

მაკა თეთრაძე

თბილისის სახელმწიფო უნივერსიტეტი
ოსუ არნოლდ ჩიქობავას სახელობის ენათმეცნიერების ინსტიტუტი

კორპუსზე დაფუძნებული ქართული ლექსიკონები

შესავალი

ლექსიკოგრაფიის მიზანი ენის ლექსიკური სიმდიდრის ობიექტური და მიუკერძოებელი აღწერაა. ლექსიკონი კი, იმის მიხედვით, თუ რა ტიპისაა, რა დანიშნულება აქვს, ენობრივ სინამდვილეს სხვადასხვაგვარად გამოხატავს. თანამედროვე ლექსიკოგრაფიას, ტრადიციულთან შედარებით, დღეს ერთი მთავარი უპირატესობა აქვს: ენობრივი ობიექტურობა გაცილებით ადვილი მისაღწევია ციფრული ტექნოლოგიების და, რაც მთავარია, ლინგვისტური კორპუსების არსებობის გამო. კორპუსის გამოყენება აიძულებს ლექსიკოგრაფს, პიროვნული დამოკიდებულება და ინტუიცია არ გახადოს გადამწყვეტი სალექსიკონო ერთეულისა და მისი მნიშვნელობების შერჩევისას (Walter 2010, 429).

პ. ჰენკსი აღნიშნავს: „ელექტრონული ლექსიკოგრაფია გზას უხსნის ყველა იმ შესაძლებლობას, რომელიც მიუწვდომელი იყო ტრადიციული ლექსიკოგრაფიისათვის: მნიშვნელობათა დადასტურების ახალი მეთოდები, აღწერის ახალი საშუალებები, მონაცემთა ორგანიზების ახალი სისტემა, ჰიპერტექსტური ბმულები...“ (Hanks 2012, 64). ამ შესაძლებლობათა შორის პ. ჰენკსს ყველაზე მნიშვნელოვნად სწორედ ჰიპერტექსტურ მონაცემთა ბაზების აგების შესაძლებლობა მიაჩნია, რომელიც აჩვენებს კავშირებს სიტყვის მნიშვნელობასა და სიტყვის გამოყენების თავისებურებას შორის (Hanks 2012, 65). ამგვარად, კორპუსული მონაცემები დღეს ლექსიკოგრაფებისათვის სამუშაოდ მთავარი რესურსია (McEnery and Hardie 2013, 729; McEnery and Hardie 2011, 24).

კორპუსზე დაფუძნებული ლექსიკონების შექმნის საფუძვლად ჯ. რ. ფირთის თეორიულ წანამძღვრებს მიიჩნევენ, რომელმაც ჯერ კიდევ 1950-იან წლებში წამოაყენა ჰიპოთეზა, რომ სიტყვის მნიშვნელობათა აღმოსაჩენად და ამ მნიშვნელობათა რიგითობის განსაზღვრისათვის აუცილებელია მნიშვნელობათა შესწავლა კონკრეტულ კოლოკაციებში, რისთვისაც დიდი ზომის ტექსტური მონაცემები გვჭირდება (Firth 1957; McEnery and Hardie 2013, 741). სწორედ ფირთის ჰიპოთეზა განავითარა ჯ. სინკლერმა, რომელიც სიტყვის მნიშვნელობის მთავარ წყაროდ კონტექსტს მიიჩნევდა – კოლოკაციების ემპირიულ კვლევას (Sinclair 1991, 39). მისთვის კორპუსლინგვისტიკა წმინდა მეთოდოლოგიური საწყისებიდან თეორიული სტატუსის მქონე მიმართულებად ჩამოყალიბდა და, შესაბამისად, კორპუსული ანალიზი ლექსიკოლოგიური კვლევების მნიშვნელობისა გახდა. ჯ. სინკლერის სახელს სამართლიანად უკავშირდება კორპუსული ლექსიკოგრაფიის განვითარება (Walter 2010, 428; McEnery and Hardie 2013, 742).

ჯ. სინკლერის კობილდის ლექსიკონიდან მოყოლებული (1987), კორპუსზე დაფუძნებული ლექსიკონები იქმნება სხვადასხვა ენისათვის. ლექსიკონთა შედგენილობა, აგების პრინციპები ესადაგება წყარო კორპუსის შინაარსს (Sharoff, Umanskaya, Wilson 2014, 1-12; Sharoff 2006, 63).

კორპუსთა ოთხ ძირითად ტიპს გამოყოფენ იმის მიხედვით, თუ რა ჟანრის ტექსტურ მასალას შეიცავს იგი: 1. წერილობითი ან ზეპირი მეტყველების კორპუსი; 2. დიალექტური (resp. რეგიონული ენობრივი ვარიანტების) კორპუსი; 3. სინქრონული ან დიაქრონული ტექსტური კორპუსი; 4. მეოთხე ტიპის კორპუსები ერთმანეთს უპირისპირდება იმის მიხედვით, ტექსტები ასახავს მშობლიური ენის თუ მეორე ენის მონაცემებს. ენის შემსწავლელთა მეტყველების კორპუსი მნიშვნელოვანი წყაროა სასწავლო ლექსიკონებისათვის (Walter 2010, 430).

ამ ოთხი მიმართულებიდან ქართული ენისათვის შექმნილია ახალი ქართული სალიტერატურო ენის, ძველი და საშუალო ქართული ენისა და ქართული ენის დიალექტური კორპუსები (<http://gnc.gov.ge/gnc/page>; <http://corpora.iliauni.edu.ge/>; <http://corpora.co>). რა თქმა უნდა, თითოეული ჯგუფი, რომელიც ემსახურება ქართული ენის კორპუსების შექმნასა და განვითარებას, დასაფასებელია, მაგრამ იმის გათვალისწინებით, თუ როგორი სისწრაფით ვითარდება კორპუსლინგვისტიკა მსოფლიოში, რა ხელსაწყოები, მეთოდები, საშუალებები ჩნდება ახალ-ახალი კორპუსების აგებისათვის, ქართული კორპუსლინგვისტიკა მსოფლიო აკადემიურ სივრცეში მოკრძალებულ ადგილს იკავებს. ამიტომაც თითოეული ნაბიჯი, რომელიც შედეგად ახალ პრო-

დუქტს გვაძლევს, ღირებულია არა მხოლოდ კორპუსლინგვისტიკის თვალსაზრისით, არამედ, საერთოდ, გამოყენებითი ენათმეცნიერების მხრივაც.

ჩვენი სტატია ეხება 2018 წელს გამოცემულ კორპუსზე დაფუძნებულ ქართულ ლექსიკონებს.

ფრანკფურტის წიგნის საერთაშორისო ბაზრობისათვის პროექტ „საბას“ ეგიდით რამდენიმე სხვადასხვა უანრის ლექსიკონი შეიქმნა, მათ შორის, კორპუსზე დაფუძნებული დიალექტური და კორპუსზე დაფუძნებული სინშირული მრავალენოვანი სასწავლო ლექსიკონები. წარმოგიდგინებ ლექსიკონების შედგენილობას, აგების პრინციპებს, მათ მნიშვნელობასა და მათი განვითარების სამომავლო გზებს.

დიალექტური კორპუსული ლექსიკონები

პირველ რიგში, მოკლედ მიმოვიხილავთ დიალექტურ ლექსიკონებსა და მის წყაროს – დიალექტური კორპუსს (corpora.co). კორპუსი 2006 წლიდან იქმნება და ივსება თსუ არნოლდ ჩიქობავას სახელობის ენათმეცნიერების ინსტიტუტის ენობრივ მონაცემთა კომპიუტერული დამუშავების განყოფილებაში (ხელმძღვანელი მ. ბერიძე). კორპუსი ამჟამად 477129 სიტყვაფორმას შეიცავს, მათ შორის ანოტირებულ მასალას (408042). კორპუსში შესულია ქართული ენის ყველა დიალექტის მონაცემები, ბოლო ასი წლის განმავლობაში შეკრებილი, გაციფრებული და ანოტირებული დიალექტური ჩანაწერები, რომელთაც მუდმივად ემატება ახალი, სამეცნიერო ექსპედიციებისას მოპოვებული მასალა. იმისათვის, რომ კორპუსმა რეპრეზენტაციულად და დაბალანსებულად ასახოს ენის გეოგრაფიული ვარიანტის თავისებურება, გასათვალისწინებელია კორპუსის ზომა და კორპუსში შესული ტექსტების უანრული მრავალფეროვნება. ქართული დიალექტური კორპუსი მასალას ყოფს დროისა და თემატური მახასიათებლების მიხედვით, ასევე, გარდა გაბმული მეტყველების ნიმუშებისა, კორპუსში შეხვდებით დიალოგებსაც, რაც ზეპირი მეტყველების თავისებურებათა კვლევისათვის უალრესად ღირებულ მონაცემებს შეიძლება გვაწვდიდეს. დიალექტურ კორპუსს ავსებს დიალექტური ლექსიკონებიც (ბერიძე, ბაკურაძე და ნადარაია 2017, 109-121). ჩვენთვის საინტერესო დიალექტებისთვის ანოტირებულ სიტყვაფორმათა რაოდენობაა: თუშური – 38496; იმერული – 38894; ფერეიდნული – 50601.

სწორედ ამ მასალაზე დაყრდნობით შეიქმნა ახალი ქართული დიალექტური ლექსიკონები სპეციალური ხელსაწყო – სალექსიკონო რედაქტორის დახმარებით (სალექსიკონო რედაქტორის აღწერა

ვრცლად იხილეთ მ. ბერიძის სტატიაში, რომელიც „კადმოსის“ ამავგ ნომერშია):

- ფერეიდნული ლექსიკონი (ბერიძე, ბაკურაძე და ნახუც-რიშვილი 2018);
- თუშური ლექსიკონი (ცოცანიძე და ბერიძე 2018);
- იმერული ლექსიკონი (ბერიძე, შარაშენიძე და კიკნაძე 2018).

სასწავლო სიხშირული ლექსიკონები

მიუხედავად იმისა, რომ სიხშირული სიების შედგენა კორპუსების ჩამოყალიბებამდეც გავრცელებული სამეცნიერო საქმიანობა იყო, კორპუსების შექმნამ ლექსიკოგრაფებს ახალი შესაძლებლობები შესთავაზა. სიხშირული სიების ერთ-ერთი ძირითადი დანიშნულება აღმოჩნდა მათი გამოყენება უცხო ენების სწავლებისას. პირველი კორპუსზე დაფუძნებული სიხშირული ლექსიკონიც ჯ. სინკლერის სახელთანაა დაკავშირებული (Hanks 2012, 24). დღესდღეობით მრავალი ენისათვის იქმნება სიხშირული სასწავლო ლექსიკონები (კორპუსზე დაფუძნებული სიხშირული ლექსიკონების ვრცელი მიმოხილვისთვის იხ. დარასელია და შაროვი 2015, 165-169).

სიხშირული ლექსიკონების შედგენილობა, შინაარსი, რა თქმა უნდა, დამოკიდებულია იმ კორპუსის ტიპზე, რომელიც გახდა სიის წყარო.

სიხშირულ სიების შექმნასთან დაკავშირებული ერთ-ერთი ყველაზე ცნობილი პროექტი არის კელის პროექტი (<http://corpus.leeds.ac.uk/serge/kelly/>). ამ პროექტის ფარგლებში კორპუსზე დაფუძნებული რამდენიმე ენის (რუსული, ინგლისური, გერმანული, არაბული, ბერძნული, შვედური, ნორვეგიული, იტალიური, პოლონური) სიხშირული სიები სასწავლო მიზნებთან შესაბამისად, ენის ფლობის დონეების ევროპული ჩარჩოს მიხედვით, სხვადასხვა დონისათვის განაწილდა. ცნობილია ასევე სასწავლო სიხშირული ლექსიკონების რუტლედჯის სერია, რომლებიც კორპუსებს ეფუძნება (<https://www.routledge.com/Routledge-Frequency-Dictionaries/book-series/RFD>).

ენის შემსწავლელისათვის ენის სინქრონიული მდგომარეობის ამსახველი ლექსიკა, სიტყვათა თანამედროვე მნიშვნელობები ენის პრაქტიკულად ათვისებისათვის არსებითია. ამიტომაც კორპუსთა ახალი მიმდინარეობა ვებკორპუსი გახდა ერთ-ერთი პოპულარული დასაყრდენი თანამედროვე ლექსიკოგრაფიისათვის. სიტყვის ყველაზე ხშირი გამოყენება და მნიშვნელობა სოციუმში შეიძლება სულაც

არ ემთხვეოდეს ლექსიკონში შესულ მნიშვნელობათა რიგითობას. ჯონ სინკლერი სწორედ ამას აღნიშნავდა: პირველი მნიშვნელობა, რაც თავში მოგდის, ხშირად არ არის ის, რომელსაც ლექსიკონში იპოვი (Sinclair 1991, 36).

ლიდზის უნივერსიტეტში შექმნილი ქართული ენის ვებკორპუსის (KaWaC) (ს. დარასელია, ს. შაროვი) სიხშირულ სიას ეფუძნება პროექტის ფარგლებში შექმნილი სამენოვანი სასწავლო ლექსიკონები: ქართულ-აფხაზურ-ინგლისური და ქართულ-ოსურ-ინგლისური სიხშირული ლექსიკონები. ვებკორპუსი მოიცავს 180 მილიონზე მეტ სიტყვას, რომელიც ასახავს თანამედროვე ოფიციალურ ქართულ ენას. კორპუსი მორფოლოგიურად ანოტირებული და ლემატიზებულია. რაც ასევე მნიშვნელოვანია, კორპუსი არის დაბალანსებული და რეპრეზენტაციული. რეპრეზენტაციულობის მისაღწევად გამოყენებულია ფუნქციური ჟანრული განზომილებების მეთოდი.

ლექსიკონების პროექტის კონცეფცია სახელმწიფო ენის დეპარტამენტში შეიქმნა (ავტორი მ. ბერიძე), უშუალოდ ლექსიკონებზე კი დეპარტამენტის, თსუ არნოლდ ჩიქობავას სახელობის ენათმეცნიერების ინსტიტუტის მეცნიერ-თანამშრომლები და კონკრეტული ენის სპეციალისტები მუშაობდნენ:

- ქართულ-აფხაზურ-ინგლისური სიხშირული ლექსიკონი (ბერიძე, თეთრაძე, დარასელია, შარაშენიძე და ფონიავა 2018);
- ქართულ-ოსურ-ინგლისური სიხშირული ლექსიკონი (ბერიძე, თეთრაძე, დარასელია, შარაშენიძე, ბეპივეი, პოპიაშვილი და ქუშარიტოვი 2018).

აღსანიშნავია, რომ თითოეულ ლექსიკონზე ლინგვისტებთან ერთად შერჩეულნი იყვნენ ისეთი სპეციალისტები, რომელთათვისაც ეს ენები მშობლიურია (მაგალითად, აფხაზური მასალის რედაქცია შეასრულა ლ. ავიძბამ). ლექსიკონების შედგენის მთავარი მიზანი იყო თანამედროვე ქართული ენის ლექსიკონის მიწოდება ენის შემსწავლელთათვის, რომელთა შორის, დიდი იმედით, აფხაზურისა და ოსურის მცოდნენიც მოიაზრებიან. ამასთან, თითოეული სიტყვაფორმის ინგლისური თარგმანი და ლათინური ტრანსლიტერაცია უფრო ადვილად გამოსაყენებელს ხდის მასალას: შემსწავლელს არ სჭირდება ქართული, ოსური ან აფხაზური ანბანის ცოდნა.

სამუშაო პროცესი ამგვარად წარიმართა: ვებკორპუსიდან აღებული 5000-სიტყვიანი სიხშირული სია, საბოლოოდ, ჩვენი მიზნების გათვალისწინებით, 3500 სიტყვამდე შევამცირეთ. გამოვრიცხეთ ქართული სალიტერატურო ენისათვის უმართებულო ფორმები (მაგ., **ესლა**

(30133) შეიცვალა სწორი ფორმით ახლა); ასევე შევამცირეთ საერთაშორისო სიტყვები: ისეთი სიტყვები, როგორებიცაა: ვიტამინი (456), ვირუსი (1530), დოლარი (38823) და მისთ. სიტყვანიდან ამოვიღეთ. თუმცა, რადგანაც სიტყვანი თანამედროვე წერილობით ოფიციალურ ენას ასახავს, ბუნებრივია, ნასესხები ლექსიკა რჩება ლექსიკონში ძირეულთან ერთად.

ინგლისური თარგმანისათვის დავეფუძნეთ კორპუსში დადასტურებულ მნიშვნელობათა სიხშირის რიგითობას, რომელიც ბევრ შემთხვევაში არ ემთხვეოდა ქეგლ-ის განმარტებებს. მაგალითად, სიტყვის **აიყვანს** I მნიშვნელობა ქეგლ-ის მიხედვით არის „ადაამიანს (ან ცოცხალ არსებას) ასწევს და დაიჭერს (ხელში)“. მხოლოდ IV და ერთგვარად გადატანილი მნიშვნელობაა „დამხმარედ, შემწედ იყოლიებს“. ჩვენს სიხშირულ სიაში კი აიყვანს ყველაზე ხშირად ჩნდება კოლოკაციაში „სამსახურში აიყვანს“, მხოლოდ ამის შემდეგ მოდის მნიშვნელობა „აწევს, ხელში აიყვანს“. მეტიც, ზოგიერთი სიტყვის ისეთი მნიშვნელობა აღმოჩნდა მაღალი სიხშირისა, რომელიც ქართული ენის განმარტებით ლექსიკონში ჯერ არ არის აღნიშნული. მაგალითად, ბმულის ძირითადი მნიშვნელობა ვებკორპუსის მიხედვით „ელექტრონული ბმულია“ (link), ამიტომაც ის სწორედ ამ მნიშვნელობით შევიყვანეთ ლექსიკონში. ბუნებრივია, ლექსიკონები ყველა მნიშვნელობას ვერ გადმოგვცემდა, ამიტომაც სალექსიკონო ერთეულს სიხშირის მიხედვით აქვს მითითებული ერთი, ორი ან სამი ყველაზე გავრცელებული მნიშვნელობა ინგლისურად, აფხაზურად ან ოსურად.

იმის გათვალისწინებით, რომ ლექსიკონები სასწავლოა, მათში შესულია არა მხოლოდ წმინდა ლექსიკური, არამედ გრამატიკული ინფორმაცია: ქართული სალექსიკონო ერთეულები წარმოდგენილია შემდეგი სახით: სახელების (მათ შორის, მასდარისა და მიმღობის) საწყისი სალექსიკონო ფორმაა სახელობითი ბრუნვის ფორმა, თითოეულს ახლავს ნათესაობითი ბრუნვის ფორმაც, რაც საშუალებას აძლევს ლექსიკონის მომხმარებელს, გაარჩიოს, როგორ შეიძლება შეიცვალოს ფუძე ფლექსიის დროს.

ზმნა წარმოდგენილია მხოლოდითი რიცხვის მესამე პირის ფორმით აწმყოსა ან მომავალში (ააგებს, არის...). რადგანაც ვებკორპუსის სიხშირულ 5000-სიტყვიან სიაში მოხვედრილი ზმნის პირიანი ფორმები შევიდა ლექსიკონში, გასათვალისწინებელია, რომ ზოგი ზმნა მხოლოდ აწმყოს მწკრივის ფორმითაა წარმოდგენილი, ზოგიც – მხოლოდ მომავლისა. თითოეულ პირიან ფორმას ახლავს მასდარი, ამით ენის შემსწავლელს გაუადვილდება ქართული ენის გრამატიკის ამ ურთულესი მონაკვეთის შესწავლა: ზმნის პირიან ფორმას შესაბამის სახელ-ზმნებთან ერთად დაიმახსოვრებს.

სალექსიკონო ერთეულების აფხაზურად თარგმნისას გასათვალისწინებელი იყო აფხაზურის ბუნება: აფხაზურს არ მოეპოვება ბრუნების კატეგორია, შესაბამისად, ქართული სალექსიკონო ერთეულის ნათესაობითი ფორმები უთარგმნელი დარჩა. რასაკვირველია, სათანადო სემანტიკის გამონახვა აფხაზურს შეუძლია სხვა საშუალებით, თუმცა ქართული ნათესაობითი ბრუნვის ფორმების შინაარსობრივი გადმოცემა აფხაზურში საჭიროდ არ მივიჩნიეთ. აფხაზური განარჩევს გონიერისა და არაგონიერის გრამატიკულ კლასს, აქედან გამომდინარე, საჭიროების შემთხვევაში სახელებში გამოვიყენეთ შემდეგი აღნიშვნები: გონიერისა (resp. ადამიანის) – [P], არაგონიერისა (resp. ნივთის) – [NP], ზოგჯერ ადამიანის კლასიც დიფერენცირებულია, შესაბამისად, გვაქვს ასეთი აღნიშვნა: მამაკაცის კლასი [M] და ქალის კლასი [F]. ზმნის პირიან ფორმებში სუბიექტად ძირითადად წარმოდგენილია ადამიანის კლასის (თუ განირჩევა მხოლოდ გონიერი [P] და არაგონიერი [NP]) ან მამაკაცის კლასის ფორმები (თუ განირჩევა კლასის სამივე ფორმა). ასევე, გადასაწყვეტი საკითხი იყო ზმნური ერთეულის აწმყოსა და მომავალი დროის ფორმების აფხაზურად თარგმნისას შინაარსობრივი და ფორმობრივი შესაბამისობის დაცვა. აფხაზურში მომავალი დროის შინაარსიც ბუნებრივად გადმოიცემა აწმყოს ფორმებით, თუმცა მეტი სიზუსტისათვის ქართულის ახლანდელი დროის ფორმები ითარგმნა აწმყოს (-yeIT /-oIT სუფიქსიანი ფორმებით) მწკრივით, მყოფადისა კი – მყოფადი I-ის მწკრივით (-II სუფიქსიანი ფორმებით). ქართული ზმნის პირიანი ფორმების შესაბამისი მასდარების აფხაზური თარგმანი ლექსიკონს ერთვის დანართის სახით. აფხაზური თარგმანისას ასევე გათვალისწინებული იყო მახვილის ადგილი სალექსიკონო ერთეულებში, რადგანაც აფხაზურში მახვილს ფონოლოგიური მნიშვნელობა აქვს.

რაც შეეხება ოსურ ვარიანტს, თარგმნისას არც სახელებთან და არც ზმნურ ფორმებთან სირთულე არ ყოფილა. ოსური სალექსიკონო ერთეული არის ქართული სიტყვაფორმის ეკვივალენტი. ოსურში სახელები წარმოდგენილია მხოლოდითი რიცხვის ფორმითა და დართული აქვს ნათესაობითი ბრუნვის ფორმები ქართულის შესაბამისად. ზმნათა თარგმანიც მიჰყვება ქართული ზმნური ფორმების შინაარსს, ისეთ შემთხვევაში, როდესაც ოსურში შეუძლებელია ესა თუ ის კატეგორია გადმოიცეს, ოსურში აღწერითი ფორმები გვაქვს.

ლექსიკონების ერთი დადებითი მხარე არის სიტყვაფორმათა თანმხლები ლათინური ტრანსლიტერაცია. გამოყენებულია საერთაშორისო ფონეტიკური ანბანი (IPA) ერთგვარი სახეცვლილებით. შევარჩიეთ ერთი სტილის ტრანსლიტერაცია ქართულისათვის, აფხაზურისა თუ

ოსურისათვის, რომ უფრო ადვილი ყოფილიყო მომხმარებლისათვის ტრანსლიტერირებული ფორმების ამოკითხვა, მიუხედავად ენათა სხვაობისა. თუმცა ოსურისათვის რთული აღმოჩნდა ტრანსლიტერაციის სტანდარტული ცხრილის შედგენა რამდენიმე გადაუჭრელი საკითხის გამო. ოსურისათვის შედგენილი საერთაშორისო ფონეტიკური ანბანი ასახავს ირონული დიალექტის ფონემურ ინვენტარს. ირონული საფუძველია სალიტერატურო, სტანდარტული ოსურისა, თუმცა ირონულშიც ზოგიერთი ფონემის ვარიაციულობა კილო-თქმათა მიხედვით ვლინდება. მაგალითად, C ირონულის კუდარულ კილოკავში (ზოგიერთი მეცნიერი მას ცალკე დიალექტადაც მიიჩნევს), რომელზეც მეტყველებენ შიდა ქართლსა და ცხინვალის რეგიონში მცხოვრები ოსები, გამოითქმის როგორ სისინა სპირანტი /s/, ხოლო ირონულის სხვა ვარიანტებში იგი შიშინა სპირანტად იკითხება /š/. ამდაგვარი სხვაობა, ძირითადად, სისინ-შიშინა ბგერებს შეეხება. მეორე, ასევე აღსანიშნი საკითხია დეზაფრიკატიზაციის პროცესი, ირონულში აფრიკატები /č/, /c/, /z/, /ʒ/, ძირითადად, სპირანტებადაა ქცეული, თუმცა ამ შემთხვევაში ტრანსლიტერაციის ტრადიციულ ვარიანტს მივყევით და შესაბამისი ნიშნებით აღვნიშნეთ.

ამგვარად, ქართულ-აფხაზურ-ინგლისურ სიხშირულ ლექსიკონში თითოეულ გვერდზე წარმოდგენილია შემდეგი ინფორმაცია:

- I – ქართული მეთაური სიტყვა;
- II – მეთაური სიტყვის ტრანსლიტერაცია;
- III – ფორმაწარმოებითი/ სიტყვაწარმოებითი ვარიანტი (vn, Gen);
- IV – ფორმაწარმოებითი/ სიტყვაწარმოებითი ვარიანტის ტრანსლიტერაცია;
- V – აფხაზური თარგმანი;
- VI – აფხაზური თარგმანის ტრანსლიტერაცია;
- VII – ინგლისური თარგმანი.

ქართულ-ოსურ-ინგლისურ სიხშირულ ლექსიკონში ზემოთ აღნიშნულ ინფორმაციას ემატება:

- VII – ოსურის ფორმაწარმოებითი/ სიტყვაწარმოებითი ვარიანტი;
- VIII – ოსურის ფორმაწარმოებითი/ სიტყვაწარმოებითი ვარიანტის ტრანსლიტერაცია;

შესაბამისად, IX სვეტი ინგლისურ თარგმანს ეთმობა.

ლექსიკონებს ახლავს: წინასიტყვაობა, ტრანსლიტერაციის ცხრილი ქართულისათვის და შესაბამისი ენისათვის, ინგლისურ თარგმან-

ში გამოყენებულ შემოკლებათა განმარტებანი და ლექსიკონებში შესულ ზმნის პირიან ფორმათა კონსტრუქციების ცნობარი, რომელიც ქართული ენის შემსწავლელთათვის გამოსადეგი უნდა იყოს იმის გათვალისწინებით, თუ რამდენად რთულია ზმნებთან არგუმენტთა შეწყობის ყალიბების დასწავლა მწკრივთა მიხედვით. დანართში მოცემულია თითოეული ზმნა 7 ძირითად მწკრივში არგუმენტთა ბრუნვის ფორმების მითითებით, მაგალითად: ჭამს ის მას, ჭამდა ის მას, შეჭამს ის მას და ა.შ. ნაცვალსახელების სახელებით მარტივად ჩანაცვლებით ქართული ენის შემსწავლელი შეძლებს კონკრეტული ზმნის გამოყენებით სწორად ააგოს წინადადება.

ამ ლექსიკონების ერთი მთავარი ნაკლი არის ის, რომ სალექსიკონო ერთეულს არ ახლავს საილუსტრაციო მასალა. დროის იმ მონაკვეთში, რომელიც გამოყოფილი იყო ამ ლექსიკონების შექმნისათვის, ტექნიკურად წარმოუდგენელი აღმოჩნდა სათანადო საილუსტრაციო მასალის დამატება. თუმცა რადგანაც ლექსიკონი ეფუძნება სინშირულ სიას, რომელიც ქართული ვებკორპუსიდან არის აღებული, ეს ფაქტი მკითხველს, ენის შემსწავლელს, უჩვენებს გზას კორპუსისაკენ, სადაც მას შეუძლია ყველა საჭირო კონტექსტის ნახვა. მომავლისათვის იგეგმება ლექსიკონებში კოლოკაციების დამატება, კონტექსტების მითითება და ამ მხრივ, მათი, როგორც სასწავლო რესურსის, დახვეწა.

თუ სინშირულ მრავალენოვან ლექსიკონებს უფრო მეტად სასწავლო დანიშნულება აქვს, დიალექტური ლექსიკონები თანამედროვე ტექნოლოგიების გამოყენებით ქართული ენობრივი მრავალფეროვნების ამსახველი ნიმუშია. საერთოდ კი, კორპუსზე დაფუძნებული ქართული ლექსიკონების შექმნით ვითარდება ქართულენოვანი ელექტრონული ლექსიკოგრაფია, ან უკეთ, კორპუსული ლექსიკოგრაფია საქართველოში.

დამოწმებანი

ბერიძე, მარინა. 2020. ქართული დიალექტური კორპუსის ლექსიკოგრაფიული კონცეფცია. *კადმოსი* 10. 2018, გვერდებს ჩვენ დავაზუსტებთ.

ბერიძე, მარინა. 2010. კორპუსის ლინგვისტიკა და თანამედროვე დისციპლინარული მატრიცა. *წახნაგი, ფილოლოგიურ კვლევათა წელიწდეული*, 216-237.

დარასელია, სოფიკო და სერგეი შაროვი. 2015. ქართული ენის სინშირული ლექსიკის განსაზღვრა ვებკორპუსის მიხედვით, აკაკი წერეთლის სახელმწიფო უნივერსიტეტის მოამბე, 2015, #1(5), 165-

- 171.
- Beridze, Marina, Maka Tetradze, Sophiko Daraselia, Nino Sharashenidze, and Natia Poniava. 2018. *Georgian-Abkhazian-English Frequency Dictionary*. Wojnowice: KEW – Kolegium Europy Wschodniej.
- Beridze, Marina, Maka Tetradze, Sophiko Daraselia, Nino Sharashenidze, Naira Bepieva, Nino Popiashvili, and Sadiat Kumaritova. 2018. *Georgian-Ossetian-English Frequency Dictionary*. Wojnowice: KEW – Kolegium Europy Wschodniej.
- Beridze, Marina, Lia Bakuradze, Zakharia Pourtskhvanidze and David Nadaraia. 2018. Interactive Visualization of Dialectal Lexis Perspective of Research Using the Example of Georgian Electronic Dialect Atlas. In *Book of Abstracts of XVIII EURALEX International Congress: Lexicography in Global Contexts*, eds. Jaka Čibej, Vojko Gorjanc, Iztok Kosem and Simon Krek, 122. Ljubljana: Ljubljana University Press. http://euralex2018.cjvt.si/wp-content/uploads/sites/6/2018/07/Euralex2018_book_of_abstracts_FINAL.pdf.
- Beridze, Marina, Lia Bakuradze and David Nadaraia. 2017. Georgian Dialect Corpus: Linguistic and Encyclopedic Information in Online Dictionaries. *Journal of Linguistics*, vol. 68, no. 2, 109-121. https://content.sciendo.com/view/journals/jazcas/68/2/article-p109.xml_.
- Beridze, Marina and David Nadaraia. 2016. Lexicographic Potential of the Georgian Dialect Corpus. In *Proceedings of the XVII EURALEX International Congress, Lexicography and Linguistic Diversity*, September 6-10, 2016, eds. Tinatin Margalitadze and George Meladze, 300-309. Tbilisi: Ivane Javakhishvili Tbilisi State University Press. <https://euralex.org/publications/lexicographic-potential-of-the-georgian-dialect-corpus/>
- Beridze, Marina, Liana Lordkipanidze and David Nadaraia. 2015a. The Georgian Dialect Corpus: Problems and prospects. In *Historical Corpora Challenges and Perspectives*, eds. Jost Gippert and Ralf Gehrke, 323–333. (The volume contains 23 papers read at the international conference “Historical Corpora 2012”, hosted by the LOEWE Research Cluster “Digital Humanities” of the State of Hesse at the University of Frankfurt on December 6-8, 2012.) Tübingen: Narr.
- Beridze, Marina, Liana Lordkipanidze and David Nadaraia. 2015b. Dialect Dictionaries in the Georgian Dialect Corpus. In *Proceedings of the 10th International Symposium on Logic, and Computation*, Tbilisi 2013, held in Gudauri, Georgia, September 23-27, 2013. Revised Selected Papers, XIV, eds. Martin Aher, Daniel Hole, Emil Jeřábek, Clemens Kupke, 82–96. Berlin: Springer.

- Firth, John. R. 1957. *Papers in linguistics, 1934-1951*. London: Oxford University Press.
- Hanks, Patrick. 2012. Corpus Evidence and Electronic Lexicography. In *Electronic Lexicography*, eds. Sylviane Granger and Magali Paquot, 57-82. Oxford: Oxford University Press.
- McEnery, Tony and Andrew Hardie. 2013. The History of Corpus Linguistics: In *The Oxford Handbook of the History of Linguistics*, ed. Keith Allan. Oxford: Oxford University Press.
- McEnery, Tony and Andrew Hardie. 2011. *Corpus Linguistics: Method, Theory and Practice*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
- Sharoff, Serge, Elena Umanskaya and James Wilson. 2014. *A Frequency Dictionary of Russian. Core Vocabulary for Learners*. London and New York: Routledge.
- O’Keeffe, Anne and Michael McCarthy, eds. 2010. *The Routledge Handbook of Corpus Linguistics*. London: Routledge.
- Sharoff, Serge. 2006. Creating General-Purpose Corpora Using Automated Search Engine Queries: In *WaCky! Working Papers on the Web as Corpus*, eds. Marco Baroni and Silvia Bernardini, 63-98. Bologna: Gedit.
- Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, John. 1987. *Collins COBUILD English language dictionary*. London: Collins.
- Walter, Elizabeth. 2010. Using Corpora to Write Dictionaries. In O’Keeffe and Michael McCarthy 2010.