

Анастасія Камараулі

Маріам Камараулі

Захарія Пуртсхвадзе

Johann Wolfgang Goethe-Universität
Institut für Empirische Sprachwissenschaft

Learner Corpora (навчальні корпуси) та їх потенціал для мультилінгвальної освіти

Анотація

В даній статті увага, передусім, концентрується на проблемі формування та використання певного виду мовного корпусу – навчального корпусу (learner corpus) – у мультилінгвальному навчанні. Дається опис всіх релевантних аспектів концептуалізації, мотивації та побудовування навчального корпусу, також наводиться приклад німецького навчального корпусу FALKO (*Fehlerannotiertes Lernerkorpus* – «корпус анотування помилок»). Крім того, ми торкаємось можливостей навчального корпусу для грузинської мови, наводячи приклади з реальних курсів грузинської мови у Франкфуртському університеті ім. Гете й Тбіліському державному університеті. У статті підкреслюється потенціал навчальних корпусів для мультилінгвальної освіти й мультилінгвальної освіти вчителів.

Ключові слова: *Learner Corpus* – навчальний корпус, мультилінгвальна освіта.

1. Учніський корпус - знайомство

Згідно з загально визнаним визначенням, учнівський корпус являє собою електронну колекцію автентичних текстів (мовного матеріалу), написаних студентами, які вивчають мову як іноземну або другу, й зібраних в електронну базу

даних.¹ Крім того, комп'ютерні учнівські корпуси є електронними колекціями автентичних FL/SL (ІМ/ДМ) текстових даних, зібраних за певним дизайнерським критерієм з конкретною SLA/FLT (ОДМ/ВІМ) метою. Корпуси кодуються

¹ Anna O'Keeffe et al: From corpus to classroom. Cambridge University Press 2007, S.23.

стандартизовано й однорідно, а також документується їх походження².

Напрочуд важливе визначення для учнівського корпусу має ідея мовної помилки, яку можна визначити як «лінгвістичну форму або комбінацію форм, які, в тому ж контексті і за подібних умов, ймовірно, не були б синтезовані носіями мови-колегами мовця»³.

Мовні матеріали можуть аналізуватися програмним забезпеченням та редагуватися. Аналіз має різні цілі. Учніський корпус являє собою новий вид мовного корпусу, який з'явився на початку 1990-х⁴ З того часу було розроблено багато учнівських корпусів для різних мов. У Католицькому Університеті Лувен було складено список з 138 різних учнівських корпусів.⁵ Список не є повним, проте в ньому знайшли місце головні учнівські корпуси, а також він надає вдалий стислий огляд. У ньому надаються різні класифікації учнівських корпусів, зокрема за мовою-реципієнтом, медіа та типом тексту. Прогнозовано, більшість з них має англійську мову як мову-реципієнта. 87 із загалу з 138 для англійської мови, 10 для

французької мови, 9 для німецької мови, 8 для іспанської мови, 3 для італійської мови, тощо. Корпусу працюють з різними медіа.. 87 з них користуються письмовими медіа (наприклад. The Advanced Learner English Corpus (ALEC), Університету Уппсала - тексти komponували/писали студенти), 33 з них користуються розмовними медіа (наприклад. The ENGLISH corpus, Університету Провансу - читання, усна мова), 11 з них користуються письмовими і розмовними медіа і 3 з них використовують мультимедіа. Загалом, ці корпуси мають лише одну мову-реципієнта, проте окрім 127 монолінгвальних корпусів список також містить 11 мультилінгвальних корпусів (наприклад. The Corpus of Young Learner Interlanguage (CYLIL) Vrije Universiteit Брюсель або The Eastern European English learner corpus Eberhard Karls університету Тюбінгена).

1.1. FALCO – приклад корпусу

Наразі, ми звернемося до детального аналізу одного з корпусів – корпусу The FALCO

(FehlerannotiertesLernerkorpus 'error

² Granger et al. 2002: 7.

³ Corder(1986:37). Corder, Stephen Pit (1986): The role of interpretation in the study. In: Corder, Stephen P. (Hrsg.): Error analysis and interlanguage. 4. impr. Oxford: Oxford University Press, S. 35–44.

⁴ Anna O'Keeffe et al: From corpus to classroom. Cambridge University Press 2007, S.23.

⁵ <http://www.uclouvain.be/en-cecl-1cworld.html>.

annotated learner corpus⁶ – «корпус анотування помилок»⁶. The FALKO було розроблено в університеті Гумбольдт Берлін, автори - Anke Lüdeling та Maik Walter у 2004 році. Головний корпус FALKO можна поділити на п'ять менших корпусів - *FalkoSummaryVL*, *FalkoSummaryL1 V1.2*, *FalkoSummaryL2 V1.2*, *FalkoEssayL1 V1.2*, *FalkoEssayL2 V2.0*.

1. Учніські тексти *Learner texts (FalkoSummaryL2)*⁷ : Колекція резюме (лінгвістичні тексти і літературні дослідження), написаних студентами, які вивчають німецьку мову на підвищеному рівні (C1-C2). Вони були написані в рамках іспиту, що є обов'язковим для іноземних студентів, які спеціалізуються в німецькій філології. Іспит проводився у Вільному університеті Берліна.
2. Тексти носія мови (*FalkoSummaryL1*): Колекція тих же текстів, що й у FSL2, але написаних носіями німецької мови.

3. Оригінальні тексти (*FalkoSummaryVL*): Колекції лінгвістичних та філологічних текстів, які слугували шаблонами для резюме. Загалом налічується 197 текстів, написаних 98 учнями.
4. *FalkoSummaryL1 1.1*⁸ : Для цього корпусу було здійснено 4 збори даних. Вони проводилися у Вільному університеті Берліна і в університеті Гумбольдт Берліна. Знову, участь у зборі даних брали лише студенти, які спеціалізуються на німецькій філології. Умови збору даних були однаковими.
5. *Falko Essay Corpus*⁹ : Цей корпус складається з двох підкорпусів.
 - a. *FalkoEssayL2*: містить колекцію творів, написаних студентами, які вивчають німецьку мову на підвищеному рівні. Для творів було запропоновано 4 різних теми, й учасники мали отримати щонайменше 60 зі 100 балів на C-тесті.

⁶ <https://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/falko>.

<https://korpling.german.hu-berlin.de/falko-suche/search.html>.

⁷ Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012, S.8ff.

⁸ Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012, S.17.

⁹ Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012, S.19f.

б. *FalkoEssayL1*: містить колекцію творів, написаних носіями мови. Учасниками виступали учні випускних класів трьох різних середніх шкіл. Запропоновані теми були ідентичними *FalkoEssayL2*, умови іспиту також не відрізнялись.

додавати більше рівнів анотування (багатошарова окрема анотація). Загалом, *Falko* містить письмові тексти, написані студентами, які вивчають німецьку мову на підвищеному рівні. Найбільш анотований підкорпус це колекція резюме.¹⁰ Анотації автоматично анотувалися Treetagger (Schmid 1994).¹¹ База даних також містить достатню інформацію про авторів, як то рівень освіти, рівень володіння мовою та багато іншої¹²

Усі підкорпуси мають різні рівні анотування, і структура *Falko* дозволяє

Table 1. Рівні анотування *Falko*¹³. ((word - слово) – Learner utterance – вираз учня, (kpos) – Part of speech – частина мови, (target-hypothesis – ціль-гіпотеза) - Assumption about proposed utterance – допущення щодо можливого виразу, (ref) – evidence reference - доказ.)

word	Dabei	Ist	es	zu	beachten
kpos	PAV	VAFIN	PPER	PTKZU	VVINP
lemma	dabei	Sein	es	zu	beachten
target_hypothesis	<i>Dabei ist zu beachten</i>				
ref	70	71	72	73	74

Рівень анотації містить цільову гіпотезу, щоб дозволяє реконструювати помилку, зроблену учнем. Помилки ідентифікуються шляхом порівняння оригінальних висловлювань з так званими реконструйованими висловлюваннями, тобто, правильними висловлюваннями зі значенням, яке мав на увазі учень.

¹⁰ Siemen et al., FALKO - Ein fehlerannotiertes Lernerkorpus des Deutschen, S.1.

¹¹ Mark Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012, S.4.

¹² Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012, S.6.

¹³ Karin Schmidt Lernerkorpora: Ressourcen für die Deutsch-als-Fremdsprache-Forschung (publikationen.uni-frankfurt.de/files/.../Schmidt_Lernerkorpora.pdf).

Table 2¹⁴. Аналіз помилок у in FALKO

word	Fraue	konnten	solche	gesellschaftlich	Zustand	veränder	
	n		n	en		n	
target_hyp_	Fraue	konnten	eine	solche	gesellschaftlic	Zustand	veränder
1	n		n	n	hen		n
-target_hyp	Fraue	konnten	solch	eine	gesellschaftlic	Zustand	veränder
_1	n			n	hen		n

Використання Falko продемонструвало, які аспекти німецької мови є складнішими для учнів (наприклад, правильне використання артиклів та прийменників) і отже, яким аспектам слід приділяти більше уваги у викладанні.

2. Мотивація, конструкція та функціонування учнівських корпусів

Головним завданням учнівського корпусу є анотування помилок. Тому тексти, написані учнями, слід порівнювати з текстами, авторами яких є носії мови. З цього робиться допущення, що, у порівнянні з помилками, зробленими учнями, існує правильна версія, надана носіями мови. Це видається легким, але фактично не існує правильного способу виразити себе першою мовою ¹⁵. Мова є напрочуд гнучкою, отже існує багато різних способів вербалізувати одне й те ж речення. До того ж, мова знаходиться в постійному процесі розвитку, отже вона постійно змінюється. Те, що видається правильним на сьогодні, може бути абсолютно невірним у майбутньому.

Проте, учнівські корпуси є важливим інструментом для дидактичних досліджень і власне для дидактики.

Мотивація для побудування учнівського корпусу може варіюватися. Наприклад, у вивченні іноземної мови деякі дієслівні конструкції можуть бути напрочуд складними для початківців. Деякі конструкції майже повністю проігноровано у навчальних матеріалах. Тому це можна б було розглядати як аргумент на користь корпусів у подібних випадках. Учнівські корпусні аналізи повсякчас піддають критиці за те, за що й критикують методичні рекомендації, які ґрунтуються на корпусах носіїв мови: вони приймають до уваги один критерій, що є важливим для викладання, й не зважають на інші. У випадку з методичними рекомендаціями, які ґрунтуються на корпусах носіїв мови,

¹⁴ Falko-Handbuch S.39.

¹⁵ Siemen et al., FALKO - Ein fehlerannotiertes Lernerkorpus des Deutschen, S. 2.

відзначалося, що таким єдиним критерієм виступає частотність у вжитку носіїв мови. Проте, учнівський корпус певним чином мотивував би учня й сприяв усвідомленню мови. Вони стимулюють студента працювати активно й самостійно, і таким чином, вони, ймовірно, підвищують як мотивацію студента, так і навчальний ефект. Отже, корпус буде застосовуватись в освіті вчителів іноземної мови, як джерело прикладів, які можна використовувати в класі, та освітніх інструментів, і допоможе підготувати інструкції й викладацькі матеріали для конкретних навчальних груп.

Мотивація лінгвістів щодо формування учнівського корпусу є різною. Головною метою може бути поліпшення дидактичних методів. Учнівські корпуси можуть ідентифікувати конкретні проблеми учнів з певною мовою. Ці уявлення можуть допомогти покращити навчальні методи для цих учнів. Доцільно зазначити, що це є важливим інструментом дидактики іноземної мови і дає підґрунтя для аналізу типології помилок певних груп учнів. Тому, це безпрограшна ситуація як для учнів, так і для учителів. Порівнюючи учнівські тексти з текстами носіїв мови, учні безпосередньо можуть вивчитися від них і поліпшити власні мовні навички, а вчителі можуть пристосувати свої методи до конкретних груп учнів. Взагалі, головні цільові аудиторії учнівських корпусів - це

учні, які вивчають іноземну мову, та вчителі, які викладають іноземні мови. Окрім них, зиск з учбових корпусів матимуть також лінгвісти і дослідники дидактичних методів.

Не зважаючи на те, що учнівські корпуси відкривають нові можливості для дидактики іноземної мови, їх все ще зрідка можна зустріти в школах і мовних класах. Однією з головних причин цього слід назвати відсутність інформації та те, що корпуси розглядаються як науковий інструмент, а не викладацький інструмент¹⁶. Тому, важливо інструктувати вчителів і навчати студентів, щоб вони могли навчитися використовувати учнівські корпуси. На цьому етапі школи й університети повинні взяти на себе ініціативу та проводити семінари. Задля того, щоб допомогти студентам вивчити нову мову, вчитель може включати учнівські корпуси у власні уроки. Наприклад, вчителі можуть пропонувати завдання, які можна зробити лише за допомогою учнівських корпусів. Значна кількість слів має широкий ряд значень, і тому вони використовуються в цілому ряді контекстів. За допомогою учнівських корпусів, студенти можуть порівняти використання цих слів у тексті носія мови та ідентифікувати різні лексичні категорії¹⁷. Або, якщо у студентів виникає певне питання, вони можуть відповісти на нього,

¹⁶ Karin Aijmer, *Corpora and Language Teaching*, 2009 John Benjamins Publishing Company, S.47f.

¹⁷ Aijmer, *Corpora and Language Teaching*, 2009 JBPC, S.50f.

шукаючи в учнівських корпусах, потенційно перетворюючись зі студентів у дослідників мови¹⁸. Учнівські корпуси можуть використовуватись як додаток до граматики, ілюструючи граматичні правила.

3. Як конструюється учнівський корпус?

Для конструювання учнівського корпусу важливо зібрати велику кількість письмових і/або розмовних матеріалів. Письмові корпуси створити легше, ніж розмовні, адже у якості джерела для них може бути використаний інтернет. В них може міститись записане мовлення, інтерв'ю, есе, іспити, тощо. Вони повинні бути записані учнями. Для порівняння ті ж матеріали мають бути доступні у письмовій й усній формі, й представлені носіями мови¹⁹.

Базовий мовний корпус може вибудовуватись із розмовних або письмових текстів і може бути використаний з комерційно доступним програмним забезпеченням, яким будь-який пересічний користувач може працювати відносно легко. Звичайно, розмовний комплекс вимагає значно більше часу для конструювання, оскільки мовлення, наприклад у відео, необхідно буде переписати і можливо закодувати

через деякі його невербальні особливості. Порівняно з цим типом, конструювання письмового корпусу є напрочуд швидким, оскільки інтернет використовується як джерело. Кожен корпус потребує принципів дизайну. Вам доведеться зайнятися не лише дизайном, але й питанням обґрунтованості, оскільки існує боротьба між тим, що доступне, що етичне або що законне. Це може бути вирішальним чинником. Також, напрочуд важливо вирішити, що представити і як представити це краще для загальної мети. У такому разі, вам доведеться обрати обсяг даних, які ви хочете зібрати і використовувати.

У випадку з розмовними корпусами, наступним кроком є запис даних. Існує вибір видів запису, включаючи аналогові касети, цифрові медіа та аудіовізуальні електронні пристрої. У традиційних аналогових носіїв, незважаючи на їх доступність, є ряд недоліків. Їх незручно зберігати, й на відміну від цифрових записів, їх не можна легко комп'ютеризувати та накласти транскрипцію пізніше. Користування цифровими пристроями дозволяє накладати звук (і зображення, якщо ви користуєтесь аудіовізуальним рекордером) на вашу транскрипцію.

Важливим аспектом є дозвіл. Дозвіл на запис повинен бути отриманий від

¹⁸ John McHary Sinclair, *How to use Corpora in Language Teaching Studies in Corpus Linguistics*, Silvia Bernardini, 2004 John Benjamins Publishing Company, S.16.

¹⁹ Anna O'Keeffe et al: *From corpus to classroom*. Cambridge University Press 2007, *How to make a basic corpus*.

мовців заздалегідь, а також мають бути підписані необхідні документи, які надають право використовувати записи для дослідження або комерційних педагогічних матеріалів, тощо. Під час отримання дозволу, можливо, необхідно буде конкретизувати, як саме записи будуть використані. Після цього, головне завдання - транскрипція, оскільки розмовні дані необхідно вручну переписати, і саме це робить корпуси розмовної мови таким серйозним викликом. Найкраще їх зберігати як «прості текстові» файли, оскільки це надає максимальної гнучкості використання з комплектами різного ПЗ. Одна година транскрибованого мовлення може займати декілька днів для переписування, залежно від складності мови. У більшості випадків, кожне слово, вокаліз, усікання, коливання, перекриття, тощо, слід транскрибувати, на відміну від почищеної версії мови спікерів. Рівень деталізації транскрипції залежить від мети вашого корпусу. Якщо ви не маєте за мету знати, де відбуваються часткові перекриття висловлювання і переривання, то немає ніякої необхідності витрачати час на транскрипцію з таким рівнем деталізації. Ця важка праця включає *зіставлення (1) із зразком, сполучення слів (2), анотацію і частини мови (3), синоніми та антоніми (4), більш складні пошуки, з використанням комбінацій попередніх видів пошуків (5),*

запити, що базуються на частоті конструкцій в різні історичні періоди і реєстри мови (6), і запити з залученням такого що налаштовується, призначеного для користувача списку (7).

Файли транскрипції слід організувати таким чином, щоб початкова інформація могла бути прослідкована. Наприклад, може бути корисною можливість відновити інформацію, як наприклад, стать, вік, число спікерів, місце народження, рід занять, рівень освіти, де стався запис, стосунки між спікерами, тощо. Ця інформація може бути збережена на початку кожної копії як інформаційний «заголовок», або в окремій базі даних, де інформація реєструється з ім'ям файлу. Інакше кажучи, корпус повинен мати детальну анотацію та надавати можливості пошуків багатьох видів лінгвістичних явищ. Вміст кожного корпусу - колекція текстів і виразів мовою. Звичайно, ми повинні розрізняти письмові і розмовні корпуси. Матеріали для письмових корпусів порівняно легко збирати, тому що все фізично доступно. Вміст розмовних корпусів, як зазначалося вище, складніший для збору і редагування. Основні матеріали для розмовних корпусів загалом надаються через звукові або аудіовізуальні рекордери.

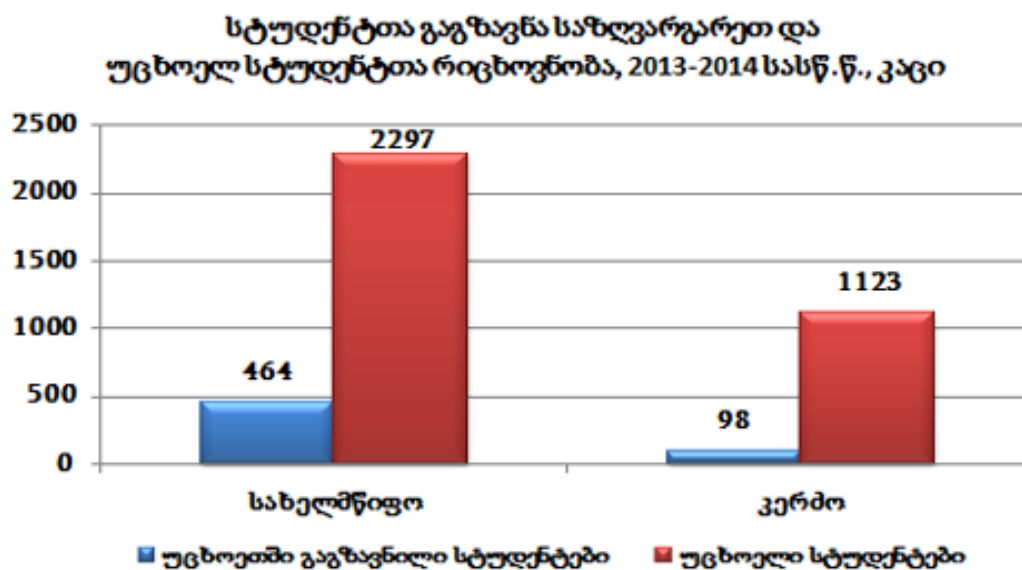


Рис 1. Національний департамент статистики
Грузії (Червоним позначена кількість
іноземних студентів в державних та

4. Зразок проекту грузинського учбового корпусу

Соціолінгвістична ситуація Грузії може характеризуватися як мультилінгвальна. У прикордонних регіонах Грузії, які межують з Азербайджаном і Вірменією, але в центральних регіонах, класи є мультилінгвальними. Це можна вважати освітнім викликом, спрямованим на розвиток відповідного мовного навчального контенту, який прицільно використовуватиме поширені помилки тих, хто вивчає грузинську мову. Один з перших кроків в цьому напрямі – збір та уніфікована документація усіх наявних помилок письмового та усного реєстрів.

Як перше джерело матеріалу, можна обрати учнівські групи в середніх школах Грузії. Згідно з офіційною статистикою,²⁰, близько 2000 носіїв грузинської мови вступають до вузів Грузії щороку.

Учителі і мовні репетитори можуть бути зобов'язані ділитися мультилінгвальним навчальним досвідом і систематизують помилки, що повторюються. Ці спостереження служать фундаментом для подальшого розвитку бази даних, що містить патерни помилок. Як ми бачили з FALKO (Ch.1.1), така база даних є необхідною для встановлення гіпотези як виду підготовки виправлення.

²⁰ GeoStat. http://www.geostat.ge/?action=page&p_id=205&lang=geo, 25.11.14, 13.00.



ქართული ენის ელექტრონული სწავლების კურსი

Gürcü dilinin elektron kursu

Վրաց լեզվի էլեկտրոնային ուսուցման դասընթաց

eLearning Course of Georgian

გზამკვლევი

MƏLUMAT KİTABÇASI

Ուղեցույց

GUIDE

Наразі в Грузії²¹ та закордоном²² існує декілька навчальних програм грузинської мови (дистанційний online курс університету Malmö, Швеція). Цільова група учнів складається з громадян Грузії, першою мовою яких є азербайджанська або вірменська. Тематична структура навчальних програм цієї програми ґрунтується на уявленні її авторів щодо можливих складнощів для учнів. Теми не підтверджені емпіричними доказами, не зважаючи на те, що акцент любої

Figure 2. eLecture for Georgian Language.

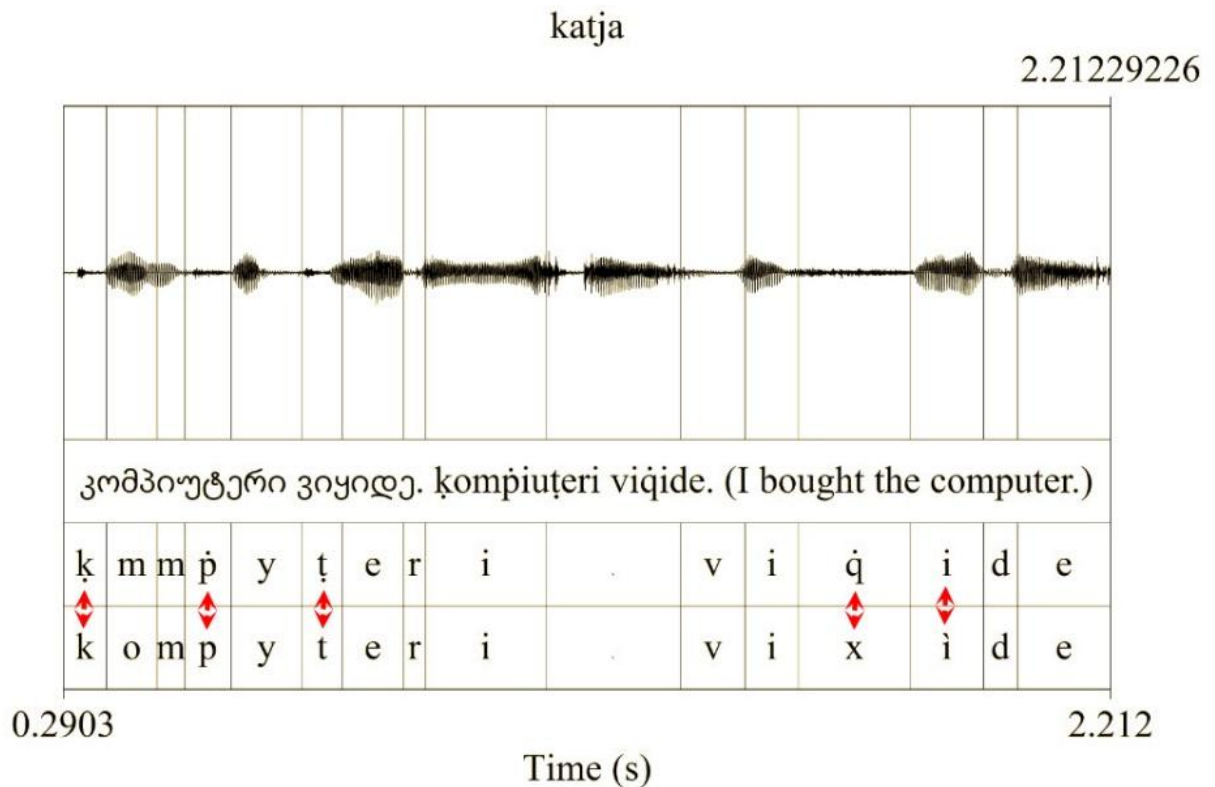
специфічної проблеми повинен бути підкріплений помилками, які учні роблять в реальному навчальному процесі.

Найбільшою складністю у вивченні грузинської (як і інших кавказьких мов) була і все ще залишається канонічна приголосна вимова. Характерною фонетичною особливістю кавказьких мов є окремі приголосні або групи приголосних. Отже, навчитися й використовувати ці приголосні є інтелектуальним та фізичним викликом для учня.

²¹ http://www.ice.ge/web/elearning_geo.html

²² <http://195.178.225.22/DiasporaGeo/Georgianonline.html>

Табл. 3. Приклад вимови учня грузинської мови. 5 помилок вимови у 2 словах.



Заплановані записи звукового матеріалу з учнями L2 служать надважливою базою даних для прискіпливого дослідження частих помилок у опануванні фонетикою та дозволяють зосередити викладацький - деякі приклади класичних мовних помилок, відмічених грузинськими викладачами ²³.

процес на цих помилках. Навіть поверхневе спостереження частих помилок виділяє проблематичні області, які мають бути центральною сферою уваги. Нижче

Table 4. Помилка порядку слів.

wording	sad	šen	iqavi?
lemma	sad_wh	šen_PPron.2Sg	qopna.be
hypothesis	šen	sad	iqavi?
Eng.	<i>Where have you been?</i>		

²³ Prof. Ketevan Gochitashvili. Tbilisi State University.

Table 5. Помилка узгодження.

wording	қargi	var.
lemma	қargi_good	қopna.be
hypothesis	қargad	var.
Eng.	<i>I am fine.</i>	

Table 6. Лексична помилка.

wording	didi	gemrieli-a
lemma	didi_big	gemrieli - қopna.be.Encl
hypothesis	zalian	gemrielia
Eng.	<i>It is very tasty.</i>	

Table 7. Помилка синтаксису, невірний порядок слів

wording	saxli	romeli	dgas	kalakši
lemma	saxli_house	romel_wh	dogma_stand	kalaki_city
hypothesis_1	saxli	romlic	dgas	kalakši
hypothesis_2	saxli	romlic	kalakši	dgas
Eng.	<i>The house, which is (standing) in the city.</i>			

wording	restorani	sad	viqavi	gušin
lemma	restorani _restaurant	sad_wh	қopna_be	gušin _yesterday
hypothesis_1	restorani	sadac	viqavi	gušin
hypothesis_2	restorani	sadac	gušin	viqavi
Eng.	<i>The restaurant I was in yesterday.</i>			

Бібліографічні посилання:

- Aijmer Karin, *Corpora and Language Teaching*, 2009 John Benjamins Publishing Company.
- Biber, Douglas; Jones, James K. (2009): *Quantitative methods in corpus linguistics*. In: Lüdeling, Corder, Steven Pit (1981): *Error Analysis and Interlanguage*. Oxford; Oxford University Press.
- Corder, Stephen Pit (1986): *The role of interpretation in the study*. In: Corder, Stephen P. (Hrsg.): *Error analysis and interlanguage*. 4. impr. Oxford: Oxford University Press.
- Diaz-Negrillo, Ana; Fernandez-Dominguez, Jesus (2006): *Error tagging systems for learner corpora*. In: RESLA 19.
- Granger, Sylviane. (2002): *A Bird's-eye View of Computer Learner Corpus Research*. In: Granger S., Kytö, Merja Anke (Hg.): *Corpus Linguistics. An International Handbook*. Vol. 2. Berlin: Mouton de Gruyter.
- O'Keeffe Anna et al: *From corpus to classroom*. Cambridge University Press 2007.
- Reznicek et al. *Das Falko-Handbuch. Korpusaufbau und Annotationen*, Version 2.01, 2012.
- Schmidt Karin, *Lernerkorpora: Ressourcen für die Deutsch-als-Fremdsprache-Forschung* (publikationen.ub.uni-frankfurt.de/files/.../Schmidt_Lernerkorpora.pdf).
- Siemen et al., *FALKO - Ein fehlerannotiertes Lernerkorpus des Deutschen*.
- Sinclair John McHary, *How to use Corpora in Language Teaching Studies in Corpus Linguistics*, Silvia Bernardini, 2004 John Benjamins Publishing Company.
- <http://www.uclouvain.be/en-cecl-lcworld.html>.
- <https://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/forschung/falko>.
- <https://korpling.german.hu-berlin.de/falko-suche/search.html>
- http://www.ice.ge/web/elearning_geo.html
- <http://195.178.225.22/DiasporaGeo/Georgianonline.html>