave *Slovníka slovenského jazyka* (Peciar, 1959 – 1968) a *Krátkeho slovníka slovenského jazyka* (Kačala, 1. vyd. 1987, 2. vyd. 1989) čerpali tvorcovia z lexikálnej kartotéky, ktorá bola vytváraná v Jazykovednom ústave L. Štúra SAV. Excerptovanie diel a vypisovanie citácií na kartotečné lístky trvalo v podstate od 40. rokov 20. storočia do roku 1991. Tak vznikla kartotéka, ktorá dnes obsahuje asi 5 miliónov kartotečných lístkov.

2. **Korpusy a lingvistika.** Nové nečakané možnosti priniesol *rozvoj korpusovej lingvistiky*, ktorá sa od 60. rokov 20. storočia mohla stále viac uplatňovať v jazykovednom výskume vďaka rozvíjajúcej sa výpočtovej technike. Po vzniku prvého elektronického korpusu – *Brownovho korpusu* (BC) – v rokoch 1961 až 1964 na Brown University v USA zaznamenávame stále väčší nárast záujmu o túto oblasť lingvistiky, vznikli a stále vznikajú desiatky korpusov pre väčšinu národných jazykov (Šulc, 1999, s. 30). Ak si všimneme len slovanské jazyky, korpusy jazykov (často i viaceré pre jeden jazyk alebo oblasť používania istého jazyka) majú v Poľsku, Slovinsku, Česku, Chorvátsku, Srbsku, Rusku a v Bulharsku. Viac o korpusoch a o korpusovej lingvistike je možné sa dozvedieť na webovej stránke Slovenského národného korpusu (SNK) <http://korpus.juls.savba.sk>, príp. na stránkach iných korpusov. Odkazy na ne tak isto nájdete na stránke SNK. (Pri príprave základných informácií v oddieloch 2 a 3 sme z týchto zdrojov čerpali aj my.)

2.1 Výraz *korpus* má niekoľko významov. V KSSJ (1997, s. 257) sa uvádzajú tri významy: 1. telo, 2. podstatná časť niečoho, 3. (odbor.) súbor skúmaných prvkov. V lingvistike sa začal

používať termín korpus práve v tomto poslednom význame. Možno teda povedať, že *korpus je špecifický súbor jazykových dát (textov), uchovávaný v elektronickej podobe* so všetkými jej výhodami (napr. prístup cez internet) a *je určený zvyčajne na vedecko-výskumné a učebné ciele*.

2.2 Druhy korpusov. V literatúre z oblasti korpusovej lingvistiky (Šulc, 1999; Čermák, 2000; Šimková, 2003) sa uvádza, že jednotlivé korpusy sa líšia mnohými parametrami, predovšetkým:

§ Jazykom: *jednojazyčné korpusy* – v súčasnosti existujú takéto korpusy pre desiatky rôznych jazykov; *dvoj- a viacjazyčné (paralelné) korpusy* – obsahujú rovnaké texty v dvoch alebo viacerých jazykoch, napr. Orwellov román *1984* a jeho preklady boli takto spracované v rámci projektu *MULTEXT* a *MULTEXT-EAST*.

§ Zaznamenanou formou jazyka: najčastejšie sú *korpusy písaných textov*, ale existujú aj *korpusy hovoreného jazyka*.

§ Veľkosťou: prvé korpusy (do r. 1975) disponovali menej ako jedným miliónom slovných tvarov, niektoré súčasné korpusy obsahujú miliardy slov.

§ Typom textov: korpusy môžu byť *všeobecné* (nešpecifické, základné) a *špecializované* na určitý typ textov či jazykových javov (ako texty jedného autora, neformálne prejavy a pod.).

§ Spôsobom uloženia: korpusy môžu byť uložené *v základnej textovej podobe* bez lingvistických informácií alebo sú *lematizované* (každé slovo má priradený základný, slovníkový tvar) a morfológicky, syntakticky, príp. sémanticky či štylisticky *anotované*.

§ Obdobím, z ktorého texty pochádzajú: korpusy konkrétneho jazyka v jeho súčasnej vývinovej fáze sa označujú ako *synchronne*, prierez určitým historickým obdobím zahŕňajú korpusy *diachronne*.

§ Vyváženosťou, resp. reprezentatívnosťou: Existujú korpusy obsahujúce akékoľvek texty, ktoré boli zhromaždené bez výberu, naopak *reprezentatívne korpusy* sa zameriavajú na prezentáciu konkrétneho národného jazyka (kladie sa veľký dôraz na zastúpenie rozličných jazykových variet, typov a žánrov textov z rozličných časových období a od rozličných autorov).

2.3 Ako sa tvorí korpus. Budovanie korpusu má viacero etáp (pozri aj Šimková, 2003, s.17):

§ *Získavanie textov a súhlasov na iné (nie komerčné) použitie diel:* Korpusy budované na vedecko-výskumné účely sú založené na nekomerčnej báze, tzn. že texty do nich sa získavajú v súlade s príslušným autorským zákonom podľa zmluvne dohodnutých podmienok. Poskytovateľmi sú autori, vydavateľstvá alebo iní držitelia autorských práv a texty sa od nich získavajú najčastejšie priamo v elektronickej podobe, príp. prenesením z internetu, zriedkavejšie skenovaním a veľmi málo ručným prepisovaním textov.

§ *Technické spracovanie korpusových dát:*

a) *čistenie textov* (odstraňovanie špecifických znakov a symbolov, odstraňovanie grafických súčastí textu – obrázky, tabuľky, grafy a pod.);

b) *prevod textov do jednotného formátu (napr. XML), ktorý zaznamenáva štruktúrne vlastnosti textu (napr. začiatky a konce odsekov, viet, zvýraznenie textu a i.);*

c) *tokenizácia, čiže segmentácia textu na najmenšie jednotky – tokeny (obyčajne tvary slov a iné znaky v texte).*

§ *Značkovanie (angl. tagging):* K textu sa dopĺňajú informácie o ňom:

a) bibliografické údaje,

b) štýlovo-žánrové charakteristiky,

c) morfologická anotácia slov (slovný druh, základný tvar slova, tzv. *lema*, gramatické kategórie), ktorá je kódovaná pomocou vopred navrhnutých značiek pre jednotlivé hodnoty. Ukážka morfologickej anotácie textov v Slovenskom národnom korpusu, ktorá obsahuje aj štruktúrne znaky (Forróová, M. – Garabík, R. – Gianitsová, L. – Horák, A. – Šimková, M., 2003, v tlači):

<w l="vináreň" t="SSfs1">Vináreň</w>

<w l="krušovice" t="SSfp1">Krušovice</w>

<w l="rovnako" t="Dx">rovnako</w>

<w l="ako" t="O">ako</w>

<w l="miestny" t="AAns1x">miestne</w>

<w l="café" t="SUns1">Café</w>

<w l="byť" t="VKepr">sú</w>

<w l="už" t="Dx">už</w>

<w l="zahalený,, t="Gtop1x">zahalené</w>

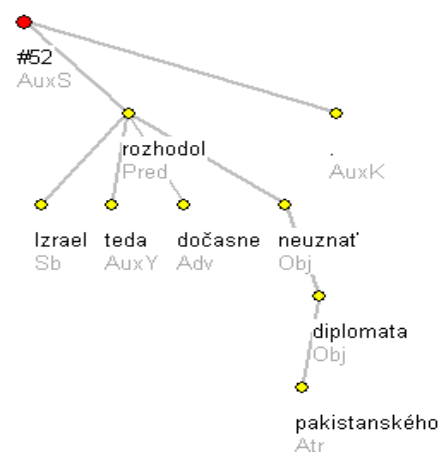
<w l="do" t="Eu2">do</w>

<w l="tma" t="SSfs2">tmy</w>

<w l="." t="Z">.</w>

d) jazykové informácie na úrovni viet (pozri stromovú štruktúru syntakticky anotovanej vety *Izrael teda rozhodol dočasne neuznať pakistanského diplomata.*);

e) informácie o významoch slova – sémantická anotácia.



2.4 Význam a využitie korpusov.

Korpusy majú široké uplatnenie v rôznych vedeckých výskumoch a aplikáciách (pozri aj Šimková, 2003, s. 17):

§ *Lingvistika* skúma a odhaľuje jazykové fenomény zisťované na základe korpusov (veľkého množstva reálnych textov). Pomocou analýzy korpusových textov môže byť overená väčšina lingvistických teórií a môžu vzniknúť aj nové hypotézy a teórie. *Lexikografia* využíva korpusy ako podporu tvorby slovníkov a lexikografi sú v súčasnosti najčastejšími používateľmi korpusov.

§ *Počítačové spracovanie prirodzeného jazyka* dokáže z korpusu vytážiť niektoré výsledky spracovania dát – zoznamy slov, spoločné umiestnenia slov (*kolokácie*), frekvenciu slov atď. Na tomto základe sa vyvíjajú systémy na spracovanie textov (automatická kontrola pravopisu, strojový preklad textov) i systémy na rozpoznávanie reči.

§ *Výučba jazykov.* Korpus môže byť vhodným zdrojom fráz a viet potrebných pri výučbe

cudzích jazykov, učebný počítačový program môže obsahovať klasický slovník spolu s menším korpusom, v ktorom možno jednotlivé slová prezerat' v kontexte, v akom sa reálne vyskytujú. Zároveň je to zdroj textov na vytváranie pravopisných, gramatických a štylistických cvičení a otvára možnosti problémového vyučovania jazyka v počítačovej učebni.

§ *Bežné využitie.* Pri práci s jazykom ho ocenia novinári, prekladatelia, iní záujemcovia o jazyk, keďže môže slúžiť ako zdroj poznania reálneho fungovania jazykových prostriedkov. Nenahrádza však kodifikačné ani gramatické príručky!

3.0 Slovenský národný korpus (SNK). Po zastavení excerpce textov do lexikografickej kartotéky sa v priebehu 90. rokov pristúpilo k elektronizácii niektorých textových zdrojov (išlo prevažne o publicistiku). Prácu charakterizovali veľmi obmedzené technické a personálne možnosti. Po náročných rokovaniach o vzniku špecializovaného korpusového pracoviska prebiehajúcich od roku 2000, do ktorých boli zainteresované Ministerstvo kultúry Slovenskej republiky, Ministerstvo školstva SR a Slovenská akadémia vied, bol v roku 2002 vládou schválený vedecko-výskumný projekt budovania elektronického korpusu v Jazykovednom ústave Ľ. Štúra SAV (Šimková, 2003, s. 17).



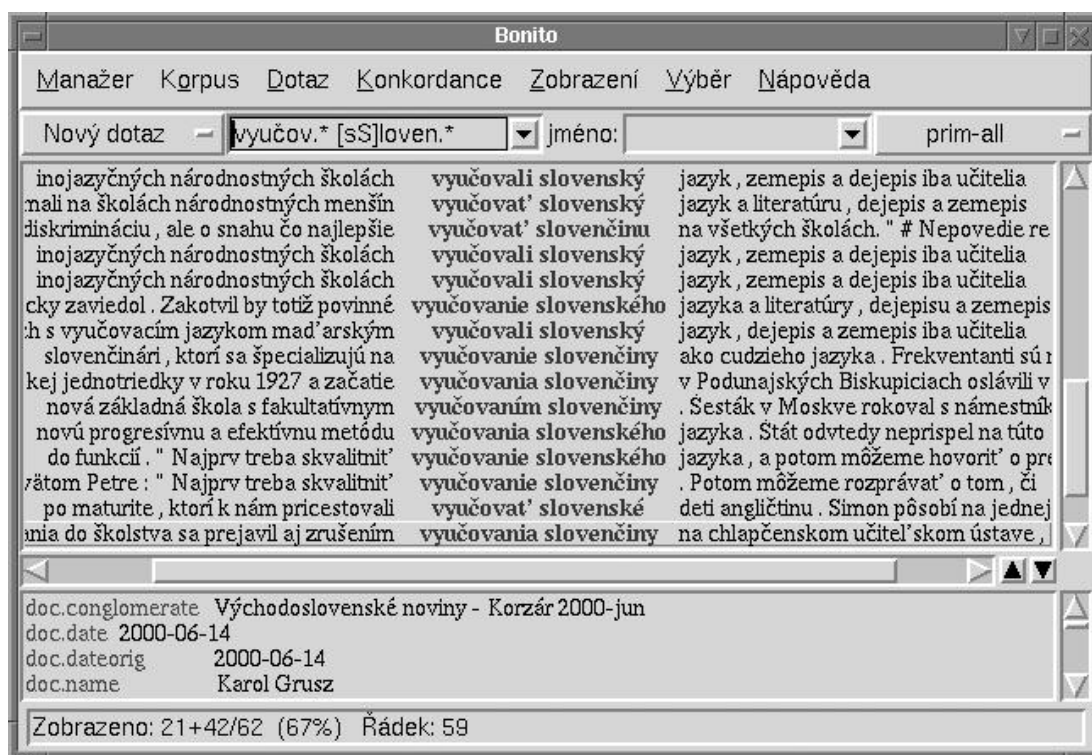
SNK ako *všeobecný jednojazyčný korpus* bude vyvážený, tzn. bude obsahovať v primeranej miere texty z rôznych oblastí používania súčasného slovenského jazyka (z rokov 1955 – 2005). Výber textov v rozsahu 200 miliónov slov sa obmedzí na písané texty a postupne sa bude rozširovať o texty z ďalších období a sfér používania slovenského jazyka (Šimková, 2003, s. 18).

3.1 Výsledky a plány. Do konca roku 2003 sa podarilo uzavrieť zmluvy a získať texty z rokov 1994 – 2003 od rôznych poskytovateľov, bola vytvorená webová stránka <http://korpus.juls.savba.sk> a korpusové dáta (približne 170 miliónov slov) s vonkajšou (bibliografickou a štýlovo-žánrovou) anotáciou boli sprístupnené používateľom. V súčasnosti sa pracuje na realizácii návrhu tagsetu (Forróová, M. – Garabík, R. – Gianitsová, L. – Horák, A. – Šimková, M., 2003, v tlači), ručnej lematizácii a morfológickom značkovaní. V nasledujúcom období sa pracovníci SNK budú zameriavať na získavanie textov spred roka 1990, na zdokonalenie www rozhrania na vyhľadávanie, na lematizáciu, morfológickú anotáciu a vybraná vzorka sa bude anotovať aj syntakticky. Súčasne sa budú budovať podkorpusy diachrónnych a nárečových textov a terminologická a lexikografická databáza (op. cit., 2003, 18 – 19).

3.2 SNK možno používať už dnes. Pri využívaní korpusových dát je možné použiť *www rozhranie* priamo na webovej stránke SNK, kde je možné vyhľadávať voľne zatiaľ len základné informácie, napr. konkrétne tvary slov. Možnosť rozšíreného vyhľadávania ponúka *korpusový manažér Manatee s klientom Bonito* pre zaregistrovaných používateľov. (Program je možné

nainštalovať po stiahnutí z www stránky SNK.) Každý, kto sa zaviaže k nekomerčnému používaniu sprístupneného korpusu, môže pracovať s testovacou verziou Slovenského národného korpusu *Prim0.2* cez vlastné konto (heslo). Heslo bude záujemcovi pridelené na základe poštou zaslaného vyplneného a podpísaného formulára, ktorý sa tak isto nachádza na webovej stránke SNK.

Toto vyhľadávanie je rozšírené, program umožňuje zobrazenie širšieho kontextu, bibliografických odkazov, frekvencie výskytu slova, kolokácií (spolu výskytov slova) a pod. Práca s programom je jednoduchá: Používateľ napíše hľadaný výraz do vyhľadávacieho okna (*Nový dotaz*) a potvrdí hľadanie (stlačením tlačidla *Enter*). Jednotlivé zobrazené riadky obsahujú citáty z textov, ktoré sa v korpuse nachádzajú a obsahujú hľadaný výraz. Širší kontext sa zobrazí v spodnom okne po stlačení ľavého tlačidla na myši. Po stlačení pravého tlačidla sa objavia v tomto okne bibliografické a štýlovo-žánrové informácie o texte, z ktorého aktuálny riadok pochádza. Napr.:



V uvádzanom príklade boli pri hľadaní použité aj *regulárne výrazy*. Ide o symboly, ktoré zastupujú znaky (písmená, číslice, interpunkčné znamienka) alebo reťazce znakov (slová), ktoré chceme vyhľadať. Svoj význam majú pri zložitejšom hľadaní, keď chceme pomocou jedného príkazu vyhľadať viacero tvarov alebo slov. Napr. pri vyhľadávaní všetkých tvarov jedného slova alebo viacerých slov s rovnakým kmeňom (*vyučov-* *sloven-*), ktoré sú zapísané s malým, príp. veľkým písmenom na začiatku (*sloven-*, *Sloven-*), keďže texty môžu obsahovať aj chyby, ktoré sa nachádzali v pôvodnom vydaní textu. Viac o používaní regulárnych výrazov a tvorení zložitejších príkazov sa možno dozvedieť aj z manuálu k *Českému národnímu korpusu* (ČNK. Úvod a příručka uživatele, 2000) alebo na webovej adrese <http://ucnk.ff.cuni.cz/manual/index.html>.

Pri práci s *Bonitom* je vhodné poznať a používať minimálne tieto regulárne výrazy (pozri aj http://korpus.juls.savba.sk/vyhľadavanie/korpusovy_manazer/):

Regulárny výraz	Význam regulárneho výrazu	Príklad použitia	Očakávaný výsledok vyhľadávania
.	BODKA nahrádza jeden ľubovoľný znak.	<i>dom.</i> <i>dom..</i>	<i>Doma, domy...</i> <i>domov, domec...</i>
*	HVIEZDIČKA určuje, že znak, ktorý je pred hviezdíčkou, sa opakuje ľubovoľný počet krát (aj nula krát).	<i>hm*</i>	<i>h, hm, hmm,</i> <i>hmmm...</i>
+	PLUS určuje, že znak, ktorý je pred týmto regulárnym výrazom, sa opakuje raz alebo viackrát.	<i>hm+</i>	<i>hm, hmm, hmmm...</i>
	ZVISLÁ ČIARA má funkciu operátora ALEBO.	<i>dom doma</i>	<i>dom, doma</i>
[]	HRANATÉ ZÁTVORKY. Definuje sa v nich množina znakov, ktoré sa môžu vyskytovať v danom výraze na mieste zátvoriek. Množina sa určuje vymenovaním znakov, ktoré sa neoddeľujú čiarkou, alebo určením intervalu napr. a-z (interval zahŕňa sled znakov bez diakritiky).	<i>[dsl]om</i> <i>[r-v]ám</i>	<i>dom, som, lom</i> <i>rám, sám, tám,</i> <i>uám, vám</i> (ak sa takéto tokeny v korpuse vyskytujú)
()	JEDNODUCHÉ ZÁTVORKY slúžia na zápis súboru požiadaviek na určitý znak vyhľadávaného výrazu pomocou regulárneho výrazu, resp. viacerých regulárnych výrazov.	<i>(D/d)om</i> <i>([Kk] [Tt])am</i>	<i>Dom, dom</i> <i>Kam, kam, Tam,</i> <i>tam</i>
(?i)	Tento regulárny výraz spôsobí, že pri vyhľadávaní sa bude ignorovať rozlišovanie veľkých a malých písmen.	<i>(?i)dom</i>	<i>Dom, dom</i>
\	Obrátený znak LOMENÉ pred znakom regulárneho výrazu znamená, že sa daný znak nebude pri vyhľadávaní chápať ako regulárny výraz, ale ako jednotka textu.	<i>napr\.</i>	<i>napr. (a nie napri,</i> <i>napre, naprd...)</i>

- § Znak v SNK sa chápe ako ľubovoľný znak okrem medzery.
- § Ak hovoríme, že regulárny výraz **BODKA** nahrádza ľubovoľný znak, znamená to, že nahrádza tiež ľubovoľnú číslicu, interpunkčné znamienko, zátvorku a pod.
- § Regulárne výrazy môžeme aj kombinovať:
 - .* (kombinácia **BODKA HVIEZDIČKA**) nahrádza ľubovoľný znak ľubovoľný počet krát. Výsledkom vyhľadávania zápisu *.*istá* budú slová zakončené sledom grafém *-istá*, ale aj samotné slovo *istá*.
 - .+ (kombinácia **BODKA PLUS**) je užitočná pri vyhľadávaní slov s určitým prefixom, sufixom, skupinou písmen a pod. Výsledok vyhľadávania zápisu *za.+* zobrazí všetky slová začínajúce sa písmenami *za-* (okrem slova *za*).
- § Pri zápise vyhľadávaného výrazu môže byť regulárny výraz použitý **na ľubovoľnom mieste**.
- § Napríklad pomocou zápisu *za.+ík* možno nájsť všetky slová začínajúce sa písmenami *za-* a končiac sa na *-ík*.
- § Naopak, zápisom *.*koreň.** možno nájsť všetky slová s bázou *koreň*. Dodatočnou úpravou tohto zápisu na *.*kore(ň/n).** sa do výsledku vyhľadávania zahrnú i slová obsahujúce alternáciu v danej báze *z ň* na *n* (napríklad i slová *koreniny, vykorenený*).

4. Možnosti využitia SNK v pedagogickej praxi. Texty v korpuse je možné využívať pomocou týchto nástrojov už dnes. Samozrejme, učiteľ by mal byť pripravený učiť sa práci s počítačom a byť ochotný skúšať nové, často netradičné a náročné spôsoby výučby. Veríme však, že súčasné možnosti, rozrastajúce sa počítačové vybavenie škôl, ktoré poskytuje aj projekt *Infovek*, budú pozitívnou výzvou.

Prácu s korpusem cez korpusový manažér *Bonito* možno uplatniť vo viacerých činnostiach učiteľa i žiaka:

- § pri individuálnej príprave učiteľa na hodinu – príprava cvičení;
- § ako motiváciu na vysvetlenie nového učiva – učiteľ zadá problémovú úlohu, na základe výsledkov hľadania žiaci dedukujú s pomocou učiteľa teoretické princípy;

- § pri výklade učiteľa, oboznamovaní sa s novým gramatickým javom – žiaci si overujú teóriu v praxi, ako sa gramatické prostriedky reálne vyskytujú v texte;
- § pri precvičovaní, upevňovaní nového učiva – vyhľadávanie príkladov priamo na hodine, učiteľ zadá úlohu a žiaci pracujú s programom sami;
- § pri opakovaní – problémové úlohy.

Vo všetkých týchto fázach môže tvorivý učiteľ využiť jeden z nasledujúcich alebo podobných príkladov, ktoré boli inšpirované témami a úlohami v učebniciach slovenského jazyka (Tibenská, 2000, 2001, 2002a, 2002b), metodickou príručkou (Tibenská – Sedláková, 1999) i osobnými skúsenosťami s výučbou slovenského jazyka na ZŠ.

Výhodou práce s korpusom je možnosť vyhľadávať cez korpusový manažér autentický textový materiál, vety alebo slová na precvičenie týchto i ďalších podobných okruhov:

4.1 Vybrané slová (od 4. ročníka):

- § Vyhľadáme vety s vybranými slovami, ktoré použijeme v diktáte.

Príkaz: *.*by.**, *.*my.**, *.*vy.** atď.; výsledkom v prvom prípade hľadania môže byť *byt*, *kobyła*, *bydliskom*, *obydlie*...

- § Dáme vyhľadávať priamo žiakom, aby sami zostavili zoznam vybraných slov.

§ Diktujeme slová, ktoré žiaci zapisujú do vyhľadávacieho okienka (nesprávny zápis, napr. *bívanie*, neúspešné hľadanie), tzv. „diktát do počítača“. Pri tom však môžeme nájsť aj homonymá (*byt* – *bit*).

4.2 Pravopis cudzích slov (8. ročník a vyššie):

- § Zadávame typické predpony a prípony pre cudzie slová.

Príkaz: *syn.**, *sym.**, *anti.**, *hyper.**, *dis.**, *dys.**, *.*izmus* atď.

§ Je možné prepojiť to s lexikologickým učivom: Zisti význam cudzieho slova z vyhľadaných konkordancií (výskytov slova, jednotlivých riadkov). Je možné pridať aj ďalšie príkazy, napr. *.*cia*, *super.** atď.

4.3 Písanie veľkých písmen (v ktoromkoľvek ročníku – podľa náročnosti):

- § Kedy píšeme v mene veľké a kedy malé písmeno?

Príkaz: *Herkules/herkules*, *Romeo/romeo*, *základn.* škol.**/*Základn.* škol.** a i.

§ „Diktát do počítača“ – Žiak sa pokúša vyhľadať zemepisný názov, názov inštitúcie dovtedy, kým ho nenapíše správne, najrýchlejší vyhráva.

4.4 Písanie čiarky a iných znamienok – čiarky môžu v textoch chýbať, najmä pred vzťažnými zámenami (*kto*, *čo*, *kde*, *kedy*), no i to sa dá využiť ako príklad na typické chyby (od 6. roč.):

§ Vyhľadaj opytovacie zámená (*Kto*, *kto*, *Čo*, *čo*, *Kde*, *kde*, *Kedy*, *kedy*, *Ako*, *ako*...). Nájdi prípad, kde zámeno slúži ako spájací výraz. Dopln chýbajúce čiarky.

4.5 Spodobovanie v predponách (prepojenie so slovtvorbou):

§ Možno nájsť predpony *vs-*, *ros-*, *bes-*, ktoré sa takto vyslovujú i zapisujú? Vyvodte závery.

Príkaz: *vs.**, *ros.**, *bes.**

4.6 Zdvojené hlásky:

§ Vyhl'adaj slová so spoluhláskami *-nn-*, *-dd-*, *-jj-*, *-šš-* atď.

Príkaz: *.*nn.**, *.*dd.**, *.*jj.**, *.*šš.** atď.

4.7 Spoluhláskové skupiny:

§ Vyhl'adaj slová so skupinami hlások *-stn-*, *-stl-*, *-ds-*, *-zdn-*. Ako sa správne vyslovujú?

Príkaz: *.*stn.**, *.*stl.**, *.*ds.**, *.*zdn.** atď.

4.8 Rytmický zákon (od 5. ročníka):

§ Vyhl'adaj slová, ktoré majú dlhé gramatické prípony a krátke prípony. O aký jav tu ide?

Príkaz: *.*ám*, *.*ách*, *.*á*, *.*ý*, *.*ých*, *.*ým*, *.*í*, *.*ieho*, *.*iemu* a *i*. alebo *.*am*, *.*ach*, *.*a*, *.*y* atď.

§ Z predchádzajúcej úlohy vyvod', kde sa najčastejšie uplatňuje rytmický zákon. Ktoré tvary slov sú výnimkou?

4.9 Tvorenie slov:

§ Nájdí všetky slová, ktoré obsahujú *-chod-*, *-uč-...* Príkaz: *.*chod.**, *.*uč.** atď.

4.10 Viacvýznamovosť slov verzus homonymia (od 5. ročníka):

§ Vyhl'adaj prídavné mená *dobrý*, *pekný*, *prázdný*, *rázovitý*. Majú vždy rovnaký význam? (Táto úloha býva tradične problémová. Žiaci nie sú často schopní sami uviesť viac významov týchto slov. Preto im tento spôsob môže pomôcť.)

§ Vyhl'adaj slová ako *kohútik*, *pero*, *oko*, *mať*, *čelo*, *kalit'*, *box*. Všimaj si ich významy a posúď, kedy ide o viacvýznamové slovo a kedy o homonymum.

4.11 Kolokácie slov a synonymá:

§ Urč, ktoré slová sa najčastejšie vyskytujú pri slove *výsledok*. Vieš medzi nimi nájsť synonymá? Sú všetky synonymné?

4.12 Určovanie slovných druhov:

§ Vyhl'adaj slovo *zima* (*teplo*, *okolo*, *hore*, *dole*, *veľa*, *blízko*, *a*, *aj*, *ale*, *bože*, *koncom*, *tuším*, *hádam*, *celkom...*). Posudzuj každú vetu zvlášť. O aký slovný druh ide?

4.13 Pády a používanie dvojtvarov:

§ Vyhl'adaj tvary *gul'ôčok/guličiek*, *prijímačok/prijímačiek*. Sleduj ich výskyt a porozmýšľaj, ktoré je správne.

§ Vyhl'adaj tvary *jamôk/jamiiek*, *sukieň/sukní*. Urč, ktorý tvar je frekventovanejší.

§ Vyhl'adaj tvary *Soni/Sone*, *Táni/Táne*, vyber len príklady na gen. sg. Ktorý tvar je správny?

§ Vyhl'adaj tvary *arzenáli/arzenále*, *lokáli/lokále*, vyber len príklady na lok. sg. Ktorý tvar je častejší?

§ Vyhl'adaj tvary slov *Kraska/Krasku, Botta/Bottu, fotela/fotelu, hotela/hotelu, faktora/faktoru, rotora/rotoru, cukre/cukry...*, ktorý tvar je častejší?

4.14 Slovesné tvary a ich gramatické kategórie (8. ročník a vyššie):

§ Vyhl'adaj zvratné *sa*. Urč, kedy ide o zámeno a kedy o súčasť slovesa.

§ Vyhl'adaj tvary slovesa *byť (som, si, sme, ste, je, sú)*, urč, kedy ide o plnovýznamové sloveso, kedy o neplnovýznamové (pomocné) sloveso a kedy o súčasť iného slovesa (súčasť gramatického tvaru).

§ Vyhl'adaj tvary *by, aby, keby*. Aký slovesný tvar (v akom spôsobe) pomáhajú vytvárať?

§ Vyhl'adaj tvary slovesa *byť* v budúcom čase (*budem, budeš, bude, budeme, budete, budú*). Aká je ich funkcia, keď nie sú plnovýznamové ani neplnovýznamové slovesá?

4.15 Vetný rozbor:

§ Učiteľ si vyberie vety vhodné na vetný rozbor (ale aj komplexný slovnodruhový rozbor).

§ Vyhl'adaj spojky *aby, žeby, keby, lebo, že, keď, pretože, keďže* alebo vzťahné zámená *kto (koho, komu, kom, kým...), čo, aký, ktorý, kde, kedy, ako, prečo* atď. a urč, ktorý vetný člen je vyjadrený vedľajšou vetou (8. ročník a vyššie).

§ Vyhl'adaj slovesné tvary ako *prší, sneží, stmieva sa, ide sa, vidno, smädí, zabilo...* a urč vety, v ktorých sa nachádzajú, podľa členitosti.

4.16 Zisťovanie štylistickej alebo teritoriálnej príslušnosti slova, atď.

5. Na záver. Práca s počítačom je pre žiakov v súčasnosti veľmi príťažlivá. A možnosť uplatniť svoje vlastné nápady a schopnosti pri práci s počítačom je tiež pozitívnou motiváciou k nezámernému učeniu.

Aj keď momentálne nie je možné využiť korpusové dáta v plnej miere, technická stránka a možnosti využitia SNK sa budú postupne zlepšovať. Veľkú zásluhu na tom bude mať iste aj prebiehajúca lematizácia a morfológická anotácia. No už v tomto období môže analýza korpusových dát posilniť komunikatívno-poznávací prístup a prirodzený pohľad na jazyk ako živý nástroj.

Ozývajú sa síce hlasy, ktoré vyčítajú, že vyučovanie s podporou počítačov spôsobuje stratu priameho kontaktu medzi učiteľom a žiakom. Podľa nášho názoru však práve táto forma výučby môže pomôcť vytvoriť pozitívnu partnerskú atmosféru na hodinách, a to aj v tom prípade, že do interakcie učiteľ – žiak vstupuje tretí člen, počítač.

Literatúra

ČERMÁK, F.: *Jazykový korpus: prostriedek a zdroj poznání*. In: Studie z korpusové lingvistiky.

Acta Universitatis Carolinae. Philologica 3 – 4. Red. F. Čermák – J. Klímová – V. Petkevič.

Praha: Univerzita Karlova – Nakladatelství Karolinum 2000. s. 15 – 37.

ČNK. *Úvod a příručka uživatele*. Red. J. Kocek – M. Kopřivová – K. Kučera. Praha: FF UK – ÚČNK 2000. 156 s.

FORRÓOVÁ, M. – GARABÍK, R. – GIANITSOVÁ, L. – HORÁK, A. – ŠIMKOVÁ, M.: *Návrh morfológického tagsetu SNK*. Medzinárodná vedecká konferencia Slovánske jazyky v počítačovom spracovaní. Bratislava 24. – 25. 10. 2003. V tlači.

Krátky slovník slovenského jazyka. Red. J. Kačala. 3. vyd. Bratislava: Veda 1997, s. 257.

ŠIMKOVÁ, M.: *Počítačové spracovanie prirodzeného jazyka a Slovenský národný korpus*. In: *Počítačová podpora prekladu*. Zborník prednášok. Budmerice 22. – 23. 5. 2003, s. 15 – 19.

ŠULC, M.: *Korpusová lingvistika. První vstup*. Praha: Karolinum 1999. 94 s.

TIBENSKÁ, E. a kol.: *Slovenský jazyk pre 5. roč. základných škôl*. Bratislava: Orbis Pictus Istropolitana 2000. 64 s.

TIBENSKÁ, E. a kol.: *Slovenský jazyk pre 6. roč. základných škôl*. Bratislava: Orbis Pictus Istropolitana 2002a. 64 s.

TIBENSKÁ, E. a kol.: *Slovenský jazyk pre 7. roč. základných škôl*. Bratislava: Orbis Pictus Istropolitana 2002b. 64 s.

TIBENSKÁ, E. a kol.: *Slovenský jazyk pre 8. roč. základných škôl*. Bratislava: Orbis Pictus Istropolitana 2001. 64 s.

TIBENSKÁ, E. – SEDLÁKOVÁ, M.: *Metodická príručka. Slovenský jazyk pre 5. ročník ZŠ*. Bratislava: Orbis Pictus Istropolitana 1999. 84 s.

URL: <http://korpus.juls.savba.sk>

URL: <http://ucnk.ff.cuni.cz/manual/index.html>