

ანასტასია ყამარაული

მარიამ ყამარაული

ზაქარია ფურცხვანაძე

გოეთეს სახელობის ფრანკფურტის უნივერსიტეტი

## „learner corpus“ და მისი პოტენციური მულტილინგვური სწავლებისას

### აბსტრაქტი

სტატია ძირითადად ეხება ენის კორპუსის ერთ-ერთი ტიპის - learner corpus-ის მულტილინგვური სწავლებისთვის გამოყენებას. სტატია აღწერს კონცეპტუალიზაციის, მოტივაციისა და კორპუსის აგებულების ყველა შესაბამის ასპექტს, მათ შორის, გერმანულ learner corpus FALKO-ს (*Fehlerannotiertes Lernerkorpus* 'error annotated learner corpus') მაგალითსაც. გარდა ამისა, ჩვენ შევეცდებით, ფრანკფურტის გოეთეს უნივერსიტეტსა და თბილისის სახელმწიფო უნივერსიტეტში არსებული ქართული ენის კურსების მაგალითებზე ვიმსჯელოთ learner corpus-ის შესაძლებლო გამოყენებაზე ქართული ენისთვის. სტატია ხაზს უსვამს learner corpora-ს პოტენციურ მულტილინგვურ სწავლებასა და მულტილინგვურ მასწავლებელთა განათლებაში.

*საკვანძო სიტყვები: Learner Corpus, მულტილინგვური სწავლება*

### 1. Learner Corpus - გაცნობა

საყოველთაო დეფინიციის მიხედვით, learner corpus არის აუთენტური ტექსტების (ენობრივი მასალის) ელექტრონული ნაკრები, რომელიც ინახება ელექტრონულ მონაცემთა ბაზაში და რომელიც გაკეთებულია იმ მოსწავლეების მიერ, რომლებიც სწავლობენ უცხო / მეორე ენას. გარდა ამისა, კომპიუტერული learner corpora

არის ელექტრონული ნაკრები აუთენტური FL/SL ტექსტუალური მონაცემებისა, რომლებიც შეგროვებულია ზუსტად შერჩეული კრიტერიუმების მიხედვით მეორე / უცხო ენის სწავლებისთვის. corpora სტანდარტიზებულიად კოდირებულია და დოკუმენტირებულია მისი წარმოშობისა და წყაროს მიხედვით.

learner corpus-ის არსი მდგომარეობს ენობრივი შეცდომის იდეაში, რომელიც შესაძლოა დახასიათდეს, როგორც "ლინგვისტური ფორმა ან ფორმების კომბინაცია, რომელიც იმავე კონტექსტსა და იმავე პირობებში, სავარაუდოდ არ დაიშვება მშობლიურ ენაზე მოლაპარაკების მიერ".

ენობრივი მასალები შესაძლოა გაანალიზდეს კომპიუტერული პროგრამის მიერ. ანალიზი სხვადასხვა მიზანს ემსახურება. learner corpus არის ენის კორპუსის ახალი ტიპი, რომელიც გამოჩნდა ადრე, 1990-იან წლებში. მას შემდეგ ბევრი learner corpora განვითარდა სხვადასხვა ენისთვის. Louvain-ის კათოლიკური უნივერსიტეტი ასახელებს 138 სხვადასხვა learner corpora-ს. სია არ არის სრული, მაგრამ მოიცავს ყველა მთავარ learner corpora-ს და კარგ სურათს იძლევა. აქ learner corpora კლასიფიცირებულია სხვადასხვა დამახასიათებელი ნიშნით, მაგალითად, მიზნობრივი ენით, მედიუმითა და ტექსტის ტიპით. როგორც მოსალოდნელია, უმეტეს მათგანს მიზნობრივ ენად აღებული აქვს ინგლისური. 138-დან 87 არის ინგლისურისთვის, 10 - ფრანგულისთვის, 9 - გერმანულისთვის, 8 - ესპანურისთვის, 3 - იტალიურისთვის და ა.შ. კორპუსი მუშაობს სხვადასხვა მედიასთან. 87 მათგანი იყენებს ბეჭდურ მედიას, მაგალითად, (The Advanced Learner English Corpus (ALEC), Uppsala University - texts composed/written by

students), 33 მათგანი იყენებს მოლაპარაკებულ მედიას, მაგალითად (The ENGLISH corpus, University of Provence - readings, oral language), 11 მათგანი იყენებს წერილობით და მოლაპარაკებულ მედიას და 3 მათგანი იყენებს მულტიმედიას. ზოგადად, ამ კორპუსს აქვს მხოლოდ ერთი მიზნობრივი ენა, მაგრამ გარდა 127 მონოლინგვური კორპუსისა, სიაში ასევე არის 11 მულტილინგვური კორპუსი (მაგალითად, The Corpus of Young Learner Interlanguage (CYLIL) Vrije Universiteit Brussel or The Eastern European English learner corpus Eberhard Karls University of Tübingen).

### 1.1. FALCO-ს მაგალითი

ახლა ჩვენ დეტალურად მიმოვიხილავთ ერთ learner corpus-ს - FALCO (FehlerannotiertesLernerkorpus 'error annotated learner corpus'). FALCO შეიქმნა ბერლინში ჰუმბოლდტის უნივერსიტეტში ანკე ლუდელინგისა და მაიკ უოლტერის მიერ 2004 წელს. მთავარი FALCO კორპუსი შეიძლება გაიყოს 5 პატარა კორპუსად: *FalkoSummaryVL*, *FalkoSummaryL1 V1.2*, *FalkoSummaryL2 V1.2*, *FalkoEssayL1 V1.2*, *FalkoEssayL2 V2.0*.

1. მოსწავლის ტექსტები (*Falko-SummaryL2*): (რეზიუმეების ნაკრები (ლინგვისტური ტექსტები და სალიტერატურო სწავლება), რომლებიც შექმნილია სწავლების მაღალ საფეხურზე მყოფი გერმანელი მოსწავლეების მიერ. ტექსტები

დაწერილია გამოცდის ფარგლებში, რომლის ჩაბარებაც სავალდებულოა უცხოელი სტუდენტებისთვის, რომლებიც სწავლობენ გერმანულ ფილოლოგიას, როგორც მთავარ საგანს. გამოცდა ჩატარდა ბერლინის თავისუფალ უნივერსიტეტში.

2. მშობლიურ ენაზე მოლაპარაკების ტექსტები (*FalkoSummaryL1*): იმავე ტექსტების ნაკრები, რომლებიც FSL2-შია, დაწერილი გერმანულენოვანი მოსწავლეების მიერ.

3. ორიგინალი ტექსტები (*Falko-SummaryVL*): ლინგვისტური და ფილოლოგიური ტექსტების ნაკრები, რომლებსაც იყენებენ როგორც მაგალითებს რეზიუმეებისთვის. სულ არის 197 ტექსტი, რომელიც დაწერილია 98 მოსწავლის მიერ.

4. *FalkoSummaryL1* 1.1<sup>1</sup>: ამ კორპუსისთვის მონაცემთა 4 ნაკრები შეიქმნა ბერლინის თავისუფალ უნივერსიტეტსა და ბერლინის ჰუმბოლდტის უნივერსიტეტებში. ამ მონაცემთა შეგროვებაშიც მხოლოდ იმ სტუდენტებმა მიიღეს მონაწილეობა, რომლებიც სწავლობენ გერმანულ ფილოლოგიას, როგორც მთავარ საგანს. ამ მონაცემთა შეგროვების პრობემებიც იგივე იყო.

5. *Falko Essay Corpus*<sup>2</sup>: ეს კორპუსი ორ ქვეკორპუსს შეიცავს:

ა) *FalkoEssayL2*: შეიცავს ესეების ნაკრებს, რომელიც დაწერილია სწავლების მაღალ საფეხურზე მყოფი გერმანელი მოსწავლეების მიერ. ამ ესეებისთვის 4 სხვადასხვა თემაა მოცემული და მონაწილეებმა C ტესტში უნდა მიიღონ არანაკლებ 60 ქულა 100 ქულიდან.

ბ) *FalkoEssayL1*: შეიცავს ესეების ნაკრებს, რომელიც მშობლიურ ენაზე მოლაპარაკების მიერ არის დაწერილი. მონაწილეები იყვნენ სამი სხვადასხვა საშუალო სკოლის დამამთავრებელი კლასის მოსწავლეები. თემები და გამოცდის პრობემებიც იყო იგივე, რაც *FalkoEssayL2* - ში.

ყველა ქვეკორპუსს აქვს მრავალშრიანი ანოტაციები და FALKO-ს სტრუქტურა საშუალებას იძლევა დამატებით უფრო მეტშრიანი ანოტაციებიც გაკეთდეს. ზოგადად, FALKO შეიცავს სწავლების მაღალ საფეხურზე მყოფი გერმანელი მოსწავლეების მიერ დაწერილ ტექსტებს. ყველაზე მეტად ანოტირებული ქვეკორპუსი არის რეზიუმეების ნაკრები. მონაცემთა ბაზა ასევე შეიცავს დაწვრილებით ინფორმაციას ავტორებზე, მაგალითად - მათ განათლებაზე, ენის ფლობის დონეზე და სხვა.

<sup>1</sup> Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012, S.17.

<sup>2</sup> Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012, S.19f.

**ცხრილი 1. Falko-ს ანოტაციის საფეხურები<sup>3</sup>.** ((word) – მოსწავლის გამონათქვამი, (kpos) – მეტყველების ნაწილი, (target-hypothesis) - შემოთავაზება-ვარაუდი, (ref) – მითითებული მტკიცებულება)

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

ანოტაცია შეიცავს გამიზნულ ნალში ადარებენ ე.წ. შესწორებულ ჰიპოთეზებს იმისათვის, რომ მოსწავლის გამონათქვამებს, რომლებიც არის სწორი მიერ დაშვებული შეცდომა შესწორდეს. გამონათქვამები მოსწავლის მიერ შეცდომების იდენტიფიცირება ხდება მინიჭებული მნიშვნელობით. შემდეგნაირად: გამონათქვამებს ორიგინ-

**ცხრილი 2<sup>4</sup>. შეცდომების ანალიზი FALKO-ში.**

word	Frauen	konnten	solchen	gesellschaftlichen	Zustand	veränder	
target_hyp_1	Frauen	konnten	einen solchen	gesellschaftlichen	Zustand	veränder	
target_hyp_1	Frauen	konnten	solch	einen	gesellschaftlichen	Zustand	veränder

FALKO-ს გამოყენებამ აჩვენა, გამოყენება და წინდებულის სწორად გერმანული ენის რომელი ასპექტები არის ხმარება...); ამასთან, რომელი ასპექტებია უფრო რთული მოსწავლეებისთვის პრიორიტეტული სწავლებაში. (მაგალითად, არტიკლების უშეცდომოდ

<sup>3</sup> Karin Schmidt Lernerkorpora: Ressourcen für die Deutsch-als-Fremdsprache-Forschung (publikationen.ub.uni-frankfurt.de/files/.../Schmidt\_Lernerkorpora.pdf).

<sup>4</sup> Falko-Handbuch S.39.

## 2. learner corpora-ს მოტივაციების, ენის სტრუქტურისა და ფუნქციის შესახებ

learner corpus-ის მთავარი ამოცანა არის შეცდომების ანოტირება. ტექსტები, რომლებიც დაწერილია მოსწავლეების მიერ, უნდა შედარდეს მშობლიურ ენაზე მოლაპარაკეთა მიერ დაწერილ ტექსტებთან. ეს გულისხმობს, რომ მოსწავლეების მიერ დაშვებულ შეცდომებთან შეფარდებით, არსებობს მშობლიურ ენაზე მოლაპარაკეთა მიერ შემოთავაზებული სწორი ვარიანტი. ერთი შეხედვით, ეს იოლი ჩანს, თუმცა, რეალურად, არ არსებობს ერთადერთი სწორი გზა, გამოხატო შენი თავი შენს მშობლიურ ენაზე. ენა არის ძალიან მოქნილი, ასე რომ, უამრავი სხვადასხვა გზა არსებობს, რომ ზუსტად იგივე აზრი გამოთქვა... გარდა ამისა, ენა მუდმივად განვითარების პროცესშია და, შესაბამისად, განუწყვეტლივ იცვლება. ის რაც დღეს სწორად ითვლება, შესაძლოა მომავალში არასწორი გახდეს. მიუხედავად ამისა, learner corpora არის მნიშვნელოვანი ინსტრუმენტი დიდაქტიკური სწავლებისა და თავად დიდაქტიკისთვის.

მოტივაციის შესაქმნელად learner corpus შეიძლება სხვადასხვაგვარი იყოს.

მაგალითად, უცხო ენის სწავლებაში ზოგიერთი ზმნის აგებულება შეიძლება რთული აღმოჩნდეს დამწყებთათვის. ზოგის სტრუქტურა თითქმის სრულიად უგულვებელყოფილია სასწავლო მასალებში. ასეთ შემთხვევაში, კორპუსი ძალიან სასარგებლოა. Learner corpus-ის ანალიზი ღია კრიტიკისთვის, მსგავსად სწავლების შესახებ რეკომენდაციებისა, რომლებიც ეფუძნება მშობლიური ენის კორპუსს და ითვალისწინებს მხოლოდ ერთ მნიშვნელოვან კრიტერიუმს სწავლებისთვის და უგულვებელყოფს სხვა დანარჩენს. მშობლიური ენის კორპუსზე დაფუძნებული სწავლების შესახებ რეკომენდაციების შემთხვევაში, ხშირად კრიტიკის საგანი ხდება სიხშირის ერთადერთ კრიტერიუმად მიჩნევა. მაგრამ learner corpus აუცილებლად მისცემს მოტივაციას მოსწავლეს და ენის ცნობადობასაც გაზრდის, მისცემს სტიმულს სტუდენტს, რომ იმუშაოს აქტიურად და დამოუკიდებლად. ეს გაზრდის როგორც სტუდენტის მოტივაციას, ისე მისი სწავლის ეფექტს. მოკლედ რომ ვთქვათ, კორპუსი გამოიყენება უცხო ენის მასწავლებელთა განათლებაში, როგორც საგანმანათლებლო იარაღი და კლასში გამოსადეგი მაგალითების მოყვანის წყარო. ამასთან, ის დაეხმარება მოსწავლეების სპეციფიკურ ჯგუფს სასწავლო მასალებისა და

ინსტრუქციების მიცემის თვალსაზრისით.

ლინგვისტებს აქვთ learner corpus-ის შექმნის განსხვავებული მოტივაციები. მთავარი მიზანი შეიძლება იყოს დიდაქტიკური მეთოდების გაუმჯობესება. learner corpus-ს შეუძლია იმ სპეციფიკური პრობლემების იდენტიფიცირება, რომლებსაც მოსწავლეები აწყდებიან რომელიმე ენის სწავლისას. ეს გარკვეული აღქმები შეიძლება დაეხმაროს მოსწავლეებს, გაიუმჯობესონ სწავლის მეთოდები. აქედან გამომდინარე, ეს არის მნიშვნელოვანი იარაღი უცხო ენის სწავლებისათვის და საშუალებას იძლევა, მოსწავლეების გარკვეული ჯგუფის შეცდომათა ტიპოლოგია გაანალიზდეს. მაშასადამე, ეს არის ორმხრივი სარგებელი: როგორც მოსწავლეებისთვის, ისე მასწავლებლებისთვის. მოსწავლეებისა და მშობლიურ ენაზე მოლაპარაკეთა მიერ დაწერილი ტექსტების ერთმანეთთან შედარებით, მოსწავლეებს თავად შეუძლიათ ისწავლონ და გაიუმჯობესონ ენობრივი უნარები. მასწავლებლებს კი შეუძლიათ თავიანთი მეთოდები მოსწავლეების სპეციფიკურ ჯგუფებს მორგონ. ზოგადად, learning corpora-ს მთავარი სამიზნე ჯგუფები არიან მოსწავლეები, რომლებიც სწავლობენ უცხო ენას და მასწავლებლები,

რომლებიც ასწავლიან უცხო ენებს. ამათ გარდა, ლინგვისტები და ის ადამიანები, რომლებიც იკვლევენ სწავლების მეთოდებს, ასევე სარგებლობენ ამ ტიპის corpora-თი.

მიუხედავად იმისა, რომ learner corpora უცხო ენის სწავლებაში ახალ შესაძლებლობებს იძლევა, ის მაინც იშვიათად გამოიყენება სკოლებსა და ენის კლასებში. ამის ერთ-ერთი მიზეზი შეიძლება იყოს ინფორმაციის ნაკლებობა და ის, რომ მას უფრო უყურებენ, როგორც სამეცნიერო ინსტრუმენტს და არა სწავლების ინსტრუმენტს. ამიტომ მნიშვნელოვანია, რომ მიაწოდო მასწავლებლებს ინფორმაცია და ასწავლო სტუდენტებს, როგორ გამოიყენონ learner corpora. სკოლებმა და უნივერსიტეტებმა უნდა გამოიჩინონ ინიციატივა და დაიწყონ ამაზე მუშაობა. იმისათვის, რომ დაეხმაროს სტუდენტს, ისწავლოს ახალი ენა, მასწავლებელს შეუძლია ჩართოს learner corpora სასწავლო პროცესში თავის გაკვეთილებზე. მასწავლებლებს შეუძლიათ, მისცენ სტუდენტებს, მაგალითად, ისეთი სავარჯიშოები, რომელთა გაკეთებაც შესაძლებელია მხოლოდ learner corpora-ს გამოყენებით. ბევრ სიტყვას აქვს სხვადასხვა მნიშვნელობა და, შესაბამისად, გამოიყენება სხვადასხვა კონტექსტში.

learner corpora-ს დახმარებით, სტუდენტებს შეუძლიათ შეადარონ ერთმანეთს ამ სიტყვების გამოყენება მშობლიურ ენაზე დაწერილ ტექსტში და მოახდინონ სხვადასხვა ლექსიკური კატეგორიის იდენტიფიცირება. ან, თუ სტუდენტს გაუჩნდა რაიმე კითხვა, მას შეუძლია იპოვოს ამ კითხვაზე პასუხი learner corpora-ს დახმარებით. ეს პოტენციურად სტუდენტებს აქცევს ენის მკვლევრებად. გრამატიკული წესების მაგალითებით ახსნით, learner corpora შეიძლება იყოს დამატებითი რესურსი გრამატიკის სწავლებისას.

### 3. როგორ შეიქმნა learner corpus?

იმისათვის რომ შეიქმნას learner corpus, მნიშვნელოვანია შეგროვდეს დიდი რაოდენობის წერილობითი და/ან ვერბალური მასალა. წერილობითი კორპუსის შექმნა უფრო ადვილია, ვიდრე ვერბალურისა, რადგანაც წერილობით კორპუსს წყაროდ შეუძლია გამოიყენოს ინტერნეტი. ის შეიძლება შეიცავდეს გამოსვლის ჩანაწერს, ინტერვიუებს, ესეებს, გამოცდებს და ა.შ. ეს ყველაფერი უნდა იყოს მოსწავლეების მიერ დაწერილი. შედარებისთვის: იგივე მასალა - ვერბალური და წერილობითი - ხელმისაწვდომი უნდა იყოს მშობლიურ

ენაზეც. ძირითადი ენის კორპუსი შეიძლება აიწყოს ვერბალური და წერილობითი ტექსტებისგან და შეიძლება გამოიყენებულ იქნეს მარტივად, კორპუსის კომპიუტერული პროგრამით. ვერბალური კორპუსის შესაქმნელად საჭიროა უფრო მეტი დრო, რადგანაც, მაგალითად, ვიდეოგამოსვლა უნდა გადმოიცეს ტრანსკრიფციით და უნდა მოხდეს მისი ზოგიერთი არავერბალური მახასიათებლის კოდირება. ინტერნეტის გამოყენებით წერილობითი კორპუსის შექმნა ძალიან სწრაფად შეიძლება. ყველა კორპუსს სჭირდება დიზაინის შექმნის გარკვეული პრინციპების ცოდნა. უნდა გაითვალისწინო არა მარტო დიზაინი, არამედ ისიც, თუ რამდენად შესაძლებელია მისი შექმნა, რადგანაც არის ხოლმე გარკვეული წინააღმდეგობები იმაზე, რა არის შესაძლებელი, რა არის ეთიკური და რა - ლეგალური. ეს ფაქტორი შეიძლება გადაამწყვეტიც კი გახდეს. ამასთან, მნიშვნელოვანია იმის გადაწყვეტა, რა და როგორ უნდა წარმოადგინო / წარმოსახო საუკეთესოდ საერთო მიზნისთვის. ამ შემთხვევაში, უნდა გადაწყვიტო, რა რაოდენობის მონაცემები გინდა რომ შეაგროვო და გამოიყენო. ვერბალური კორპუსის შექმნაში შემდეგი საფეხური არის მონაცემების ჩაწერა. ჩანაწერების

გასაკეთებლად არსებობს უამრავი სხვადასხვა ვარიანტი, მაგალითად, კასეტების ანალოგი, ციფრული მედია, აუდიოვიზუალური ციფრული ჩამწერები და სხვა. ტრადიციული ანალოგი იაფია, თუმცა, ბევრი უარყოფითი მხარე აქვს. ის არის ძალიან დიდი და მოუხერხებელი, ციფრული ჩამწერებისგან განსხვავებით, მათი კომპიუტერიზება და გადაწერა ადვილი არ არის. ციფრული ხელსაწყოების გამოყენება საშუალებას იძლევა, გადაწერასთან ერთად, გასწორდეს ხმა (აუდიოვიზუალურ ჩამწერის შემთხვევაში - გამოსახულებაც).

მნიშვნელოვანი საკითხია ნებართვის ქონა. ჩაწერაზე ნებართვის აღება წინასწარ უნდა მოხდეს სპიკერებთან. კვლევისთვის, კომერციული პედაგოგიური მასალებისთვის ან სხვა მიზნებისთვის ჩანაწერების გამოყენებაზე ნებართვა უნდა გაფორმდეს. შესაძლოა საჭირო გახდეს, განისაზღვროს, როგორ იქნება გამოყენებული ჩანაწერები მათზე ნებართვის მიღებისას. ამის შემდეგ მთავარი ამოცანაა გადაწერა, რადგანაც ვერბალური ინფორმაცია საჭიროებს ხელით გადაწერას. სწორედ ეს წარმოადგენს სალაპარაკო ენის კორპუსისთვის ასეთ სირთულეს, ერთგვარ გამოწვევას. ისინი დაცულია, როგორც "უბრალო" ტექსტების

ფაილები, რათა მაქსიმალურად მოსახერხებელი იყოს მათი გამოყენება სხვადასხვა კომპიუტერული პროგრამით. ერთსაათიანი გამოსვლის ჩანაწერის გადაწერას შეიძლება დღეები დასჭირდეს, ეს დამოკიდებულია იმაზე, რამდენად რთულია ენა. უმეტეს შემთხვევაში, თითოეული სიტყვა, ხმოვანი სიგნალები, ხმაური, გადაფარვა და ა.შ. - ყველაფერი იწერება, განსხვავებით სპიკერების მიერ ნათქვამის გასწორებული / დამუშავებული ვერსიისა. გადაწერის სკურპულოზურობა დაკავშირებულია შენი კორპუსის მიზანთან. თუ არ გაქვს იმის მოთხოვნა, რომ იცოდე, სად უნდა იყოს კომენტარები და სად - ჩანართები, მაშინ არ აქვს აზრი ასეთ დეტალურ გადაწერაში დრო დაკარგო... ეს რთული სამუშაო მოიცავს: 1) ნიმუშების დაწყვილებას, 2) სიტყვათა თანაარსებობას, 3) ანოტაციასა და სიტყვით გამოსვლის ნაწილს, 4) სინონიმებსა და ანტონიმებს, 5) არსებული კვლევების გათვალისწინებით, უფრო კომპლექსური კვლევების ჩატარებას, 6) საკითხებს, რომლებიც დაკავშირებულია ენის სტრუქტურის ცვლილებასთან ენის განვითარების სხვადასხვა ეტაპზე, 7) და საკითხებს, რომლებსაც მომხმარებელი განსაზღვრავს.



გადაწერილი ფაილები ისე უნდა იყოს ორგანიზებული, რომ შესაძლებელი იყოს საინფორმაციო წყაროს პოვნაც. მაგალითად, კარგი იქნება, თუ შესაძლებელია ისეთი ინფორმაციის ნახვაც, როგორცაა სქესი, ასაკი, სპიკერების რაოდენობა, დაბადების ადგილი, საქმიანობა, განათლება, სად მოხდა ჩაწერა და სხვა. ეს ინფორმაცია შეიძლება იყოს მოცემული თითოეული ტრანსკრიპტის დასაწყისში, როგორც "სათაური" ან მოცემული იყოს ცალკე მონაცემთა ბაზაში, სადაც ინფორმაცია ჩატვირთულია ფაილის სახელით. მოკლედ რომ ვთქვათ, კორპუსი უნდა იყოს გამდიდრებული ანოტაციებით და საშუალებას უნდა გაძლევდეს ლინგვისტური ფენომენის მრავალი ტიპი მოძებნო. ყველა კორპუსის შინაარსი წარმოადგენს ენაში ტექსტებისა და გამოთქმების ნაკრებს. რა თქმა უნდა, ჩვენ უნდა განვასხვაოთ ერთმანეთისგან წერილობითი და ვერბალური კორპუსები. წერილობითი კორპუსისთვის მასალები არის გაცილებით უფრო ადვილი შესაგროვებლად, რადგანაც ყველაფერი ფიზიკურად ხელმისაწვდომია. ვერბალური კორპუსი შინაარსობრივად, როგორც ზემოთ უკვე აღვნიშნეთ, უფრო რთულია, მასალების შეგროვებისა და რედაქტირების თვალსაზრისით. ძირი-

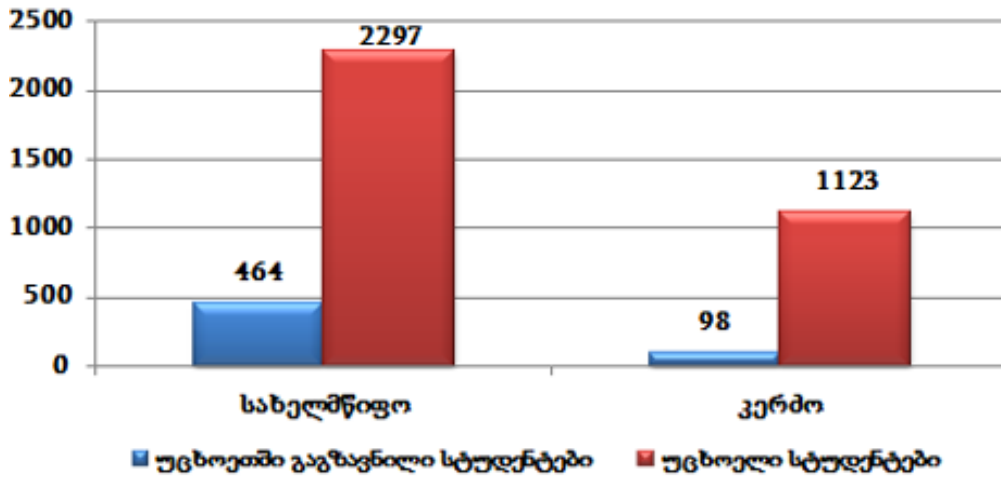
თადი მასალები ვერბალური კორპუსისთვის, როგორც წესი, მოცემულია აუდიო ან აუდიოვიზუალური ჩამწერებით.

#### 4. learner corpus-ის შესაძლო ქართული ვარიანტი

საქართველოში სოციოლინგვისტური სიტუაცია შეიძლება დახასიათდეს, როგორც მულტილინგვური. საქართველოს საზღვრისპირა რეგიონებში - აზერბაიჯანისა და სომხეთის საზღვრებთან, საკლასო გარემო არის მულტილინგვური. განათლებაში გამოწვევას ენის ისეთი სასწავლო მოდულების შექმნა / განვითარება წარმოადგენს, რომელშიც მრავალმხრივ იქნება გამოყენებული ქართული ენის შესწავლისას ხშირად დაშვებული შეცდომები. ამ მიმართულებით ერთ-ერთი პირველი ნაბიჯი ყველა შესაძლო შეცდომის თავმოყრა და დოკუმენტირებაა - როგორც წერილობით, ისე ვერბალური სახით.

საქართველოს სკოლების მოსწავლეთა ჯგუფების ჩართულობა შეიძლება იყოს მასალის პირველწყარო. ოფიციალური სტატისტიკით, ყოველწლიურად დაახლოებით 2000 არაქართულენოვანი იღებს უმაღლეს განათლებას საქართველოში.

სტუდენტთა გაგზავნა საზღვარგარეთ და უცხოელ სტუდენტთა რიცხოვნობა, 2013-2014 სასწ.წ., კაცი



სქემა 1. საქართველოს სტატისტიკის ეროვნული სააგენტო (წითელი სვეტი გვიჩვენებს უცხოელი სტუდენტები რაოდენობას სახელმწიფო და კერძო უნივერსიტეტებში).

მასწავლებლები და ენის ტრენერები შეიძლება იძულებულნი გახდნენ, თავიანთი მულტილინგვური სწავლების გამოცდილება გაუზიარონ ერთმანეთს და პერიოდულად დაშვებული შეცდომების სისტემატიზაცია მოახდინონ. ამგვარი დაკვირვებები ქმნის საფუძველს, შეიქმნას ისეთი მონაცემთა ბაზა, რომელიც მოიცავს შეცდომათა ნიმუშებს. როგორც FALKO-ს მაგალითზე ვნახეთ, ასეთი მონაცემთა ბაზა საჭიროა ჰიპოთეზების შესაქმნელად, როგორც წინასწარი კორექციისთვის.

ამჟამად საქართველოსა და სხვა ქვეყნებში ქართული ენის შესასწავლი პროგრამები ძალიან ცოტაა (ონლაინ დისტანციურ კურსს გვთავაზობს, მაგალითად, მალმოს უნივერსიტეტი

შვედეთში, ასევე - არნ. ჩიქობავას სახ. ენათმეცნიერების ინსტიტუტი). სასწავლო კურსების მიზნობრივი ჯგუფები არიან საქართველოს მოქალაქეები, რომელთა მშობლიური ენაა აზერბაიჯანული და სომხური. პროგრამის სილაბუსებში ასახულია პროგრამის ავტორების ხედვები იმ შესაძლო სირთულეებთან დაკავშირებით, რომლებიც შეიძლება მოსწავლეს შეხვდეს. გარკვეულ საკითხებს ემპირიული საფუძველი არ გააჩნიათ, მიუხედავად იმისა, რომ ნებისმიერ სპეციფიკურ საკითხზე აქცენტი უნდა იყოს მკაცრად ორიენტირებული სასწავლო პროცესში მოსწავლეების მიერ დაშვებულ შეცდომებზე.



ქართული ენის ელექტრონული სწავლების კურსი

Gürcü dilinin elektron kursu

ქართული ენის ელექტრონული სწავლების კურსი

eLearning Course of Georgian

გამკვლევო

MƏLUMAT KİTABÇASI

ინტერფეისი

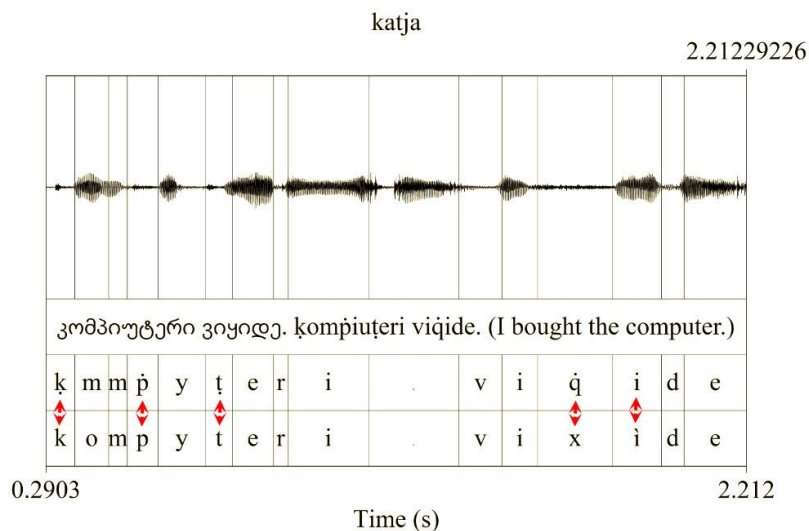
GUIDE

## სქემა 2. ქართული ენის ელექტრონული სწავლების კურსი.

ქართული ენის (სხვა კავკასიურ ენათა მსგავსად) შესწავლისას ყველაზე დიდი სირთულე თანხმობების წარმოთქმას უკავშირდება. არსებობს თანხმობები, რომლებიც მხოლოდ

კავკასიური ენებისთვისაა დამახასიათებელი. ასე, რომ მოსწავლისთვის თანხმობის გამოთქმა მნიშვნელოვანი გამოწვევაა, როგორც ინტელექტუალურად, ისე ფიზიკურად.

ცხრილი 3. მეტყველების მაგალითი (ქართული შემსწავლელი). 5 მცდარი გამონათქვამი 2 სიტყვაში.



მოსწავლეებისთვის, რომლებიც მეორე ენას სწავლობენ, აუდიომასალები მნიშვნელოვან მონაცემთა ბაზას წარმო-

ადგენს იმისთვის, რომ მიაგნონ ფონეტიკურ ნაწილში ყველაზე ხშირად დაშვებულ შეცდომებს და გაითვალის-

წინონ ეს შემდეგ სწავლების პროცესში. ამ ხშირად დაშვებულ შეცდომებზე ზედაპირული დაკვირვებაც კი საკმარისია, რომ გამოიკვეთოს მნიშვნელოვანი პრობლემები. ქვემოთ

მოცემულია ქართული ენის მასწავლებელთა მიერ მოყვანილი ენობრივი შეცდომების ტიპური მაგალითები.

**ცხრილი 4. სიტყვათა წყობის შეცდომები.**

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

**ცხრილი 5. სიტყვათშეთანხმების შეცდომები.**

wording	kargi	var.
lemma	kargi_good	qopna.be
hypothesis	kargad	var.
Eng.	I am fine.	

**ცხრილი 6. ლექსიკური შეცდომები.**

wording	didi	gemrieli-a
lemma	didi_big	gemrieli - qopna.be.Encl
hypothesis	3 alian	gemrielia
Eng.	It is very tasty.	

**ცხრილი 7. სინტაქსური შეცდომები**

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.	The house, which is (standing) in the city.			

wording	reštórani	sad	viqavi	gušin
lemma	reštórani _restaurant	sad_wh	qopna_be	gušin _yesterday
hypothesis_1	reštórani	sadac	viqavi	gušin
hypothesis_2	reštórani	sadac	gušin	viqavi
Eng.	The restaurant I was in yesterday.			

### ლიტერატურა:

- აიჰმერი, 2009 - Aijmer Karin, Corpora and Language Teaching, 2009 John Benjamins Publishing Company.
- ბიბერი, დუგლასი..., 2009 - 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.
- გრანდერი, სილვიანე, 2002 - 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.
- დიაზ-ნეგრილიო, ფერნანდეს-დომინგუეზი, 2006 - Diaz-Negrillo, Ana; Fernandez-Dominguez, Jesus (2006): Error tagging systems for learner corpora. In: RESLA 19.
- კორდერი, 1986 - 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.
- ო'კიფი, 2007 - O'Keeffe Anna et al: From corpus to classroom. Cambridge University Press 2007.
- რეზნიჩეკი, 2012 - Reznicek et al. Das Falko-Handbuch. Korpusaufbau und Annotationen, Version 2.01, 2012.
- სიმენი, 2014 - Siemen et al., FALKO - Ein fehlerannotiertes Lernerkorpus des Deutschen.
- სინკლერი, მაკჰარი..., 2004 - Sinclair John McHary, How to use Corpora in Language Teaching Studies in Corpus Linguistics, Silvia Bernardini, 2004 John Benjamins Publishing Company.
- შმიდტი, 2014 - Schmidt Karin, Lernerkorpora: Ressourcen für die Deutsch-als-Fremdsprache-Forschung (publikationen.ub.uni-frankfurt.de/files/.../Schmidt\_Lernerkorpora.pdf).
- <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://www.ice.ge/web/elearning_geo.html)
- <http://195.178.225.22/DiasporaGeo/Georgianonline.html>