

ROMI – první rozsáhlá databanka romského etnolektu češtiny

Jazykové korpusy slouží především k zachycení podoby jazykové komunikace vytvářené na základě různých kritérií. Těmi může být vymezení skupiny jejich uživatelů, časové období, konkrétní jazyková forma (např. psaný projev žáků v korpusu SCRIPT, přepis nahrávek vyučovacími hodinami v korpusu PEDAGOG, novinové články v korpusu PDT atd.). Jazykové databanky jsou primárně zamýšleny jako zdroj dat pro zkoumání jazykového komunikátu z lingvistického hlediska, případně k uchování jazykových projevů k výzkumům pozdějším. Řada jazykových korpusů ale představuje cenný zdroj bádání i v jiných oblastech, například k analýze sociologické, diskurzní, etnografické, genderové a řadě dalších.

Mezi českými korpusy s velkým výzkumným potenciálem vyniká aktuálně budovaný korpus ROMI, který vzniká jako subkorpus rozsáhlé databanky CZESL, jejímž cílem je zachytit češtinu cizinců a Romů.¹ Databanka má sloužit především pro pedagogické účely – v první řadě jako zdroj pro analýzu jazykových kompetencí těchto skupin uživatelů češtiny a její využití ve výuce: ke zmapování hlavních

problémů,² při sestavování cvičebnic, při explanaci konkrétních jazykových jevů, vyhledání relevantních příkladů a podobně. Tomuto zaměření odpovídá jednak uživatelská přístupnost databanky (snadné vyhledávání, široce uživatelsky zaměřený přepis nahrávek a textů ad.), jednak věkové vymezení mluvčích a další zohledňované parametry zaměřené na školní prostředí.

ROMI představuje naprosto ojedinělý, rozsáhlý soubor textů a nahrávek romských mluvčích češtiny od předškolního věku do cca 26 let, který přináší poprvé v dějinách české lingvistiky a romistiky takto rozsáhlý jazykový materiál romského etnolektu. ROMI bude přístupný ve formě přepisu textů a nahrávek širší odborné veřejnosti (podobně jako např. Český národní korpus), nahrávky a originály textů pak v určitých případech (např. za účelem fonetického rozboru). Jazykový materiál je unikátní nejen svým rozsahem (k 17. 3. 2011 obsahuje 2 466 písemných textů a 497 zhruba 10-20minutových nahrávek, přičemž sběr dat probíhá od října 2009 a pokračovat bude cca do října 2011), ale také celorepublikovým zaměřením (dosud jediná systematická studie romského etnolektu M. Bořkovcové [*Romský etnolekt češtiny*. Signeta, Praha 2006]) se zaměřuje v první řadě na jednu kon-

¹ Databanka CZESL vzniká v rámci projektu *Inovace vzdělávání v oboru čeština jako druhý jazyk*, jenž je podporován ESF (reg. číslo CZ.1.07/2.2.00/07.0259). Na jeho realizaci se podílejí Technická univerzita v Liberci, Univerzita Karlova a Asociace učitelů češtiny jako cizího jazyka (více viz www.c2j.cz); ROMI vzniká v Ústavu českého jazyka a teorie komunikace FF UK a podílejí se na něm především Karel Šebesta, Kateřina Šormová a Zuzanna Bedřichová.

² Podstatným impulzem pro vznik ROMI byla potřeba zachytit hlavní problémy jazykových projevů romských žáků, neboť obtížné zvládnutí vyučovacího předmětu „český jazyk“ bývá uváděno jako nejčastější příčina přesunu romských žáků na základní školy praktické, viz např. studie GAC *Vzdělanostní dráhy a vzdělanostní šance romských žáků a žáků základních škol v okolí vyloučených romských lokalit*, Praha 2009. Dostupné z www.gac.cz/.../GAC_Vzdelanostni_drahy_a_sance_romskych_deti.pdf.

krétní komunitu obývajících v době výzkumu pražský Smíchov) a pestrostí zkoumaných prostředí. Do projektu se zapojila řada základních škol všech typů (běžné základní školy, základní školy speciální a praktické), ale také řada individuálních spolupracovníků z neziskových organizací romských i neromských, romští pedagogičtí asistenti i individuální výzkumníci. Jako unikátní zdroj nejen jazykových dat se osvědčila spolupráce s jedním romským sdružením, jehož členové dosud nahráli několik desítek nahrávek přímo ve své komunitě (a v nahrávkách pokračují). Cílem projektu totiž není jen shromáždit materiál „zvenku“, ale zapojit do vybudování databanky i romské mluvčí.

Přestože na okolnosti sběru jazykového materiálu nebyly kladeny žádné specifické požadavky (naopak, záměrem bylo shromáždit materiál v co nejširším spektru, od školních slohových prací po polořízené rozhovory), řada „sběračů“ (tazatelů) kladla svým respondentům (často i systematicky) otázky, které představují cenné informace sociologického a etnografického charakteru, od znalostí romštiny a jejího užívání po projevy diskriminace, popis prostředí, v němž respondenti žijí a s nímž jsou konfrontováni, chování pedagogů ve škole, způsob slavení různých svátků, oblíbenou hudbu a řadu dalších.

Všechny nahrávky a texty přinášejí základní informace o respondentech (věk, pohlaví, třída, pokud navštěvují školu), většina z nich, především texty, ještě řadu dalších parametrů: informace o znalosti romštiny a okolnostech jejího užívání v okolí respondenta, o typu školy ze sociologického hlediska a podle zřizovatele, o velikosti obce, o nářečním území, kde respondent žije, o sdružení, pokud sběr materiálu

probíhal v rámci jeho působení, o tom, zda jde o sociálně vyloučenou lokalitu,³ a také upřesňující informace o vzniku jazykového materiálu (zadané téma, časový limit apod.). Při přepisu nahrávek byl záměrně zvolen uživatelsky přístupný přepis (na základě pravidel pro přepis mluveného projevu v ČNK), jenž zachycuje alespoň základní výslovnostní prvky romského etnolektu, především výrazný přízvuk a krátké/dlouhé kvantity. Již v této fázi budování korpusu je zřejmé, že jedním z nejpřínosnějších výsledků z jazykového hlediska je vztah mezi psaným a mluveným projevem, který se projevuje výraznou disproporcí. Zatímco psané projevy se vyznačují velkou chybovostí (psaní několika slov dohromady, chybějící interpunkce, neohraničování vět apod.) a stručností, mluvený projev je naopak často z jazykového i obsahového hlediska bohatý. Způsob přepisu obou typů projevu umožňuje snadnou orientaci v textu, díky níž je možné korpus využívat k řadě jiných než čistě jazykových analýz. Jako příklady uvedme byť jen pár z mnoha možných okruhů bádání: první představuje etnografické zkoumání popisu slavení různých zvyků a významných událostí v životě rodiny/komunity (Vánoce, křest, v tomto případě Velikonoce, ukázky č. 1 a 2), dokumentace trávení volného času (ukázka č. 5), diskurzivní analýza (ukázka č. 3 a 4) a způsob, jakým se členové komunity vymezují vůči svému neromskému okolí (ukázka č. 6, 7, 8).

ROMI by měl být dokončen do konce roku 2011 a doufáme, že jeho zveřejnění v roce 2012 vyústí nejen v jeho využití

³ Na základě mapy sociálně vyloučených lokalit vytvořené GAC pro MPSV, dostupné z <http://www.esfcr.cz/mapa/index.html>.

k jazykové analýze či pedagogické činnosti, ale bude provázáno i zájmem ze strany badatelů v mnoha dalších, prosociálně orientovaných disciplínách.

Ukázka č. 1 (písemný projev): Co (D) dělám o|a velikanocích

Tak kdycki chodíme prni naku[p s tečkou] ovat vajíčka šusenki čokolady. Beranka vyno, a Tata|tata z Mamínkou nam daj peníze aby sme sly ponlaski. a v patek nejí ježíš vtal z mrtvích a v nedeli chodime do kostela vajíčka dame co mame a vsecko do košíčku a hlaně jak je ježíš|jožíš št[a]sní ze svojí rodínou a jak sou zdravý a stasný<°>

Ukázka č. 2 (písemný projev): Co znamená jí Velikonoce

Velikonoce znamenají že ježíš kristus byl v pátek ukřižovaný a v neděli|neděli vstal z mrtvích a na velikonoce se doržuje křesťanský svátek a ten svátek se musí dodržovat a na ten svátek se podává beránek chlebičky nebo vajíčka malovaný no tak todle znamenají velikonoce.

Ukázka č. 3 (písemný projev): Rodina – fungující

máma Tak máma musí vařit a uklízet ale musí i tu|ty děti ji pomáhat a poslouchat aby ta máma se nezhroutila nebo abu|aby se jí něco nestalo nebo nezamotala se jí hlava a taky ten táta musí pomáhat aby ta máma nebyla nato sama protože kdyby te|ka kdyby chodí děti do školi tak ten táta musí pomáhat musí uklízet a nebo luxovat nebo něco takového AO a když děti přijdou XXX ze školi tak taky musí pomáhat taky luxovat nebo zametat nebo něco takového no tak todle je rodina.

Ukázka č. 4 (písemný projev): Rodina – nefungující

Rodina nefungující je taková že třeba když táta bereti peníze tak de to prochlastat a takovy no třeba ti děti neposlouchaj nebo se perou v druhém pokoji nebo kdyžta máma řekne a dost tak voni neposlouchaj tak Ti spolubydlíci to sliši a takového a třeba ti sousedi můžou taky zavolat na socialku a tu|ty děti můžou sebrat do dětského domova no tak todle je nefungující Rodina. Nebo děti seperou ve škole astoho je největší průser třeba přijedou policie tak co si myslíte budou ti rodiče šťastný nebudou samozřejmě nejsou no tak maj ti děti zaracha|zuracha a takovi no tak tak nefungující rodina i ti děti.

Ukázka č. 5 (písemný projev): Jak sem prožil víkend

V sobotu sem hrál hru na počítači a pak sem šel z bratrancem na Formanku a tam sme hráli kopanou a fobbal a pak jsme šli na karlovo náměstí a pak jsme jeli na michle domu k bratrancovi a tam jsme hráli novou hru závodní a hráli jsme do 10:30 hodin no a pak jak sme vstali tak jsme šli k druhému bratranci na Černý most<ni> XXX sami sme jeli no a pak sem šel domů.

Ukázka č. 6 (mluvený projev):

<0> co romština, ovládá jí dobře už?

<1> ano.

<0> mluvíš i doma?

<1> ano, s babičkou, s mámou, s tátou.

<0> a venku?

<1> taky.

Ukázka č. 7 (mluvený projev):

<0> hmm, jaký, jak se k vám chovali vaši dobrovolníci?
 <1> dobře. akurát ty, co byli v tom bazenu, byli úplee hnusný.
 <0> to sou který? to sou ty bílý nebo ty černý?
 <1> ty bílý. šli sme do šatny a chtěli sme se nasvačit a voni na nás nadávali, že jim tam našlapem, potom, dyž to utřem, tak na tom budou čmouhy.
 <0> hmm, a jenom tohle, jo?
 <1> no. a potom kamarád to tam musel vytírat.
 <0> a jinak vám nenadávali, jenom tak, jenom todlencto?
 <1> no, a potom, že nesmíme skákat ze stran, že to je zakázaný, nesmíme se honit ...
 <0> hmm.
 <1> ...topit se.
 <0> a co nějak vulgárně vámš tam nenapadali?
 <1> ne.
 <0> dobře. jak se vám teda líbili plavčík. takže říkáš, že plavčík byl v týdle sortě zlej.
 <1> byl.
 <0> čím se tě nejmíc plavčík dotknul?
 <1> (dlouhá pauza) ještě jednou prosím?
 <0> čím se tě nejmíc dotknul?
 <1> hmm, že na nás přiš*, že sem spadl ze strany, skočil sem do vody ...
 <0> hmm.
 <1> ... a von hned přišel ke mně a začal na nás nadávat ...
 <0> jak vám nadával?
 <1> ... že nemáme skákat. ještě jednou skočíš ze strany a pudeš si sednout na lavičku.
 <0> hmm, a ještě něco se ti tam nelíbilo? v tom bazenu teda.
 <1> ne, jinak šechno dobrý.
 <0> a co teda ty romský dobrovolníci?
 <1> a ty byli dobrý.

Ukázka č. 8 (mluvený projev):

<0> zeptám se trošku jinak, hele, máš i bílý kamarády nebo jenom ty tmavej?
 <1> jenom v romských dětech.
 <0> jenom v romský komunitě. proč se třeba neskamarádíš s i s těma bílejma?
 <1> (dlouhá pauza) ne, ptože třeba chceme si zazpívat romsky a voni to neuměj.
 <0> hmm.
 <1> voni zpíva furt jenom disko takovýto.
 <0> a nemyslíš si třeba Honzo, že ty český děti by nechtěli se naučit náš jazyk romskej.
 <1> chtěli, ale třeba Mar* paní Nováková udělala tedkon ten kurz romštiny.
 <0> hmm.
 <1> a tam český děti choděj.
 <0> no.
 <1> já sem tam byl jednou ale je to hezký.
 <0> ale co když třeba chtějí i ty starší, větší děti, který si myslím, že už nechoděj se naučit tu, ten náš jazyk, naši kulturu a naše písne?
 <1> tak ať tam přijdou a my je to naučíme.
 <0> co když třeba se chtěj s váma kamarádit venku?
 <1> no tak ať s náma choděj ven.
 <0> přimete je?
 <1> přimeme.

*Zuzanna Bedřichová, Karel Šebesta
& Kateřina Šormová*