

# ელექტრონული კორპუსები: ისტორია, ტიპოლოგია, მნიშვნელობა, გამოყენება

## 1.1. კორპუსის ისტორია

ელექტრონული კორპუსების შექმნის ისტორია მეოცე საუკუნის 80-იანი წლებიდან იწყება. ენათა ეროვნული კორპუსების შექმნამ მნიშვნელოვნად შეცვალა მიდგომები ენების კვლევებისადმი. ამ პერიოდიდან აქტიურად იწყება მსოფლიოს ერებისა და ენებისათვის ნაციონალური კორპუსების შექმნა, რაც ნიშნავს ამ ენებზე არსებული ზეპირი, წერითი თუ ბეჭდური ინფორმაციის ელექტრონულ ფორმატში გადაყვანას. ამ მასალების საფუძველზე შეიქმნა და კვლავ აქტიურად იქმნება ელექტრონული ბიბლიოთეკები, რომლებიც აღჭურვილია მაღალტექნოლოგიური საძიებო და საანალიზო სისტემებით. ასეთი მაღალტექნოლოგიური სისტემების არსებობა კი ძალიან აადვილებს და აჩქარებს ამა თუ იმ დოკუმენტირებული ენობრივი ფაქტისა თუ ინფორმაციის მოძიებას და შემდეგ მათ მიზნობრივად გამოყენებას.

კორპუსი ლათინური სიტყვისგან „corpus“ წარმოსდგება და „სხეულს“, „ტანს“ ნიშნავს. თანამედროვე ლინგვისტიკაში იგი ელექტრონული ტექსტების მასივის / კრებულის აღსანიშნად გამოიყენება, ხოლო ნაციონალური კორპუსი ეწოდება ელექტრონული ტექსტების ისეთ ერთობლიობას, რომელიც მხოლოდ და მხოლოდ ერთი რომელიმე კონკრეტული ენისათვის იქმნება.

კორპუსი მეტყველების / ტექსტის მეოთხე ფაქტურას წარმოადგენს. თუ აქამდე გვექონდა სამი ფაქტურა: 1. ზეპირი; 2. წერილობითი და 3. ბეჭდური, კორპუსული ლინგვისტიკის განვითარების კვალდაკვალ მას დაემატა მეოთხე ფაქტურა - კორპუსული / ელექტრონული.

კორპუსებზე მსჯელობის დროს აუცილებლად უნდა გავითვალისწინოთ ერთი რამ. მართალია, რაც უფრო დიდი და მოცულობითია კორპუსი, მით უფრო მეტი ინფორმაციის მიღება შეგვიძლია მისგან და უფრო ეფექტურად შეიძლება ამ

კორპუსის გამოყენება ენის კვლევის საქმეში, მაგრამ ამავდროულად მნიშვნელოვანია მხედველობაში მივიღოთ კორპუსების ერთ-ერთი მნიშვნელოვანი მახასიათებელი-რეპრეზენტაციულობა, ანუ იგივე პროპორციულობა. თუ გავითვალისწინებთ იმ ფაქტს, რომ კორპუსი არის ენის შემცირებული მოდელი, მივხვდებით, რომ კორპუსულ ლინგვისტიკაში გადამწყვეტი მნიშვნელობა ენიჭება არა ელექტრონულ ფორმატში გადაყვანილ ტექსტთა რაოდენობას, არამედ რეპრეზენტაციულობას, რაც გულისხმობს მოცემული ენის დაბალანსებულ, პროპორციული რაოდენობის, სხვადასხვა პერიოდის, ჟანრის, სტილის, ავტორის და ა.შ. ტექსტების მასივს.

ელექტრონული კორპუსის ძირითადი მიზანი არის მასში არსებული ელექტრონული ფორმატის ტექსტების საფუძველზე სხვადასხვა ტიპის ლინგვისტური ამოცანების გადაჭრა. ამისათვის კორპუსში მხოლოდ ტექსტების ატვირთვა საკმარისი არ არის, არამედ საჭიროა დამატებითი ლინგვისტური თუ ექსტრალინგვისტური ინფორმაცია. ყველა ელექტრონული კორპუსი აღჭურვილია საგანგებო საკვლევი ინსტრუმენტებით, რომელსაც კორპუსის მენეჯერს ვუწოდებთ. კორპუსის მენეჯერი დაკომპლექტებულია ისეთი საკვლევი ინსტრუმენტებით, როგორცაა მონიშვნა, ანოტირება, რომელშიც თავის მხრივ გამოიყოფა ანოტირების სხვადასხვა ტიპები, როგორცაა: ფონეტიკური, სინტაქსური, სემანტიკური და ლექსიკური დონეები. თეგირება და ა.შ. ეს ის საკვლევი ინსტრუმენტებია, რომელთა გარეშეც ელექტრონული კორპუსების შექმნას აზრი არ ექნებოდა და წინააღმდეგ შემთხვევაში ისინი ჩვეულებრივ, ციფრულ ბიბლიოთეკას დაემსგავსებოდნენ.

## 2.2. კორპუსის ტიპოლოგია

ელექტრონულ კორპუსებში რამდენიმე ძირითად ტიპს განასხვავებენ. მაშასადამე, კორპუსი შეიძლება იყოს:

- პატარა (რამოდენიმე ათასი ან მილიონი სიტყვა) და დიდი (50-100 მილიონი, მილიარდი სიტყვა). რაც უფრო დიდია კორპუსი, მით უფრო სრულად არის ასახული / წარმოდგენილი მოცემული ენის ბუნება;
- თემატური და რეპრეზენტაციული: რეპრეზენტაციულ კორპუსში სათანადო მოცულობით უნდა იყოს წარმოდგენილი შემდეგი: მხატვრული ნაწარმოებები, სამეცნიერო ხასიათის ტექსტები, ჟურნალისტური ნაშრომები, ინტერნეტ ფორუმები, რელიგიური ხასიათის და ა.შ ტექსტები, თემატური კორპუსი შეიძლება მოიცავდეს მხოლოდ ერთ-ერთ მათგანს ზემოთ ჩამოთვლილიდან.
- სინქრონული (თანამედროვე ენა) და დიაქრონული (ისტორიული კორპუსები); კორპუსების უმეტესობა არის სინქრონული. დიდი მოცულობის დიაქრონული კორპუსი მოიცავს უფრო ძველ პერიოდებს - ძველი რუსული ენა, ძველი ქართული ენა და ა.შ.
- სტანდარტული (სალიტერატურო) ენა და დიალექტური.
- პროზისა და პოეზიის;
- ერთენოვანი და მრავალენოვანი (პარალელური);
- წერილობითი და მულტიმოდალური (აუდიო, ვიდეო და ა.შ.).

### 2.3. კორპუსის გამოყენების სფეროები

კორპუსული ლინგვისტიკა ფართო ცნებაა და წარმოადგენს მეთოდოლოგიას, რომლის გამოყენებაც ენობრივი კვლევის მრავალი ასპექტით შეიძლება. კორპუსული ლინგვისტიკა შეიძლება განვიხილოთ როგორც მეთოდების, პროცედურებისა და რესურსების კრებული, რომელთაც ემპირიულ მონაცემებთან აქვთ საქმე.

ელექტრონული კორპუსები დღესდღეობით მრავალ სფეროში გამოიყენება და დარგის განვითარებასთან ერთად მათი რაოდენობაც შესაბამისად იზრდება. ქვემოთ მოცემულია არასრული სია იმ დარგებისა, რომლებიც აქტიურად იყენებენ

კორპუსული ლინგვისტიკის მიღწევებს კვლევით-სამეცნიერო საქმიანობაში. ერთ-ერთი მათგანია შედარებით-ისტორიული ენათმეცნიერება. იგი ერთ-ერთი მთავარი მიმართულებაა, რომელმაც გავლენა მოახდინა კორპუსული ლინგვისტიკის განვითარებაზე. ლინგვისტები, რომლებიც დაკავებული იყვნენ გამოეკვლიათ ენასთან დაკავშირებული ისტორიული თუ თანამედროვე ტექსტები, ყოველთვის მიმართავდნენ ამ უზარმაზარ და ინფორმაციულად მნიშვნელოვან კრებულებს, როგორც ძირითად დამადასტურებელ მასალას.

ელექტრონული კორპუსები ფართოდ გამოიყენება ლექსიკოგრაფიაში. თამამად შეიძლება ითქვას, რომ კორპუსის ლინგვისტიკა მნიშვნელოვან როლს ასრულებს ლექსიკონის შედგენაში. დროის უმცირეს მონაკვეთში ლინგვისტს შეუძლია სიტყვებისა თუ შესიტყვებების მაგალითების დაძებნა მილიონობით სიტყვა-ფორმის შემცველ ელექტრონულ კორპუსებსა თუ წერილობითი სახის ტექსტებში.

ლექსიკოგრაფიისა და კორპუსული ლინგვისტიკის ურთიერთმიმართებისა და განვითარების საქმეში ადამ კილგარიფი (Adam Kilgarriff) და დევიდ თაგუელი (David Tugwell) გამოყოფენ ოთხ პერიოდს:

- 1) პირველი პერიოდი, კომპიუტერის გამოგონებამდე, ანუ ეს არის ის პერიოდი, როდესაც ლექსიკოგრაფები აგროვებდნენ / აბარებდნენ ციტატებს, როგორც ეს დიდი ოქსფორდის ლექსიკონის მაგალითზე მოხდა.
- 2) მეორე პერიოდი არის ქობილდის ლექსიკონის შედგენის პერიოდი ადრეულ 1980-იან წლებში, როდესაც ლექსიკოგრაფები აგროვებდნენ ყველა მონაცემებს კორპუსიდან.
- 3) მესამე პერიოდი კი არის გვიანი 1990-იანი წლები, როდესაც კორპუსი გახდა იმდენად დიდი, რომ მის მომხმარებელს სჭირდებოდა კორპუსიდან მიღებული ვრცელი მონაცემების დაჯამება.
- 4) მეოთხე პერიოდი კი არის უკვე კომპიუტერული პროგრამებით აღჭურვილი კორპუსის პერიოდი.

კიდევ ერთი მნიშვნელოვანი დარგი, რომელიც თავის კვლევებში აქტიურად იყენებს ელექტრონულ კორპუსებს, არის სოციოლინგვისტიკა. ელექტრონული კორპუსები შედგება რეალური ტექსტებისაგან, წერილობითი და სასაუბრო ენის რეალური მაგალითებისაგან. ეს ყველაფერი სოციოლინგვისტებისათვის ღირებულ ინფორმაციას იძლევა ენის სხვადასხვა ვარიაციის შესასწავლად.

თარგმანთმცოდნეობა კიდევ ერთი დარგია, რომელიც ელექტრონულ კორპუსებს ძალიან აქტიურად იყენებს. ისეთი კორპუსები, რომლებიც შეიცავს ტექსტებსა და მონაცემებს რამდენიმე ენაზე, განსაკუთრებით ღირებულია მთარგმნელებისთვის. ასეთი კორპუსებია შედარებითი და პარალელური კორპუსები. ასეთი კორპუსების საშუალებით ადვილი ხდება ამა თუ იმ სიტყვისა თუ შეისიტყვების ზუსტი თარგმნითი ექვივალენტის, სინონიმის და ა.შ. განსაზღვრა ან ამ სიტყვის პრაქტიკული გამოყენების სფეროს დადგენა.

ელექტრონული კორპუსები ასევე გამოიყენება ენის შესწავლა / სწავლების საქმეში, დიალექტოლოგიაში და სხვა. ტექსტები, რომლებიც ამა თუ იმ კონკრეტულ კორპუსშია მოთავსებული, მოიცავს დიალექტებს და ლინგვისტს აძლევს ძვირფას ინფორმაციას გეოგრაფიული ვარიაციების შესახებ.

## 2.4. მნიშვნელოვანი კორპუსები

დღესდღეობით არსებულ ელექტრონულ კორპუსთაგან დიდი მნიშვნელობით გამოირჩევა შემდეგი კორპუსები:

ბრაუნის კორპუსი - თანამედროვე ამერიკული ინგლისური ენის ბრაუნის სტანდარტული კორპუსი (the Brown Standard Corpus of Present-Day American English). ბრაუნის კორპუსი შეიქმნა 1963-64 წლებში, ბრაუნის უნივერსიტეტში ჰენრი კუჩერასა

და ნელსონ ფრენსისის მიერ და მუშაობა დასრულდა 1967 წელს. ბრაუნის კორპუსი შედგება დაახლოებით 500 ინგლისური ენის ტექსტისაგან, რომელიც 15 ჟანრად არის დაყოფილი და ჯამში შეადგენს 1 მილიონ სიტყვას. ბრაუნის კორპუსმა დიდი როლი შეასრულა სხვა კორპუსების შექმნაში, რომელთა შორის აღსანიშნავია LOB-ის (Lancaster, Oslo, Bergen) კორპუსი. ეს ელექტრონული კორპუსი ლანკასტერის უნივერსიტეტში იქნა შედგენილი 1978 წელს. ბრაუნის კორპუსი თანამედროვე კორპუსებთან შედარებით არის პატარა, მაგრამ ყოველთვის არის ცნობილ და მნიშვნელოვან კორპუსებთან რიგებში, რადგან ეს არის პირველი თანამედროვე და ელექტრონული კორპუსი (იხ. სურათი 1).

|   | Genre group         | Category     | Content of category                 | No. of texts |
|---|---------------------|--------------|-------------------------------------|--------------|
| I. Informative prose (374)<br>(see Figure 2)  | Press (88)          | A            | Reportage                           | 44           |
|   |                     | B            | Editorial                           | 27           |
|   |                     | C            | Review                              | 17           |
|   | General Prose (206) | D            | Religion                            | 17           |
|   |                     | E            | Skills, trades and hobbies          | 36           |
|   |                     | F            | Popular lore                        | 48           |
|   |                     | G            | Belles lettres, biographies, essays | 75           |
|   |                     | H            | Miscellaneous                       | 30           |
|   |                     | J            | Science                             | 80           |
|   |                     | Learned (80) |                                     |              |
| II. Imaginative prose (126)<br>(see Figure 3) | Fiction (126)       | K            | General fiction                     | 29           |
|   |                     | L            | Mystery and detective Fiction       | 24           |
|   |                     | M            | Science fiction                     | 6            |
|   |                     | N            | Adventure and Western               | 29           |
|   |                     | P            | Romance and love story              | 29           |
|   |                     | R            | Humor                               | 9            |
| <b>Total</b>                                  |                     |              |                                     | <b>500</b>   |

სურათი 1.

როგორც ზემოთ აღვნიშნეთ, ბრაუნის კორპუსი ყოველთვის დარჩება ერთ-ერთ მნიშვნელოვან კორპუსად, რადგან ის პირველი კორპუსი იყო, რომელიც შეიქმნა. უკვე შემდეგ ეტაპზე შექმნილ და მოცულობით უფრო დიდ კორპუსებს შორისაა:

ბრიტანული ნაციონალური კორპუსი, (British National Corpus - BNC) რომლის კონსორციუმი შეიქმნა 1990 წელს და მას ოქსფორდის უნივერსიტეტის გამომცემლობა ხელმძღვანელობდა. აღნიშნულ პროექტზე მუშაობა 1991 წელს დაიწყო და სამ წელიწადში 100 მილიონი თანამედროვე ბრიტანული ინგლისური სიტყვა შევიდა ბრიტანულ ნაციონალურ კორპუსში, აქედან 90% არის წერილობითი და 10% სასაუბრო. პირველი გამოცემა გამოვიდა 1994 წელს, ხოლო ევროპელი მკვლევრებისათვის ბრიტანეთის ეროვნული კორპუსი ხელმისაწვდომი გახდა 1995 წლის თებერვალში, კორპუსის რედაქტირებული ვერსია “BNC World“ გამოვიდა 2001 წელს, ხოლო 2007 წელს კი გამოვიდა მესამე რედაქცია. ბრიტანეთის ნაციონალური კორპუსი შედგება ორი ქვეჯგუფისაგან: *BNC Sampler* (1 მილიონი წერილობითი და 1 მილიონი სასაუბრო ინგლისური) და *BNC Baby* (4 მილიონი სიტყვა ოთხი განსხვავებული ჟანრიდან). ბრიტანეთის ნაციონალური კორპუსის მესამე გამოცემა მოიცავს 4049 ტექსტს, რაც შეადგენს 5,228,040 კილობაიტს, ან დაახლოებით 5.2 გეგაბაიტს. საერთო ჯამში კი მოიცავს 100 მილიონ სიტყვას, უფრო ზუსტად 96,986,707 სიტყვას.

შემდეგი მნიშვნელოვანი კორპუსია ოქსფორდის ინგლისური ენის კორპუსი (Oxford English Corpus). ეს არის დიდი კორპუსი ინგლისური ენისა, რომელიც შეადგენს 2 მილიარდ სიტყვას. კორპუსში მოცემული ტექსტები ყველა ტიპის წყაროდან არის აღებული, ლიტერატურული ნაშრომებიდან დაწყებული ინტერნეტ-ფორუმებით დამთავრებული. ოქსფორდის ინგლისური ენის კორპუსი გამოიყენება ოქსფორდის უნივერსიტეტის გამომცემლობის მიერ ლინგვისტური კვლევის დეპარტამენტისა და ოქსფორდის ინგლისური ლექსიკონების შემდგენლების მიერ.

ამერიკული ინგლისურის მხრივ მნიშვნელოვანი ელექტრონული კორპუსია ამერიკული ნაციონალური კორპუსი (American National Corpus). უნდა აღინიშნოს ის, რომ ამერიკული ნაციონალური კორპუსი არის მიმდინარე პროექტი, მასზე მუშაობა ჯერ არ დამთავრებულა, როდესაც საბოლოოდ დასრულდება აღნიშნულ პროექტზე მუშაობა, მასში შევა 100 მილიონი სიტყვა („ახალი” ენის: ვებ-ბლოგების, ვებ-გვერდების, „ჩატების”, ელ-ფოსტებისა და რეპის ტექსტების ჩათვლით). ამჟამად ამერიკულ ნაციონალურ კორპუსში შესულია 22 მილიონი სიტყვა. ამერიკული ნაციონალური კორპუსი იქნება ამერიკული ინგლისურის ყველაზე სრულყოფილი სურათი, რაც კი აქამდე შექმნილა და ამავე დროს დიდი რესურსი ლინგვისტური და ლექსიკოგრაფიული კვლევებისათვის.

ინგლისური ენის საერთაშორისო კორპუსი (International Corpus of English) მოიცავს მთელ რიგ კორპუსებს, რომლებიც საკუთრივ მოიცავს ინგლისური ენის ვარიაციებს იმ ქვეყნებიდან, სადაც ინგლისური ენა არის პირველი ან მეორე სახელმწიფო ენა. თითოეული ჯგუფი ინგლისური ენის საერთაშორისო კორპუსისა მოიცავს 1 მილიონ სიტყვას 1989 წლიდან. აღნიშნულ პროექტზე მუშაობა კი მიმდინარეობს 1990 წლიდან.

რუსული ნაციონალური კორპუსი (Russian National Corpus - RNC) ძირითადად მოიცავს ტექსტებს XVIII საუკუნეების შუა პერიოდიდან XXI საუკუნის ადრეულ პერიოდამდე. ეს პერიოდი წარმოადგენს რუსული ენის წარსულსა და აწმყოს ფართო სოციოლინგვისტური თვალსაზრისით: ლიტერატურული, სასაუბრო, მოძველებული, ნაწილობრივ დიალექტური. კორპუსი მოიცავს ორიგანალ (გადაუთარგმნელ) ნაშრომებს: პროზას, დრამასა და პოეზიას, რომლებიც არის კულტურული ღირებულების მქონე და საინტერესოა ლინგვისტური თვალსაზრისით. მხატვრული ნაწარმოების გარდა, კორპუსი მოიცავს დიდი მოცულობით როგორც წერილობითი ენას, ასევე სასაუბროს: მემუარები, ესეები, ჟურნალისტური ნაშროები, სამეცნიერო ლიტერატურა, საჯარო გამოსვლები, წერილები, დღიურები, დოკუმენტები და ა.შ.



ძირითადი კორპუსი, რომელიც მოიცავს სტანდარტული რუსულის ტექსტებს შეიძლება დაიყოს 3 ნაწილად, რომელთაგან თითოეულ მათგანს აქვს განმსახვავებელი თავისებურებები: თანამედროვე წერილობითი ტექსტები (1950-დან დღემდე), სასაუბრო რუსულის ქვეკორპუსი (ამავე პერიოდის ზეპირი მეტყველების ჩანაწერი) და ადრეული ტექსტები (18 საუკუნის შუა პერიოდიდან 20 საუკუნის შუა პერიოდამდე). რუსული ნაციონალურ კორპუსში ძებნა შესაძლებელია როგორც სამივე ქვეკორპუსში, ასევე ცალ-ცალკე.

## 2.5. ქართული ენის კორპუსები

საქართველოში კორპუსული ლინგვისტიკის განვითარება, სხვა ქვეყნებთან შედარებით, გვიან დაიწყო. მიუხედავად ამისა, ქართული ენისათვის დღესდღეობით არსებობს შედეგი მნიშვნელოვანი კორპუსები. ესენია:

ტიტუსი - არმაზი - კავკასიური ენები და კულტურები: ელექტრონული დოკუმენტაცია. არმაზის სერვერი არის ფოლკსვაგენის ფონდის მიერ პროგრამის „ერთობლივი გზები ევროპისაკენ“ ფარგლებში დაფინანსებული ერთობლივი პროექტი, რომელიც მიზნად ისახავდა საქართველოს ჰუმანიტარული დარგის წამყვანისამეცნიერო-კვლევითი ინსტიტუტებისა და კვლევით-საგანმანათლებლო დაწესებულებების საერთაშორისო საინფორმაციო ქსელში (WWW) გაერთიანებას მონაცემთა ბანკის („არმაზი“) სახით. ეს პროექტი აერთიანებს ქვეპროექტებს და წარმოადგენს ზემოთ აღნიშნული ინსტიტუტებისა და დაწესებულებების ხელთ არსებული ინფორმაციის ციფრულ რესურსს მათი თანამედროვე ტექნოლოგიური საშუალებებით კვლევის შესაძლებლობებით. შეიძლება ითქვას, რომ ეს არის პირველი ქართულენოვანი ტექსტების აკადემიური მონაცემთა ბაზა, რომელიც აერთიანებს სხვადასხვა პერიოდის ქართული ლიტერატურის ტექსტურ მასალას და მისი ხელმძღვანელია პროფ. იოსტ გიპერტი (იხ. სურათი 2).



სურათი 2.

მეორე ასევე არანაკლებ მნიშვნელოვანი ქართული კორპუსია GEKKO - თანამედროვე ქართული ენის კორპუსი. იგი მოიცავს ორ ქვეკორპუსს: თანამედროვე ქართული ენის კორპუსს (124 055 170 ერთეული) და ქართულ ლიტერატურულ კორპუსს, მორფოლოგიური ანოტირებით (20 903 850 ერთეული). პროექტის ხელმძღვანელია პაულ მოირერი (Paul Meurer, Uni Computing) (იხ. სურათი 3 და 4).

Corpus: ქართული | Basic search | Advanced search
Query: ხისტი
Run Query | Load saved:
Done. Real time: 0.0298 sec. (0.003

Hit 1 - 30 of 495 | next 30 hits | Go to: | Download ( Excel mode) | Type: kwic | Show line filter | Show: |

Table with 4 columns: count, cpoc, match. Contains concordance results for the word 'ხისტი'.

სურათი 3

Concordance snippet showing text with HTML tags like <div>, <div>, </div> and words like 'ხისტი', 'პოლიტიკის', 'ეკონომიკის'.

სურათი 4

საქართველოს ლინგვისტური პორტრეტი - ქართული დიალექტური კორპუსი - ქართულ დიალექტურ კორპუსში წარმოდგენილია სამი ქართველური ენა (ქართული, სვანური, ზანური) და მათი დიალექტები. კორპუსი თავისი სტრუქტურით იძლევა ქართული ენობრივი სინამდვილის სრულყოფილად წარმოჩენის შესაძლებლობას,

მეტიც, კორპუსის მოდელი შეიძლება გამოყენებული იქნეს ნებისმიერი მულტილინგვისტური ენობრივი არეალის აღსაწერად. პროექტის ხელმძღვანელია მარინე ზერიძე (იხ. სურათი 5).



|   |   |
|---|---|
| <p><u>მთავარი გვერდი</u></p> <p><u>ლექსიკონები</u></p> <p><u>კორპუსის შესახებ</u></p> <p><u>პროექტის ამოცანები</u></p> <p><u>კორპუსის შედგენილობა</u></p> <p><u>კორპუსის გამოყენება</u></p> <p><u>დაფინანსების წყაროები</u></p> <p><u>მუშაკები</u></p> <p><u>გალერეა</u></p> <p><u>გუნდი</u></p> <p><u>კონტაქტი</u></p> <p><u>მიმღენა</u></p> | <p>ქართული დიალექტები კორპუსი წარმოადგენს <b>ქართული ლინგვოკულტურული სივრცის</b> ელექტრონული პორტრეტის შექმნის ცდას.</p> <p><b>საქართველოს ლინგვისტური პორტრეტი</b> წარმოადგენს <b>სამი ქართველური ენითა (ქართული, სვანური, ზანური)</b> და მათი მრავალრიგობიანი <b>დიალექტებით</b>.</p> <p>კორპუსი თავისი სტრუქტურით იზღვევს ქართული ენობრივი სინამდვილის სრულყოფილად წარმოჩენის შესაძლებლობას, ნეტც, კორპუსის მოდელი შეიძლება გამოყენებული იქნეს ნებისმიერი მულტილინგვისტური ენობრივი არეალის აღსაწერად.</p> <p>კორპუსი შეიცავს არა მხოლოდ ტერიტორიული დიალექტების, არამედ ენის სხვადასხვა სიმუხრონიული თუ დიაქრონიული ვარიაციის აღწერის პოტენციურ შესაძლებლობასაც.</p> <p>ქართული ენობრივი პორტრეტი კორპუსში აღიწერება რეგიონალური და ლინგვისტური მახასიათებლებით: რეგიონი _ რაიონი _ სოფელი; ენა _ დიალექტი/ქვესისტემა _ თემა (ქვედიალექტი).</p> <ul style="list-style-type: none"> <li>• <b>ქართული ტერიტორიული დიალექტები:</b> <ul style="list-style-type: none"> <li>◦ ხევსურული</li> <li>◦ ფშავური</li> <li>◦ თუშური</li> <li>◦ შიშვეთური</li> <li>◦ მთიულურ-კუდაშავური</li> <li>◦ ქართლური</li> <li>◦ კახური</li> <li>◦ იმერლოური</li> <li>◦ ფერეიდნული</li> <li>◦ მესხური (სამეხურ-ჯეჯახური)</li> <li>◦ რაჭული</li> <li>◦ იმერული</li> <li>◦ ლაზურული</li> <li>◦ გურული</li> <li>◦ აჭარული</li> <li>◦ ჩვენებურების ქართული</li> </ul> </li> <li>• <b>სვანური ტერიტორიული დიალექტები</b></li> <li>• <b>ზანური (მეგრულ-კანური) ტერიტორიული დიალექტები</b></li> </ul> <p>ქართული ეთნოოკულტური სინამდვილის წარმოსაჩვენად კორპუსში გამოყენებულია მახასიათებლები: ეთნიკური წარმომავლობა და კეთილური წარმომავლობა; მაგალითად: ეთნიკური წარმომავლობა _ ქართველი / კეთილური წარმომავლობა _ სვანი. ეთნიკური წარმომავლობა _ ქართველი / კეთილური წარმომავლობა _ იმერელი...</p> <p>მოქმედებს და ტექსტის აღწერის მახასიათებლები ჭეშნავ ქართული ლინგვოკულტურული სივრცის ისტორიულად ჩამოყალიბებულსა და ურყევ სურათს. მის ცენტრში დგას "ენობრივი პიროვნება", რომელიც არის:</p> <ul style="list-style-type: none"> <li>• <b>ინფორმაციის რეგიონული ცენტრის სახელით</b></li> </ul> |
|---|---|

სურათი 5.