

## **Archív nárečí Slovenského národného korpusu. Východiská jeho tvorby, súčasný stav a perspektívy**

**Agáta KARČOVÁ**

### **Abstrakt:**

V príspevku sa venujeme východiskám tvorby, teoretickým a metodologickým postupom a predpokladom nevyhnutným pre tvorbu nárečového archívu. Aplikáciu týchto postupov demonštrujeme na pripravovanom Archíve nárečí Slovenského národného korpusu (AN SNK), ktorý vzniká v oddelení SNK JÚLŠ SAV od r. 2013 vďaka spolupráci s univerzitnými pedagógmi a s ďalšími poskytovateľmi. AN SNK v súčasnosti nadobúda jasné črty reprezentatívnej a praktickej databázy, ktorá je koncipovaná ako prehľadná a rozsiahla báza využiteľná na dialektologické bádania, ale aj na širší humanitne orientovaný výskum.

### **Kľúčové slová:**

dialektológia, slovenské nárečia, terénny výskum, pramenný zdroj, zvukové nahrávky, transkripcia nárečových prehovorov, zvukový nárečový archív, nárečové korpusy

### **Kontakt:**

Jazykovedný ústav Ľ. Štúra Slovenskej akadémie vied  
agatak @ korpus.juls.savba.sk

### **1 Úvod**

Hodnotu a prínos dialektológie ako integrálnej súčasť lingvistiky netreba osobitne zdôrazňovať. Slovenské nárečia mali nezastupiteľnú úlohu pri formovaní sa predpisovných kultúrnych útvarov slovenčiny ako predpokladu vzniku jej kultivovanej a kodifikovanej celospoločenskej formy. Ich význam sa naplnením tejto úlohy nevyčerpal, pretože „aj dnes sú vzťahy medzi týmto súborom nárečí a spisovnou slovenčinou veľmi živé, plné originálnej synergie, pestrých kontaktov a vzájomného ovplyvňovania“ (Kačala, 1999, s. 107–108).

Dialektológia bola už od čias svojho konštituovania založená na terénnych výskumoch hovoreného jazyka a nepochybne túto tradíciu terénneho výskumu založila (porov. Ripka, 1999, s. 100). Súčasný terénny výskum sa už málokedy vykonáva bez nahrávania rečových prejavov, ktoré prináša nesporne veľa výhod – možnosť pracovať s materiálom aj s odstupom času, pri jeho zodpovedajúcej kvalite možnosť analyzovať komunikáty na

všetkých jazykových úrovniach a mať kedykoľvek k dispozícii pramenný materiál v jeho autentickej podobe.

K zhromažďovaniu a spracovaniu pramenného materiálu pochádzajúceho z terénnych výskumov prispieva v súčasnosti aj oddelenie Slovenského národného korpusu Jazykovedného ústavu Ľ. Štúra Slovenskej akadémie vied (ďalej *SNK JÚĽŠ SAV*). Medzi jeho úlohy okrem mnohých ďalších<sup>1</sup> patrí tvorba nárečového korpusu. Prvá verzia textového Korpusu nárečí SNK (ďalej *KN SNK*) *dialekt-1.0*<sup>2</sup> bola sprístupnená v septembri 2014 s rozsahom 73 855 textových jednotiek a je bezplatne prístupný na vyhľadávanie všetkým registrovaným používateľom SNK. V súčasnosti sa pripravuje jeho novšia rozšírená verzia. Tento korpus je významným príspevom k modernému dialektologickému výskumu, pretože umožňuje spoľahlivé, rýchle a efektívne vyhľadávanie v transkripciách nárečových prejavov a uľahčuje tak prácu s textovými nárečovými pramennými zdrojmi.

Od počiatku riešenia úloh zameraných na spracovanie dialektologických materiálov sa v *SNK JÚĽŠ SAV* uvažovalo rovnako o možnosti vytvoriť archív nahrávok, príp. korpus, ktorý by obsahoval zvukové nárečové nahrávky prepojené a spárované s príslušnými transkripciami. V snahe vytvoriť komplexný a reprezentatívny zdroj nárečových dát je ambíciou tvorcov zhromaždiť čo najväčšie množstvo nárečových nahrávok z celého územia Slovenska. Zároveň sme si v súvislosti s napĺňaním tejto úlohy stanovili ďalší zámer – uprednostniť pri zhromažďovaní nárečového materiálu jedinečné a dosiaľ nepublikované dáta, ktoré by sa v prípade, že by sa včas nespracovali a nezdigitalizovali, mohli znehodnotiť či úplne zničiť. Budovanie Archívu nárečí Slovenského národného korpusu (ďalej *AN SNK*), získavanie, zhromažďovanie a spracúvanie nárečových nahrávok, príslušných transkripcií a súvisiacich materiálov tak výrazným dielom prispieva k zachovaniu kultúrneho dedičstva Slovenska.<sup>3</sup>

V prvej časti nášho príspevku sa budeme venovať všeobecne východiskám tvorby, teoretickým a metodologickým postupom ponúkajúcim sa pre prácu s rozmanitými nárečovými zdrojmi. V druhej časti opíšeme praktickú aplikáciu postupov, ktorá bola predpokladom pre vytvorenie súčasnej podoby Archívu nárečí SNK.<sup>4</sup> Čiastočne zanalyzujeme doteraz poskytnutý a spracovaný materiál a naznačíme niektoré problémy

---

<sup>1</sup> Podrobné informácie o aktuálne riešených projektoch SNK sú na stránke <http://korpus.juls.savba.sk/projects.html>.

<sup>2</sup> *Korpus nárečí SNK. Verzia dialekt-1.0*. Dostupné na internete: <http://korpus.juls.savba.sk>.

<sup>3</sup> Koncepcia pre *AN SNK* sa tvorila pod vedením M. Šimkovej v nadväznosti na *KN SNK* a s využitím doterajších skúseností a celého zázemia oddelenia *SNK JÚĽŠ SAV*. Na jeho realizácii sa podieľajú A. Karčová, R. Garabík, K. Gajdošová, technickú podporu zabezpečujú spolupracovníci R. Brída, A. Gogora a J. Mášik.

<sup>4</sup> Interná verzia *Archívu nárečí Slovenského národného korpusu* sa nachádza na dátovom úložisku *SNK JÚĽŠ SAV*.

a možnosti spracovania, ktoré vznikli pri konfrontácii s týmto materiálom v priebehu našej práce.

## **2 Východiská a ciele nárečového zvukového archívu**

Tvorba nárečového zvukového archívu pozostáva z niekoľkých spravidla za sebou nasledujúcich fáz, napĺňanie ktorých predpokladá nielen odborné a praktické znalosti z dialektológie, ale aj manažérske schopnosti, zabezpečenie právnej stránky projektu, technickú podporu a vybavenie a v neposlednom rade skúsenosti s tvorbou korpusov a databáz. Prvej fáze uskutočňovania projektu predchádza vytvorenie premysleného konceptu, stanovenie si cieľov a zamerania výsledného celku, určenie, aké charakterové črty má mať databáza, ktoré materiály má primárne obsahovať, akým spôsobom majú byť utriedené, pre koho má byť výsledná databáza určená a čo má používateľom poskytnúť. Zhrňujúco a abstrahovane sa proces tvorby databázy nárečových zdrojov dá prehľadne zhrnúť do nasledujúcich fáz:

- získavanie zdrojov, zabezpečenie právnej stránky projektu
- technická fáza – digitalizácia a ďalšie spracovanie dát
- verifikácia a doplnenie metadát, selekcia a štrukturácia relevantného materiálu
- vytvorenie zvukovo-textového korpusu (predpokladá sa existencia alebo vytvorenie relevantných transkripcií).
- 

### **2.1 Získavanie zdrojov, zabezpečenie právnej stránky projektu**

Najdôležitejším predpokladom úspešnosti tvorby archívu nárečí je spolupráca s poskytovateľmi nárečových nahrávok a súvisiacich materiálov v prípade, že tvorcovia počítajú so zhromaždením a spracovávaním už existujúcich externých zdrojov. Prvá a zároveň nevyhnutná fáza riešenia úlohy spočíva v takomto prípade vo vyhľadani a oslovení potenciálnych poskytovateľov takých prác, ktoré sú vhodné na zaradenie do pripravovaného celku. Okrem nahrávok je potrebné získať tiež materiály, príp. vyžiadať si informácie o okolnostiach vzniku zvukových záznamov, o respondentoch, výskumnom zámere tvorcu, možné transkripcie a pod. Táto fáza predchádza ostatným, vo väčšine prípadov by mala kontinuálne vyústiť do dlhodobých spoluprác, preto sa rieši spravidla paralelne aj s ďalšími fázami budovania archívu.

Každý poskytnutý zdroj – zvuková nahrávka, publikovaný alebo nepublikovaný text, by mal byť ošetrený dvojstrannou licenčnou zmluvou, ktorou poskytovateľ vyjadruje svoj súhlas so spracovaním zdroja, jeho zaradením do archívu nárečí, resp. korpusu nárečí a so sprístupnením tretím osobám na vedecko-výskumné účely. Prijímateľ sa v nej zaväzuje, že materiál spracuje a sprístupní v súlade s platnými zákonmi danej krajiny. Takáto zmluva zaručuje korektnú spoluprácu, dodržanie príslušných zákonných ustanovení a legitimizuje sprístupnenie archívu záujemcom o prácu s nárečovými zdrojmi.

## **2.2 Technická fáza – digitalizácia a ďalšie spracovanie dát**

Po prevzatí zapožičaných alebo darovaných zdrojov nasleduje fáza ich digitalizácie a technických úprav. V prípade rôznych typov nosičov (MC kazety, MG kazety, diskety, CD a ďalšie) to znamená nahrávanie do pamäte počítača pomocou prídavného zariadenia (prehrávača kaziet, disketovej mechaniky a i.), v prípade prevzatia nahrávok v digitálnej podobe ich skopírovanie alebo stiahnutie do vybraného archívneho priečinku dátového úložiska preberajúceho pracoviska.

Nasleduje overenie spustiteľnosti každej jednotlivej nahrávky, konverzia nevhodných, nečitateľných či neštandardných formátov alebo videoformátov (napr. MPG, M4A, AMR) do vhodných bezstratových formátov (WAV, MP3, MP4, WMA a pod.). Chybné nahrávky sa už v tejto fáze vyradujú, resp. sa opätovne vyžadujú od poskytovateľa v použiteľnej podobe.

Ďalšie technické spracovanie zvukového materiálu je spravidla limitované technickým vybavením a možnosťami pracoviska. Medzi potrebné úpravy patrí takzvané odšumovanie nahrávok, t. j. odfiltrovanie nežiaducich ruchov, predovšetkým konštantného šumu v pozadí, vyrovnávanie hlasitosti jednotlivých nahrávok a ďalšie. Tieto nevyhnutné úpravy sa môžu realizovať s minimálnymi finančnými nákladmi pomocou bezplatných softvérových nástrojov (Audacity, pavucontrol a pod.). Textové zdroje rôzneho charakteru sa digitalizujú, práce dodané v elektronickej podobe sa v prípade potreby upravujú (napr. nesprávne nastavené kódovanie). Podobne sa postupuje pri existujúcich transkripciách celých nahrávok alebo ich častí. Po digitalizácii je potrebné transkripcie rozpoznať a skontrolovať ich podľa predlohy.

## **2.3 Verifikácia a doplnenie metadát, selekcia a štrukturácia relevantného materiálu**

Veľmi dôležitá a časovo náročná je fáza spracovávania a dopĺňania metadát, t. j. podstatných informácií o hovoriacich – zvyčajne ide o 1 explorátora a 1 alebo viacerých respondentov – a o samotných nahrávkach. Bez doplnenia údajov o okolnostiach vzniku nahrávky, základných sociolingvistických údajov o hovoriacich a bez poznania výskumného zámeru a stratégie jej tvorcu je nahrávka prakticky nepoužiteľná, aj keby spĺňala atribúty technicky vyhovujúceho zdroja. Predpokladá sa, že zvukové nahrávky získavané z rôznych inštitúcií, napr. škôl, múzeí, ako aj od súkromných osôb – profesionálnych dialektológov, študentov či laikov – nie sú vytvorené rovnakým postupom a nerovnaký bude aj spôsob zachytávania metadát, pričom sa môže stať, že niektoré základné údaje chýbajú. Tieto informácie môžu byť z hľadiska tvorcov archívu nekompletné a je potrebné dodatočne ich vyhľadávať, aby sa zabezpečili podľa možnosti všetky metadáta podľa vytvoreného jednotného systému. Rovnako treba mať na pamäti, že spoľahlivosť zachytených údajov a ich interpretácia závisí od charakteru dodaného zdroja (práca odborníka, skontrolovaná práca študenta, neskontrolovaná práca študenta).

Už počas získavania zdrojov, ich prvotného spracovania a dopĺňania metadát alebo bezprostredne po týchto úkonoch nevyhnutne nasleduje selekcia a triedenie materiálu, ktoré sa uskutočňujú s prihliadaním na zameranie budúceho archívu ako celku, ale aj na množstvo a povahu poskytnutých zdrojov. V tejto fáze sa z nazhromaždeného materiálu vylúčia technicky nevyhovujúce záznamy, zreteľne nenárečové útvary (ak nahrávka neobsahuje žiadnu časť v nárečí), materiál s absentujúcou podstatnou časťou metadát a pod.

Z povahy rôznorodého materiálu od viacerých poskytovateľov, nerovnakých stratégií, výskumných metód a cieľov zbierania nárečového materiálu vyplýva viacero možných prístupov k selekcii, triedeniu a usporiadaniu materiálu. Pri tejto fáze je podstatná otázka, ako štruktúrovať nerovnorodé dáta a ako uchopiť materiál tak, aby hotový celok – archív – poskytoval nielen technicky vyhovujúce nahrávky doplnené relevantnými a overenými metadátami, ale aby sa s týmto celkom dalo efektívne pracovať a spoľahlivo v ňom vyhľadávať podľa vybraných kritérií.

Triedenie materiálu sa môže uskutočňovať podľa najrozličnejších kritérií: zvukové záznamy sa dajú triediť podľa ich technickej kvality, podľa autora nahrávky (dialektológ, študent, laik), podľa miesta a času nahrávania, veku respondenta, konkrétneho dialektu, v ktorom je zachytený hovorený prejav, podľa miery autenticity prejavu či stupňa zachovanosti pôvodnej podoby nárečia. Nahrávky sú rôzne aj z hľadiska témy a formy prehovorov, môžu byť monologické alebo dialogické, riadené či spontánne, nasimulované alebo čítané atď.

Z uvedeného vyplýva veľa možností zachytávania a triedenia materiálu, načrtávame tu iba niektoré možnosti a opozície:

<b>poskytovateľ</b>	stratégia, výskumný zámer inštitúcia / jednotlivec dialektológ / laik
<b>respondent</b>	autochtón / neautochtón vek pohlavie vzdelanie
<b>prehovor</b>	miesto nahrávania čas nahrávania typ nárečia

„čistá“ podoba nárečia / nivelizovaná podoba nárečia

téma

miera spontánnosti

Načrtnuté spôsoby triedenia sa dajú kombinovať. V prípade, že je archív koncipovaný ako celok určený nielen na dialektologický výskum, ale jeho záber je širší a predpokladá sa jeho využitie na humanitne zamerané bádania, je dobrým riešením zachovať v ňom všetok vyhovujúci materiál bez vytvárania podcelkov, ale s možnosťou jednoduchého aj pokročilého vyhľadávania podľa dostupných metadát. Výber kritérií vyhľadávania potom bude závisieť od požiadaviek konkrétneho používateľa a jeho preferencií.

Dopĺňanie metadát, ktoré sa spravidla zachytávajú v osobitných súboroch, je potrebné realizovať jednotne a prehľadne tak, aby bolo možné bezproblémové vyhľadávanie. Pri ich výbere je potrebné mať na pamäti, že nie všetky načrtnuté kritériá sú rovnako zistiteľné a spoľahlivo určiteľné. Niektoré z nich sú jednoznačné (napr. miesto nahrávania), iné môžu byť v opozícii (monologický vs. dialogický prejav). Ďalšie z nich do istej miery závisia od subjektívneho posúdenia jednotlivca, často bez možnosti overenia (napr. členenie nahrávok na osi *spontánnosť – pripravenosť prehovoru*) či s neostrými hranicami (*riadený rozhovor* môže prejsť do *voľného rozhovoru*; hovoriaci môžu na úrovni jedného komunikátu z rôznych dôvodov prepínať *jazykové kódy*, napr. *prejav v štandardnej variete jazyka – prejav s nadnárečovými prvkami – prejav s niekoľkými nárečovými prvkami – čisto nárečový prejav* a i.). Výber typu a množstva zachytávaných metadát závisí v neposlednom rade od časových možností jeho tvorcov a náročnosti získavania týchto doplňujúcich informácií.

V zásade sa môžu metadáta členiť na povinné a nepovinné, ako aj na verifikované a neverifikované. Snaha vyjsť v ústrety používateľovi a uľahčiť mu prácu s archívom by mala smerovať k realizácii vyhľadávača, prostredníctvom ktorého by sa mohol v archíve orientovať a vyhľadať potrebné dáta, napríklad zadaním požiadavky parafrázovanej slovami:

*Nájdí všetky nahrávky v šarišskom nárečí nahraté pred rokom 2000 v Prešovskom okrese, v ktorých hovoria respondentky ženského pohlavia narodené pred rokom 1950 a žijúce v Prešovskom okrese.*

#### **2.4 Vytvorenie zvukovo-textového korpusu**

Z dôkladne spracovaného archívu, príp. jeho vybranej časti sa môže vytvoriť zvukovo-textový korpus ako špecifický súbor jednotne spracovaných dát. Korpus obsahujúci zvukové nahrávky a ich príslušné transkripcie, prirodzene, poskytuje ďalšie možnosti

výskumu, okrem iného umožňuje napríklad vyhľadávanie podľa konkrétnych tvarov slov alebo ich častí. Vytvorenie takéhoto korpusu predpokladá existenciu dostatočného množstva relevantných transkripcií a prehĺbenú spoluprácu s dialektológmi, prípadne prípravu jednotného transkripčného systému, spolupracovníkov vykonávajúcich odborný prepis, ďalšie technické spracovanie dát a mnohé ďalšie kroky.

Pri možnom riešení tejto fázy tak zákonite vznikajú otázky: akým spôsobom zaradiť do homogénneho celku prepisy realizované nerovnakou transkripčnou metódou, aký spôsob prepisu zvolíť na transkripciu zatiaľ neprepísaných nahrávok alebo ich častí. Vhodný výber transkripčnej metódy je témou na samostatnú štúdiu, preto sa mu v rozsahu tohto príspevku nebudeme podrobne venovať. Na stručné osvetlenie tejto problematiky použijeme slová R. Krajčoviča: „Základným predpokladom úspechu dialektologickej heuristiky je, aby nárečové texty boli bezprostrednými živými prejavmi spoľahlivého informátora a aby boli zapísané (prepísané) podľa jednotných, vopred dohodnutých zásad“ (Krajčovič, 1988, s. 191). Autor zároveň vymedzuje 2 základné typy prepisu, a to odborný a čitateľský, ktoré dopĺňa kombinovaný prepis vhodný na didaktické účely.

Jednotliví dialektológovia a pracovné kolektívy si volia ten typ prepisu, ktorý zodpovedá ich výskumnému zameraniu, pričom sa spravidla v každej práci odborného charakteru uvádza legenda so zoznamom použitých transkripčných znakov a zvolený spôsob zachytávania rôznych výslovnostných javov, akými sú spodobovanie, mäkkosť konsonantov, výslovnosť zdvojených spoluhlások, splyvavá výslovnosť a pod. Na ilustráciu situácie na Slovensku stačí porovnať niekoľko odborných diel, napr. *Praktická dialektológia* (2012), *Slovník slovenských nárečí* (I. zv., 1994; II. zv., 2005) a *Atlas slovenského jazyka* (IV, 1984), aby sme demonštrovali rozdielnosť zvolených transkripčných metód. Rovnako ďalšie dialektologické práce, ktoré vznikajú ako samostatné autorské výstupy, majú spravidla nerovnaký spôsob transkripcie. Študentom filológií sa v rámci zadania seminárnej práce z dialektológie spravidla odporúča prepisovať nahrávky kombinovanou metódou. Táto metóda poskytuje priestor na vlastné vymedzenie transkripcie, preto sú študentské prepisy navzájom nekompatibilné.

S prihliadaním na nerovnakosť transkripčných systémov prichádza do úvahy reálna možnosť zaradiť do potenciálneho korpusu existujúcej transkripcie profesionálnych dialektológov (monografie, zborníky, dizertačné a habilitačné práce a pod.), ako aj študentské prepisy vznikajúce pod odborným vedením skúsených pedagógov (bakalárske a diplomové práce) bez toho, aby sa zasahovalo do spôsobu transkripcie, ktorý bol pre danú prácu zvolený (do úvahy pripadajú len menšie úpravy, napr. zjednotenie zápisu *ch*, *dz*, *dž* a pod.). Rovnako sa do korpusu môžu spracovať aj nahrávky neprofesionálov, ktoré spĺňajú potrebné kritériá, pričom by bolo potrebné prepísať ich jednotným vopred pripraveným kombinovaným prepisom so zachytávaním vybraných javov.

### 3 Súčasný stav AN SNK, jeho tvorba, ciele a perspektívy

AN SNK je súčasťou úlohy elektronizácie nárečového materiálu, ktorého prvým výstupom na báze publikovaných textov je Korpus nárečí SNK v textovej podobe. Na pracovných stretnutiach tvorcov KN SNK sa postupne ukazovalo, že na slovakistických pracoviskách sú ďalšie nárečové materiály v zvukovej podobe, ktoré by bolo potrebné spracovať. To viedlo pracovný tím k rozšíreniu pracovnej úlohy a príprave koncepcie nárečového archívu.

V predchádzajúcej kapitole sme stručne opísali abstrahovaný proces budovania nárečového archívu rozdelený do viacerých fáz, vychádzajúc zo skúseností kolektívu SNK realizujúceho viacero databáz a rôznych typov hovorených aj písaných korpusov. Teraz sa zameriame na opis súčasnej podoby konkrétneho Archívu nárečí SNK<sup>5</sup> vznikajúceho od r. 2013 v oddelení SNK JÚLŠ SAV s významným príspevom všetkých jeho poskytovateľov. Vyberieme tiež niektoré konkrétne problémy a výzvy, ktoré reálne poskytnuté dáta prinášajú.

#### 3.1 Poskytovatelia AN SNK

Na Slovensku je pomerne malý počet dialektológov, väčšina z nich pôsobí na akademickej pôde – na vysokých školách a v Slovenskej akadémii vied. Vzhľadom na túto skutočnosť sme už v predprípravnej fáze riešenia našej úlohy zisťovali záujem o spoluprácu na všetkých relevantných vysokoškolských pracoviskách, pričom nevyhnutnosť spolupráce s nimi sme považovali za predpoklad vzniku AN SNK v jeho požadovanej rozsiahlej a reprezentatívnej podobe. Komunikácia s dialektológmi na Slovensku prebiehala vo všetkých prípadoch na profesionálnej úrovni, u oslovených univerzitných pedagógov sme sa stretli s porozumením pre napĺňanie úlohy tohto charakteru. Prvá fáza tvorby AN SNK sa úspešne realizovala vďaka ústretovosti dialektológov, ktorí boli ochotní bezplatne poskytnúť potrebné zdroje, venovali triedeniu, verifikácii a zapožičaniu študentských aj vlastných materiálov svoj osobný čas a energiu a potvrdili svoju spoluprácu z právneho hľadiska prostredníctvom licenčných zmlúv.

V súčasnosti úspešne spolupracujeme s dialektológmi pôsobiacimi na univerzitách:

- doc. Mgr. Jaromír Krško, PhD., Mgr. Alexandra Chomová, PhD.  
Filozofická fakulta Univerzity Mateja Bela v Banskej Bystrici
- doc. PhDr. Marianna Sedláková, PhD.  
Filozofická fakulta Univerzity Pavla Jozefa Šafárika v Košiciach
- Mgr. Tomáš Bánik, PhD.  
Filozofická fakulta Konštantína Filozofa v Nitre
- PhDr. Peter Karpinský, PhD., Mgr. Veronika Perovská  
Filozofická fakulta Prešovskej univerzity v Prešove

---

<sup>5</sup> Stránka AN SNK je dostupná na adrese: [http://korpus.juls.savba.sk/dialect\\_recordings.html](http://korpus.juls.savba.sk/dialect_recordings.html)



- PhDr. Viera Kováčová, PhD.  
Filozofická fakulta Katolíckej univerzity v Ružomberku
- PaedDr. Andrej Závodný, PhD.  
Pedagogická fakulta Trnavskej univerzity v Trnave

So všetkými uvedenými poskytovateľmi sme nadviazali dlhodobú spoluprácu, pričom v AN SNK sa v súčasnosti nachádzajú spracované študentské nahrávky, seminárne práce, čiastkové transkripcie a ďalšie súvisiace materiály zo všetkých uvedených vysokoškolských pracovísk. Vďaka našim poskytovateľom sme doteraz získali do AN SNK 521 jedinečných študentských nárečových nahrávok od 310 jednotlivých autorov.

Z vlastnej výskumnej praxe dialektológov výrazne obohacujú AN SNK viaceré odborné príspevky, nahrávky a príslušné transkripcie Andreja Závodného (nahrávky zo Smrdák a i.), dizertačná práca Viery Kováčovej *Sotácke nárečia na západoslovensko-východoslovenskom jazykovom pomedzí* (2005), diplomová práca Jozefa Bilského *Súčasný stav používania nárečia v Dlhej Lúke* (2014) vypracovaná pod vedením Marianny Sedlákovej, ako aj významná publikácia kolektívu autoriek: Gabriely Múcskovej, Kataríny Muzikovej a Viery Wambach *Praktická dialektológia. Vysokoškolská príručka na nárečovú interpretáciu* (2012), ktorá obsahuje nahrávky z 28 lokalít Slovenska reprezentujúce najtypickejšie nárečové areály aj s príslušnými transkripciami. Ďalšie materiály získavame priebežne, pričom spoluprácu nám prisľúbili ďalší lingvisti. Obsah archívu dopĺňajú nahrávky z radov laikov: záznam šurianskeho nárečia na základe našich inštrukcií nahral Michal Komžík, viaceré nahrávky so zameraním na opis folklórnych tradícií sme získali vďaka Dagmar Šimunovej od ich autorky pani Zuzany Knapekovej z Modry-Kráľovej. Za vyhľadanie a odporúčenie niektorých vhodných zdrojov, ako aj za odborné konzultácie a cenné pripomienky ďakujeme Kataríne Ballekovej, vedúcej Dialektologického oddelenia JÚLŠ SAV, a ďalším pracovníkom tohto oddelenia.

### 3.2 Ďalšie technické spracovanie poskytnutých zdrojov

Archív nárečí SNK v čase odovzdania tohto príspevku (marec 2015) obsahuje 561 samostatných zvukových súborov, z ktorých podstatnú časť (521) tvoria študentské nárečové nahrávky. Z textových zdrojov disponuje AN SNK okrem spomínaných odborných prác dialektológov aj 80 seminárnymi prácami, z ktorých niektoré obsahujú čiastkové transkripcie príslušných nahrávok.

Materiál dodaný na starších typoch nosičov (203 MG kaziet, 18 MC kaziet) sa digitalizoval na pracovisku SNK, seminárne práce a ďalšie textové materiály dodané v tlačenej podobe sa oskenovali v potrebnej kvalite a vytvorili sa z nich samostatné súbory. Rovnako sa naskenovali dodané transkripcie, pričom sa rozpoznal ich text a zrekonštruoval sa v elektronickej podobe tak, aby sa dal použiť v prípade tvorby korpusu.

Pri väčšine poskytnutých nahrávok, ktoré sa digitalizovali z pôvodných nosičov, ale aj pri novších nahrávkach dodaných na CD alebo vo forme súborov prijatých priamo v elektronickej podobe sa ukázala nevyhnutnosť odfiltrovať v nich šumy a iné ruchy, niektoré upraviť aj z hľadiska hlasitosti, príp. realizovať ďalšie technické úpravy. Využili sa pri tom skúsenosti a zručnosti vytvorené v SNK v procese tvorby Slovenského hovoreného korpusu<sup>6</sup> (ďalej SHK). Väčšina doteraz dodaných študentských nahrávok už prešla touto úpravou, pričom kvalita každej z nich je zaznamenaná na 3-bodovej stupnici (*veľmi dobrá / dobrá / nízka*). Táto informácia o kvalite bude podstatná pri výbere nahrávok do korpusu, ako aj pri práci s archívom.

### 3.3 Štruktúra zvolených metadát a ich dopĺňanie, selekcia relevantného materiálu

S nekompatibilitou zachytávania a umiestňovania metadát, ako aj ich nekompletnosťou z hľadiska istej zvolenej štruktúry sme počítali už pred začatím tvorby AN SNK. Hneď na začiatku sme preto na základe existujúcich súborov metadát vytvorených pre potreby Slovenského hovoreného korpusu<sup>7</sup> a Korpusu nárečí SNK vypracovali jednotný formulár aj pre AN SNK na zaznamenanie informácií o hovoriacich a nahrávke a rozposlali ho dlhodobým poskytovateľom. V nezmenenej podobe ho uvádzame v nasledujúcej tabuľke.

**Tabuľka 1: Formulár s metadátami pre AN SNK**

#### **NAHRÁVAJÚCI** (explorátor):

name: MENO A PRIEZVISKO  
sex: POHLAVIE (f – žena, m – muž)  
birthdate: DÁTUM NARODENIA (vo formáte rok – mesiac – deň; napr. 1945-12-01)  
birthplace: MIESTO NARODENIA (napr. Krupina)  
location: MIESTO TRVALÉHO POBYTU (napr. Vrútky, okr. Martin)

**NAHRÁVANÝ** (respondent): *v prípade viacerých účastníkov rozhovoru vyplniť údaje každého z nich*

name: MENO A PRIEZVISKO  
sex: POHLAVIE (f – žena, m – muž)  
birthdate: DÁTUM NARODENIA (vo formáte rok – mesiac – deň)

---

<sup>6</sup> *Slovenský hovorený korpus – s-hovor-4.0*. Bratislava: Jazykovedný ústav Ľ. Štúra SAV 2012. Dostupné na internete: <http://korpus.juls.savba.sk>.

<sup>7</sup> Zoznam metadát SHK je na stránke <http://korpus.juls.savba.sk/shk.html>. Bližšie informácie o metadátach používaných v SHK sú v príspevku *Slovenský hovorený korpus (2008 – 2012)* (Gajdošová – Šimková, 2014).

birthplace: Miesto narodenia (napr. Bardejov)  
location: Miesto trvalého pobytu (napr. Dlhá Lúka, okr. Bardejov)  
longestplace: Miesto najdlhšieho pobytu (na ktorom strávil respondent najviac času, napr. Dlhá Lúka, okr. Bardejov)

**ÚDAJE O NAHRÁVKE:**

name: Téma nahrávky (napr. O pôvode miestnych názvov; Spracovanie ľanu; O zvykoch na svadbe)  
explordate: Dátum nahrávania (vo formáte rok – mesiac – deň; napr. 2014-02-12)  
explorplace: Miesto nahrávania (napr. Dlhá Lúka, okr. Bardejov)  
dialect: Určenie nárečia (podľa rozčlenenia na základné areály) (napr. šarišské nárečie)

Tieto základné metadáta sú predpokladom jednoduchého alebo pokročilého vyhľadávania podľa konkrétneho kritéria alebo kombinácie viacerých kritérií. Pri tvorbe jednotného vzoru sme sa snažili nielen o výber najpodstatnejších metadát, ale aj o stručnosť, prehľadnosť a jednotnú formu ich zachytávania, pričom ich vyplnenie nie je časovo náročné. Ďalšie podstatné okolnosti, napr. profesionalitu explorátora, časové trvanie nahrávky, jej kvalitu a podobne môžeme doplniť dodatočne a samostatne na základe komunikácie s poskytovateľom, ako aj po analýze samotného poskytnutého zdroja.

Široko koncipovaný projekt zahŕňa okrem vedeckých výstupov profesionálov aj študentské práce a nahrávky laikov, preto si vyžaduje zvýšenú pozornosť predovšetkým v tejto fáze riešenia. Údaje zachytené a interpretované študentom alebo neprofesionálom musia prejsť dôkladnou kontrolou. Ide hlavne o presnú identifikáciu nárečia respondenta a určenie významných jazykových vplyvov na jeho reč. Medzi relevantné údaje sa tieto metadáta môžu dostať až po dôkladnej kontrole, príp. so zaznamenaním, že neboli verifikované (pomocou dohodnutého znaku, napr. otáznika).

V SNK aktuálne zaznamenávame nevyhnutné informácie o technických parametroch nahrávky (spustiteľnosť, technická kvalita), zároveň zachytávame, v akej forme (na akom type nosiča) bola nahrávka poskytnutá, v názve každej nahrávky je informácia o jej formáte. Overujeme, či je nahrávka spontánna, predpripravená alebo vznikla ako záznam čítaného už existujúceho textu, čo sa však dá spoľahlivo určiť len pri explicitnej informácii v priloženom zdroji (napr. seminárnej práci), príp. zisťovaním, či daný text už nefiguruje v publikovanej literatúre. Tento aspekt rečových prehovorov zachytených v nahrávkach sa v budúcnosti pokúsime analyzovať na základe určujúcich znakov.

Vďaka spolupráci s univerzitnými pedagógmi sme získali niektoré práce odovzdané v uplynulom školskom roku aj s potrebnými metadátami vyplnenými podľa vzoru vypracovaného v SNK, čo výrazne uľahčuje a zefektívňuje ďalšiu prácu. Metadáta

k študentským prácam z predminulých rokov je potrebné pracne vyhľadávať v seminárnych prácach, príp. v samotných zvukových záznamoch a aj napriek zvýšenej snahe tvorcov archívu a ochotnej spolupráci poskytovateľov nie sú vo všetkých prípadoch kompletne.

Vzhľadom na mnohostranné zameranie AN SNK a zachovanie možností pre rôzne typy výskumu sme sa pri selekcii rozhodli vyradiť iba najnutnejšiu časť nevyhovujúceho materiálu, a to nahrávky na veľmi zlej technickej úrovni a prehovory, ktorých žiadna časť nie je v nárečí. V AN SNK sme ponechali všetky ostatné rôznorodé zdroje: prejavy obohatené o piesne (ktoré sa primárne nehodia na dialektologický výskum, ale môžu poslúžiť napr. folkloristom), spontánne rečové prejavy, ale aj štylizované a nasimulované rozhovory, nahrávky s autochtónnymi obyvateľmi, ale aj mladšími respondentmi, ktorí používajú nárečie v zreteľne redukovanej podobe a pod. Vďaka informáciám, ktoré systematicky zachytávame, sa pri vyhľadávaní v AN SNK a prípadnej tvorbe KN SNK budú môcť vyselektovať najvhodnejšie nahrávky pre dialektologický výskum (prejavy autochtónov, čisto nárečové spontánne prehovory, prejavy zaujímavé z hľadiska témy, z lokalít, ktorým sa zatiaľ nevenovala pozornosť a pod.).

### **3.4 Špecifiká študentských nárečových nahrávok**

Publikované aj nepublikované nahrávky, ich transkripcie a súvisiace štúdie a práce dialektológov sú pre AN SNK a nárečový výskum významným prínosom. Tieto materiály sú spravidla v zodpovedajúcej technickej kvalite, s dôkladným a presným odborným alebo čitateľským prepisom s presne formulovanými a v celej transkripcii dodržiavanými pravidlami, s jasným výskumným zámerom a zachytenými metadátami.

Väčšiu časť AN SNK však už v súčasnosti tvoria študentské nahrávky, realizované študentmi slovenskej filológie na spomenutých univerzitách. Už od začiatku budovania AN SNK sme sa cielene zameriavali aj na tento typ nárečových zdrojov, pretože ide o jedinečné nahrávky, ktoré sú po dôkladnom spracovaní vhodné na výskumné účely. Vznikali spravidla ako súčasť jednorazového projektu s cieľom splnenia študijných povinností bez ambície ich tvorcov publikovať ich, resp. ďalej podrobnejšie spracovávať a analyzovať. Ich zaradenie do AN SNK im tak „predlžuje život“ a umožňuje ich ďalšie využitie. Prinášajú množstvo pozitív, ale aj zvýšenú náročnosť pri ich spracúvaní a analýze, na ktoré treba upozorniť.

Niektoré pozitíva študentských nárečových nahrávok môžeme prehľadne zhrnúť takto:

- jedinečnosť nahrávok
- veľké množstvo nahrávok
- reprezentatívnosť nahrávok z hľadiska ich geografického zaradenia
- možnosť porovnávacieho štúdia
- rozmanitosť respondentov
- štýlová rôznorodosť (monológy, dialógy, riadené rozhovory, anketa...)

- tematická pestrosť
- spontánnosť prehovorov (explorátor je často v príbuzenskom vzťahu s respondentom, preto sa rozhovory odohrávajú v uvoľnenej a dôvernej atmosfére).

Jedinečnosť nahrávok zaručujú ich autori – študenti slovakistiky, pretože na väčšine vysokých škôl na Slovensku každý školský rok odovzdávajú spolu so seminárnou prácou z dialektológie aj svoju vlastnú nahrávku zaznamenávajúcu prejav respondenta, ktorého si ľubovoľne vyberajú. Vznikajú tak originálne zvukové záznamy na rôznych miestach, s rozličnými ľuďmi a na rôzne témy, čo v prípade spolupráce s danými školami predstavuje nezanedbateľné možnosti rozširovania Archívu nárečí SNK aj v budúcnosti.

Na ilustráciu uvádzame prehľad niektorých údajov, ktoré sme vybrali z doteraz poskytnutých zdrojov z FF PU v Prešove. V 1. stĺpci je údaj o veku respondenta, v 2. lokalita, v ktorej sa nahrávanie realizovalo, a v 3. nárečie, ktoré respondent používa. V prípade prázdneho miesta ide o neuvedený údaj, ktorý je potrebné podľa možnosti doplniť.

**Tabuľka 2: Prehľad vybraných údajov zo študentských nahrávok FF PU v Prešove**

Vek resp.	Lokalita	Nárečie	Vek resp.	Lokalita	Nárečie
72	Ľubotín	šarišské	73	Čaklov	zemplínske
89	Torysa	šarišské	83	Ďapalovce	zemplínske
69	Okružná	šarišské	82	Moravany	zemplínske
23	Giraltovce	šarišské		Lučkovce	zemplínske
43	Prešov	šarišské	78	Giglovce	zemplínske
80	Abranovce	šarišské	20	Kochanovce	sotácke
83	Sedlice	šarišské	43	Papín	sotácke
80	Krivany	šarišské	60	Žakarovce	spišské
84	Ľutina	šarišské	23	Čirč	rusínske
	Hažlín	šarišské	67	Torysky	rusínske
25	Bartošovce	šarišské			
93	Raslavice	šarišské	67	Honce	gemerské

Z uvedenej vzorky 23 nahrávok môžeme ľahko určiť, že vekový priemer respondentov je 64 rokov, po odrátaní 4 nahrávok s respondentmi vo veku rovesníkov študentov – explorátorov (23, 25, 20, 23) tento priemer predstavuje až 73 rokov. Súčasťou tohto celku je napríklad aj viac ako polhodinová nahrávka s osobnými spomienkami 93-ročnej respondentky z Raslavic. Študenti si vo väčšine prípadov zámerne vybrali starších respondentov a v prípade, že išlo o autochtónnych obyvateľov uvedených lokalít, tak vytvorili vzácne záznamy použiteľné na dialektologický výskum (pri splnení ďalších atribútov).

Okrem 1 nahrávky, ktorá bola zachytená v krajskom meste Prešov, boli všetky nahrávky realizované v obciach, a viac ako 95 % z nich bolo nahratých v lokalitách spadajúcich do územia Prešovského kraja, v rámci ktorého sa nachádza aj Prešovská univerzita. Iba 1 odovzdaná nahrávka z horeuvedenej vzorky bola nahratá na území stredoslovenského makroareálu. Medzi nahrávkami nájdeme aj také, ktoré zachytávajú nárečie nepatriace medzi slovenské (ich určenie – *rusínske* – uvádzame tak, ako ich zaznamenali tvorcovia), ale dotvárajú obraz o nárečiach východného Slovenska, lebo ich nositeľmi sú príslušníci etnika tvoriaceho integrálnu súčasť obyvateľstva žijúceho na tomto území. Aj z priestorového hľadiska tak dané nahrávky predstavujú zaujímavú vzorku vhodnú na výskum východoslovenských nárečí. Po analýze mnohých ďalších súborov nahrávok prijatých z ďalších vysokých škôl na Slovensku sme sa rovnako stretli s tým, že prevažná časť študentských nahrávok bola zachytená v makroareáli zhodnom s tým, v ktorom sídli vysoká škola. Spolupráca so všetkými oslovenými univerzitami sa v tomto kontexte ukazuje ako kľúčová, lebo zaručuje reprezentatívnosť a komplexnosť zhromažďovaných dát.

Pri väčšom objeme zdrojového materiálu, akým je napríklad rozsiahly archív doteraz poskytnutých nahrávok z FF KU v Ružomberku (134 zvukových nahrávok), sa takmer zákonite vyskytujú také, ktoré sa realizovali v tej istej lokalite, ale boli nahraté v nerovnakom čase a s rozličnými respondentmi. V danom súbore dát sa napríklad nachádzajú viaceré nahrávky z Likavky (8), Liptovských Sliačov (7), Oravskej Lesnej (6), Oravského Veselého (3) a i. Pri nahrávkach z totožných lokalít, ktoré vytvorili explorátori v rôznom čase, sa tak otvára priestor na porovnávací výskum nárečí na časovej osi.

Pozitívne môžeme vnímať rozsiahly študentský materiál aj z hľadiska jeho žánrovej rôznorodosti. Mnohé nahraté rozhovory sa uskutočnili s rodinnými príslušníkmi študentov, čo mnohokrát vyústilo (niekedy po prvotnej krátkej tréme z nahrávania) do spontánneho prejavu s vyššou mierou dôvernosti, ako by tomu bolo pri nahrávaní cudzej osoby.

Nie všetky študentské nahrávky sa však vyznačujú originalitou. Medzi mnohými jedinečnými nahrávkami sa nájdú aj také, ktoré obsahovali nasimulovaný rozhovor či iným spôsobom predpripravený prejav, ale aj čítania už publikovaného nárečového textu. Nahrávka čítaného archaického alebo prevzatého textu aj v prípade, že je študent

nositeľom daného nárečia, je tak nerelevantná pre ďalší výskum (nahrávka neodráža jazykovú kompetenciu študenta, z tohto pohľadu je neadekvátne z hľadiska lexikálneho, štylistického atď.). V niektorých prípadoch autori takýchto čítaných nahrávok uviedli túto informáciu priamo v texte priloženej seminárnej práce, inokedy sa táto neoriginalita výpovede dá vydedukovať (ale nie úplne spoľahlivo) zo suprasegmentálnych javov prejavu – intonácie, tempa reči a pod.

Ako sme už spomenuli vyššie, chybné môžu byť v niektorých prípadoch aj interpretácie nahrávok – stretávame sa s nesprávnym určením nárečia, podcenením jazykových vplyvov u respondenta atď. Tieto chyby môžu vzniknúť nielen z nepozornosti študenta, ale aj pre nedostatok jeho skúseností a schopnosti orientovať sa v zložitej problematike rozčlenenia slovenských nárečí, pretože základné nárečové areály charakterizuje viacero nielen spoločných, ale aj odlišných jazykových znakov.<sup>8</sup> Metadáta pri mnohých študentských nahrávkach, hlavne tých, ktoré boli nahrávané dávnejšie, úplne chýbajú (z niektorých archívov sú k dispozícii len zvukové záznamy a seminárne práce sa z rôznych dôvodov nezachovali). V prípade, že k nahrávkam existujú transkripcie, tieto sú bez kontroly často neúplné a nepresné, napriek tomu môžu pomôcť pri prepisoch odborníkmi.

Na riešenie týchto výziev, ktoré prinášajú študentské materiály, je vypracovaná metodológia zahŕňajúca dôkladné overovanie a verifikáciu metadát (predovšetkým pri určení nárečia, ale aj miery spontánnosti), ktorá predpokladá spoluprácu odborníkov – dialektológov. Vo významnej miere prispievajú k verifikácii spolupracujúci pedagógovia, ktorí pomáhajú pri overovaní a dopĺňaní metadát, ale aj kontrole seminárnych prác, dodaných transkripciami atď.

#### 4 Záver

Originálny nárečový prejav je vynikajúcim výskumným materiálom z viacerých hľadísk. P. Žigo v posudku publikácie s výstižným názvom *Zo studnice rodnej reči* (p. Balleková a kol. 2014) vyzdvihuje hodnotu nárečového textu takto: „... nárečový text prekračuje hranice jazykovedy a zasahuje oblasť etnológie, národnej kultúry a histórie a stáva sa výrazovým prostriedkom duchovného dedičstva a kultúrnej pamäti.“ Archív nárečí SNK ako rozsiahly homogénny celok má ambíciu stať sa spolu s Korpusom nárečí SNK užitočnou, reprezentatívnou a dostupnou bázou na praktický vedecký výskum v oblasti dialektológie, ale aj na širšie filologické bádania, pričom svoje uplatnenie môže nájsť aj v ďalších humanitných vedách, akými sú história, sociológia, antropológia, folkloristika. Jeho bezplatná prezenčná dostupnosť<sup>9</sup> všetkým zaregistrovaným používateľom výrazne

---

<sup>8</sup> Túto rozmanitosť v rámci tradične vyčleňovaných nárečových areálov reflektuje L. Bartko (1997, s. 198) použitím plurálovej podoby, keď hovorí o *abovských a spišských nárečiach*.

<sup>9</sup> V prípade AN SNK ide o možnosť vyhľadávania v prezenčnej forme, t. j. na prácu s ním je z hľadiska ochrany zdrojov nevyhnutná osobná prítomnosť používateľa v SNK JÚLŠ SAV. V prípade vytvorenia

rozširuje možnosti jeho uplatnenia. Môže slúžiť ako bohatý zdroj informácií o histórii, folklóre a spôsobe života našich predkov v rôznych regiónoch Slovenska a poskytnúť ďalšie cenné informácie pracovníkom kultúrnych inštitúcií, tvorcom kroník, kníh či stránok o obciach Slovenska, ako aj širokej odbornej verejnosti. Uplatniť sa môže rovnako aj pri vyučovaní dialektológie na rôznych stupňoch vzdelávacieho systému na Slovensku.

Súčasná podoba AN SNK nadobúda jasné rysy reprezentatívnej a praktickej databázy, jeho tvorcovia z oddelenia SNK JÚLŠ SAV neustále pracujú na tom, aby boli zvukové záznamy technicky vylepšené a doplnené dostupnými a presnými metadátami. Databáza sa plynulo rozširuje o ďalšie relevantné zdroje tak, ako ich postupne získavame od našich poskytovateľov. Zároveň sa naša snaha zameriava na tvorbu nástrojov a prostredia vhodného na jednoduché aj pokročilé vyhľadávanie čo najširšieho repertoáru nárečových javov.

Veríme, že pilotná verzia AN SNK bude v najbližšom čase pripravená na sprístupnenie v prezenčnej podobe pre všetkých záujemcov a používateľov SNK. Rovnako očakávame, že vďaka kontinuitnej spolupráci s doterajšími spolupracovníkmi aj ďalšími oslovenými poskytovateľmi sa archív rozšíri o nahrávky s odbornými transkripciami, z ktorých bude možné vytvoriť zvukovo-textový korpus ako dostupný zdroj s možnosťou vzdialeného prístupu. V súčasnosti sa v rámci tvorby zvukovo-textového Korpusu nárečí SNK pripravuje jednotný transkripčný systém, ktorý by mal byť jednoduchý, ale nie zjednodušujúci a ktorým sa v textovej podobe adekvátne zachytia vybrané nárečové komunikáty.

Za úspešný začiatok a dlhodobú účasť pri realizácii tvorby NA SNK patrí vďaka najmä všetkým poskytovateľom, ktorí prispeli pri realizácii úlohy aj svojimi postrehmi, radami a skúsenosťami, vyseletovali pre nás vhodné zdrojové materiály, vytriedili ich, kontrolovali, priraďovali k nim potrebné metadáta, svojim študentom sprostredkovali formulár s metadátami a odporučili nám ďalšie vhodné zdroje. S prispením každého poskytovateľa nárečových záznamov sa v databázach SNK spracuje materiál z celého Slovenska, ktorý bude spätne k dispozícii na výskum pre každého záujemcu zo spolupracujúcich slovakistických pracovísk.

#### **Literatúra:**

BALLEKOVÁ, K. a kol. (2014): *Zo studnice rodnej reči 2*. Bratislava: VEDA, Vydavateľstvo Slovenskej akadémie vied. (Pavol Žigo – výber z posudku).

BARTKO, L. (1997): Abovské a spišské nárečia. In: J. . Mlacek (ed.) *Studia Academica Slovaca*. 26. Prednášky XXXIII. letného seminára slovenského jazyka a kultúry. Bratislava:

---

zvukovo-textového nárečového korpusu by bol možný vzdialený prístup (podobne ako pri Slovenskom hovorenom korpuse).



Stimul – Centrum informatiky a vzdelávania FF UK, s. 198–205.

BILSKÝ, J. (2014): *Súčasný stav používania nárečia v Dlhej Lúke*. [Diplomová práca]. Košice: Filozofická fakulta Univerzity P. J. Šafárika.

GAJDOŠOVÁ, K. — ŠIMKOVÁ, M. (2014): Slovenský hovorený korpus (2008 — 2012). In: K. Gajdošová — A. Žáková (eds.) *Jazykovedné štúdie XXXI. Rozvoj jazykových technológií a zdrojov na Slovensku a vo svete (10 rokov Slovenského národného korpusu)*. Bratislava: Veda, Vydavateľstvo Slovenskej akadémie vied, s. 65–84.

HABOVŠTIAK, A. (1984): *Atlas slovenského jazyka. 4. zv. Tvorenie slov. Časť prvá. Mapy. Časť druhá. Úvod – Komentáre – Dotazník – Indexy*. Bratislava: Veda, Vydavateľstvo Slovenskej akadémie vied.

KAČALA, J. (1999): Slovenské nárečia v novej jazykovej situácii. In: A. Ferenčíková (ed.) *Nárečia a národný jazyk*. Bratislava: VEDA, Vydavateľstvo Slovenskej akadémie vied, s. 105–112.

KOVÁČOVÁ, V. (2005): *Sotácke nárečia na západoslovensko-východoslovenskom jazykovom pomedzí*. [Dizertačná práca]. Bratislava: Slavistický ústav Jána Stanislava SAV.

KRAJČOVIČ, R. (1988): *Vývin slovenského jazyka a dialektológia*. Bratislava: Slovenské pedagogické nakladateľstvo.

MÚCSKOVÁ, G. – MUZIKOVÁ, K. – WAMBACH, V. (2012): *Praktická dialektológia. Vysokoškolská príručka na nárečovú interpretáciu*. Wien: Facultas Verlags- & Buchhandels AG Wien.

RIPKA, I. (1999): Dialektika dialektológie. In: A. Ferenčíková (ed.) *Nárečia a národný jazyk*. Bratislava: VEDA, Vydavateľstvo Slovenskej akadémie vied, s. 99–104.

*Slovník slovenských nárečí. I (A – K)*. (1994) Ved. red. I. Ripka. Bratislava: VEDA, Vydavateľstvo Slovenskej akadémie vied.

*Slovník slovenských nárečí II (L – P)*. (2006) Ved. red. I. Ripka. Bratislava: VEDA, Vydavateľstvo Slovenskej akadémie vied.

*Archív nárečí Slovenského národného korpusu*. Bratislava: Jazykovedný ústav Ľ. Štúra SAV 2013. Interná verzia na dátovom úložisku SNK.

*Internetová stránka oddelenia Slovenského národného korpusu*. [Cit. 2015-03-20.] Dostupné na: <http://korpus.juls.savba.sk>.

*Korpus nárečí Slovenského národného korpusu – dialekt-1.0*. Bratislava: Jazykovedný ústav Ľ. Štúra SAV 2014. Dostupné na internete: <http://korpus.juls.savba.sk>.

*Slovenský hovorený korpus – s-hovor-4.0*. Bratislava: Jazykovedný ústav Ľ. Štúra SAV 2012. Dostupné na internete: <http://korpus.juls.savba.sk>.

**Résumé**

**An Archive of Dialects of the Slovak National Corpus**

The preparation of the Archive of Dialects of the Slovak National Corpus (AN SNK) started in 2013 at the Slovak National Corpus Department of L. Štúr Institute of Linguistics, Slovak Academy of Sciences. The collaboration with the Slovak universities, dialectologists and other contributors made it possible to build an extensive database containing more than 560 unique sound recordings of various Slovak dialects, majority of them recorded by students. The recordings and corresponding written resources included in the AN SNK have been sorted, their sound quality improved and supplied by metadata describing the researcher, informants and the recording. The pilot version of the archive will be made available soon as a unique and representative resource freely accessible to all SNK users for research and scientific purposes.