vzťahu k národnému jazyku, čo zároveň predpokladá pestovanie pozitívneho vzťahu k spisovnému jazyku (Findra, 2001, s. 93). Cieľom i výsledkom by mala byť schopnosť občana tvoriť diferencované texty vzhľadom na špecifiká komunikačných sfér a komunikačných situácií. Vyplýva z toho kvalitatívna rovnocennosť jednotlivých variet, ktorá nie je závislá od ich teritoriálneho rozšírenia a funkčnej diferenciácie.

**Literatúra**

FIELD, Fredric W.: Linguistic borrowing in bilingual context. Amsterdam: Benjamins 2002. 260 s.

FINDRA, Ján: Jazyk v komunikácii: dynamika vývinu, či ohrozenie? In: Slovenská reč, 2001, roč. 66, č. 2, s. 91 – 107.

LEHISTE, Ilse: Pojem interferencie. In: Antológia bilingvizmu Ed. J. Štefánik. Bratislava: Academic Electronic Press 2004, s. 82 – 87.

MÚCSKOVÁ, Gabriela: Nárečie a spisovný jazyk v bežnej hovorenej komunikácii obyvateľov Gelnice. [Dizertačná práca]. Bratislava: Filozofická fakulta Univerzity Komenského 2006. 242 s.

ŠTEFANKO, Ondrej: Ľudová kultúra Slovákov v Rumunsku. In: Atlas ľudovej kultúry Slovákov v Rumunsku. Zodp. red. O. Štefanko. Nadlak: Vydavateľstvo kultúrnej a vedeckej spoločnosti Ivana Krasku 1998, s. 29 – 53.

## Metadáta v Slovenskom hovorenom korpuse

Katarína Gajdošová

Slovenský národný korpus,
Jazykovedný ústav Ľ. Štúra Slovenskej akadémie vied, Bratislava

Oddelenie Slovenského národného korpusu Jazykovedného ústavu Ľ. Štúra SAV v Bratislave (ďalej SNK) poskytuje svojim používateľom predovšetkým databázu písaných textov (http://korpus.juls.savba.sk), momentálne v rozsahu 339 miliónov tokenov – textových jednotiek. Súčasne s mnohými inými aktivitami sa v oddelení SNK začína utvárať hovorený korpus slovenčiny. Cieľom projektu Slovenského hovoreného korpusu (ďalej SHK) je priniesť používateľom SNK dostatočne rozsiahly materiál na výskum hovorenej podoby slovenčiny. Jadrom korpusu budú spontánne nahrávky z celého územia Slovenska.

V našom príspevku sa zameriame na metadáta pripravovaného hovoreného korpusu slovenčiny a porovnáme ich s inými podobne zameranými projektmi. Východiskom metadát navrhovaných v SHK sú klasické sociolingvistické parametre, avšak vzhľadom na špecifiká pripravovaného projektu a jeho ciele obsahujú niektoré osobitné údaje.

V roku 1963 sa na Slovensku začal a v roku 1965 dokončil výskum, ktorý poznáme pod názvom Hovorená podoba spisovnej slovenčiny. Výskum bol zameraný na jazyk miest Slovenska a uskutočňoval sa formou nahrávania na kotúčové magnetofónové pásky a dotazníka distribuovaného medzi respondentov tam, kde sa nahrávky realizovali. Celý projekt bol zastrešený Slovenským jazykovedným združením (dnešná Slovenská jazykovedná spoločnosť pri SAV) na čele s Eugenom Paulinym. Východiská, ciele a ich naplnenie bilancovala rovnomenná konferencia z roku 1965. Písomné materiály z nej boli prístupné len interne v rámci Združenia a po viac ako 40 rokoch boli oficiálne vydané tlačou v roku 2007 (Hovorená podoba spisovnej slovenčiny, 2007).

Keďže projekt bol zameraný na mapovanie hovorenej podoby spisovnej slovenčiny v mestách, aj jeho metódy odrážali tento cieľ. Okrem informácií o respondentovi zaznamenávali explorátori aj informácie o meste, v ktorom sa robili nahrávky. Metadáta spojené s výskumom hovorenej podoby spisovnej slovenčiny môžeme zaradiť do štyroch skupín (Matejčík, 2007, s. 25):

1) *Informácie o respondentovi*:
   a) vek: 30 – 40 rokov – táto hranica bola stanovená z jednoduchého dôvodu: *„V tomto veku je človek najviac jazykovo ustálený“* (Matejčík, tamže);
   b) vzdelanie a ne/záväznosť používania spisovného jazyka:
      - základné vzdelanie – nepovinné používanie spisovného jazyka; respondentmi boli továrenskí robotníci, predavačky v obchodoch, údržbári, kaderníčky, upratovačky, zamestnanci komunálnych služieb;
      - stredné vzdelanie (maturita) – pasívny kontakt so spisovným jazykom; medzi respondentov boli vybraní pracovníci v administratíve, zdravotníctve a školstve;
      - vyššie vzdelanie (maturita alebo vysoká škola) – povinné používanie spisovného jazyka na pracovisku; vybraní respondenti vykonávali povolania ako lekári, umelci, redaktori, inžinieri, učitelia a právnici.
2) *Informácie o meste (mieste uskutočnenia zvukových záznamov)*:
   a) údaje o starobylosti mesta,
   b) údaje o hospodárskom (zväčša priemyselnom) rozvoji,
   c) údaje o vzraste obyvateľstva za posledných 20 rokov, údaje o prílieve ľudí z okolitých i vzdialenejších území a podobne.
3) *Inventarizácia (nielen) jazykových javov mesta*:
   a) údaje o sociálnom zložení obyvateľstva,
   b) údaje o jazykových oblastiach ľudí dochádzajúcich za prácou do mesta a odchádzajúcich za prácou z mesta,
   c) historické údaje o kolonizácii mesta,
   d) údaje o profesiách obyvateľov mesta v minulosti a súčasnosti,
   e) údaje o sociálnych vrstvách mesta,

f) údaje o hláskosloví (jeho inventári v meste), lexike, morfológii, syntaxi,
g) údaje o skupinách jazykových javov (napr. výslovnosť hlások *ä, ľ*),
h) príp. údaje o jednom jave (výslovnosť hlásky *ľ* vo všetkých pozíciách).

4) *Informácie o prejave.* Každý respondent musel rozprávať plynule alebo prostredníctvom odpovedí na otázky na dve témy. Prvá – povinná – téma bola spoločensky záväzná. Druhá téma bola ľubovoľná. Výpovede respondenta o oboch témach trvali spolu 20 minút.

Žiaden iný spomenutému výskumu podobný komplexný projekt sa na Slovensku dodnes nerealizoval. V zahraničí sú hovorené korpusy jednotlivých jazykov známe už celé desaťročia. Z mnohých spomenieme výberovo len štyri, pričom sa v rámci témy nášho príspevku sústredíme na ich metadáta.

Britský národný korpus (http://www.natcorp.ox.ac.uk/) obsahuje 90 % písaných textov a 10 % textov je hovorených. Štruktúru hovorenej časti Britského národného korpusu tvoria nahrávky formálnych obchodných a vládnych stretnutí, ktoré sa realizovali priamo v rádiách alebo sa prezentovali v rádiách prostredníctvom telefonických rozhovorov účastníkov diskusií. Respondenti boli delení vyvážene na základe veku, bydliska, sociálneho postavenia.

Český národný korpus ponúka v súčasnosti tri hovorené korpusy češtiny: Pražský hovorený korpus, Brniansky hovorený korpus a korpus ORAL2006 (všetky dostupné na www.korpus.cz). Štruktúra metadát Pražského hovoreného korpusu a Brnianskeho hovoreného korpusu je totožná. Skladá sa z kategórií:

a) pohlavie: muž – žena;
b) vek: mladší (nad 20 rokov) – starší (nad 35 rokov);
c) vzdelanie: nižšie (ZŠ, maturita) – vyššie (vysokoškolské);
d) situácia, v ktorej bola nahrávka realizovaná: formálna (riadený rozhovor) – neformálna (spontánne prehovory).

Dôvodom na zvolenie týchto kategórií bola jednoduchosť spracovávania dát pre vyhľadávanie prostredníctvom korpusového manažéra Manatee s klientom Bonito a možnosť zdanlivej vyváženosti korpusov (k zložitosti výpočtov porovnávajúcich konkrétny jazykový jav alebo prostriedok pri nerovnakom naplnení záznamov v jednotlivých kategóriách porov. Esvan, 2006, s. 97).

V súčasnosti najaktuálnejší hovorený korpus češtiny – ORAL2006 – mierne modifikuje sociolingvistické parametre zaznamenávané v predchádzajúcich hovorených korpusoch češtiny. Tvoria ho kategórie:

a) pohlavie: muž – žena,
b) vek: do 35 rokov – nad 35 rokov,
c) vzdelanie: základné – stredoškolské – vysokoškolské (aj začaté).

Pridané informácie o respondentoch sú:

a) presný vek,
b) vzdelanie,
c) nárečová oblasť, v ktorej respondent žil v detstve.

Všetky nahrávky, ktoré tvoria hovorený korpus ORAL2006, vznikli nahrávaním spontánnych prejavov respondentov, sú teda neformálne a na rozdiel od predchádzajúcich korpusov sa tvorcovia ORAL2006 snažia dokumentovať celé územie ČR.

Korpus DIALOG (http://www.ujc.cas.cz/oddeleni/index.php?page=DIALOG) je českým korpusom verejných záznamov televíznych relácií, ktoré sa realizovali ako rozhovory, debaty, polemiky, diskusie, talkshow so známymi osobami spoločenského života. Korpus tvoria prepisy nahrávok, ktoré sa spracúvajú v Ústave pro jazyk český AVČR v Prahe s technickou spoluprácou Ústavu formálnej a aplikovanej lingvistiky MFF UK v Prahe. Zo sociolingvistického hľadiska sa v korpuse DIALOG uvádzajú len tieto informácie:

a) meno účastníka diskusie,
b) pohlavie (dopĺňané do hovoreného korpusu technicky, čo pri niektorých menách typu Janků nie je presné),
c) účastník diskusie – moderátor,
d) funkcia hovoriaceho (čo má napísané v titulku vedľa svojho mena na obrazovke, ako bol predstavený).

**Metadáta v SHK**

Koncepcia pripravovaného hovoreného korpusu slovenčiny počíta s rozšírením tradičných sociolingvistických parametrov o niektoré pre výskum významné zložky či ich časti.

Skupinu **demografických faktorov** tvorí delenie respondentov podľa pohlavia (muž – žena). Ďalšou zložkou je vek účastníkov rozhovoru. Na rozdiel od iných korpusov, najmä českých, sme zvolili bohatšie škálovanie vekových skupín:

a) 20 – 30;
b) 31 – 40;
c) 41 – 50;
d) 51 – 60;
e) 61 – 70;
f) 71 – 80.

Podrobnejšia diferenciácia vekových kategórií je pre výskum výskytu niektorých javov v hovorených prejavoch potrebná, na základe mnohých lingvistických ohlasov priam nevyhnutná (porov. napr. Esvan, 2006, s. 96 – 99). Realizácia takéhoto triedenia nie je o nič náročnejšia ako iná, no nesie v sebe v kombinácii s inými zložkami sociolingvistických informácií väčší potenciál výpovednej hodnoty, na druhej strane však aj väčšie nároky pri technickom spracovaní analýz.

Kategóriu vzdelanie delíme na 4 skupiny:

a) vysokoškolské,
b) stredoškolské,
c) odborné,
d) nižšie.

Na základe skúseností z iných sondážnych výskumov sa nazdávame, že osobitnú výpovednú hodnotu má okrem tradičného členenia na nižšie, stredoškolské a vysokoškolské vzdelanie aj položka odborné vzdelanie. Do nej patria respondenti s učňovským vzdelaním alebo stredoškolským vzdelaním bez maturity.

**Geografická zložka** metadát má tri časti:

a) miesto narodenia,
b) miesto najdlhšieho pobytu,
c) miesto súčasného pobytu.

Všetky tri údaje sú pre lingvistický výskum dôležité, keďže v prejave respondenta sa môže odraziť nárečovej oblasti jeho najdlhšieho pôsobiska, v prípade sťahovania aj miešanie dialektov rôznych oblastí jeho pohybu. Súčasťou sociolingvistických údajov bude aj informácia o ne/aktívnom používaní konkrétneho nárečia respondentom.

Z hľadiska získavania materiálu bude zaujímavá aj informácia o respondentoch, ktorí majú slovenskú štátnu (príp. národnú) príslušnosť, ale ich materinským jazykom nemusí byť slovenčina napriek tomu, že ju ovládajú. Rovnako to môžu byť respondenti, ktorí sú prirodzene bilingválni.

Na výslovnosť a rečový prejav hovoriaceho môže do istej miery vplývať aj ovládanie cudzieho jazyka, príp. viacerých cudzích jazykov. Preto bude jednou z položiek metadát hovoreného korpusu slovenčiny aj tento údaj, ako aj informácia o prípadnom dlhšom pobyte respondenta v zahraničí.

Medzi ďalšie zložky metadát patrí informácia o mieste získania nahrávky, ako aj spresnenie týkajúce sa prostredia, kde nahrávka vznikla (napr. trieda v škole, ulica, čajovňa...), a času (dátum) realizácie nahrávky.

Relevantné zložky metadát SHK ďalej obsahujú informácie:

- privátnosti – neprivátnosti prehovoru,
- pripravenosti – nepripravenosti prehovoru,
- známosti – neznámosti účastníkov prehovoru,
- rovnocennosti – nerovnocennosti partnerov rozhovoru,
- formálnosti – neformálnosti prejavu,
- počte respondentov (aj moderátorov) v prejave (nahrávke), napr. dialóg, trialóg...
- téme rozhovoru (ne/určená) – ak je to zachytiteľné, téma sa konkretizuje.

Tzv. doplnkovú komentárovú časť metadát hovoreného korpusu slovenčiny budú tvoriť informácie o javoch a okolnostiach, ktoré akýmkoľvek spôsobom môžu ovplyvňovať prejav respondenta. Tie informácie vyplní explorátor sám na základe rozhovoru s respondentom. Pôjde o informácie typu piercing na jazyku, šušlanie ako fyziologická porucha, vyrastanie v hluchonemej rodine, hrubý hlas, chrčanie ako fyziologická porucha a pod.

Metadáta SHK poslúžia používateľom na získanie relevantných informácií o jazykovej situácii respondenta, ako aj na porovnávanie realizácie jazykových

javov v rozličných skupinách hovoriacich. Dynamiku slovenčiny tak bude možné skúmať nielen v rámci základného delenia respondentov podľa pohlavia, ale aj s dôrazom na ďalšie špecifické informácie zachytené v podobe predstavených metadát. V porovnaní s výskumom Hovorenej podoby spisovnej slovenčiny zo 60. rokov minulého storočia štruktúra SHK bude obsahovať nielen „mestské nahrávky“, ale aj mediálne prejavy, diskusie a polemiky, ako aj formálne prejavy či neformálne rozhovory. Podrobnejším štruktúrovaním metadát sa snažíme vyjsť používateľom SHK v ústrety tak, aby im ponúkaný materiál poslúžil čo najviac.

**Literatúra**


Britský národný korpus. Dostupný z WWW: http://www.natcorp.ox.ac.uk/

Český národný korpus. Dostupný z WWW: korpus.cz

ESVAN, Françoise: Srovnávací rozbor mluvených korpusů (PMK a BMK): metodologické problémy a první výsledky. In: Korpusová lingvistika: Stav a modelové přístupy. Ed. F. Čermák – R. Blatná. Praha: Ústav Českého národního korpusu 2006, s. 95 – 117.

Hovorená podoba spisovnej slovenčiny. Ed. S. Ondrejovič. Bratislava: Veda 2007. 228 s.

Korpus DIALOG. Dostupný z WWW: http://www.ujc.cas.cz/oddeleni/index.php?page=DIALOG

ČERMÁK, František: Mluvené korpusy. In: Korpusová lingvistika: Stav a modelové přístupy. Ed. F. Čermák – R. Blatná. Praha: Ústav Českého národního korpusu: 2006, s. 53 – 67.

MATEJČÍK, Ján: K metóde výskumu. In: Hovorená podoba spisovnej slovenčiny. Ed. S. Ondrejovič. Bratislava: Veda 2007, s. 25 – 33.


## Ochranná známka (s ohľadom na chrématonymá mliekarenského priemyslu)


Lenka Garančovská

Katedra slovenského jazyka a literatúry,
Fakulta humanitných vied Univerzity Mateja Bela, Banská Bystrica


Meno má v procese výrobca – výroba – spotrebiteľ významnú hodnotu. Spolupôsobí na tom, že výrobok sa môže, ale aj nemusí stať predajne atraktívnym. Vhodne zvolený názov predáva, je súčasťou obchodnej propagácie a reklamy, psychologicky vyvíja tlak na zákazníka. Tvorca pomenovania sa ho snaží obohatiť o určité asociačné príznaky, hlavne konotácie pôsobiace na kúpyschopnosť konzumenta. Okrem charakterizačnej funkcie, dať spotrebiteľovi základnú informáciu o vlastnostiach a použití výrobku, vstupuje do popredia konotačná, aso-