

Велика кількість таблиць БД «корпусСловника» зумовлена автоматичним способом вилучення текстової інформації та імпорту текстових даних, що спричинило дублювання текстової інформації у деяких полях таблиць, тому на другому етапі конструювання електронної лексикографічної системи – розроблення програмних запитів та створення консольного інтерфейсу – кількість таблиць буде зменшена. Сьогодні колектив проекту працює над цим завданням. Найближчим часом електронна версія словника «Активні ресурси сучасної української номінації» буде представлена на порталі mova.info. Учасники проекту щиро дякують колективу авторів словника «Активні ресурси сучасної української номінації» за наданий електронний варіант тексту Словника.

Література

1. Активні ресурси сучасної української номінації: Ідеографічний словник нової лексики / Є. А. Карпіловська, Л. П. Кислюк, Н. Ф. Клименко, В. І. Критська, Т. К. Пузирєва, Ю. В. Романюк; Відп. ред. Є. А. Карпіловська. Київ: ТОВ «КММ», 2013. 416 с.
2. Граматичний словник української літературної мови. Словозміна / В. І. Критська, Т. І. Недозим, Л. В. Орлова, Т. К. Пузирєва, Ю. В. Романюк; Відп. ред. Н. Ф. Клименко. Київ: Видавничий дім Дмитра Бураго, 2011. 760 с.
3. Граматичний словник української мови. URL: <http://www.mova.info/grmasl.aspx>.
4. Колодяжная Л. И. Структура словарного текста в аспекте машинной лексикографии: автореф. дис. ...канд. филол. наук. Москва, 1986. 23 с.

Наталія Клименко
(*м. Покровськ*)

ПРИНЦИПИ УКЛАДАННЯ СЛОВНИКА СХІДНОСТЕПОВИХ ГОВІРОК

Останнім часом активізувалося вивчення словникового складу новостворених українських діалектів. І чимала заслуга в цьому належить дослідникам слобожанських говірок. Ще на початку ХХІ ст. побачили світ такі лексикографічні праці, як: «Словник східнослобожанських українських говірок» (автори: К. Д. Глуховцева, В. В. Леснова, І. О. Ніколаєнко, Т. П. Терновська, В. Д. Ужченко), «Словник діалектної лексики Луганщини» (відп. редактор – З. С. Сікорська). Приблизно в цей же час з'являється низка публікацій А. А. Сагаровського, присвячених укладанню діалектного словника Харківщини: («Лексика Центральної Слобожанщини (Харківщини) як об'єкт словництва» (2001), «Фрагмент діалектного словника Харківщини»

(2001), «Із досвіду укладання діалектного словника Центральної Слобожанщини (Харківщини)» (2003) та ін. У доробку лексикологів південно-східного наріччя української мови маємо й інші праці: «Фразеологічний словник говірок Нижньої Наддніпрянщини» В. А. Чабаненка (2001), «Фразеологічний словник східнослобожанських і степових говірок Донбасу» В. Д. Ужченка (2002), «Словник українських говорів Одещини» (2011) та ін.

Найменш лексикографічно опрацьованими дотепер залишаються східностепові говірки Донеччини. У 2006 році побачив світ перший випуск «Матеріалів до словника східностепових українських говірок», видання яких було заплановане як поетапне відображення роботи над укладанням Словника східностепових українських говірок. Наразі йдеться про настійну необхідність активізувати подальшу роботу над укладанням словника говірок Донеччини. Актуальність цього полягає в тому, що східностепові говірки вже досить давно «привертають увагу дослідників історією формування та розвитку, своєрідною мовною структурою, хоча й досі залишаються чи не найменш вивченими серед інших південно-східних говорів» [3, с.3]. Крім того, збереження найповнішої за обсягом інформації про східностепові українські говірки є важливим, доки вони не зазнали глибинних структурних перетворень унаслідок не лише інтенсивної міждіалектної взаємодії, а й через воєнні дії та окупацію, що охопили з 2014 року частину Донецької області [2, с.136].

Метою цієї розвідки є висвітлення основних принципів укладання Словника східностепових українських говірок, що передбачає ілюстрацію лексичного матеріалу, який увійде до Словника, та опрацювання його в лексикографічних статтях.

Для укладання Словника, а відповідно й другого випуску «Матеріалів до словника східностепових українських говірок», використано записи діалектного матеріалу, здійснені в 60 обстежених населених пунктів Донецької області, що відображають діалектні особливості лексичного складу одного з новостворених периферійних говорів української мови. Основою Словника є багатотисячна картотека, збагачення якої новими записами з різних населених пунктів Донецької області триває.

Пропонований Словник є диференційним, що не охоплює всієї лексики, яка функціонує в східностепових говірках Донеччини, а подає лише традиційну діалектну лексику, яка не увійшла до основного словникового складу української літературної мови й репрезентує всі фонетичні, словотвірні, семантичні різновиди діалектної лексики, уживаної в донецьких говірках.

У Словнику буде зібрано як традиційну лексику повсякденного спілкування, так і назви, пов'язані з різними видами діяльності, переважно господарською та професійною. До реєстру Словника включено: назви їжі,

напоїв, кухонного начиння, сільськогосподарську лексику, метеорологічні, астрономічні номени, назви людських хвороб і понять, пов'язаних з ними, ботанічну лексику, назви людини та її рис, транспортну лексику, назви одягу, взуття, прикрас, головних уборів, меблів та інших хатніх речей, рибацьку лексику, назви бджільництва.

Реєстр словника охоплює: **специфічно східностепові слова** (наприклад, *све^рбл'ачка* 'кропива' (Є М), *с'орбанка* 'дуже рідкий суп' (Бл Амв), *л'опанка* 'рідка картопля' (Н-укр М), *кладунка* (Кам Д), *се^ррун'ц'і* 'штаниці для немовлят' (Є М), *ки^сл'атни^к* 'приплюснутий, поколотий ножем хлібець з кислого тіста, спечений на сковороді' (Бил Сл), *л'апанець* 'т. с.' (Грод), *троакар* 'довгий верхній піджак без підкладки, пошитий із чотирьох частин, для носіння восени або навесні' (С, Сонц, Мих П, Ол М, Роя, Олексіївка ВН), **регіоналізми**, тобто слова, поширені в інших південно-східних говорах (*буханка* 'хлібина', *це^рбєрка* 'відро', *бутил'ок* 'трилітрова скляна банка', *розорит* 'розібрати на частини', *пособит* 'допомогти', *баран*, *баранчик* 'середина кавуна', *годіна*, *годінка* 'добра погода', *гробкі* 'поминальне свято', *сапєтка* 'низький широкий плетений кошик з двома ручками', *потанц'і* 'страва зі шматочків хліба, зволжених водою та приправлених олією', *тормозок* '1) їжа, яку беруть шахтарі з собою в шахту; 2) будь-яка їжа, яку беруть із собою в дорогу').

Крім лексичних, до Словника східностепових українських говірок увійдуть і семантичні діалектизми, що, як правило, є лексемами відомими іншим українським говорам, однак у східностепових говірках мають іншу семантичну структуру: *мамка* 'велика хижа риба роду окунеподібних, яка містить багато ікри й живе в прісній та морській воді' (Бз НА); *токсикоз* 'отруєння бджіл отрутохімікатами, які застосовують для боротьби зі шкідливими комахами, хворобами сільськогосподарських культур та бур'янами' (Гр); *р'абчик* 'моряцька смугаста спідня сорочка, звичайно трикотажна' (М-Кач НА), *лапс'ордак* 'ірон. назва будь-якого верхнього одягу' (у більшості східностепових говірок); *сукн'а* 'спідниця з сукна' (Лис Амв); *д'ід* 'дротяний гачок гаплика' (Вільне, Шевч ВН), *закот* 'верхній зимовий одяг рівного крою, з овечої шкіри, з широким виложистим коміром' (Н-петр ВН), *калитá* 'кишеня, яку жінки носили під фартухом' (Шах Яс).

Реєстрові слова подано спрощеною фонетичною транскрипцією й розташовано в алфавітному порядку. Тлумачення значення реєстрових слів здійснено переважно описово, хоча використано й антонімічний, синонімічний та комбінований (тлумачно-синонімічний) способи.

Так само, як і в першому випуску «Матеріалів до словника східностепових українських говірок», послідовно подано *граматичні ремарки*. При відмінюваних іменниках, записаних у називному відмінку однини, указано закінчення родового відмінка однини та рід. Якщо іменник уживається тільки в однині, це позначено скороченням **одн.**, якщо слово

вживається у множині, на це вказує ремарка **мн.**, збірні іменники позначено скороченням **збірн.** Реєстрові слова-прикметники подано в називному відмінку однини чоловічого роду й зазначено закінчення родового відмінка однини. Дієслова репрезентовано у формі інфінітива; наведено закінчення I особи одн. та 3 особи мн. дієслів теперішнього часу. При словах, що належать до інших частин мови, є вказівка на цю частину мови. Незначну групу дієслів подано в безособовій формі, у таких випадках зазначено вид і форму дієслова. Наприклад, *розгодини^елося* док., безос., *визв'їзділо*, док., безос.

Обов'язковим елементом кожної словниковій статті є ілюстративний матеріал, який розкриває особливості функціонування слова в говірці. Ілюстрації вживання реєстрового слова передано спрощеною фонетичною транскрипцією після викладу значення слова. Паспортизацію діалектної лексики наведено за значенням та розміщено за кожним ілюстративним матеріалом.

Кожне реєстрове слово, крім односкладових, подається з наголосами – одним, рідше двома. Два наголоси поставлено на одному слові, коли в досліджуваних говірках зафіксовано два акцентуаційні варіанти (*р^озп^ірка* (*р^озп^ірка*, *р^озп^ерка*), **-и, ж.** *Розп^ірка спереду чоловічих штанів*). Наголоси позначаються й у формах відмінюваних слів (*п^і'інж^ак* (*п^і'ін'ж^ак*, *п^і'ж^ак*, *п^і'ндж^ак*), **-а, ч.** *Легкий чоловічий піджак на підкладці, який носили восени або навесні*).

За умови функціонування в східностепових говірках фонетичних або словотворчих варіантів реєстрового слова, їх наведено в одній словниковій статті. *Фонетичні варіанти* подано в дужках поруч з реєстровим словом, а *словотворчі варіанти* розміщено в кінці словникової статті й позначено горизонтальною стрілкою – порівняймо:

з^іркало (з^іркало), **-а, с.** Спеціально виготовлений предмет побуту з відшліфованою поверхнею, що відображає предмети, розташовані перед ним (у більшості східностепових говірок).

→з^іркал'це

При значній кількості реєстрових слів наприкінці словникової статті наводяться *синонімічні назви*, що позначаються зірочкою (*):

байц^ір^ія, **збірн.** Нитки, вистрєпані з одягу (С)

по:бр'із^ій от^е байц^ір^ія / ск'іки^е вон^о б^уд^е"т'іл'іп^ац':а // (С)

*стр'іп^ія, стр'іп^ки, тр'іп^ки, об:ів^ки, пац'ірк^ки

У багатозначних реєстрових словах кожне окреме значення позначено цифрою й наведено з абзацу. На перше місце винесено найбільш загальне значення, а також те, що має найбільший ареал поширення. *Омоніми* подано як окремі реєстрові слова.

Словник відображає деякі *стилістичні явища*, особливості вияву парадигматичних та синтагматичних відношень у системі лексики східностепових українських говірок. Наприклад,

лапс'орда́к (лапс'ерда́к, лапс'урда́к, лапсарда́к),-а́, ч. *Ірон.* назва будь-якого верхнього одягу (у більшості східностепових говірок)

Кожне реєстрове слово паспортизується літерним скороченням назв населених пунктів, у яких воно зафіксоване. Список деяких скорочень назв населених пунктів, зазначених у Словнику й продубльованих у доповіді, наводимо нижче. Якщо слово охоплює весь ареал поширення досліджуваних говірок, то в дужках після значення вказано (в усіх східностепових говірках) або (у більшості східностепових говірок).

Перспективи подальших наукових досліджень убачаємо не тільки в завершенні роботи над другим випуском «Матеріалів до до словника східностепових українських говірок», а й укладанні повного Словника східностепових українських говірок.

Список скорочень обстежених населених пунктів Донецької області

Є М – Єлизаветівка Мар'їнського району, Бл Амв – Благодатне Амвросіївського району, Н-укр М – Новоукраїнка Мар'їнського району, Кам Д – Кам'янка Добропільського району, Бил Сл – Билбасівка Слов'янського району, Грод – Гродівка Покровського району, С – Селидове, Сонц – Сонцеве Покровського району, Мих П – Михайлівка Покровського району, Ол М – Олександрівка Мар'їнського району, Роя – Роя Мар'їнського району, Олексіївка ВН – Олексіївка Великоновосілківського району, Бз НА – Безіменне Новоазовського району, Гр – Гришине Покровського району, М-Кач НА – М – Качкари Новоазовського району, Лис Амв – Лисиче Амвросіївського району, Вільне – Вільне Волноваського району, Шевч ВН – Шевченко Великоновосілківського району, Н-петр ВН – Новопетриківка Великоновосілківського району, Шах Яс – Шахова Ясинуватського району.

Література

1. Вікторіна О. М. Принципи укладання словника лексики та фразеології народної медицини та лікувальної магії Кіровоградщини. *Сучасні проблеми мовознавства та літературознавства*. Ужгород, 2008. № 12. С. 94-97.
2. Клименко Н. Б. Діалектні тексти як джерело дослідження генези та сучасного стану українських говірок Донеччини. *Лінгвістичні студії / Linguistic Studies*: Вінниця: ДонНУ імені Василя Стуса, 2018. Вип. 35. С. 135-139.
3. Омельченко З. Л., Клименко Н. Б. Матеріали до словника східностепових українських говірок. Донецьк: Вид-во ДонНУ, 2006. 114 с.

4. Сагаровський А. А. Лексика Центральної Слобожанщини (Харківщини) як об'єкт словництва. *Сучасні проблеми мовознавства та літературознавства*. Ужгород, 2001. В. 4. С. 465-469.
5. Сагаровський А. А. Фрагмент діалектного словника Харківщини. *Волинь-Житомирщина*. Житомир, 2001. № 6. С. 249-254.
6. Сікорська З. С. Про особливості Словника діалектної лексики Луганщини. *Волинь-Житомирщина*. Житомир, 2005. № 14. С. 192-196.
7. Словник українських східнослобожанських говірок / К. Глуховцева, В. Леснова, І. Ніколаєнко, Т. Тернавська, В. Ужченко. Луганськ, 2002. 234 с.
8. Хобзей Н. В. Сучасне українське словництво. *Україна: культурна спадщина, національна свідомість, державність*. Львів, 2006. № 15. С. 819-826.

**Валентина Робейко, Микола Сажок, Руслан Селюх, Дмитро Федорин,
Олександр Юхименко
(м. Київ)**

МОНІТОРИНГ НОВИНИХ YOUTUBE-КАНАЛІВ ЗА ДОПОМОГОЮ СИСТЕМИ РОЗПІЗНАВАННЯ МОВЛЕННЯ

Оперативний моніторинг телерадіопрограм є важливим суспільним завданням, що має на меті здійснення запису, зберігання та аналізу аудіовізуальних даних медіа-сервісів на всій території України. На сьогодні велика кількість світових телеканалів ведуть своє мовлення засобами веб-сервісу YouTube. Цей сервіс дає змогу користувачам Інтернету переглядати трансляції як наживо, так і окремі випуски телепередач або сюжетів у записі. Українські телеканали віднедавна активно опановують новий спосіб донесення медійної інформації до користувача. В недалекому минулому провести аудіовідеозапис трансляцій телеканалів можливо було через антену або з місцевих сервісів цифрового ТБ. Тепер же до переліку національних новинних каналів, що максимально використовують можливості YouTube, входять канали *Еспресо*, *5 канал*, *112*, *24*, *Zik* та ряд інших. Крім того окремі передачі новин, такі як *ТСН*, *Факти ICTV* тощо, створюють власні YouTube-канали, викладають туди записи своїх випусків та інколи транслюють їх наживо.

Перша з розроблених в Україні систем, що давала змогу проводити оперативний моніторинг аудіоінформації, забезпечувала надійне перетворення на текст записів спонтанного мовлення, в яких наявні лише незначні шуми і завади — [2]. Сюжети, знайдені за ключовими словами, користувач переглядав синхронно з розпізнаним текстом із метою аналізу або редагування.

На сьогодні рівень розробок у галузі розпізнавання мовлення