

Страшко І. В.

Національний педагогічний університет імені М. П. Драгоманова

КОРПУС ESLO ЯК ДЖЕРЕЛО СОЦІОЛІНГВІСТИЧНИХ ДАНИХ

Розглядається специфіка створення, розвитку й експлуатації усного корпусу для французької мови, який має назву «Les Enquêtes sociolinguistiques à Orléans» – ESLO. Вивчаються принципи збирання лінгвального матеріалу, критерії відбору інформантів, використання системи опитування, узагальнюється практичний досвід експлуатації корпусу. Основний зміст дослідження становить аналіз структурних і функціональних характеристик мовленнєвого корпусу ESLO. На підставі аналізу робіт французьких науковців установлено, що він був зібраний у два етапи з інтервалом у сорок років, що призвело до появи двох діахронічно диференційованих частин – ESLO 1 і ESLO 2. Підкреслюється, що ESLO 1 є найважливішим свідченням про французьку мову 70-х рр., його формування відбувалось на зорі соціолінгвістики та розпочиналося як дидактичний проєкт. Зазначається, що таке соціолінгвістичне закріплення корпусних даних, зібраних у реальних ситуаціях комунікації, не тільки відкрило нове предметне поле їх вивчення, а й актуалізувало проблеми аналізу дискурсивності міської території, переплетення мовних і соціальних змін, мовної варіативності. Встановлено, що ESLO 2 створювався на методологічному й епістемологічному підґрунті цифрового повороту в галузі гуманітарних і соціальних наук і відображає сучасний стан розвитку французької мови. ESLO 1 і ESLO 2 дозволяють проаналізувати мову у синхронічному та діахронічному аспектах, дають можливість формувати власний субкорпус відповідно до обраних соціологічних параметрів. Гетерогенність мовних структур, подана на основі записів інформантів різної статі, віку, соціальних і професійних груп, робить корпус ESLO унікальним джерелом соціолінгвістичної інформації та інструментом наукових досліджень. Соціолінгвістичні опитування в Орлеані, які набули корпусної форми, легітимізують розмовну французьку мову як таку в усій її різноманітності, тим самим надаючи їй статусу національного надбання.

Ключові слова: усний корпус, соціолінгвістика, соціологія, французька мова, розмовне мовлення, звукозапис.

Постановка проблеми. Корпусна лінгвістика – один із найбільш динамічних напрямів сучасного мовознавства. Попри те, що репрезентативні усні корпуси дають можливість проведення масштабних і статистично достовірних досліджень, у когорті корпусних ресурсів переважають текстові, а кількість усних баз, побудованих для різних мов, донині незначна. Інтерес до створення мовленнєвих корпусів був значною мірою ініційований розробками в галузі автоматичного розпізнавання мовлення [1; 4; 5; 25], проте їх експлуатація сьогодні виникає у багатьох галузях. Незважаючи на вагомий доробок у дослідженні корпусних ресурсів, питання формування й експлуатації корпусів для французької мови дотепер залишаються маловивченими. Серед розмаїття усних баз даних заслуговує на увагу корпус ESLO (фр. Les Enquêtes sociolinguistiques à Orléans).

Актуальність пропонованої розвідки визначається, по-перше, необхідністю заповнення лакун у вивченні дистинктивних характеристик мовленнєвого корпусу ESLO, по-друге, його безпере-

чною значимістю як джерела соціолінгвістичної інформації та інструменту наукових досліджень.

Постановка завдання. Метою статті є розгляд принципів збирання лінгвального матеріалу та критеріїв відбору інформантів у практиці створення корпусу ESLO.

В основу дослідження покладені методи дескриптивного та структурного аналізу, класифікації та інтерпретації.

Аналіз останніх досліджень і публікацій.

Робота ґрунтується на теоретичній базі, зібраній шляхом систематизації положень праць українських і зарубіжних дослідників. Так, аналіз засад створення усного корпусу ESLO проведений на основі робіт французьких науковців (L. Abouda, G. Bergounioux, O. Baude, C. Dumont та ін.). Окремо аналізувалися джерела щодо мовної варіативності, репрезентовані роботами J. Boutet, F. Gadet, C. Moïse. Розглядалися також дослідження Л. Крисіна, Б. Ларіна, С. Vauvois, Ph. Blanchet, Т. Bulot, L.-J. Calvet. Дотичними є праці Р. Bourdieu, С. L. Briggs. Серед останніх

досліджень, релевантних для нашої розвідки, слід виділити роботи I. Eshkol, G. Cloiseau, C. Dugua, L. Kanaan-Caillol, N. Rendulić, що вивчають мовний матеріал крізь призму соціолінгвістики. У цій статті ми спираємося також на розвідки українських і зарубіжних науковців (В. Жуковська, В. Захаров, Н. Леміш, О. Кривнова), чії теоретичні положення ми використовували для аналізу структурних, функціональних характеристик мовленнєвого корпусу ESLO.

Виклад основного матеріалу. Корпус ESLO (Les Enquêtes sociolinguistiques à Orléans) займає унікальне місце серед французьких усних баз даних. Окрім досить великого обсягу (близько 7 000 000 слововживань), який дозволяє скласти збалансований субкорпус, особливість ESLO полягає в самій його структурі: цей «резервуар» даних був зібраний у два етапи з інтервалом у 40 років, що призвело до появи двох діахронічно розмежованих частин – завершеного корпусу ESLO 1, розроблення якого тривало у 1968–1974 рр., та відкритого ESLO 2, розпочатого у 2008 р., поповнення джерельної бази якого продовжується [9; 10; 13; 16; 27].

ESLO знаменує дві важливі віхи у розвитку французького мовознавства: ESLO 1 створювався на зорі соціолінгвістики, а ESLO 2 розвивався на методологічному й епістемологічному підґрунті цифрового повороту в галузі гуманітарних і соціальних наук [14].

У рамках соціолінгвістики дослідницькі підходи до методології складання й експлуатації корпусів набувають іншого значення, оскільки остання, за П. Енкреве, «прагне <...> зафіксувати гетерогенний набір слів, який зможе сформувавши мову спільноти» [Цит. за: 14, с. 2]. Саме у соціальному характері мови, у зв'язку між мовою та соціально кваліфікованою мовною спільнотою дослідники (У. Вайнрайх, У. Лабов) вбачають першоджерело та рушійну силу мовних змін. Отже, спостерігати за мовою потрібно саме в соціальній динаміці, а не в ілюзорному пошуку стабілізованої мови в межах гомогенної спільноти. Пропонуючи теоретичні рамки, в яких чітко визначена соціальна природа мови, соціолінгвістика центрується наперед навколо питання даних, місце яких стає першорядним [18; 23]. Соціолінгвістичні опитування в Орлеані, які набули корпусної форми, дають можливість оцінити рамки цього позиціонування на основі конкретних наукових проєктів, серед яких: дослідження мови на лексико-семантичному, фонетико-фонологічному рівні (наприклад, аналіз фактів

фонетичної реалізації слова *омлет* (фр. *omelette*) у типовому форматі діалогу), аналіз лінгвокогнітивного рівня мовної особистості у кулінарному дискурсі (див.: [8; 11; 17]). Зокрема, побудований субкорпус кулінарних рецептів (тривалістю 1 год 40 хв) на основі 96 відповідей на питання анкети *Comment est-ce qu'on fait une omelette? Pourriez-vous m'expliquer comment on fait? Pouvez-vous me donner la recette de l'omelette? (Як приготувати омлет? Чи не могли б ви пояснити, як його готують? Можете дати рецепт омлету?)* Ці питання, зазначає О. Бод, стали своєрідним «вимірним приладом», лінгвокультурологічним «експериментом», у якому представлені спільні когнітивні уявлення про кулінарний рецепт (опис серії послідовних операцій із певною кількістю інгредієнтів), конкретні лексичні одиниці, регульовані форми спільного обміну та суб'єктивне вираження смаків і думок [12]. Появу у 2014 р. квантитативно і квалітативно збалансованого субкорпусу ESLO-MD (фр. ESLO Micro-Diachronie), який нараховує 1 млн слововживань (більше 80 годин записів), можна вважати наступним етапом розвитку корпусних досліджень ESLO: цей задокументований і анотований корпус, побудований для компаративного аналізу вживання *le futur simple – le futur périphrastique*, за твердженням укладачів, може слугувати базою даних для відстеження варіативності усного мовлення, мікродіахронічної еволюції мовних явищ [11; 30].

Створення усного корпусу під назвою «Les Enquêtes sociolinguistiques à Orléans» (пізніше ESLO 1) в руслі соціолінгвістики передбачало спостереження за мовою в її повсякденному використанні, не відокремлюючи її від соціальних ситуацій, у яких вона виробляється. Цей корпус є найважливішим свідченням про французьку мову 70-х рр.: він нараховує приблизно 4 500 000 слововживань, 470 записів загальною тривалістю понад 300 годин. Його створення розпочиналося як дидактичний проєкт: вмотивована прагненням створити аудіовізуальний метод викладання мови на основі «живих», автентичних документів, група британських учених, лінгвістів за освітою, залучених до викладання французької як іноземної, поставила за мету зібрати системно об'єднаний цілісний набір живих матеріалів на основі опитування, проведеного у середньостатистичному французькому місті, позбавленому надмірно виражених мовних характеристик. Щоб виключити неконтрольовані змінні, потрібне було місто розміром з Орлеан, у якому можна було б відновити динаміку наявних мовних форм, із

визнаними та зафіксованими варіантами, що елімінує гіпотезу про вибір Орлеана як міста, у якому стандартизована мова представлена невизначеним регіональним різновидом [16]. Орлеан відповідає також усім необхідним критеріям: місцезнаходженню, незначній діалекталізації й акценту, історичній традиції, розвитку регіону.

Визначення загальних мовних практик не на основі записів мовлення типових інформантів, а через неоднорідність мовних практик певної місцевості означало прийняття соціальних відмінностей через наперед встановлені есенціалістські категорії (стать, вік, походження, соціально-професійну категорію). У певному сенсі такий підхід був частиною іншого бачення, іншого виміру соціолінгвістики, який ґрунтувався на врахуванні соціально зумовлених мовних відмінностей, на аналізі ідентифікованих фактів «мовної індивідуалізації» (P. Bulot) мешканців міста. Це врахування різноманітності не виключало, однак, пошуку мови, якою користується конкретна мовна спільнота: йшлося про те, щоб спостерігати та фіксувати в певний час, у невеликому місті динаміку мовних практик його мешканців.

Спираючись на офіційні статистичні реєстри, група опитування отримала від Національного інституту статистики й економічних досліджень (фр. L'Institut national de la statistique et des études économiques – INSEE) координати шести сотень осіб, по 100 на кожну із шести соціально-професійних категорій відповідно до класифікації INSEE. Згодом співпраця із групою CSE П. Бурдьє призвела до її перегляду та проектування нової сітки – «шкали АМ» (названої на честь її автора Алекса Мулліно), що включала п'ять соціальних категорій і слугувала доповненням до попередньої. У ній робиться спроба, поряд із загальноприйнятими критеріями, врахувати потенційну географічну мобільність і культурні практики респондентів, співвіднести їхній культурний рівень (який, за теорією П. Бурдьє [21], розглядався у зв'язку з рівнем освіти, аналізом ієрархічних і статусних структур і сегрегації) із соціально-професійними категоріями. Як наслідок, в усному опитуванні взяли участь 156 осіб (80 чоловіків і 76 жінок), розподілених на три вікові групи: 18–30, 31–50, 51 і старші. Виключивши з числа респондентів дітей, представників іммігрантських кіл і вихідців із села – щоб усунути діалектну варіативність, – укладачі склали диверсифіковану вибірку інформантів, мешканців Орлеану і його околиць, зіставляючи їх за віком, статтю, професійним середовищем і рівнем освіти. Ця

вибірка не була соціологічно репрезентативною ані щодо населення міста загалом, ані щодо відносного розподілу соціально-професійних категорій, однак для лінгвістичного дослідження вона забезпечує достатньо рівну кількість інформантів у кожній із категорій [16].

Усне опитування було розроблене на основі відкритої, «соціолінгвістичної» та закритої анкети. Перша (відкрита) полягала в аналізі особистого досвіду респондента через використання різних типів дискурсу. Як зазначають дослідники [16], окремі запитання були сформульовані таким чином, щоб спровокувати мовців на вживання конкретної форми, конкретного типу дискурсу. Напр.: *Quelqu'un frappe à la porte de cette pièce. Qu'est-ce que vous lui dites? (Хтось стукає у двері цієї кімнати. Що ви йому/їй скажете?)*. Друга, напівзакрита, під назвою «соціолінгвістична анкета» містить добірку висловлювань респондентів про мовну та культурну норму, їхнє ставлення до певного виду мовленнєвої діяльності, напр.: *Où parlez-vous le mieux le français? Est-ce que vous trouvez que d'une façon générale, on parle de mieux en mieux ou de plus en plus mal ou est-ce que c'est sans changement? Qu'est-ce que vous pensez du français? Est-ce que d'après vous il y a des différences dans la façon de parler selon le milieu social? Est-ce que vous seriez favorable à une réforme de l'orthographe? (Де найкраще говорять французькою? Загалом, на вашу думку, ми говоримо дедалі краще, дедалі гірше чи нічого не змінилося? Що ви думаєте про французьку? На вашу думку, чи існують відмінності у способі мовлення залежно від соціального середовища? Чи підтримали б ви реформу орфографії?)* Вплив соціологічних поглядів П. Бурдьє (див.: [21]) також зустрічається у третій анкеті, закритій, яка стосується цивільного стану, культурних звичок респондентів. Вона нараховувала 26 питань, вісім із яких стосувалися їхньої особистої ідентифікації: місця проживання, соціального походження, віку. Напр.: *Depuis combien de temps habitez-vous Orléans? Qu'est-ce qui vous a amené à vivre à Orléans? Vous êtes à Orléans depuis toujours? (Як довго ви живете в Орлеані? Що змусило вас переїхати жити в Орлеан? Ви завжди жили в Орлеані?)*; сім – навчання в освітніх закладах, включаючи більш деталізовані питання щодо вивчення конкретних предметів, зокрема латини (напр.: *Qu'est-ce qu'on devrait apprendre surtout aux enfants à l'école? Dans quelles matières aimeriez-vous que vos enfants soient forts? Est-ce que d'une façon générale vous pensez qu'il soit utile que les enfants apprennent le latin à l'école? – Чого насамперед*

слід навчати дітей у школі? У яких предметах ви б хотіли, щоб ваші діти були найбільш обізнані? Як ви вважаєте, чи корисно було б, якби діти вивчали латину в школі?), десять – ставлення респондентів до засобів масової інформації, питання стосовно політики та політичних подій (Напр.: *Est-ce que vous avez un poste de radio? Ecoutez-vous la radio? Votre chaîne préférée? Quel genre d'émissions est-ce que vous choisissez? Est-ce que à votre avis la radio la télévision le cinéma ce sont des instruments de culture? Est-ce qu'il y a un parti politique qui représente vos opinions? Que pensez-vous des événements de mai 68? – У вас є радіо? Ви слухаєте радіо? Який ваш улюблений канал? Які передачі ви обираєте? На вашу думку, чи є інструментами культури радіо, телебачення та кіно? Чи існує політична партія, яка представляє ваші погляди? Що ви думаєте про події травня 1968 р.?). Анкета включала також питання щодо роботи, відпочинку, освіти (напр.: *Qu'est-ce que vous faites comme travail?, Et votre femme, est-ce qu'elle travaille aussi? Pourquoi (pas)?, Et vos enfants, que font-ils? / métier?, est-ce que dans votre profession il y a beaucoup de femmes qui exercent? Qu'est-ce que vous faites de votre temps libre – soirées, week-end? Ким ви працюєте? А ваша дружина, вона теж працює? Чому (ні)? Чим займаються ваші діти? / Яка у них професія? У вашому професійному середовищі багато жінок? Чим ви займаєтесь у вільний час – вечорами, у вихідні?)).**

Варто зазначити, що на рівні соціолінгвістичної анкети, автором якої був соціолог Бернар Верньє, чітко простежується належність до концепції, розробленої в Центрі європейської соціології, очолюваному тоді П. Бурдьє: гіпотеза, покладена в її основу, базується на соціальній теорії передачі культури та мови [15; 16; 21]. Надання еквівалентної ваги культурному й економічному капіталу, внесення в анкету питань щодо соціальної мобільності населення не тільки перемістило суто економічну соціологію у площину аналізу культурних звичок респондентів, а й дало змогу визначити їхню мовну компетентність, ставлення до мовних стандартів, отримати оціночні судження про мовленнєве виробництво. Крім того, питання були сформульовані таким чином, щоб встановити відповідність між формою / формулюванням і прагматикою, мовною грою і мовленнєвим актом, зв'язок між мовними формами та соціальними, культурними та ситуативними факторами комунікації.

Разом із проведеним соціологічним опитуванням, яке становило ядро корпусу ESLO 1, були зібрані записи спонтанного мовлення у різних

комунікативних ситуаціях: від ситуативних, неформальних (неформальних телефонних розмов, записів із «прихованим мікрофоном» у магазинах, на вулиці, громадському транспорті) до формалізованих (круглих столів, дискусій, публічних дебатів, зустрічей із науковцями, представниками муніципалітету, політичними діячами, профспілковими лідерами тощо).

Підтримання функціонування ESLO 1 і можливість повторного використання зібраних матеріалів здійснювалися шляхом поширення корпусу, збереження записів та анкет, каталогізації та індексації, транскрибування репрезентативних уривків, які охоплювали усі категорії респондентів (за класифікацією INSEE та AM). У 1974 р. був опублікований описовий та аналітичний каталог, у якому перераховувалися звукозаписи, надано їх короткий зміст, індексація питань, структура анкети, точна соціологічна категоризація інформантів та опис ситуації опитування.

Застосування усних даних ESLO 1 для дидактичних потреб було мінімальним: реалізовано лише два навчальні проєкти щодо викладання французької мови як іноземної [16]. Наукова експлуатація ESLO 1 французькими дослідниками тривалий час теж майже відсутня [10; 14], що пов'язане насамперед із трудомісткістю оперування даними: використання корпусу суттєво ускладнювалося через наявність лише друкованого каталогу, рукописної документації, існування транскрипційних записів тільки для невеликої частини усних даних, що спричинене, зокрема, відсутністю теоретичного підґрунтя, оскільки наукова діяльність щодо транскрибування, розроблення транскрипційних схем для розмовної французької почнеться з Клер Бланш-Бенвеніст у 1980-х рр. Відсутність експлуатації ESLO 1 пояснюється й епістемологічними причинами: символічне домінування писемної форми мови унеможливило рецепцію усної форми як об'єкта лінгвістичних досліджень, тож усне мовлення, не будучи легітимізованим у французькому мовознавстві, довго залишалося поза увагою науковців. До того ж французька лінгвістика, будучи побудованою на сосюрівській дихотомії, винесла «за рамки» опис варіативності загалом і усних форм зокрема [10].

У рамках проєкту ELILAP / ELICOP (ELILAP 1980–83 рр., згодом ELICOP – 1993–2001 рр.) дослідники з університетів у Лувені й Амстердамі транскрибували частину корпусу ESLO, зробили частиномовну і синтаксичну розмітку, розробили

конкордансер. У 1993 р. науковці Центру антропологічних і лінгвістичних досліджень в Орлеані провели оцифрування вихідних документів, включаючи аудіоматеріали на магнітній стрічці, транскрибували і зробили розмітку усього корпусу [10; 13; 26].

Сьогодні первинні дані ESLO 1 (аудіокасети, документація щодо усіх видів записів) знаходяться у службі звукових архівів Національної бібліотеки Франції (<http://archivesetmanuscripts.bnf.fr/ead.html?id=FRBNFEAD000095934>), яка забезпечує їх збереження, оцифрування, опис і доступ із локального сервера [18].

Розвиток інформаційних технологій, мобілізація інструментів, використання унікальних перспектив цифрового поля і пов'язані з ним нові дослідницькі практики стимулювали розробку корпусу ESLO 2, розпочатого у 2008 р. у лабораторії LLL (фр. Laboratoire Ligérien de Linguistique) при університеті Орлеана. ESLO 2 демонструє розвиток джерельного наповнення, удосконалення методології збору усних даних, що відобразилося у корпусній архітектурі (детальніше див.: <http://eslo.huma-num.fr/index.php/pagecorpus/pagepresentationcorpus>).

Нині доступ до звукозаписів, транскрипцій, метаданих корпусів ESLO надається на веб-сайті проекту лінгвістичної лабораторії, розміщеному в мережі на платформі Huma-Num за адресою: <http://eslo.huma-num.fr>. Цей сайт також дає можливість ознайомитися з каталогами (за записами, інформантами та транскриптами), генерувати субкорпуси, завантажувати дані та метадані, читати синхронізовані зі звуком транскрипти (сегментовані на просодично, синтаксично та семантично узгоджені одиниці, з одночасною наявністю трьох версій транскрипції – А, В і С), вести пошук і відбір інформації, для чого створений відповідний інтерфейс, поданий на рис. 1.

Варто зазначити, що транскрипція корпусу ESLO – нульова транскрипція, яку розробники розглядають як першу анотацію, що супроводжує навігацію у звуковому сигналі, звучання якого відображає орфографічний запис. Транскрипти синхронізовані зі звуком за допомогою програмного забезпечення Transcriber, що дозволяє здійснювати сегментацію на декількох рівнях [28, с. 20] (рис. 2).

Ось приклад мовленнєвого фрагмента, поділеного на сегменти [28, с. 23] (рис. 3).



Рис. 1. Скріншот сторінки пошуку у каталогах для корпусу ESLO

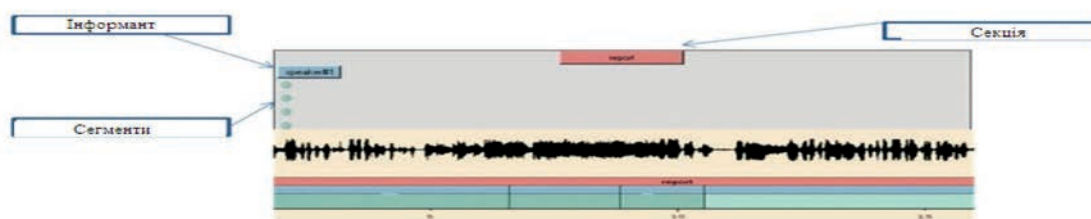


Рис. 2. Приклад сегментації за секціями, за інформантами та за сегментами



Рис. 3. Приклад поділеного на сегменти мовленнєвого фрагмента

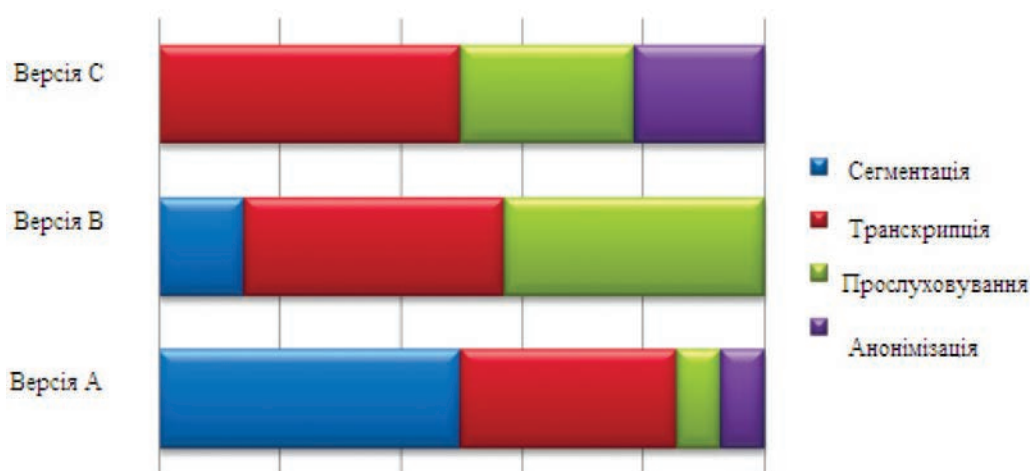


Рис. 4. Приклад розподілу завдань залежно від версії транскрипції

Наведемо ще один ілюстративний приклад розподілу завдань залежно від версії транскрипції (А – необроблена версія, В – коректурна, С – перевірена версія, яка є доступною для усіх користувачів) [28, с. 18] (рис. 4).

Попри те, що більшість документів і записів є у відкритому доступі, з усім корпусом можна ознайомитися тільки після підписання відповідної угоди.

Необхідність урахування сучасних соціологічних теорій і нових соціальних реалій унеможливила застосування в ESLO 2 класифікації АМ. Тому інформанти просто класифіковані за вісьмома соціально-професійними категоріями відповідно до нової номенклатури INSEE.

Корпуси ESLO 1 і ESLO 2 як один із найбільших резервуарів даних французького розмовного мовлення дозволяють проаналізувати мову у синхронічному та діахронічному аспектах, надають можливість формувати власний підкорпус відповідно до обраних соціологічних параметрів (гендерних,

вікових, професійних, освітніх). Водночас якісна звукова нерівноцінність деяких записів у ESLO 1 і ESLO 2, яка проте вирішується завдяки їхній акустичній градації, може призвести до обмежень дослідницьких завдань, наприклад, фонетичних або орфоепічних, однак на рівні лексичних і граматичних запитів їх експлуатація виявляється досить продуктивною.

Соціолінгвістичне закріплення даних, зібраних у реальних ситуаціях комунікації, та обсяг розмовної французької мови не тільки відкрили нове предметне поле їх вивчення, а й актуалізували проблеми аналізу дискурсивності міської території, переплетення мовних і соціальних змін, мовної варіативності. Сьогодні, як і раніше, корпус ESLO слугує плідною базою досліджень у галузі лінгвістики, дидактики, соціології, історії, легітимізуючи розмовну французьку мову як таку в усій її різноманітності, тим самим надаючи їй статусу національного надбання.

Список літератури:

1. Богданов Д. С., Кривнова О. Ф., Подрабинович А. Я. Современный инструментарий для разработки речевых технологий. *Информационные технологии и вычислительные системы*. 2004. Т. 2.
2. Жуковська В. В. Ресурси корпусної лінгвістики у дослідженні історичної динаміки мови. *Матеріали міжнародної наукової конференції «Слово і речення: синтактика, семантика, прагматика»*. Київ : Київ. ун-т ім. Б. Грінченка, 2013. С. 151–156.
3. Захаров В. П. Корпусная лингвистика : учебно-методическое пособие. Санкт-Петербург, 2005. 48 с.
4. Кривнова О. Ф. Фонетическое обеспечение для построения речевого корпуса. *Акустика речи. Медицинская и биологическая акустика*. 2003. Т. 3. С. 118–122.
5. Кривнова О. Ф., Захаров Л. М., Строкин Г. С. Речевые корпуса (опыт разработки и использование). *Сборник трудов Международного семинара Диалог*. 2001. С. 230–236.
6. Леміш Н. Корпусна метарозмітка спеціальних текстів з лінгвоантропогенезу. *Науковий часопис Національного педагогічного університету імені М. П. Драгоманова. Серія 9: Сучасні тенденції розвитку мов*. 2017. Вип. 15. С. 139–152.
7. Крысин Л. П. К социальным различиям в использовании языковых вариантов. *Вопросы языкознания*. № 3. 1973.
8. Abouda L. Le temps des omelettes. Une nouvelle valeur pour le présent? *Revue de Sémantique et Pragmatique*. 2012. С. 71–94.
9. Abouda L., Baude O. Constituer et exploiter un grand corpus oral: choix et enjeux théoriques. Le cas des ESLO. *Corpus en Lettres et Sciences sociales, Des documents numériques à l'interprétation*. 2007. С. 161–168.
10. Abouda L., Baude O. Du français fondamental aux Eslo. *Grand corpus de français parlé, Bilan historique et perspectives de recherche*. 2005. Т. 33. №. 2. С. 131–146.
11. Abouda L., Skrovec M. Pour une micro-diachronie de l'oral: le corpus ESLO-MD. *SHS Web of Conferences. EDP Sciences*, 2018. Т. 46. С. 11004.
12. Baude, Olivier. Corpus d'Orléans: ESLO1: la recette de l'omelette (Version 1). Laboratoire Ligérien de Linguistique. 2015. <https://doi.org/10.34847/COCOON.262C6F48-FC06-375B-9313-DD4142D55732> [en ligne] (02.02.2021).
13. Baude O., Dugua C. (Re) faire le corpus d'Orléans quarante ans après: quoi de neuf, linguiste? *Corpus*. 2011. № 10. С. 99–118.
14. Baude O., Dugua C. Les ESLO, du portrait sonore au paysage digital. *Corpus*. 2016. № 15.
15. Bergounioux G. Le français oral en variation: les enseignements de l'Enquête Sociolinguistique à Orléans. 2018.
16. Bergounioux G., Baraduc J., Dumont C. L'étude sociolinguistique sur Orléans (1966-1991): 25 ans d'histoire d'un corpus. *Langue française*. 1992. №. 93. С. 74–93.
17. Bergounioux G., Introduction. Champion. Linguistique de corpus : une étude de cas. La recette de l'omelette, Champion. *Linguistique de corpus: une étude de cas*. Paris : Champion. 2016. С. 7–13.
18. Bibliothèque nationale de France: <https://www.bnf.fr/fr> [en ligne]. (23. 01. 2021).
19. Bulot, T., Bauvois, C. et Blanchet, P. (dir.). Sociolinguistique urbaine (Variations linguistiques: images urbaines et sociales). Cahiers de sociolinguistiques 6. Rennes : Presses universitaires de Rennes. 2001.
20. Boutet J., Gadet F. Pour une approche de la variation linguistique. *Le français aujourd'hui*. 2003. № 4. С. 17–24.
21. Bourdieu P. Le marché linguistique. *Questions de sociologie*. Paris : Editions de Minuit. 1984.
22. Briggs C. L. Interviewing, power/knowledge, and social inequality. *Handbook of interview research: Context and method*. 2002. С. 911–922.
23. Calvet L. J. Aux origines de la sociolinguistique la conférence de sociolinguistique de l'UCLA (1964). *Langage & société*. 1999. Т. 88. № 1. С. 25–57.
24. Claudine Moïse. Pour quelle sociolinguistique urbaine? *Pratiques langagières urbaines, enjeux identitaires, enjeux cognitifs*. 2002. С. 75–87.
25. Eshkol, I. Interpréter le contexte dans un corpus oral: fonctions et limites du traitement automatique des données linguistiques. 2015. С. 67–80.
26. Eshkol-Taravella I., Baude, O., Maurel, D., Hriba, L., Dugua, C., & Tellier, I. et al. Un grand corpus oral «disponible»: le corpus d'Orléans 1 1968–2012. *Traitement automatique des langues*. 2011. Т. 53. № 2. С. 17–46.
27. ESLO. URL: <http://eslo.huma-num.fr> [en ligne]. (23. 01. 2021).
28. Guide du transcripateur et du relecteur des ESLOs. URL: http://eslo.huma-num.fr/images/eslo/pdf/GUIDE_TRANSCRIPTEUR_V4_mai2013.pdf [en ligne]. (24. 01. 2021).
29. Labov William, *Le Parler ordinaire*, Paris, Minuit, 1978.
30. Rendulić N., Kanaan-Cailloil L. Je crois que, je pense que: valeurs et variation dans un corpus oral diachronique. *SHS Web of Conferences. EDP Sciences*, 2016. Т. 27. С. 02014.

Strashko I. V. THE ESLO CORPUS AS A SOURCE OF SOCIOLINGUISTIC DATA

The paper describes the specifics, the development, and the use of the ESLO corpus. The principles of lingual material selection, the criteria for choosing the informants, the use of the polling system, and practical information about the corpus exploitation were analyzed as well. The main content of the research was the analysis of the structural and functional characteristics of the ESLO speech corpus. On the analysis of French scientists papers, it was found that the speech corpus was constituted in two stages with an interval of forty years, which resulted in the emergence of two diachronically differentiated parts – the ESLO 1 and the ESLO 2. The ESLO 1 was created at the dawn of sociolinguistics, its constitution began as a didactic project. It is the most important evidence of the French language of the seventies. The creation of this corpus in the angle of sociolinguistics provided for language observation in its daily use, without separating it from the social situations in which it is produced. It is emphasized that such sociolinguistic anchorage of corpus data collected in real situations of communication has not only opened a new field for their study but also actualized the problems of discourse-analysis of urban territory, intertwining of linguistic and social changes, language variability. It was revealed that the ESLO 2 was created on the methodological and epistemological basis of the digital technologies use, and reflects the current state of the French language. The ESLO 1 and the ESLO 2 make it possible to analyze language in synchronic and diachronic aspects, provide an opportunity to form a subcorpus accordingly to the selected sociological parameters. The heterogeneity of language structures, represented by oral speech recordings of the informants of different sexes, ages, social and professional groups, makes the corpus a unique source of sociolinguistic information and a tool of scientific research. It was noted that sociolinguistic surveys in Orleans, which have acquired a corpus form, legitimize the spoken French language as such in all its diversity, thereby giving it the status of a national treasure.

Key words: *speech corpus, sociolinguistics, sociology, French language, conversational speech, recording.*