

Eurohujera známe od roku 2004

S lingvistou Václavem Cvrčkem z Českého národního korpusu o češtině coby moderním jazyku pro 21. století

KATEŘINA MÁZDROVÁ / FOTO MATĚJ STRÁNSKÝ

Nejprve pro neznalé: Co vlastně jazykový korpus je?

Jde o elektronický soubor autentických textů, tištěného i mluveného slova, v jednotném formátu, aby byl dobře prohledávatelný. Korpus jazyk systematicky mapuje.

Proč se psaný a mluvený jazyk sbírá?

Velký objem takového materiálu reprezentuje to, jak skutečně jazyk používáme. Lze v něm pak vyhledávat jazykové jevy, zejména slova a slovní spojení, srovnávat je, zkoumat jejich okolí, zjišťovat trendy... Korpus je tak schopen – nejen lingvistům, ale i běžným uživatelům – říct něco o užívání jazyka, o jazykových jevech v jejich autentickém kontextu.

Proč je důležitý autentický jazyk?

Pro výzkum je třeba mít data, platí to stejně pro popis jazyka jako pro popis přírody. Zároveň je ale každý soubor textů výpovědí o nás a našem způsobu vnímání reality. To, jaký jazyk používáme, jak strukturujeme chápání světa kolem nás, v tom je každý jazyk jedinečný.

Co přesně se sbírá?

Psaný a mluvený jazyk, a to spontánně mluvený či třeba veřejné projevy. Co se týče tištěného slova, jeho sběr je snadnější, protože materiály jsou snáz dostupné. Pověšinou již existují v elektronické podobě a v korpusu českého jazyka v současnosti představují objem asi čtyři miliardy slov. Oproti tomu korpus mluveného slova je chudší, čítá okolo sedmi milionů slov, a je to logické: sběr je náročnější. Audionahrávky se musí přepisovat, kontrolovat a pro vědecké účely anonymizovat.

Jen na okraj: zakladatelem vůbec prvního lingvistického korpusu na světě, tzv. Brown Corpus, byl v roce 1964 český emigrant Henry, původně Jindřich, Kučera. Jeho korpus tehdejší americké angličtiny obsahoval sice „pouhý“ milion slov, v každém případě byl ale jeho počín přímo kolumbovský a nepředstavitelně komputačně náročný, pokud si uvědomíme, že to bylo v době, kdy počítače byly ještě v plenkách a vše fungovalo přes děrné štítky.

Jak si stojí český korpus ve srovnání s korpusy jiných jazyků?

Myslím, že v tomhle ohledu můžeme být relativně pyšní; necháme-li stranou angličtinu a němčinu jakožto



VÁCLAV CVRČEK (38)

Vystudoval bohemistiku, lingvistiku a fonetiku na FF UK. Je vědeckým pracovníkem Ústavu Českého národního korpusu FF UK. Český národní korpus je přístupný na www.korpus.cz.

korpusově nejvíc pokryté jazyky, jsme zhruba asi třetí na světě. Přitom korpus českého jazyka vznikl až v roce 1994. Mimochodem, jeho zakladatel, jazykovědec František Čermák, narazil tenkrát na silný odpor tehdejšího vedení Ústavu pro jazyk český AV ČR, kam by vědecký projekt tohoto ražení nejspíš patřil. Nakonec našel spřízněnou duši na Filozofické fakultě UK, pod kterou byl korpus začleněn a působí zde dodnes.

Nejvýznamnější hodnota

Komunisté letos v lednu podali návrh zákona, ve kterém stálo, že „český jazyk je nejdůležitějším znakem svébytnosti českého národa, nejvýznamnější hodnotou jeho kulturního dědictví a výrazem suverenity České republiky“ a že z toho důvodu zákon stanovuje, že „státním jazykem na území České republiky je český jazyk“. Návrh byl nakonec smeten ze stolu, nicméně – jsou jazykové zákony třeba?

V takových bezzubých podobách samozřejmě ne, je to podobné, jako kdybyste vytvářela zákon o počasí. Jazyk stejně jako počasí je věc, kterou nejde zcela regulovat, a regulaci nepomohou ani žádné drakonické sankce. I když se o to nejen poslanci, ale často i lingvisté – zvláště ti čeští v tom mají dlouhou tradici – snaží.

To, co obdobné zákony na ochranu jazyka přitom vesměs kritizují, jsou cizí vlivy, to znamená přejímky z jiných jazyků. Je tu snaha o jazykový purismus, o očištění češtiny od těchto „nánosů“. Nicméně ta základní představa, že jazyk obohatíte či vylepšíte tím, že z něj odstraníte nějaké množství slov, je absurdní.

Souvisí občasně snahy o jazykovou čistotu s tím, že jsme malá země, a že tedy i český jazyk je „malý“?

To je taková naše představa, že čeština je malý jazyk. Není to pravda: když se podíváte na počet mluvčích, je čeština střední jazyk.

Vstupem do Evropské unie před patnácti lety se český jazyk podle novinových titulků ocitl v „babylonu jazyků“. Co mu to přineslo? A co vzalo?

Nemyslím, že by samotný oficiální vstup představoval radikální jazykovou změnu. V Evropě jsme geograficky byli vždycky a třeba evropská legislativa na nás dopadala už před vstupem. O vstup jsme usilovali několik let, takže jsme si spoustu termínů průběžně osvojovali. Samozřejmě v souvislosti s jazykem a Evropskou unií dochází k vytváření specifického „bruselského žargonu“. Jsme víc vystaveni těmto terminologickým vlivům, na druhou stranu se to ale týká relativně malého okruhu lidí, vesměs v administrativě.

Nicméně právě ze srovnávání starších a novějších dat v korpusu psaného jazyka vyplývá, že bezprostředně po vstupu do Evropské unie stoupla četnost používání slov s ní spojených, což není nic překvapivého, bylo to dáno zájmem o evropská témata. Poté několik let frekvence uvádá, ale za poslední roky opět stoupá, což je dáno tématy, jež hýbou společností, jako migrační krize apod. Noviny prostě o EU píší častěji.

Jak jsou na tom slova jako proevropský a protievropský a jejich tvary?

U obou slov je zaznamenatelná rostoucí tendence, používáme je tedy častěji než před dvaceti lety. To lze interpretovat tak, že se názory pro a proti postupně radikalizují. V době vstupu se samozřejmě u nás mluvilo o proevropském a protievropském směřování, dnes se ale o obou

z nich píše nesrovnatelně častěji, protože zřejmě mají lidé tendenci se víc vymezovat.

V každém případě tu nová slova máme. Jednak složeniny jako euroměna, euroregion, eurorealista, eurooptimista či euroskeptický, euroatlantický a další, pak, řekněme další nová slova jako tuzemák.

Ano, složeniny s euro- tvoří velkou skupinu slov a zaznamenáváme jejich nárůst; jsou odrazem faktu, že jsme součástí Evropské unie. Jejich četnost byla velká v roce 2004, tehdy bylo všechno euro: eurookna, europárky... Dnes už nové složeniny tolik nevytváříme, ale ty etablované užíváme čím dál častěji. Vedle toho najdeme i případy složenin, jako třeba euroměna, euroatlantický, které téměř vymizely. Nebo eurorealista a eurooptimista: vrcholy měly v roce 2001, což souvisí s představiteli ODS a jejich tehdejším programem, v němž zdůrazňovali, že nejsou euroskeptici, ale eurorealisté.

Nicméně myslím, že některé složeniny bychom používali nejspíš i v případě, kdybychom do EU nevstoupili. Jen bychom třeba v souvislosti s „eurofederalistou“ nebo „euroskeptikem“ nemluvili o českém, ale německém či francouzském politikovi. No a co se týče „tuzemáku“, to je spíše jednotlivost a jde o vysloveně anekdotickou záležitost.

Kdybychom ovšem součástí EU nebyli, třeba takového „eurohujera“ bychom neznali.

Ano, to je asi specificky české, odkazuje totiž na šprta Hujera ze známé komedie, takže eurohujer v tomto kontextu je takový evropský servilní šprt. Z korpusu lze vyčíst, že se tohle slovo začalo používat poprvé v roce 2004, vrchol mělo v roce 2014.

Je čeština v éře globalizace moderní, progresivní jazyk pro 21. století?

Rozhodně. Všechna slova, která přejímá z cizích jazyků spolu se schopností je adaptovat, svědčí o tom, že lidé neztrácejí důvěru v češtinu coby komunikační prostředek, který je vhodné využívat. Adaptovaná slova jako „googlovat“ a „lajkovat“ jsou totiž jasným dokladem toho, že čeština je prostě živým jazykem, který se neustále vyvíjí a reaguje na potřeby mluvčích. Protože pokud by se jazyk nevyvíjel, byl by oprávněně považován za mrtvý.

Kdesi jste zmínil, že je to nejen moderní a funkční jazyk, ale také že není v ohrožení. Co jste tím myslel?

Řekl bych to takhle: právě jazykový purismus, všechny ty snahy o „čistotu jazyka“, třeba i pomocí zákonů, vždy pramení z pocitu ohrožení ve smyslu politickém. Není tomu tak jen u nás, ale všude ve světě. Největší vzruch obdobných snah jsou vesměs v obdobích politických nebo kulturních útlaků, za okupace, v letech nesvobody kvůli politickému systému... V momentě, kdy tu žádné takové ohrožení není, tyto snahy okamžitě poleví a vyskytují se víceméně ojediněle. Proto se ty převážně komunistické návrhy jazykových zákonů nesetkávají s větší podporou, protože lidé rozhodně necítí žádné ohrožení ve smyslu ohrožení národa. Což je skvělá zpráva. ☺