

Báseň a počítač



AKADEMIE VĚD
ČESKÉ REPUBLIKY

věda 12

kolem
nás
Výzvy
a otázky

Ústav pro českou literaturu AV ČR, v. v. i. (<http://ucl.cas.cz>), je jediným neuniverzitním pracovištěm v České republice, které se věnuje základnímu výzkumu v oblasti literární vědy. Je garantem výzkumu národní identity v oblasti literatury a nejsilnější institucí svého vědního oboru – literárněvědné bohemistiky ve světovém měřítku. Předmětem jeho výzkumů jsou historie i současnost české literatury v jejím jazykovém i teritoriálním vymezení, její vztahy s literaturami jiných jazyků či regionů a rovněž literární život v českých zemích od jeho počátků až po dnešek. Zabývá se přitom jak literaturou uměleckou, tak i populární či triviální. Kulturní dědictví země, specifická situace české literatury a literární vědy jsou reflektovány také v bádání, překladatelské a editorské činnosti členů ÚČL na poli teorie. Zde ÚČL nejčastěji překračuje hranice národní filologie směrem k obecným otázkám literatury a k problematice literatury srovnávací. Výzkum zasahuje i do oblasti teorie výtvarného umění, filozofie a kulturní antropologie. Pro literárněvědnou bohemistiku v národním i mezinárodním kontextu vytváří ÚČL organizační centrum a informační zázemí. Poskytuje studijní pobyty kolegům ze zahraničí, vydává ústřední oborové periodikum, provozuje veřejně přístupnou knihovnu. Zaměřuje se rovněž na informační služby pro odbornou, studentskou a občanskou veřejnost, poskytované prostřednictvím digitálních knihoven, bibliografických a slovníkových databází. Někteří pracovníci ÚČL se účastní literárního života, nejčastěji jako kritici či recenzenti, mnohdy i vlastní uměleckou tvorbou. ÚČL má dvě pracoviště, pražské a brněnské, jejich badatelský program je společný a oddělení procházejí napříč těmito pracovišti. Členové obou pracovišť se zapojují do vysokoškolské výuky, přičemž ÚČL sám je školicím pracovištěm postgraduálního studia.

Báseň a počítač?!

Báseň a počítač? Jde takové spojení vůbec dohromady? Nestojí poezie a výpočetní technika v protikladu? Využití výpočetní techniky v přírodních a exaktních vědách nikoho nepřekvapí, chápe se jako samozřejmé. Jako samozřejmé se ale postupně začalo chápat také to, že počítače se staly nedílnou součástí našeho života, a pronikly tedy do všech oblastí, kterými se člověk zabývá. Užití výpočetní techniky v humanitních vědách má svou tradici – velmi plodně se rozvíjelo zejména v šedesátých letech 20. století, kdy počítače (tehdy obrovské mašiny, které umělo obsluhovat jen pár specialistů) pomáhaly zpracovávat velké datové soubory. Nikdy to ale nefungovalo tak, že by – a trochu přeháníme – literární vědec vytrhl báseň ze sbírky a podobně jako ve filmu *Jáchyme, hod' ho do stroje* (1974) ji „hodil“ do počítače, který za okamžik vyplivne výsledek. Literární vědec samozřejmě musel počítači data připravit – odměnou mu bylo, že počítač data seřadil a umožnil v nich pohodlně vyhledávat. Novou etapu ve využití počítačů znamenala devadesátá léta, období prudkého nástupu výpočetní techniky a jejího ovládnutí světa. Se zdokonalením počítačů a rozšířením jejich možností také rostlo jejich využití v humanitních vědách. Využití počítačů v literární vědě se váže k tomu, že objekt výzkumu této vědy, tj. literární texty, existuje v elektronické podobě. Vše, o čem bude v této brožurce dále řeč, by nikdy nemohlo vzniknout, pokud by neexistovala Česká elektronická knihovna, která byla vytvořena v Ústavu pro českou literaturu AV ČR.

V následujících kapitolách pojednáme o tom, co všechno se dá s literárními texty za využití moderní výpočetní techniky dělat. Většina kapitol se týká verše, nejprve se tedy stručně seznámíme s disciplínou, která se veršem zabývá (kapitola Co je to versologie?). V další kapitole si povíme, co je potřeba udělat, aby z elektronických textů vznikl elektronický korpus, tj. takový soubor textů, který je připraven ke složitějšímu vyhledávání a k zadávání složitějších dotazů (kapitola Korpus českého verše). Poté se seznámíme s nejrůznějšími aplikacemi, které umožňují pracovat s daty z korpusu: detailní informace o českém verši 19. století poskytuje Databáze českých meter (viz stejnojmennou kapitolu); která slova se v 19. století rýmovala se slovem „láska“, ukáže aplikace Gunstick (kapitola Láska : páska); odpověď na otázku, která česká báseň je nejeufoničtější, podá aplikace Eufonometr (kapitola Dá se spočítat eufonie?); o čem je Máchův *Máj* a které podstatné jméno se v něm objevuje nejčastěji, prozradí aplikace Hex a Frekvenční slovníky (viz stejnojmenné kapitoly); a nakonec se možná dozvíte i to, jak zjistit, kdo napsal *Rukopisy...*

Co je to versologie?

Versologie (teorie verše) je literárněvědná disciplína, která se zabývá veršovanými texty. Jejím cílem je verš těchto textů popsat, klasifikovat a vysvětlit jeho vztah k ostatním složkám díla (jazykové, tematické, kompoziční atd.). Pochopení veršovaného textu se bez porozumění jeho verši neobejde – vždyť to, jakým veršem je báseň složená, zda je rýmovaná a strofická, podmiňuje autorův výběr slov.

Dejme tomu, že se autor rozhodne napsat rýmovaný sonet ve čtyřstopém jambu (vysvětlení pojmů viz například Ibrahim–Plecháč–Říha 2013). Příkladem

takového textu budiž sonet Jaroslava Seiferta (Seifert 2005, s. 109), jehož první strofa zní:

Krví se zalkla její hlína,
když drtil ocelový pás
dlaň náměstí a lámal ...
uliček pod ... Týna.

Třemi tečkami jsme označili vynechaná slova. Které slovo bude na konci třetího verše aneb co lámal „ocelový pás“? „Krk“, „šjí“, nebo „vaz“? Samozřejmě „vaz“, protože jedině toto slovo je zároveň jednoslabičné a zároveň se rýmuje se slovem „pás“ z druhého verše. A jak bude vypadat třetí verš – „uliček pod věžičkami Týna“, nebo „uliček pod věžemi Týna“? Ano, „pod věžemi“, protože jinak se nevejdeme do devíti slabik (první a čtvrtý verš, které se rýmují, mají také stejný počet slabik).

Předpokládáme, že konkrétní užití jednotlivých druhů verše (např. volba mezi trochejem a jambem) není zcela náhodné a obsahuje informaci využitelnou při celkové interpretaci díla. Podívejme se, co může versologie říct o Máchově *Máji* (1836). *Máj* jsme zvolili nejen kvůli tomu, že je známý, ale také kvůli tomu, že se v něm střídají různé druhy verše. Jejich střídání a výběr jsou dány rozhodnutím autora a úkolem versologa (teoretika verše) je navrhnout, proč se tak děje, jaký to má význam, funkci a účinek.

Nejprve malý test: Jakými verši začíná Máchův *Máj*? Pokud jste odpověděli „Byl pozdní večer – první máj...“, nemáte tak úplně pravdu. Takto začíná 1. zpěv, kterému předchází dedikační (věnovací) báseň. Dedikační báseň se od zbytku skladby výrazně liší a literární vědci se nemožou shodnout, zda byla myšlena vážně, nebo zda se jedná o ironii. Podívejme se, jak vypadá (všechny ukázky citujeme podle České elektronické knihovny – Svadbová (ed.) 2005):

Čechové jsou národ dobrý!
Nešťastný, jenž v nouzi lká,
necht se k Čechovi obrátí,
ten mu rychlou pomoc dá;
byť i Čecha nepřitelem,
nešetří Čech jeho vin.
Čechové jsou národ dobrý,
a Ty Čechů věrný syn!

Sbor

Věrný syn i bratr náš,
dobré Čechů srdce máš!

Čechové jsou národ statný,
nepřemožen český voj.
Necht se s vítězstvím rozloučí,
proti němuž český boj.

Čech kde stojí, krutá bitva
tamť, i mnohý slavný čin.
Čechové jsou národ statný,
a Ty Čechů věrný syn!

Sbor

Věrný syn i bratr náš,
statné Čechů srdce máš!

Čechové jsou národ věrný,
věrnost jejich první čest;
vlasti své i svému králi
Čech i v smrti věrný jest.
Bůh můj – král můj – vlasti moje!
poslední je Čecha vzdech.
Čechové jsou národ věrný,
a Ty syn jich – věrný Čech!

Sbor

Věrný Čech i bratr náš,
věrné Čechů srdce máš!

Věrný syn jsi Čechů kmene,
věrný bratr bratřím svým;
jazyk český je i Tobě
otců drahým dědictvím.
České hory – české doly –
české luhy – český háj –
šírá vlast – ta česká země
nejmilejší Tobě ráj.

Sbor

Věrný Čech jsi – vlastenec,
protož vděčný u věnec
květ Ti vije Čecha máj!

Dedikační báseň oslavuje český národ a její odlišnost od zbytku *Máje* se projevuje i ve formě: člení se do strof a obsahuje sborový refrén, což obojí souvisí s tím, že to je píseň (jedná se o žánr společenské písně, která se měla recitovat nebo zpívat všude tam, kde se shromáždila nějaká skupina Čechů – stejného žánru je i naše národní hymna, která vznikla ve stejné době jako *Máj*). Odlišný je i její verš: jedná se o čtyřstopé trocheje.

Proč Mácha dedikační báseň napsal právě čtyřstopým trochejem? Výběr verše podtrhuje tematiku a žánr. Čtyřstopý trochej se v té době chápal jako typicky český verš. Píseň oslavující český národ je tedy napsána českým „národním“ veršem. Žádné překvapení, naopak přesné dodržení konvence a naplnění čtenářova

očekávání. Volbou tohoto verše mohl Mácha sledovat určitý kompoziční cíl. Po zcela konvenční (tematicky i formálně) dedikační básni přijde šok: první zpěv je napsán jamby.

A nejen to: zatímco dedikační báseň byla strofická a její verš velmi pravidelný, zbytek básně je nestrofický a verše obsahují mnoho nepravidelností. A ještě ke všemu je složen jambem, veršem patrně převzatým z německé literatury a v té době poměrně neobvyklým. Kompoziční zlom nemohl být větší. Trochej a jamb byly vnímány jako protiklady a v průběhu 19. století se ustavily jejich významové opozice, takže básník generace májovců (50.–60. léta 19. století) a lumírovců (70.–80. léta 19. století) už počítal s tím, že s trochejem si čtenář spojí přídavná jména v levém sloupci tabulky a s jambem v pravém.

Trochej	Jamb
přirozený	umělý
domácí	cizí
tradiční	moderní
prostý	literární
lidový	vzdělanecký
kolektivní	individuální
spontánní	intelektuální

Významové opozice českého trocheje a jambu (19. století)

Právě zažitost těchto spojení mohla ve zvláštních případech přinést překvapivé významové efekty: nechá-li Josef Václav Sládek promlouvat své venkovské hrdiny jambem, může z toho čtenář usuzovat, že jsou prostřednictvím verše, „významově s vesnickou tematikou kontrastujícího, stavění do jedné řady s osobnostmi světové humanistické tradice i současného intelektuálního výkvětu lidstva“ (Červenka – Sgallová 1988, s. 69).

Ale zpátky k *Máji*. První zpěv začíná popisem májové krajiny, idylická atmosféra je postupně stále více rozrušována. Zatímco v přírodě vše spěje k erotickému splynutí, lidé prožívají tragédii odloučení. Nesoulad narůstá a vrcholí v promluvě-prokletí plavce ústící v Jarmilinu sebevraždu.

„Viz,“ plavec k ní
 tichými slovy šepce praví,
 „tam při jezeru vízka ční
 nad stromů noc; její bílý stín
 hlubokoť stopen v jezera klín;
 však hlouběji ještě u vodu vryt
 je z mala okénka lampy svit;
 tam Vilém myšlenkou se baví,
 že příští den jej žití zbaví.

On hanu svou, on tvoji vinu
se dozvěděl; on svůdce tvého
vražďe zavraždil otce svého.
Msta v patách kráčí jeho činu. –
Hanebně zemře. – Poklid mu dán,
až tváře, jenž co růže kvetou,
zbledlé nad kolem obdrží stán,
až štíhlé oudy v kolo vpletou.
Tak skoná strašný lesů pán! –
Za hanu jeho, za vinu svou
měj hanu světa, měj kletbu mou!“

Obrátí se. – Utichl hlas –
Po skále slezl za krátký čas,
při skále člun svůj najde.
Ten rychle letí, co čápa let,
menší a menší, až co lilie květ
mezi horami po vodě zajde.

Tiché jsou vlny, temný vod klín,
vše lazurným se pláštěm krylo;
nad vodou se bílých skví šatů stín,
a krajina kolem šepce: „Jarmilo!“
V hlubinách vody: „Jarmilo! Jarmilo!!“

Je pozdní večer – první máj –
večerní máj – je lásky čas.
Zve k lásky hrám hrdliččin hlas:
„Jarmilo! Jarmilo!! Jarmilo!!!“

V místě plavcova monologu také došlo ke změně verše: dosud vše plynulo ve čtyřstopých jambech, od této chvíle je až do čtyřveršového refrénu, který 1. zpěv uzavírá, disharmonie podtržena také užitím nepravidelných veršů (Červenka 1995, s. 8)...

Druhý zpěv začíná několika čtyřstopými trocheji. Ty už se objevily v dedikační básni, a tak se můžeme ptát, zda mají na obou místech stejnou funkci nebo význam. Řečeno obecněji: Jsou funkce nebo význam jednoho druhu verše vždy stejné? V žádném případě, nepopisujeme totiž neživou přírodu fungující podle přírodních zákonů, ale individuální výtvořiny lidského ducha, v nichž platí zákony pravděpodobnosti. Abychom se dobrali možné funkce a významu užití čtyřstopého trocheje na začátku 2. zpěvu, musíme si nejprve uvědomit, že trochejský verš byl řazen k veršům s tzv. sestupným spádem, což se odvozovalo od formy trochejské stopy, která je dvouslabičná, přičemž první slabika je důrazná a druhá nedůrazná. Právě tento fakt byl dáván do souvislosti s tematikou pasáže, v níž je tohoto verše užit: jedná se o popis padající hvězdy. Sestupný spád trochejského verše je v souladu s klesavým pohybem hvězdy (Jakobson 1938/1995: 462).

Mácha bývá někdy vykládán jako Byronův epigon (nepříliš tvůrčí následovník). Právě v oblasti verše se ovšem projevila Máchova výjimečnost. Byron sice používal jamb (a v tomto se skutečně mohlo jednat o inspiraci), ale v jeho poemách nedocházelo ke střídání různých druhů veršů. To je v žánru poémy Máchův vynález (inspirovalo ho knižní drama – Mickiewiczovy *Dziady* a Goethův *Faust* – viz Červenka 1995, s. 24). Zcela zvláštní je podoba Máchova jambu: četnost užití tříslabičných slovních celků na začátku a jednoslabičných slov na konci verše nemá v české poezii obdobu (Červenka 1998).

Několika příklady z *Máje* jsme se snažili ukázat, že při četbě veršovaného textu je nutné přihlížet k jedné z jeho nejdůležitějších a jeho specifických zakládajících vlastností – a sice k jeho verši. Toto „přihlížení“ není snadné, kromě toho, že vyžaduje některé speciální znalosti, spočívá jeho obtížnost také v nemalé intelektuální investici vynaložené na odhalení souvislostí mezi veršem a ostatními složkami díla.

Korpus českého verše (<http://versologie.cz/kcv.html>)

Chtěli jste si přečíst nějakou českou básnickou klasiku, třeba Nerudovy *Písně kosmické* nebo Máchův *Máj*, a v knihovně jste ji nenašli, protože byla vypůjčená? Zkusili jste internet, ale chtěli byste víc než nedůvěryhodné kopie, u kterých nemáte jistotu, že jsou kompletní a bez chyb? V takovém případě je tu pro vás Česká elektronická knihovna, největší virtuální knihovna české poezie na světě, kterou vytvořili odborní pracovníci Ústavu pro českou literaturu AV ČR (zdarma dostupná na <http://www.ceska-poezie.cz/cek/>). Na jednom místě je kdykoli a odkudkoli a z konce 18. století po poslední básně Viktora Dyka nebo Josefa Svatopluka Machara z první třetiny 20. století. A to vše zdarma a s různými možnostmi filtrování a vyhledávání.

Když jsme před několika lety začali přemýšlet o tom, že bychom chtěli z hlediska verše popsat všechny texty v České elektronické knihovně, bylo jasné, že vlastními silami to nezvládneme. Bude nám muset pomoci počítač. Žádný program, který by dokázal automaticky rozpoznávat české verše, neexistoval a lidé okolo nás pochy-

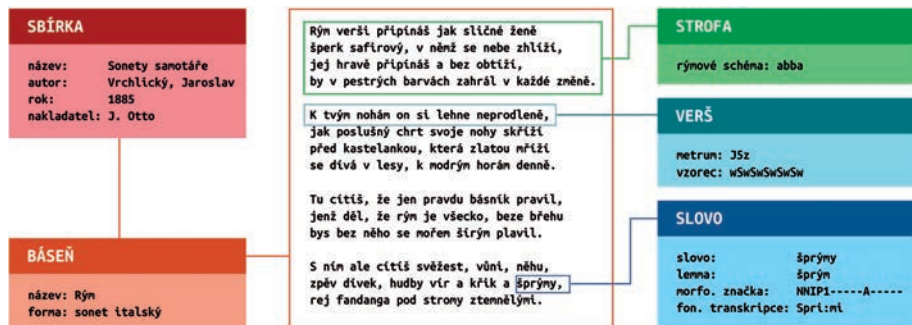


Schéma záznamu v Korpusu českého verše

bovali, že by existovat mohl. Cožpak to jde, naučit počítač odlišit čtyřstopý mužský trochej od pětistopého ženského jambu? Naštěstí jsme věděli, že v zahraničí se o automatickou analýzu verše pokouší hned několik týmů. To nám dodalo odvalu...

První verze programu (jehož autorem je Petr Plecháč), která dokázala rozlišit jamb od trocheje (daktyl, daktylotrochej a podobné vymoženosti jsme zatím nechali stranou), vznikla poměrně rychle. Program jsme pojmenovali Květa, podle Květy Sgallové, významné české versoložky, která se dávno před námi zasazovala o využití moderní techniky při analýze veršovaných textů. Pak následovala druhá verze, pak třetí atd. A začínalo být čím dál tím jasnější, že se úkol podaří. Byli jsme nadšeni, protože výhody automatické analýzy jsou obrovské. Je rychlá, nedochází v ní k chybovosti způsobené nepozorností toho, kdo analýzu provádí, všechny texty jsou zpracovány jednotně podle stejných kritérií, výsledky se dají snadno zpřístupnit a dále zpracovávat atd. Výhoda automatické analýzy spočívá také v oboustranném ovlivňování počítače a člověka: člověk počítači diktuje, co má dělat, zároveň ale počítač člověka nutí k tomu, aby se vyjadřoval přesně a aby pracoval s jasně definovanými podmínkami. Když byla hotova poslední verze programu, která už dokázala určit i daktylotrocheje s předrážkou, bylo jasné, že Květa nám splnila sen: automaticky zpracovat všech 2,5 milionu veršů České elektronické knihovny.

Zároveň s prací na Květě začal vznikat Korpus českého verše. Korpus je soubor počítačově zpracovaných textů, v němž se dají pomocí speciálních programů vyhledávat nejrůznější věci. Existuje např. korpus českého jazyka, který obsahuje miliardy slov a nabízí neuvěřitelné možnosti nejen pro jazykovědce, ale pro všechny, kteří se zajímají o češtinu (<http://ucnk.ff.cuni.cz/>). Korpus českého verše, který vychází z České elektronické knihovny, je svým rozsahem daleko skromnější (15 milionů slov). Rozdíl Korpusu českého verše oproti České elektronické knihovně spočívá v tom, že texty jsou lematizovány, foneticky, morfologicky, metricky a stroficky anotovány. Co si pod těmito pojmy představit? Ke každé slovní jednotce v korpusu je připojena informace o jejím základním slovním tvaru, kterému se říká lemma – lemma je tvar, který najdeme ve slovníku, např. lemmatem slovního tvaru „stolem“ je slovo „stůl“. Ke každému slovu je dále připojena informace o fonetickém přepisu a gramatických kategoriích, u každého verše je určeno metrum (jamb, trochej...), rozsah (kolik má stop), typ klauzule (mužská, ženská...) a metrický vzorec. Na vyšších rovinách jsou pak anotovány rýmové dvojice a pevné formy (sonet apod.).

Databáze českých meter (<http://versologie.cz/dcm/>)

Leckde se můžeme setkat s označením Máchy coby „zakladatele“ nebo „vynálezce“ českého jambu. O *Máji* se často traduje, že v něm byl poprvé v dějinách českého verše použitý jamb. Je to pravda? Odpověď můžeme nalézt pomocí on-line aplikace Databáze českých meter, která zpřístupňuje metrický a strofický popis všech textů obsažených v Korpusu českého verše. V aplikaci je možné nejen procházet jednotlivé záznamy v databázi, ale za pomoci různých filtrů také využívat statistické funkce.

Když zvolíme jako vyhledávací filtry „jamb“ a „trochej“, zjistíme, že v první polovině 19. století co do počtu básní trochej skutečně zřetelně převládá nad jambem. Situace se mění až s nástupem májovců (50.–60. léta 19. století) a v poezii lumí-

rovců (70.–80. léta 19. století) už získává zřetelnou převahu jamb. Na první pohled je ale jasné, že jambické verše se objevují i před rokem 1836, tedy před prvním vydáním Máchova *Máje*.



Poměr počtu jambických a trochejských básní v jednotlivých letech

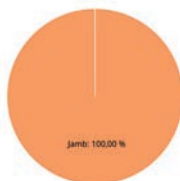
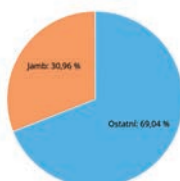
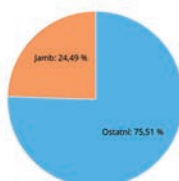
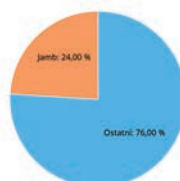
Vzhledem k tomu, že aplikace umožňuje procházet jednotlivé záznamy, můžeme snadno zjistit, že nejstarší doklady jambů v naší databázi pocházejí už z prvního almanachu Antonína Jaroslava Puchmajera *Sebrání básní a zpěvů* z roku 1795. Jambem je napsána epická báseň *Dvě sestry* z pera Šebestiána Hněvkovského a několik epigramů.

Dvě sestry kvetly v mladosti
tak krásně jako máj,
z nichž starší v marné radosti
si oblíbila ráj.

Však druhá v ctnosti, v čistotě
svůj život konala,
bydlela dlouho v samotě,
v ní rozkoš mívala.

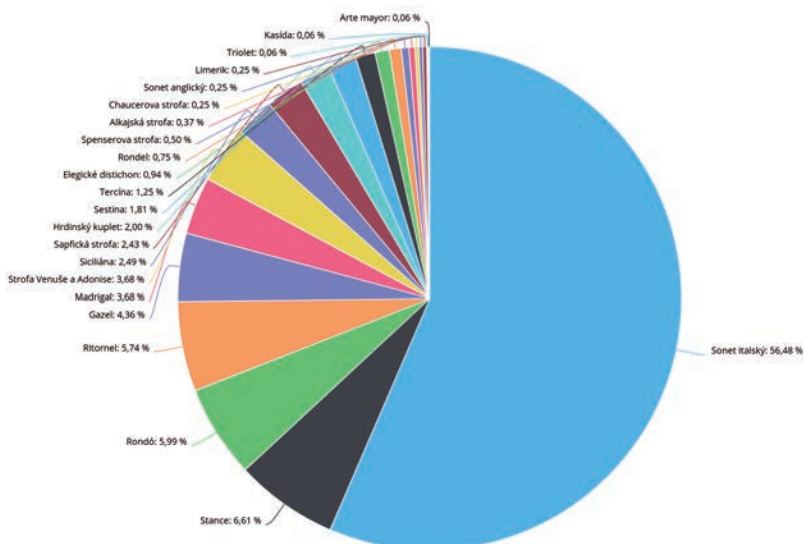
...

V úvodní kapitole jsme zmínili, že pro soudobé čtenáře Máchova *Máje* bylo užití jambu šokující. Jak se to ale slučuje se zjištěním, že české jambické verše existovaly už dávno před tím? Odpověď můžeme hledat v četnosti jejich užití. Podíváme-li se na šest knih vydaných do roku 1836, v nichž mají jambické verše nejvyšší zastoupení, zjistíme, že dva tituly výrazně vybočují. Prvním z nich je epická skladba *Zuzana Babylonská* od Bohuslava Tablice z roku 1803 napsaná výlučně jambem. Jedná se ale o tzv. alexandriny (verše zřetelně odkazující nikoliv k německé, ale k francouzské literatuře). Druhým titulem je právě *Máj* s téměř 90 % jambických veršů. V dalších dvou básnických sbírkách se už zastoupení jambů pohybuje jen okolo jedné třetiny, v posledních dvou sledovaných pouze okolo jedné čtvrtiny všech veršů.

Bohuslav Tablic: Zuzana Babylonská
(1803)Karel Hynek Mácha: Máj
(1836)Jiří Opočenský: Písně k světku tolerančním...
(1831)Václav Hanka: Dvanáctero písní 2
(1816)Antonín Josef Puchmajer: Nové básně 3
(1814)Šebestján Hrněvkovský: Básně drobné
(1820)

Šest titulů s nejvyšším zastoupením jambických veršů do roku 1836

Statistiky lze v Databázi českých meter využít také například u tzv. pevných forem (sonet, tercína atd.). Poznáte z grafu (aniž se podíváte na popisek), který z českých básníků měl tak bohatý repertoár?



Zastoupení jednotlivých pevných forem v díle Jaroslava Vrchlického

Láska : páska (<http://versologie.cz/gunstick/>)

Skládali jste někdy veršovaný text, který se měl rýmovat, a nemohli jste najít vhodný rýmový pár? Inspirace nepřicházela a vy jste začali být zoufalí? V takové chvíli by mohl pomoci rýmový slovník. Shodou okolností u nás jeden takový před pár lety vyšel (Cvrček–Cvrčková 2011). Představme si verš: „Poutá mě k tobě velká láska“, ke kterému hledáme vhodné doplnění. Kolik slov rýmuji-cích se ke slovu láska vás napadne? Abychom v tomto případě shodu považovali za rým, musí se jednat o sudoslabičná slova, která končí na ÁSKA, ASKA, ÁSKÁ, ASKÁ.

Slovník ke slovu láska (výběr omezuje výše uvedenými podmínkami, vynecháváme také vlastní jména) nabízí 23 slov (anglosaská, černovláska, hláska, hnědovláska, cháska, kráska, laská, maska, mlaská, páska, plavovláska, poukázka, praská, práská, rusovláska, samohláska, saská, sázka, sebeláska, sedmikráska, tmavovláska, vazká, vráska). Našli jste jich víc?

Slovník nám sice poradí, která slova připadají v úvahu, ale neposkytne žádné příklady z praxe, tedy ukázky toho, jak verše s těmito páry mohou vypadat a zároveň doklady toho, které z možností češtiny básníci skutečně využili.

Podrobné vyhledávání v existujících textech české poezie nabízí on-line aplikace Gunstick (pojmenovaná podle básníka prerie strýčka Gunsticka, který se objevuje v knize *Poklad na střibrném jezeře* a několika filmech o Vinnetouovi a který mluví jen v rýmech). Gunstick je aplikace, která zpřístupňuje výsledky automatické analýzy rýmových shod v básnických sbírkách obsažených v Korpusu českého verše a vydaných do roku 1920 (databáze obsahuje přes jeden milion rýmových párů). Do vyhledávače jednoduše vložíme slovo, např. „láska“, a odpovědí je seznam všech rýmových párů k tomuto slovu, které jsou obsaženy v téměř 1200 sbírkách. Dotaz ještě můžeme zpřesnit zadáním autora (zajímat nás mohou např. doklady jen u Karla Hynka Máchy) nebo období, na které se má hledání omezit (např. od roku 1830 do roku 1836). Výsledek je zobrazen formou interaktivních grafů a tabulek a ve chvílce víme, který pár byl ve hledaném období nejoblíbenější a který byl naopak na okraji zájmu. U každého výskytu rovněž najdeme jméno jeho autora, název sbírky i básně, citaci veršů, v nichž se pár objevuje, a odkaz na plný text v České elektronické knihovně.

Zkusme to třeba právě se slovem láska.

Pět nejčastěji se vyskytujících dvojic:

láska : praská (17 %)

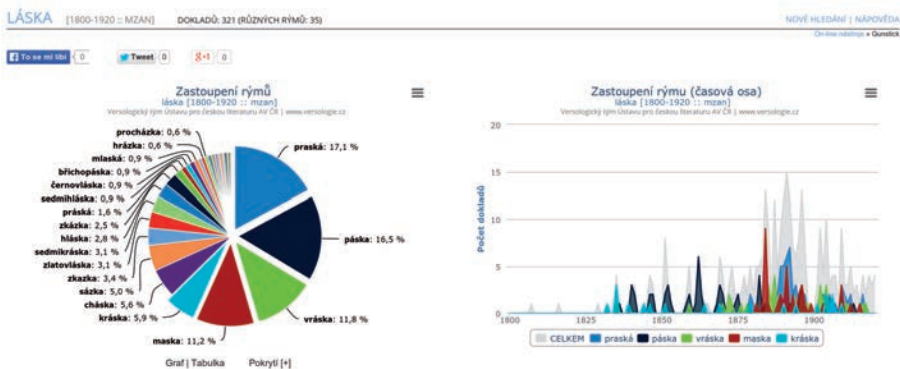
láska : páska (17 %)

láska : vráska (12 %)

láska : maska (11 %)

láska : kráska (6 %).

Pokud bychom chtěli vybírat z těch nejméně frekventovaných, pak můžeme použít výrazy láska : mlaská (1 %) nebo láska : břichopáska (1 %). Výsledky dotazu se zobrazují na časové ose, je tedy dobře vidět, jak se obliba jednotlivých párů mění.



Rýmy ke slovu láška (česká poezie 19. století)

Z databáze také můžeme zjistit, kdy a v které básnické sbírce se poprvé objevil jeden z nejoblíbenějších rýmů české poezie 19. století, a sice spojení láška : páška. Bylo to v roce 1836 v básni Čtvero kvítků od Václava Věnceslava Rába:

Nevinnost, ctnost, věrnost, láska
zdob tě v každé době,
jest to silná krásy páška,
vnadná v každé době.

Ráb ale nebyl první, kdo „lásku“ zrýmoval s „páskou“. Již v roce 1814 nalézáme v posledním Puchmajerově almanachu spojení lásky : pásky (Tu milé lásky : své vábné strojí pásky), které se později objeví i v Máchově *Máji* (bloudila blankytnými pásky : planoucí tam co slzy lásky). Obliba tohoto spojení ale nutně vedla k tomu, že postupně začalo být pocíťováno jako kliše. Graf vpravo ukazuje, že ve druhé polovině století, respektive v době nástupu lumírovců (Jaroslav Vrchlický, Julius Zeyer ad.) začali básníci upřednostňovat daleko dramatičtější a zlověstnější spojení, jako např. láska : práška (Krutá láska : struna práská, J. Vrchlický), láska : vráska (Zvěstí tvou je mír a láska : Však proč tajná kalí vráska, S. Čech) nebo láska : maska (To nebyla přec pravá láska : jen její pozlacená maska, J. Vrchlický).

Až tedy budete příště potřebovat najít nějaké zajímavé rýmové spojení, neváhejte se obrátit na strýčka Gunsticka. Zaručeně vám poradí...

Dá se počítat eufonie?

(<http://versologie.cz/eufonometr/eufonometr.html>)

Libozvučnost (eufonie) je poněkud subjektivní pojem, a tak nás možná překvapí, že se dá měřit. Abychom byli přesní: měřit se dá opakování hlásek ve verši, respektive to, zda se četnost jejich opakování vymyká poměrům běžným v normálním jazyce. První algoritmus pro měření „eufonie“ navrhl v 60. letech československý lingvista

Gabriel Altmann (Altmann 1966). Jeho pozměněná podoba je využívána aplikací Eufonometr. Eufonometr umožňuje vypočítat eufonický koeficient libovolného textu. Stačí vložit nebo napsat text (ale pozor: zadávání příliš krátkých textů nemusí vést k smysluplným výsledkům) a automaticky se zobrazí eufonicky relevantní hlásky a jejich eufonický koeficient v jednotlivých řádcích. Zároveň se zobrazí i celkový eufonický koeficient vloženého textu a autor, jehož průměrný eufonický koeficient je nejbližší eufonickému koeficientu vloženého textu. Jako malou ukázkou jsme zvolili známý úryvek z Máchova *Máje*, který bývá často uváděn jako příklad eufonického textu:

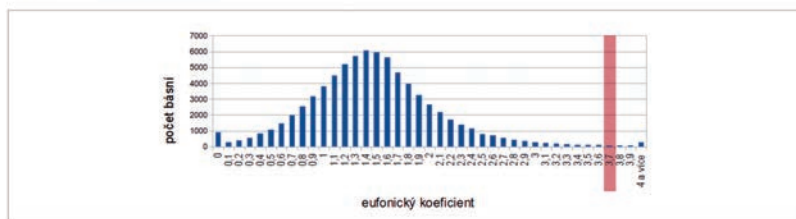
Ach zemi krásnou, zemi milovanou,
kolébku mou i hrob můj, matku mou,
vlast jedinou i v dědictví mi danou,
šírou tu zemi, zemi jedinou! –

<p>vstupní pole</p> <p>1: Ach zemi krásnou, zemi milovanou, 2: kolébku mou i hrob můj, matku mou, 3: vlast jedinou i v dědictví mi danou, 4: šírou tu zemi, zemi jedinou! –</p>	<p>SPUSTIT</p>	<p>eufonicky relevantní hlásky v jednotlivých řádcích</p> <p>1: ou=3.1 2: k=1.16 m=0.03 3: ou=2.19 d=4.95 i=0.12 4: ou=3.5</p>
---	----------------	--

celkový eufonický koeficient vloženého textu / autor s nejbližším průměrným eufonickým koeficientem

E = 3.76460668953035

Zachar, Jiří = Lešehrad, Emanuel
1. nejefoničtější autor v Korpusu českého verše
(průměrný eufonický koeficient = 2.2227)



Eufonický koeficient úryvku z Máchova *Máje*

Eufonický koeficient úryvku (ale připomínáme, že je příliš krátký) je 3,8. Pro srovnání: nejvyšší průměrný eufonický koeficient v Korpusu českého verše má básník Emanuel Lešehrad (2,2). Máchův průměr je 1,5, což stačí na 187. místo mezi více než 350 autory. Podíváme-li se na žebříček autorů, zjistíme, že vysokých hodnot dosahují symbolisté nebo autoři symbolismem ovlivnění, což potvrzuje tradovanou orientaci symbolistů na „hudebnost verše“. Žebříčky (které pochopitelně nic nevyovídají o kvalitě) nemají jen autoři, ale i texty. Podívejme se na dva texty, z nichž jeden se umístil na předních příčkách (tj. byl vyhodnocen jako velmi „eufonický“) a druhý se naopak umístil úplně dole mezi texty s hodnotou 0. Jistě poznáte, který je který:

František Leubner: Zvony, sb. <i>Na okrajích kancionálu života</i> (1897)	Josef Svatopluk Machar: Má minulost, sb. <i>Pêle-Mêle</i> (1927)
<p>Vím vám dum doby: Dumou mou duní nad Prahou zvony, a kvílí – kvílí – kvílí – Ví-li, co kvílí? – kdy v hloubi touha zdlouha napotom chvílí v nich údery bolu lká na hony stony...</p> <p>Dóm, – rudý kol západ, – hrad, – mlhy clony,... přes jezy ječí Vltava, v temnou dálku pílí, a onde ostrovů dolů sklony se do vody chýlí, kam dům, dům se zrcadlí bílý.</p> <p>Od Sionu shodou tonů tonu v rekvíem hran... Jich pádný hláhol neumlká, lká a lká v chladný vzduch a stíny honí od zvonice jak havranů hejno, hejno vran.</p> <p>Pak klekání vlání ustává pěti, – snad budí je ráno sám Pán Bůh... Noc orlice černá, od západu letí tmou – tmou – tmou – –</p>	<p>Má minulost – pár očazených hranic, kde roky moje shořely, krom upomínek nezbylo z nich pranic, a ty jsou rovny popeli.</p> <p>Z nich zžíravinu duše má jen cítí, když přehrabuje tyto kupy šedé – a člověk neví: byl přec šťasten v žití? či vzpomínat si jenom nedovede?...</p>

Ano, Leubnerův text byl vyhodnocen jako vysoce „eufonický“, Macharův jako naprosto „neefonický“.

Hex (<http://versologie.cz/hex/>)

V roce 2015 si připomeneme 600 let od upálení Jana Husa v Kostnici. Dejme tomu, že budeme chtít zjistit, jak byla osobnost Jana Husa tematizována v české poezii 19. století. Můžeme postupovat tak, že do vyhledávače v České elektronické knihovně zadáme výraz „Hus“ a projdeme všech 584 výskytů ve 227 sbírkách. To chvíli potrvá, a to jsme prošli jen výskyty výrazu „Hus“, čeká nás ještě „Husa“ (272 výskytů ve 149 sbírkách), „Husovi“ (43 výskytů ve 21 sbírkách) atd. A to nemluvíme o tom, že takový dotaz nedokáže rozlišit mezi Husem a husou... Rychlejší a smysluplnější odpověď získáme, využijeme-li aplikaci Hex (pojmenovanou podle stejnojmenného počítače v knize Terryho Pratchetta *Zajímavé časy*). Aplikace Hex byla vyvinuta k vyhledávání tzv. klíčových slov v textech obsažených v Korpusu českého verše (pokud chcete zjistit klíčová slova v nějakém jiném, třeba vlastním, textu, doporučujeme aplikaci KWords, dostupnou na stránkách Ústavu Českého národního korpusu: <http://kwords.korpus.cz/>). Klíčová slova napovídají, o čem text pojednává. Algoritmus označí jako klíčová slova taková, jejichž frekvence v daném textu statisticky významně převyšuje jejich frekvenci v celém Korpusu českého verše. Hex pracuje s lemmaty, tj. slovníkovými tvary, zadáme-li tedy do vyhledávače výraz „Hus“, nebude se vyhledávat jen tvar „Hus“, ale také všechny ostatní tvary tohoto slova: „Husa“, „Husovi“ atd. (a zároveň se nám do vyhledávání nepřiplete „husa“). Lemma „Hus“ bylo vyhodnoceno jako klíčové ve 33 textech. Výsledky jsou zobrazeny pomocí interaktivních grafů a tabulek, můžeme si je seřadit podle jména autora nebo třeba roku vydání. Z tabulky je také možné příslušný text otevřít a přečíst si ho.

Na ukázkou jsme vybrali začátek z Macharovy básně Dva soudy (sb. *Tristium Praga*, 1926):

Byl šestý červenec. Den Husův. V dílnách klid,
na četných táborech se kupil český lid
a slavil heroa svých dějin. Řečníci
na hlavy spoceně proud ohněm sršící
svých řečí pouštěli: byl znovu souzen Řím
a vesměs odsouzen – tak vděčným srdcem svým
na pouť svou výroční jde národ do dějin
za pravdou poznanou, jak šel kdys jeho syn.

Aplikace Hex ovšem funguje oběma směry: jednak umožňuje vyhledávání klíčových slov v textech obsažených v Korpusu českého verše, jednak umožňuje procházet tyto texty (jednotlivé sbírky nebo básně) a zobrazit jejich klíčová slova. Udělejme si malý test s několika Erbenovými baladami. Do vyhledávače jsme zadali Erben + Kytice (s omezením vyhledávání na podstatná jména) a ke každé baladě obdrželi seznam jejích klíčových slov. Poznáte z tohoto seznamu (neuvádíme všechna klíčová slova, abychom moc nenapovídali), které balady to jsou?

Balada 1: dítě, dveře, matka

Balada 2: stříbro, zlato, dítě, pahorek, chýže, žena

Balada 3: květ, matka, dcera, holoubátko

Balada 4: pán, matka, lov, sluha

Balada 5: hrob, milý, lampa, komora, růženec

Balada 6: jezero, voda, dceruška, dítě

Balada 7: hrad, král, oko, ruka, noha, tělo

Balada 8: děvče, půlnoc

Poznali jste balady Polednice, Poklad, Dceřina kletba, Lilie, Svatební košile, Vodník, Zlatý kolovrat a Štědrý večer?

A jestlipak poznáte, která známá česká báseň obsahuje následující klíčová přídavná jména?

bledý, růžový / kamenný, strašný, širý, pouhý, zítřejší, temný, noční, dlouhý, daleký, hluboký / širý, bílý, jediný, mladistvý, temný, růžový, rychlý, modrý, krásný / smutný, bledý, večerní, hluboký?

Uhodli jste, že se jedná o Máchův *Máj*? Lomítky jsou odděleny jednotlivé zpěvy.

Frekvenční slovníky (<http://versologie.cz/slovniky/slovniky.html>)

Doposud jsme mluvili o klíčových slovech. Trochu odlišnou perspektivu nabízí frekvenční slovník. Frekvenční slovník je seznam slov seřazených podle frekvence (četnosti výskytu) v daném textu. Na výše uvedené adrese nabízíme frekvenční slovníky všech autorů a sbírek Korpusu českého verše. K dispozici jsou jak slovníky lemmat, tak i slovníky jednotlivých slovních tvarů. Zkuste si tipnout prvních deset nejfrekventovanějších podstatných jmen Máchova *Máje*. Jak moc se vaše tipy liší od následujícího seznamu?

1. hora
2. čas
3. vězeň
4. jezero
5. kolo
6. hlas
7. láska
8. tvář
9. noc
10. den

Nebo si můžeme z hlediska frekvence slov porovnat dvě sbírky jednoho autora, např. Čelakovského *Ohlas písní ruských* a *Ohlas písní českých* (vybíráme prvních deset nejfrekventovanějších podstatných jmen):

Ohlas písní ruských	Ohlas písní českých
kůň	panna
matička	oko
země	srdce
mládec	milý
hora	hora
Volha	milá
pole	pán
město	láska
druh	kůň
otec	bůh

Deset nejfrekventovanějších podstatných jmen v Čelakovského *Ohlasech*

Jaké rozdíly mezi českou a ruskou duší z toho vyvodíte, necháme už na vás.

Kdo napsal Rukopisy?

Napsal všechny básně obsažené v *Slezských písních* Vladimír Vašek? Je autorem básní připisovaných Josefu Barákovi Jan Neruda? A kdo napsal *Rukopisy*?

Snahy o určení autorství na základě statistických metod nejsou v literární vědě žádnou novinkou. Snad nejčastěji diskutovaným problémem je autorství díla připisovaného Williamu Shakespearovi. Například už v roce 1901 americký badatel Thomas Corwin Mendenhall na základě zkoumání frekvence slov různé délky zpochybnil hypotézu, že skutečným autorem díla je Francis Bacon (jeho metoda byla později kritizována C. B. Williamsem). Mendenhall uvádí, že počítání slov zabralo tehdy dvěma najatým asistentkám několik měsíců. Od té doby se mnohé změnilo. Digitalizované texty a anotované korpusy obsahují nepřehledné množství proměnných, u nichž lze za pomoci počítače moderními statistickými metodami zkoumat, zda něco prozrazují o autorovi textu, době vzniku apod. Takovou proměnnou se může stát leccos: metrum verše (srov. kapitolu Databáze českých meter), rýmové páry (srov. kapitolu Lásky : pásky), klíčová slova nebo frekvenční distribuce lemmat (srov. kapitolu Hex a Frekvenční slovníky). Pro současnou (nejen) českou versologii tak otázky určování autorství představují jednu z velkých výzev do budoucnosti.

Literatura:

Altmann, Gabriel. The measurement of euphony. In: *Teorie verše. I. Sborník brněnské versologické konference*. Brno: Universita J. E. Purkyně 1966, s. 263–264; Cvrček, Václav – Cvrčková, Ludmila. *Velký slovník rýmů*. Praha: NLN 2011; Čapek, Karel. Máchovy kantilény. *Slovo a slovesnost* 2, 1936, s. 69–70; Červenka, Miroslav. Polymetrie v české epice 19. století. In: Miroslav Červenka – Květa Sgallová: *Z večerní školy versologie. 3. Polymetrie. Metrika překladu*. Praha: Ústav pro českou literaturu AV ČR 1995, s. 7–100; Červenka, Miroslav. Máchovo místo ve vývoji českého verše. *Česká literatura* 46, 1998, č. 4, s. 439–443; Červenka, Miroslav – Sgallová, Květa. Český verš. Sémantika metra v poezii lumírovců. In: *Słowiańska metryka porównawcza. 3. Semantyka form wierszowych*. Wrocław: Zakład Narodowy im. Ossolińskich 1988, s. 55–104; Ibrahim, Robert – Plecháč, Petr – Říha, Jakub. *Úvod do teorie verše*. Praha: Akropolis 2013; Jakobson, Roman. K popisu Máchova verše. In: Roman Jakobson: *Poetická funkce*. Jinočany: H & H 1995 (1938), s. 427–476; Mendenhall, Thomas Corwin. A mechanical solution to a literary problem. *Popular Science Monthly* 9, 1901, s. 97–110; Seifert, Jaroslav. Praha. In: Jaroslav Seifert: *Dílo*. 8. Praha: Akropolis 2005, s. 10–122; Svadbová, Blanka (ed.). Česká elektronická knihovna. Poezie 19. a počátku 20. století. Praha: Ústav pro českou literaturu AV ČR 2005; Williams, Carrington Bonsor. Mendenhall's Studies of Word-Length Distribution in the Works of Shakespeare and Bacon. *Biometrika* 62, 1975, s. 207–212.

Autoři na publikaci pracovali s podporou grantu GA ČR (P406/11/1825) a s podporou na dlouhodobý koncepční rozvoj výzkumné instituce 68378068.

Oddělení teorie ÚČL

(<http://ucl.cas.cz/cs/oddeleni/oddeleni-teorie>)

Badatelská činnost oddělení se realizuje jednak v podobě individuálních vědeckých projektů jednotlivých členů oddělení, jednak v rámci kolektivních úkolů, které rozšiřují a doplňují celoustavní výzkumné projekty.

Jedním z hlavních úkolů oddělení teorie je práce na tematickém *Slovníku literárněvědného strukturalismu*. Pracovníci oddělení se také podílejí na řešení několika mezinárodních projektů a spolupracují se zahraničními institucemi. Podstatnou součástí aktivit oddělení teorie je pořádání odborných přednášek, mezioborových kolokvií a konferencí.

Versologický tým ÚČL (<http://versologie.cz/>)

Versologický tým vznikl na konci roku 2012 a sdružuje pracovníky zabývající se teorií verše napříč ústavními odděleními. Tým navazuje na někdejší versologický tým Ústavu pro českou a světovou literaturu ČSAV, který tvořili pracovníci tehdejšího textologického oddělení (Miroslav Červenka, Stanislava Mazáčová, Milada Chlěbcová ad., externě s týmem spolupracovala Květa Sgallová). Badatelská činnost týmu se realizuje jak formou individuálních, tak kolektivních výzkumných projektů. V současné době je hlavním vědeckovýzkumným úkolem versologického týmu řešení grantového projektu *Dějiny a teorie českého verše 19. století*.

Z PUBLIKACÍ VERSOLOGICKÉHO TÝMU:

Ibrahim, Robert – Plecháč, Petr – Říha, Jakub. *Úvod do teorie verše*. Praha: Akropolis 2013. (<http://www.akropolis.info/kniha/1165/Uvod-do-teorie-verse>)

Ibrahim, Robert. *Teorie literatury: učebnice pro střední školy*. Plzeň: Fraus 2014. (<http://ucebnice.fraus.cz/teorie-literatury-pro-ss/>)

PhDr. ROBERT IBRAHIM, Ph.D.

V ÚČL AV ČR se zabývá dějinami a teorií českého verše, zejména 19. století. Patnáct let pracoval jako středoškolský učitel češtiny (dvanáct let na Anglickém gymnáziu v Praze). Jako vedoucí autorského kolektivu se podílel na publikaci *Interpretace textů (nejen) ke státní maturitě* (Akropolis, Praha 2010), s Petrem Plecháčem a Jakubem Říhou napsal *Úvod do teorie verše* (Akropolis, Praha 2013). Je také autorem středoškolské učebnice *Teorie literatury* (Fraus, Plzeň 2014).

Mgr. PETR PLECHÁČ, Ph.D.

V ÚČL AV ČR se zabývá kvantitativní a korpusovou versologií. Je autorem počítačového programu Květa a on-line nástrojů přístupných na webových stránkách versologického týmu (<http://versologie.cz>). S Robertem Ibrahimem a Jakubem Říhou napsal *Úvod do teorie verše* (Akropolis, Praha 2013).

V EDICI VĚDA KOLEM NÁS PŘIPRAVUJEME:

Ondřej Kučera: **Elektromagnetická pole živých buněk**

Radomír Vlček: **Josef Macůrek**

Milan Řípa: **Historie výzkumu termojaderné fúze**

Adam Doležal: **Bioetika**

DOSUD VYŠLO:

Pavel Žďárský: **Alois Musil**

Magdalena Bendová: **Eduard Hála**

J. Marešová, K. Mezihoráková, K. Uhlíková: **Akademické soupisy
uměleckých památek**

Ondřej Beránek: **Oriental Institute (English edition)**

Edice Věda kolem nás | Výzvy a otázky
Báseň a počítač | Robert Ibrahim, Petr Plecháč

Vydal Ústav pro českou literaturu Akademie věd ČR, v. v. i.,
Na Florenci 3/1420, 110 00 Praha 1, v Nakladatelství Academia,
Středisko společných činností AV ČR, v. v. i., Vodičkova 40, 110 00 Praha 1.
Grafickou úpravu a obálku navrhl Jakub Krč, studio Lacerta.
Technická redaktorka Monika Chomiaková. Odpovědná redaktorka
Petra Královcová. Vydání 1., 2014. Ediční číslo 11748.
Sazba a tisk **SERIFA**®, s. r. o., Jinonická 80, 158 00 Praha 5.

Další svazky získáte na:
www.vedakolemnas.cz | www.academiaknihy.cz | www.eknihy.academia.cz