

## ბათუმის ლინგვოკულტუროლოგიური დიგიტალური არქივი: დოკუმენტირების ახალი პერსპექტივები

რამაზ ხალვაში (ბათუმი)

მეცნიერული კომუნიკაციის ახალი ფორმების ევოლუციამ და ვებ-ტექნოლოგიების განვითარებამ უპრეცედენტო შესაძლებლობები შეუქმნა მეცნიერებს გლობალური კავშირების მეშვეობით სწრაფად და ეფექტურად გამოიყენონ ელექტრონულ საცავებში დაცული დიგიტალური რესურსები, რომელთა რაოდენობაც დღითიდღე იზრდება. ღია რესურსების (Open-Resource, OR) შექნამ და რესურსების გაცვლამ (Resource Exchange, RE) დიგიტალური არქივების ინფრასტრუქტურის განვითარება გამოიწვია, რამაც, თავის მხრივ, სამეცნიერო რესურსების სწრაფი და ეფიციენტური გამოყენების შესაძლებლობები შექმნა და თვისობრივად შეცვალა XXI საუკუნის ჰუმანიტარიაში კვლევის პროცესი – მეცნიერებს საშუალება მიეცათ საკუთარი რესურსების გახსნისა და გაერთიანების გზით (ანუ „მსოფლიო ბიბლიოთეკის“ შექმნით) შეემუშავებინათ საერთაშორისო სამეცნიერო კომუნიკაციის ახალი გზები.

ენობრივი რესურსების დიგიტალიზაციისა და არქივირების აუცილებლობა დღის წესრიგში დადგა ამერიკის ლინგვისტური საზოგადოების მიერ 1991 წელს ჩატარებულ სიმპოზიუმზე, რომელიც საფრთხეში მყოფი ენების საკითხს მიეძღვნა. სტატისტიკურ მონაცემებზე დაყრდნობით<sup>1</sup> სიმპოზიუმზე შეკრებილმა მეცნიერებმა ენათა გაქრობის შემაშფოთებელი დინამიკა ივარაუდეს: მიხეილ კრაუსის პროგნოზით, წარმოდგენილი სტატისტიკური მონაცემებიდან გამომდინარე, ენათა მხოლოდ 10% გააგრძელებს არსებობას მომავალში, ენათა 50% გაქრობის საფრთხე ემუქრება, ხოლო დანარჩენი ენები საუკუნის დასასრულისთვის სავარაუდოდ გაქრება. მიუხედავად იმისა, რომ ენათა მოსალოდნელი გაქრობის საშიშროების შესახებ მოგვიანებით გამოთქმული პროგნოზი ნაკლებად პესიმისტურად გამოიყურებოდა, ენათა გაქრობის შესახებ კრაუსის მიერ გამოთქმული ვარაუდი, რომელიც მან 1992 წელს გამოქვეყნებულ სტატიაშიც გაიმეორა,<sup>2</sup> ყველაზე ხშირად ციტირებული შეფასებაა საფრთხეში მყოფი ენების საკითხთან დაკავშირებით გამოქვეყნებულ სამეცნიერო ლიტერატურაში. საგულისხმოა, რომ გაქრობის საშიშროება უმეტესად ინდიგენური ხალხების ენებს ემუქრება, რომელთაგან უმრავლესობა ჯერ კიდევ არ არის დოკუმენტირებული, რაც იმას ნიშნავს, რომ ამ ენებთან ერთად გაქრება არა მარტო ენა, როგორც სისტემა, არამედ ინდიგენური მოსახლეობის სულიერი კულტურაც.

### 1. ბათუმის ლინგვოკულტუროლოგიური დიგიტალური არქივი

ენისა და კულტურის ციფრული დოკუმენტირება და არქივირება კულტურული მემკვიდრეობის კვლევის აუცილებელ წინაპირობასაა და XXI საუკუნის გადაუდებელ ამოცანას წარმოადგენს. 2015 წელს ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტში ფრანკფურტის გოეთეს სახელობის უნივერსიტეტთან თანამშრომლობით შეემუშავებული პროექტი „ბათუმის ლინგვოკულტუროლოგიური დიგიტალური არქივი“, რომელიც განხორციელდა 2016–2019 წლებში

<sup>1</sup> სტატისტიკური ინფორმაცია ეფუძნებოდა ენის მატარებელთა რაოდენობას, რომელიც განისაზღვრა ბიბლიის თარგმანების საჭიროების გათვალისწინებით (*Ethnologue* 1988).

<sup>2</sup> Krauss 1992: 4–10.

შოთა რუსთაველის ეროვნული სამეცნიერო ფონდის დაფინანსებით (გრანტის ნომერი DI 2016-46), მიზნად ისახავდა საქართველოში მეცნიერული კომუნიკაციის ახალი ფორმების შემოტანას – ენობრივი რესურსების შეგროვებასა და დამუშავებას ენისა და კულტურის დოკუმენტირების თანამედროვე მეთოდებისა და ტექნოლოგიების გამოყენებით, კერძოდ, აჭარის ლინგვოკულტუროლოგიური დიგიტალური არქივის შექმნას მულტიმედია-ციფრული კვლევების ხელშეწყობის მიზნით.

ის ფაქტი, რომ პროექტის საკვლევო არეალი ისტორიული აჭარის რეგიონით განისაზღვრა, არ იყო შემთხვევითი. საკვლევო არეალის არჩევანი განპირობებული იყო მისი ლინგვოკულტუროლოგიური თავისებურებებით: საქართველოს ამ რეგიონში ოთხი საუკუნის განმავლობაში ერთმანეთის გვერდით ცხოვრობენ ქრისტიანული და მაჰმადიანური აღმსარებლობის ადამიანები, რომელთა თანაცხოვრების საფუძველზეც წარმოიქმნა კულტურათა ერთგვარი სიმბიოზი, რომელშიც დღემდე არის დაცული ენობრივი და კულტურული მრავალფეროვნება. აჭარის რეგიონის ლინგვოკულტუროლოგიური დოკუმენტირების აუცილებლობა გლობალიზაციისა და მიგრაციული პროცესების შედეგებით იყო გამოწვეული: ენობრივი, რელიგიური, ეთნიკური, ეკონომიკური და სოციალური ცვლილებების შედეგად საშიშროება შეექმნა აჭარის უნიკალურ მოზაიკურ სურათს.

### 1.1 პროექტის წევრები და კომპეტენცია

საგრანტო პროექტში მონაწილეობდნენ: მანანა თანდაშვილი (პროექტის სამეცნიერო ხელმძღვანელი, ფრანკფურტის გოეთეს უნივერსიტეტი), რამაზ ხალვაში (პროექტის ხელმძღვანელი, ბათუმის სახელმწიფო უნივერსიტეტი), ხათუნა ბერიძე (პროექტის კოორდინატორი, ბათუმის სახელმწიფო უნივერსიტეტი), მზია ხახუტაიშვილი (მკვლევარი, ბათუმის სახელმწიფო უნივერსიტეტი), ნანა ცეცხლაძე (მკვლევარი, ბათუმის სახელმწიფო უნივერსიტეტი), რუსუდან პაპიაშვილი (მკვლევარი, თბილისის სახელმწიფო უნივერსიტეტი), გრიგოლ კახიანი (მკვლევარი, ბათუმის სახელმწიფო უნივერსიტეტი) და შორენა ფუტყარაძე (მკვლევარი, ბათუმის სახელმწიფო უნივერსიტეტი).

პროექტის სამეცნიერო ხელმძღვანელის მანანა თანდაშვილის მრავალწლიანი გამოცდილება და კომპეტენცია ენის დოკუმენტირების სფეროში დიდად დაეხმარა სამეცნიერო ჯგუფს არა მხოლოდ ენობრივი რესურსების შეგროვებასა და ტექნოლოგიურ დამუშავებაში, არამედ ბათუმის ლინგვოკულტუროლოგიური დიგიტალური არქივის კონცეპტუალური და ტექნოლოგიური ჩარჩოს შემუშავებაშიც. ფრანკფურტის უნივერსიტეტსა და ბათუმის უნივერსიტეტს შორის გაფორმებული მემორანდუმის ფარგლებში არაერთი აქტივობა განხორციელდა დიგიტალური ჰუმანიტარიის მიმართულებით: დაარსდა ბათუმის საერთაშორისო საზაფხულო სკოლა დიგიტალურ ჰუმანიტარიაში, რომელიც 2012 წლიდან მოყოლებული ყოველ წელს ტარდებოდა, მათ შორის ენის დიგიტალური დოკუმენტირების საკითხებზე. 2015 წელს ბათუმის სახელმწიფო უნივერსიტეტში დაინერგა საბაკალავრო მაინორი პროგრამა დიგიტალურ ჰუმანიტარიაში, რომელიც მოიცავდა საფრთხეში მყოფი ენების დიგიტალური დოკუმენტირების საკითხებს (პროგრამის ხელმძღვანელები მანანა თანდაშვილი და რამაზ ხალვაში). გარდა ამისა, სამეცნიერო ჯგუფის წევრთა უმრავლესობას უკვე ჰქონდა

საერთაშორისო სამეცნიერო პროექტებში (ECLinG და SSGG) ჯგუფური მუშაობის<sup>3</sup> გამოცდილება: რამაზ ხალვაში, მზია ხახუტაიშვილი, ნანა ცეცხლაძე და შორენა ფუტყარაძე 2001–2016 წლებში მონაწილეობდნენ არაერთ ექსპედიციაში, რომლებიც თურქეთში მცხოვრები ქართველების ლინგოკულტუროლოგიურ შესწავლას ისახავდა მიზნად. რამაზ ხალვაშის ხელმძღვანელობით შეიქმნა თანამედროვე ქართული პრესის (გაზეთ „სამუსლიმანო საქართველოს“) კორპუსი, ხოლო ხათუნა ბერიძის ხელმძღვანელობით შეიქმნა დიგიტალური რესურსი „თარგმნილი პოლიტიკური კორპუსი“. რუსუდან პაპიაშვილი 2002–2015 წლებში მონაწილეობდა ECLinG-ისა და SSGG-ს პროექტებში – ახორციელებდა სამეცნიერო მივლინებებში მოპოვებული მასალის ტექნიკურ დამუშავებას. შესაბამისად, პროექტის წევრებს უკვე ჰქონდათ გამოცდილება და მკაფიო წარმოდგენა დასახული მიზნის მისაღწევად: დიგიტალური ლინგოკულტუროლოგიური არქივის შესაქმნელად.

## 1.2. პროექტის განხორციელების ეტაპები.

პროექტის განხორციელების პირველ ეტაპზე შემუშავდა პროექტის კონცეპტუალური, მეთოდოლოგიური და თეორიული ჩარჩო: ჩატარდა სემინარები და ვორკ-შოპები, და შექმნა თეორიული ნაშრომები დოკუმენტაციის მეთოდოლოგიური და ტექნოლოგიური საფუძვლების შესახებ.

ბათუმის სახელმწიფო უნივერსიტეტში ლინგოკულტუროლოგიური დიგიტალური არქივის (DigiArchive) შექმნა სამ ეტაპად განხორციელდა:

I ეტაპი – მოსამზადებელი პერიოდი, რომელიც 6 თვეს მოიცავდა და ითვალისწინებდა:

1. დოკუმენტირებისა და არქივირების თეორიული და მეთოდოლოგიური საფუძვლების გაცნობას;
2. დიგიტალური არქივის ტექნოლოგიური ჩარჩოს შემუშავებას;
3. პროექტის სამეცნიერო ხელმძღვანელის სემინარებსა და ტრენინგებს კვლევისათვის რელევანტური ტექნოლოგიური ინსტრუმენტების ასათვისებლად;<sup>4</sup>
4. პროექტში მონაწილე პერსონალის მივლინებებს ფრანკფურტის უნივერსიტეტსა და ნეიმეხენის მაქს პლანკის ინსტიტუტში კვალიფიკაციის ასამაღლებლად.

II ეტაპი – საკვლევი რესურსების მოპოვებისა და ტექნიკური დამუშავების პერიოდი, რომელიც 24 თვეს გრძელდებოდა და ითვალისწინებდა:

1. ინტერდისციპლინური ლინგოკულტუროლოგიური რესურსების მოპოვებას;
2. მოპოვებული მასალების სისტემატიზაციას, ლაბორატორიულ დამუშავებასა და არქივირებას;
3. პროექტის ორენოვანი (ქართულ-ინგლისური) ვებგვერდის კონცეპტუალიზაციას, ვებგვერდისა და სამომხმარებლო ინტერფეისის საპილოტე ვერსიის შექმნას; (იხ. სურათი 1)

<sup>3</sup> ხალვაში 2018: 148–154.

<sup>4</sup> ISLE – <http://www.mpi.nl/ISLE/>; Toolbox – <http://www-01.sil.org/computing/toolbox/>; TEI – <http://www.tei-c.org/index.xml>; EAF – <http://www.file-extension.org/de/extensions/eaf/>; IMDI – <https://tla.mpi.nl/imdi-metadata/>; TLA – <https://corpus1.mpi.nl/ds/asv/>; DOBES – <http://dobes.mpi.nl/dobesprogramme/>.

4. DigiArchive-ის ვებგვერდის კონტენტით შევსებას, სამომხმარებლო ინტერფეისის ვერიფიცირებას, სერვისის ოპტიმიზაციას, გამოცდასა და გახსნას ფართო წოდებისათვის.

III ეტაპი – რესურსების არქივირებისა და თეორიული კვლევის პერიოდი, რომელიც გაგრძელდა 6 თვეს და ითვალისწინებდა:

1. DigiArchive-ის ტექნოლოგიური და მეთოდოლოგიური ჩარჩოს შექმნას;
2. თეორიული ნაშრომებისა და DigiArchive-ის გზამკვლევის შექმნას.



სურ. 1: ბათუმის ლინგუოკულტუროლოგიური დიგიტალური არქივის ვებგვერდი

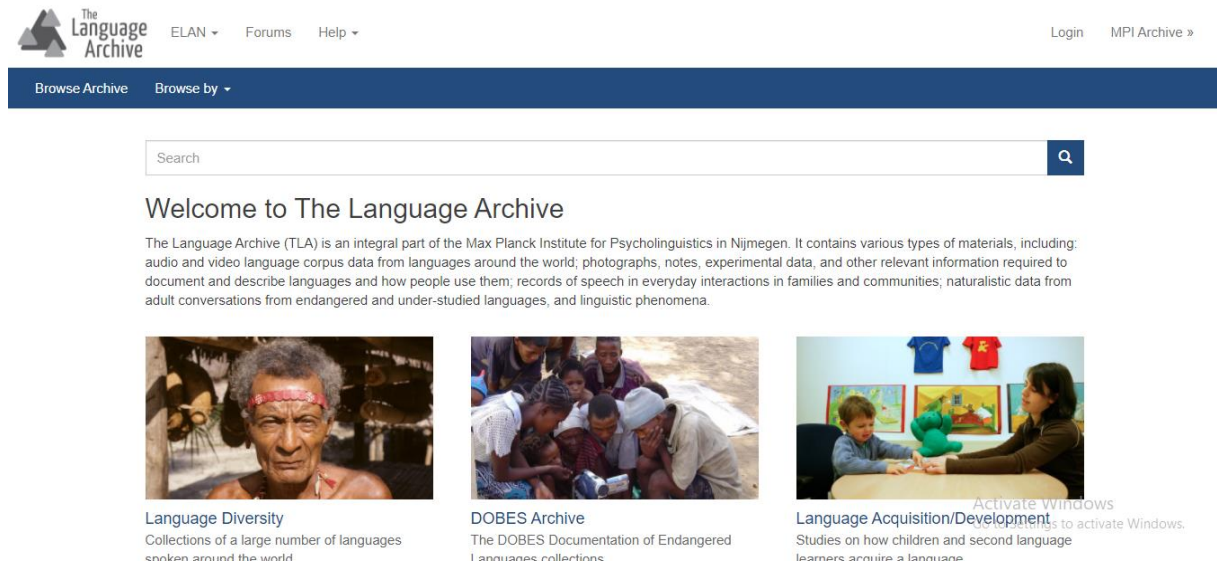
### 1.3 კვლევითი თემის აქტუალობა

XXI საუკუნეში ჰუმანიტარული მეცნიერება უახლესი ტექნოლოგიური საშუალებებით დოკუმენტირებულ და არქივირებულ ინტერდისციპლინურ კვლევით რესურსებს ეფუძნება. ენისა და კულტურის დოკუმენტირების თანამედროვე გამოცდილებამ საფუძველი ჩაუყარა ახალ დისციპლინას – დოკულინგვისტიკას, რომელსაც დღეისათვის გააჩნია საკუთარი თეორიული ბაზისი (დიგიტალური დოკუმენტირებისა და არქივირების ძირითადი პრინციპები და მეთოდები) და ჩარჩო-პროგრამა. გარდა რესურსების მოპოვების (მასალის ჩაწერის), აღრიცხვის (რეგისტრაციის) და ტექნიკური დამუშავებისა (დიგიტალიზაციის), მნიშვნელოვანი ადგილი უჭირავს რესურსების შენახვისა (არქივირების) და შემდგომი მოვლის (დაცვის) მეთოდოლოგიური და ტექნოლოგიური საშუალებების დანერგვას. ამიტომ პროექტის პერსონალი იმთავითვე აცნობიერებდა, რომ საქართველოში დოკულინგვისტიკის მეთოდებისა და ტექნოლოგიების გამოყენების დანერგვა არსებითად ახალი სამეცნიერო მიმართულების – დიგიტალური ქართველოლოგიის განვითარებას ნიშნავდა საქართველოში.

დოკულინგვისტიკის ფარგლებში დამუშავებული მეთოდი – დიგიტალური დოკუმენტირება და არქივირება ახლებურად აყენებს საკითხს: დოკუმენტირების ობიექტი არის ენა არა, როგორც ნიშანთა სისტემა, არამედ ენა მის ბუნებრივ

მოცემულობაში – კონკრეტულ დროსა და სივრცეში, კონკრეტული სოციუმის ასაკობრივი, სოციალური და გენდერული ბალანსის გათვალისწინებით.<sup>5</sup>

პროექტის განხორციელების დროს მოპოვებული რესურსების დიგიტალური დოკუმენტირებისათვის გამოყენებულ იქნა საერთაშორისო სტანდარტი, რომელიც დამუშავდა ფოლკსვაგენის ფონდის პროგრამის DOBES-ის<sup>6</sup> ფარგლებში (იხ. სურ. 2).



სურ. 2: ნეიმეხენის მაქს პლანკის სახელობის ინსტიტუტის ენის არქივის ვებგვერდი

DOBES-ის პროგრამა მიზნად ისახავდა საფრთხეში მყოფი ენების დოკუმენტირებას მათი არქივირებისა და შესაძლო რევიტალიზაციის მიზნით. პროგრამის ფარგლებში მოპოვებული რესურსები განთავსდა ნეიმეხენის მაქს პლანკის სახელობის ფსიქოლინგვისტიკის ინსტიტუტის ენის არქივში TLA (The Language Archive). ენობრივ მონაცემთა ელექტრონული არქივირებისათვის აქ საგანგებოდ შეიქმნა საარქივო მულტიმედიური ფორმატი EAF, რომელიც წარმოადგენს სამი სახის მედიარესურსის – ვიდეო-, აუდიო- და ტრანსკრიბირებული ტექსტის სინქრონიზაციას ერთ ფაილში. აღნიშნული ფაილი შეიცავს ტრანსკრიბირებულ ტექსტს, რომელიც სინქრონიზებულია ვიდეო-გამოსახულებასთან და აუდიო-სიგნალთან. ენობრივი რესურსის ამგვარად მომზადება-არქივირება საგრძნობლად ამადლებს პროდუქტის ხარისხს, ვინაიდან: ა) იგი იძლევა მასალის ვერიფიცირების საშუალებას და ბ) რესურსს ანიჭებს ინტერდისციპლინური კვლევის რესურსის სტატუსს.<sup>7</sup>

<sup>5</sup> Himmelmann 1998: 161–195.

<sup>6</sup> <https://dobes.mpi.nl/>.

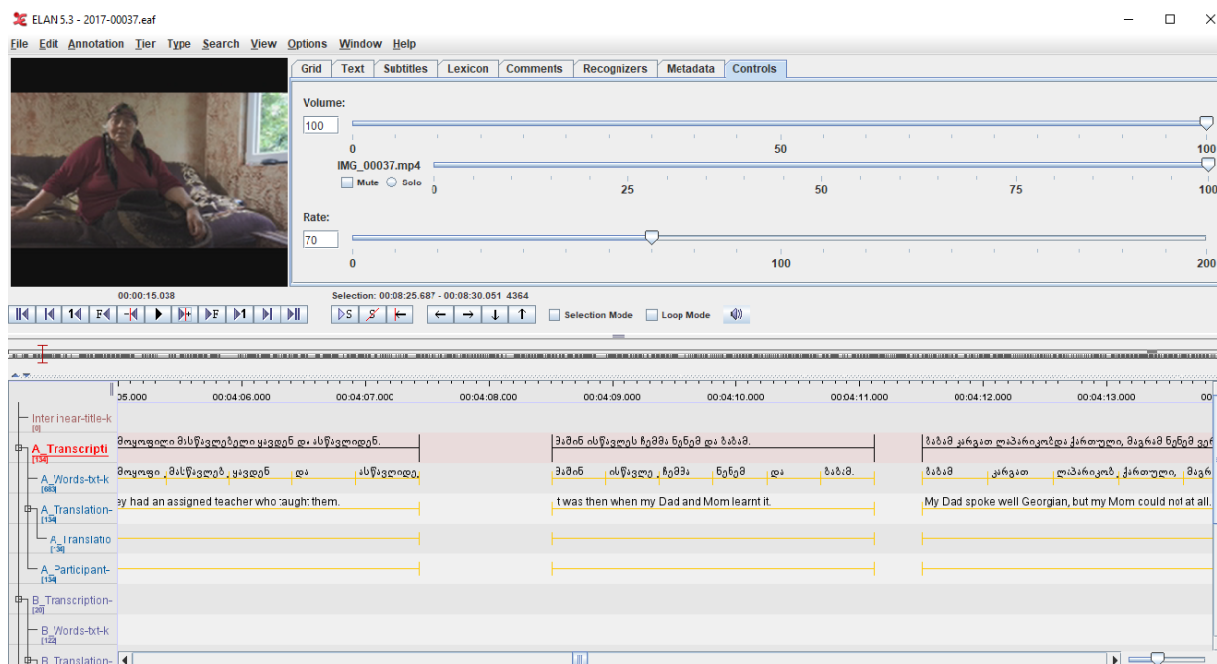
<sup>7</sup> თანდაშივილი 2020: 198.

2. პროექტის ეტაპობრივი განხორციელება

2.1. ძირითადი პრინციპები და მეთოდები

პროექტის პირველ ეტაპზე განხორციელდა საველე ექსპედიციის კონცეპტუალური-ზაცია, რომელიც შედგებოდა შემდეგი ამოცანებისაგან:

1. სამეცნიერო ექსპედიციის კონცეპტუალური დაგეგმვა – ექსპედიციის გეგმა-გრაფიკის შემუშავება და მისი განხორციელების ოპტიმირება რისკ-ფაქტორების (სახელმწიფო-პოლიტიკური, კლიმატური, რელიგიური, საზოგადოების სოციალური სტრუქტურა და ა. შ.) გათვალისწინებით;
2. მასალების საველე პირობებში მოპოვება დოკუმენტაციის ძირითადი პრინციპების გათვალისწინებით, როგორცაა: ენობრივი კომპეტენციის განსხვავებული საფეხურების განსაზღვრა, ენის/დიალექტის გამოყენების სფეროების დადგენა ბილინგვალიზმის პირობებში, კოდების ცვლისა და ენის/დიალექტის საფრთხეში მყოფობის ინდიკატორების კვლევა, ასაკობრივი ცენზისა და გენდერული ბალანსის დაცვა, მასალის თემატური მრავალფეროვნება;
3. მასალის გადაღების/ფიქსაციის პროცესში ტექნიკური პარამეტრების დაცვა (გადაღების ობიექტის განსაზღვრა, კადრების მონაცვლეობის სიხშირე, გამოსახულების ოპტიმირება, გადაღების პლანის ცვლა და ა.შ.).



სურ. 3: პროექტის ფარგლებში მოპოვებული მასალა მულტიმედიალურ ფორმატში (EAF)

მოპოვებული მასალის ლაბორატორიული დამუშავება განხორციელდა შემდეგი პრინციპების დაცვით (იხ. სურათი 3):

1. მოპოვებული მასალების აღრიცხვა განხორციელდა საერთაშორისო სტანდარტის TEI (Text Encoding Initiative) მიხედვით.
2. მოპოვებული მასალების გადამუშავებისა და მათი ინტერდისციპლინარული კვლევის რესურსად გარდაქმნის დროს, ასევე, დაცული იყო საერთაშორისო სტანდარტები, რაც გულისხმობს:

- a) მოპოვებული მასალების დიგიტალიზაციას AVI (Audio Video Interleave) ფორმატში;
- b) AVI ფაილების დაჭრა-დახარისხებას (Session-Making);
  - მეტამონაცემებით აღჭურვა საერთაშორისო სტანდარტის საგანგებო ფორმატის IMDI (ISLE [International Standard for Language Engineering] Meta Data Initiative) მიხედვით;
  - მასალების ტრანსკრიბირება კოდირების საერთაშორისო კონსორციუმის Unicode-ის სტანდარტების შესაბამისად;
  - ტრანსკრიბირებული მასალების გადატანა მონაცემთა მართვის, ტექსტის გარჩევისა და ანალიზის მულტიმედია ფორმატში Flex (Fieldworks Language Explorer);
  - ვიდეო, აუდიო და მულტიმედია ანოტირებული ფორმატების სინქრონიზაცია (Annotation Format) ფაილში.

## 2.2 პროექტის რესურსების არქივირება და დაცვა

ბათუმის ლინგოკულტუროლოგიური დიგიტალური არქივის ერთ-ერთ მნიშვნელოვან ამოცანას წარმოადგენდა პროექტის ფარგლებში გადამუშავებული რესურსების არქივირებისა და შემდგომი დაცვის უზრუნველყოფა, რაც გულისხმობდა საარქივო რესურსების შენახვისა და დაცვის უზრუნველყოფას, რაც შემდეგი სახით განხორციელდა:

- ა) შეიქმნა არქივის დუბლიკატი სარკისებური სერვერის სახით;
- ბ) უზრუნველყოფილ იქნა სერვერის ავტომატური სათადარიგო კოპირება (Backup).

XXI საუკუნეში საერთაშორისო სამეცნიერო საზოგადოება კვლევითი რესურსების ინტერდისციპლინარულობისა და ინტეგრაციული კვლევებისაკენ მიისწრაფვის. ამის დასტურია, მაგ., სამეცნიერო კონსორციუმის CLARIN-ის (Common Language Resources and Technology Infrastructure) შექმნა, რომელიც ევროპული სამეცნიერო პროექტის სახით განხორციელდა და დღეს მაღალი რანგის სამეცნიერო ფორუმია ენობრივი რესურსებისა და ტექნოლოგიური ინფრასტრუქტურის განვითარების, სტანდარტიზაციისა და გაცვლის პერსპექტივებით. პროექტის შემდგომი მიზანია ბათუმის ლინგოკულტუროლოგიური დიგიტალური არქივის რესურსების ინტეგრაცია საერთაშორისო არქივებში სამეცნიერო მიმოქცევის მიზნით. ბათუმის დიგიტალური ლინგოკულტუროლოგიური არქივის შემდგომი განვითარების პერსპექტივა საერთაშორისო რესურსის სტატუსს მიანიჭებს პროექტის ფარგლებში მოპოვებულ მასალებს და წვდომადს გახდის მას ინტერდისციპლინური კვლევებისათვის საერთაშორისო დონეზე.

პროფესიული გადამზადების მიზნით, პროექტის ფარგლებში დაიგეგმა და განხორციელდა ახალგაზრდა მეცნიერთა მივლინებები ფრანკფურტის უნივერსიტეტსა (გერმანია) და ნეიმეხენის მაქს პლანკის ინსტიტუტში (ჰოლანდია), იქ არსებული რესურსების, არქივირების პროგრამული საშუალებებისა და ტექნოლოგიური მეთოდების გაცნობის მიზნით. ასევე დაიგეგმა ფრანკფურტის უნივერსიტეტის თანამშრომლის მივლინებები საქართველოში სემინარის, ტრენინგებისა და საჯარო ლექციების ჩასატარებლად.

## 2.3 პროექტის შედეგები

პროექტი რამდენიმე ეტაპად განხორციელდა და შესრულდა შემდეგი სახის სამუშაოები:

1. დოკუმენტებისა და არქივების მეთოდების თეორიული და მეთოდოლოგიური საფუძვლების გაცნობა;
2. არქივების ტექნოლოგიური და პროგრამული ინსტრუმენტების ათვისება;
3. რესურსების სავსელე პირობებში მოპოვება;
4. ახლადმოპოვებული რესურსების ტექნიკური მომზადება არქივებისათვის:
  - ა) ტრანსკრიბირებული ტექსტების გადატანა არსებული ფორმატიდან Flex-ის ფორმატში;
  - ბ) ტექსტების შემდგომი დამუშავება გლოსირებისა და ხაზთაშორისი ანალიზის თვალსაზრისით;
  - გ) რესურსების თარგმნა ინგლისურად;
5. რესურსების სისტემატიზაცია და არქივირება:
  - ა) საარქივო რესურსების მეტამონაცემებით აღჭურვა;
  - ბ) საარქივო რესურსების გადატანა EAF ფორმატში.

3. კვლევის შედეგების სამეცნიერო ღირებულება და პრაქტიკული გამოყენება.

პროექტის ფარგლებში დაარქივებული რესურსები დამუშავების ხარისხის მიხედვით ოთხ ტიპად დაჯგუფდა:

- A ტიპის რესურსი: არქივირებული აუდიო- და ვიდეომასალა – 611 ფაილი 100 საათამდე მოცულობის (MP3 და AVI ფორმატი);
- B ტიპის რესურსი: არქივირებული აუდიო- და ვიდეომასალები ტრანსკრიბირებული ტექსტის თანხლებით – 205 ფაილი 30 საათამდე მოცულობის (Flex-ში სინქრონიზებული რესურსები);
- C ტიპის რესურსი: არქივირებული ვიდეომასალები მულტიმედიური ანოტირებით – ტრანსკრიბირებული, გლოსირებული და ინტერლინიალიზებული – 69 ფაილი 10 საათამდე მოცულობის (EAF ტიპის რესურსი);
- D ტიპის რესურსი: დიგიტალურად დოკუმენტირებული და არქივირებული ვიდეომასალა მულტიმედიური ანოტირებით და ინგლისური თარგმანით – 35 ფაილი 5 საათამდე მოცულობის (EAF ტიპის რესურსი).

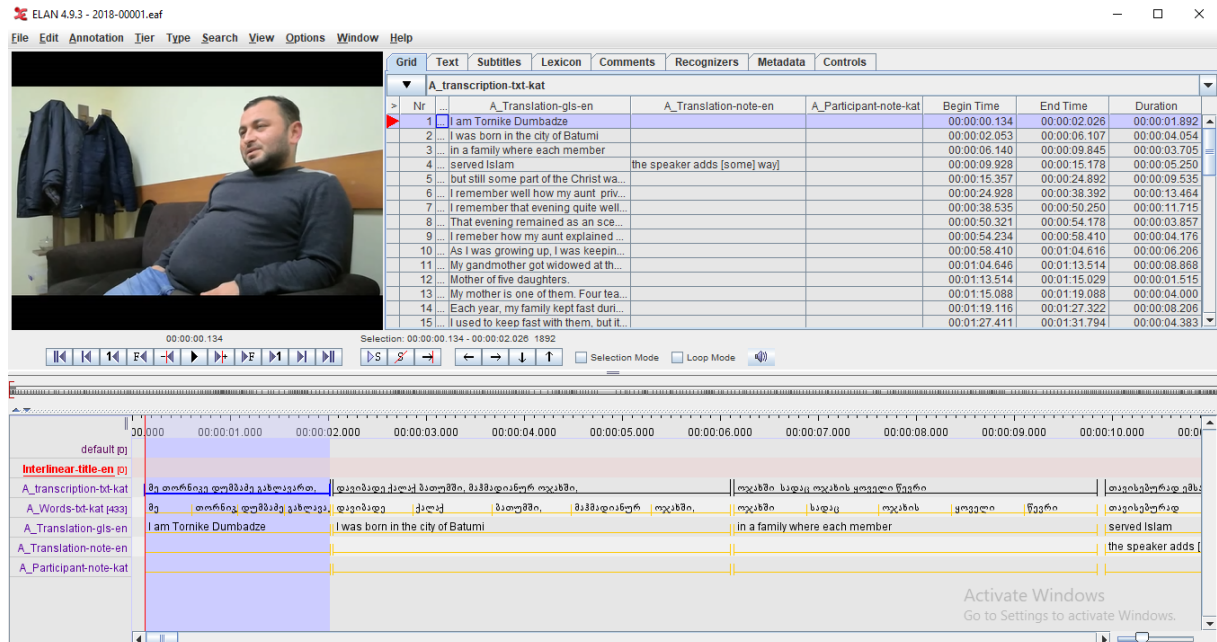
4. დიგიტალური ლინგვოკულტუროლოგიური არქივის მნიშვნელობა

ბათუმის ლინგვოკულტუროლოგიური დიგიარქივის შექმნამ მდგრადობა მიანიჭა დიგიტალური დოკუმენტირებისა და არქივირების განვითარებას საქართველოში. მაგალითად, 2021 და 2023 წლების ბათუმის დიგიტალური ჰუმანიტარიის საზაფხულო სკოლა მთლიანად მიეძღვნა ლინგვოკულტურო-ლოგიური დიგიტალური არქივის (DigiArchive) ახალი მასალით შევსებას და მიზნად ისახავდა აჭარის ლინგვოკულტუროლოგიური პორტრეტის შექმნაში სტუდენტების (ბაკალავრების, მაგისტრანტებისა და დოქტორანტების) ჩართვას. საკვლევი არეალი განისაზღვრა სხალთა-ხისხანის ხეობით. სტუდენტების მიერ მოპოვებული და დამუშავებული მასალა (20 სთ.) განთავსდა DigiArchive-ის ვებგვერდზე.<sup>8</sup>

<sup>8</sup> <https://digiarchive.bsu.edu.ge/>.



DigiArchive-ის ფარგლებში რესურსების მოპოვების, ლაბორატორიული დამუშავებისა და არქივირების თანამედროვე სტანდარტების გამოყენებამ, ერთი მხრივ, და კვლევის ინოვაციური მეთოდების თეორიული და პრაქტიკული სახელმძღვანელოების შექმნამ, მეორე მხრივ, მყარი საფუძველი შეუქმნა ჰუმანიტარულ კვლევებს საქართველოში.



სურ. 4: დიგიტალური არქივის რესურსი – ინტერვიუ სინქრონიზებული წინადადებების ინგლისურ თარგმანთან

პროექტის ფარგლებში მომზადდა და გამოქვეყნდა შემდეგი სამეცნიერო ნაშრომები:

1. მანანა თანდაშილის მონოგრაფია „ენისა და კულტურის დოკუმენტირების თანამედროვე მეთოდები და ინსტრუმენტები: ლინგვოკულტუროლოგიური დიგიტალური არქივის მაგალითზე“ (გამომცემლობა „უნივერსალი“, თბ., 2020; შემდგომში: თანდაშილი 2020). ნაშრომში განხილულია ენისა და კულტურის დოკუმენტირების თანამედროვე მეთოდები და ინსტრუმენტები, განხილულია რესურსების შექმნის ისტორია და მიმოხილულია თანამედროვე დიგიტალური არქივები და მათი შექმნის ისტორია. ნაშრომში ასევე განხილულია დოკუმენტაციის სფეროსათვის რელევანტური ზოგადთეორიული საკითხები, როგორცაა კულტურის გაგების საკითხი, მეხსიერების ფორმები, კულტურული ტრანსმისია და მისი მნიშვნელობა; ასევე წარმოდგენილია პროექტის ფარგლებში მოპოვებული ემპირიული მასალების ანალიზი.
2. მანანა თანდაშილის, რამაზ ხალვაშის, რუსუდან პაპიაშვილის, მზია ხახუტაიშვილის, გრიგოლ კახიანისა და ნანა ცეცხლაძის მომზადებული მეგზური „ენისა და კულტურის ელექტრონული დოკუმენტირების თანამედროვე ინსტრუმენტები“ (გამომცემლობა „უნივერსალი“, თბ., 2020; შემდგომში: თანდაშილი და სხვ. 2020). მეგზური შეიქმნა პროექტის „ბათუმის ლინგვოკულტუროლოგიური დიგიტალური არქივის“ (შიფრი №DI 2016-46) ფარგლებში. მეგზურში დეტალურადაა განხილულია ენისა და კულტურის

დოკუმენტირების თანამედროვე ინსტრუმენტების ELAN-ის და FLEx-ის გამოყენების კონკრეტული გზები.

გარდა ამისა, პროექტის ფარგლებში სამეცნიერო ჯგუფმა გამოაქვეყნა რამდენიმე სტატია.<sup>9</sup>

#### 5. ლინგვოკულტუროლოგიური არქივის რესურსების პირველადი დოკუმენტირება

ლინგვოკულტუროლოგიური არქივი, როგორც ზემოთ უკვე აღვნიშნეთ, მიზნად ისახავს ინტერდისციპლინური რესურსების თავმოყრას, დიგიტალიზებას და არქივირებას. დიგიტალური არქივები, შესაბამისად, განსხვავებული თემატიკის რესურსებს აერთიანებს. ფართო დიაპაზონის რესურსების აღწერისათვის გამოიყენება მეტამონაცემების უნიფიკაციის სტანდარტი Dublin Core Metadata Initiative (DCMI), რომელიც ორ დონედ იყოფა: მარტივი (Simple) დონე შეიცავს 15 ელემენტს, ხოლო კომპლექტური დონე – დამატებით 3 ელემენტს: 1. Title, 2. Creator, 3. Subject, 4. Description, 5. Publisher, 6. Contributor, 7. Date, 8. Type, 9. Format, 10. Identifier, 11. Source, 12. Language, 13. Relation, 14. Coverage, 15. Rights. დუბლინის ბირთვის კვალიფიკაციური (კომპლექტური) ელემენტების ნაკრებში დამატებით სამი მახასიათებელია შეტანილი: Audience, Provenance, RightsHolder.<sup>10</sup>

დოკუმენტირების პროცესის უმნიშვნელოვანესი ელემენტია მასალისთვის უნიკალური იდენტიფიკატორის (ID) მინიჭება. პროექტის ფარგლებში შევიმუშავეთ ID-ის გენერირების ამგვარი პრინციპი „წელი+ხუთციფრიანი ნომერი“, მაგალითად: 2017-00001, 2017-00002 და ა. შ. საჭიროა აღინიშნოს, რომ ერთი და იგივე უნიკალური იდენტიფიკატორი ავტომატურად განმეორდება მასალის ELAN ფორმატში ტრანსპორტირების მიზნით, რის შედეგადაც მივიღებთ ერთი და იმავე სახელწოდების რამდენიმე ფაილს სხვადასხვა გაფართოებით, მაგალითად: 2017-00001.eaf; 2017-00001.pfsx; 2017-00001.etf და ა. შ.

არქივირებული მასალა მულტიდისციპლინური გამოყენებისთვისაა გამიზნული. იგი ერთნაირად საინტერესოა ენათმეცნიერთათვის, ეთნოლოგებისათვის, ისტორიკოსებისთვის, კონფლიქტოლოგებისთვის, ლიტერატურათმცოდნეებისთვის, ეთნოლინგვისტებისთვის, სოციოლინგვისტებისთვის და ა.შ.

დიგიტალური რესურსების ერთ-ერთ უპირატესობას წარმოადგენს ის გარემოება, რომ იგი საშუალებას გვაძლევს მოვახდინოთ მასალის კომპლექსური ვერიფიცირება (გადამოწმება და დაზუსტება), რისი საშუალებაც პროექტის განხორციელების პირველივე ეტაპზე მოგვეცა, როცა შევუდექით ჯგუფის ხელთ არსებული ვიდეომასალების აღრიცხვას, სელექციასა და რეგისტრაციას. ამ მიზნით პირველ რიგში დავიწყეთ პროფ. შუშანა ფუტკარაძის მიერ 1989–1992 წლებში ჩაწერილი ვიდეოფაილების (3 ვიდეოკასეტა) გაციფრება. ვიდეომასალა მაგნიტურ ფირზე იყო ჩაწერილი, რის გამოც მისი დიდი ნაწილი, სამწუხაროდ, უკვე დაზიანებულიყო. ჩვენ მიერ გაციფრებული (გადარჩენილი) ნაწილი კი უაღრესად საინტერესო აღმოჩნდა სწორედ მასალის ვერიფიცირების თვალსაზრისით. მაგალითად, ერთ-ერთ კასეტაზე ჩაწერილი იყო იმერხეველი (სოფ. იფხრეველი) ხალხური მთქმელის – მევლუდ ღაღანიძის სიმღერა, რომელიც ისამაილ ყარას (შავიშვილის) მიერაა გადაღებული 1988 წლის 28 მაისს. აღნიშნული მასალა პირველად 1993 წელს გამოაქვეყნა შორენა ფუტკარაძემ წიგნში „ჩვენებურების

<sup>9</sup> Tandaschwili et al. 2017; Beridze et al. 2019; Tandaschwili et al. 2020.

<sup>10</sup> ხალვაში, 2018: 88–89.

ქართული“. საგულისხმოა, რომ შ. ფუტკარაძის მიერ გამოქვეყნებული და ჩვენ მიერ გაცემული მასალის შედარებამ ცხადყო, რომ შუშანა ფუტკარაძის პუბლიკაცია განსხვავდება ვიდეომასალაზე დაფიქსირებული ვარიანტისაგან.<sup>11</sup> ეს ფაქტი იმაზე უნდა მეტყველებდეს, რომ შ. ფუტკარაძემ იგივე სიმღერა იმავე ავტორისაგან სხვა დროსაც ჩაიწერა და ამ კონკრეტულ შემთხვევაში არ უსარგებლია ი. ყარას (შავიშვილის) ვიდეოჩანაწერით. ასეა თუ ისე, ვიდეომასალაზე შემონახულია იმავე მთქმელის – მევლუდ დადანიძის იგივე სიმღერა, რაც ტექსტის ვერიფიცირების შესაძლებლობას იძლევა. ყველაზე მთავარი კი ისაა, რომ ვიდეომასალა ინახავს არა მხოლოდ ტექსტს, არამედ გამოსახულებასა და ხმას, რის გამოც დაარქივებული ფაილი უკვე საინტერესოა მუსიკოლოგიური თვალსაზრისით და მასალას სძენს ინტერდისციპლინური კვლევითი ობიექტის მნიშვნელობას.

## 6. პროექტის ფარგლებში მოპოვებული რესურსების თემატური აღწერა

პროექტის მიზანი, როგორც ზემოთ უკვე აღვნიშნეთ, იყო აჭარის რეგიონში ზეპირმეტყველების დოკუმენტირება და არქივირება, რათა საშუალება შეეუქმნათ მეცნიერებს ლინგვოკულტუროლოგიური კვლევა დიგიტალური რესურსებისა და მეთოდების გამოყენებით განახორციელონ. აღნიშნული რეგიონი გამორჩეულ რეგიონს წარმოადგენს არა მხოლოდ საქართველოში, არამედ კავკასიის მასშტაბითაც. მის გამორჩეულობას განაპირობებს ის ფაქტი, რომ ეს რეგიონი ბოლო ხუთი საუკუნის განმავლობაში პოლიტიკური გარიგებების ობიექტად იყო ქცეული და ხელიდან ხელში გადადიოდა, რაც შესაბამისად აისახა კიდევ ამ რეგიონზე – გარდა ეთნიკური სიჭრელისა, აჭარა მულტილინგურობითა და მულტირელიგიურობითაც ხასიათდება.

პროექტის ფარგლებში მოძიებული (აღრე გადაღებული მასალები) და საკუთრივ პროექტის ფარგლებში ჩაწერილი მასალა მრავალფეროვანია როგორც თემატურად, ასევე დარგობრივად და ჟანრობრივადაც. მასალაში კარგად ჩანს აჭარელთა სამეურნეო-ყოფითი თავისებურებანი (აბრეშუმის ჭიის მოყვანა, ბოსტნეული კულტურები და ა.შ.), ტრადიციები (საქორწინო, დაკრძალვის, მენავეობის, სტუმარ-მასპინძლობისა და სამზარეულოს წესები: ნაცრის კვერი, კეცის მჭადი, ბეთმეზი, ლეჩერი, დაწნული ყველი და ა. შ.) და რელიგიური რწმენა-წარმოდგენები, კულტურული და ისტორიული მემკვიდრეობა. ფიქსაციისა და არქივირების პროცესში საგულისხმო მასალა დაგროვდა ისეთი საკითხების შესახებ, როგორცაა: მუჰაჯირობა – რუსეთ-თურქეთის ომის (1877–1878 წწ.) დროს ორი იმპერიის შეთანხმების საფუძველზე ქართველთა გადასახლება თურქეთის სხვადასხვა პროვინციებში; საზღვრების დემარკაციის დროს (1929 წ.) თურქეთსა და საბჭოთა იმპერიას შორის მოქცეული ქართული სოფლების „გახლეჩა“; კომუნისტური რეპრესიები (1924, 1937). შესაბამისად, ლინგვოკულტუროლოგიური დიგიტალური არქივში დაცული მასალა მნიშვნელოვან ბაზას ქმნის ინტერდისციპლინური კვლევის თვალსაზრისით.

## 7. დასკვნა

პროექტის „ბათუმის ლინგვოკულტუროლოგიური დიგიტალური არქივი“ ძირითადი შედეგია ის, რომ საქართველოში პირველად შეიქმნა დოკუმენტირებისა

<sup>11</sup> შდრ.: ფუტკარაძე 1993: 280–281.

და არქივირების თანამედროვე ჩარჩო-სტანდარტი, რომელიც ხელს შეუწყობს ინტერდისციპლინური კვლევების განვითარებას არა მხოლოდ საქართველოში, არამედ კავკასიის რეგიონშიც. ამასთანავე, პროექტის ფარგლებში მოპოვებული და არქივირებული რესურსები ღია რესურსებს წარმოადგენს, ინტერნაციონალური მოხმარების პროდუქტია, და, შესაბამისად, ხელმისაწვდომია არაქართულენოვანი მომხმარებლისათვისაც.

#### ლიტერატურა

- თანდაშილი (2020): მანანა თ., ენისა და კულტურის დოკუმენტირების თანამედროვე მეთოდები და ინსტრუმენტები: ლინგვოკულტუროლოგიური დიგიტალური არქივის მაგალითზე. თბილისი: უნივერსალი. <https://rustaveli.org.ge/res/docs/e6a82ece21d145ec651a51990efaa881530df0a6.pdf>.
- თანდაშილი და სხვ. (2020): მანანა თანდაშილი, რამაზ ხალვაში, რუსუდან პაპიაშილი, მზია ხახუტაიშვილი, გრიგოლ კახიანი, ნანა ცეცხლაძე, ენისა და კულტურის ელექტრონული დოკუმენტირების თანამედროვე ინსტრუმენტები: მეგზური. თბილისი: უნივერსალი. <https://rustaveli.org.ge/res/docs/15221d2c35a4e0aafa12df872d8f0a326102e5d1.pdf>.
- ფუტკარაძე (1993): შუშანა ფ., ჩვენებურების ქართული. I. ბათუმი: აჭარის ჟურნალ-გაზეთების გამომცემლობა. <https://dspace.nplg.gov.ge/handle/1234/454756>.
- ხალვაში (2018): რამაზ ხ., დიგიტალური ჰუმანიტარიის შესავალი. ბათუმი: ბათუმის შოთა რუსთაველის სახელმწიფო უნივერსიტეტის გამომცემლობა. <https://www.academia.edu/35956471/>.
- Beridze et al. (2019): Khatuna B., Manana Tandaschwili, Marina Giorgadze, Ramaz Khalvashi, “Translation of the Batumi Linguocultural Digital Archive: The methodology of Glossing and Tagging Social, Political and Cultural Information in Translation of DRSP”, *International Journal of Multilingual Education* 13, 2019, 25–38. [https://www.multilinguaeducation.org/storage/uploads/articles\\_contents/4.%20Khatuna-DigiArchive.pdf](https://www.multilinguaeducation.org/storage/uploads/articles_contents/4.%20Khatuna-DigiArchive.pdf).
- Ethnologue* 1988: *Ethnologue. Languages of the World*. Ed. by Barbara Grimes. 11<sup>th</sup> edn. Dallas, Texas: Summer Institute of Linguistics. <https://books.google.de/books?id=UWdiAAAAMAAJ>.
- Hale et al. (1992): Ken H., Michael Krauss, Lucille J. Watahomigie, Akira Y. Yamamoto, Colette Craig, LaVerne Masayeva Jeanne, Nora C. England, “Endangered Languages”, *Language* 68/1, 1–42. <https://doi.org/10.2307/416368>; <https://www.jstor.org/stable/416368>.
- Himmelman (1998): Nikolaus P. H., “Documentary and descriptive linguistics”, *Linguistics* 36, 161–195. <https://doi.org/10.1515/ling.1998.36.1.161>.
- Krauss (1992): Michael K., “The world’s languages in crisis”, *Language* 68/1, 4–10 [part of Hale et al. 1992].
- Tandaschwili et al. (2017): Manana T., Ramaz Khalvashi, Khatuna Beridze, Mzia Khakhutaishvili, Nana Tsetschladze, “Batumi Linguocultural Digital Archive (Contemporary Technological Achievements for the Database Arrangement of the Folklore Resources)”, *International Journal of Multilingual Education* 10, 2017, 52–68. [https://www.multilinguaeducation.org/storage/uploads/articles\\_contents/6-tandashvili.pdf](https://www.multilinguaeducation.org/storage/uploads/articles_contents/6-tandashvili.pdf).
- Tandaschwili et al. (2020): Manana T., Ramaz Khalvashi, Grigol Kakhiani, Mzia Khakhutaishvili, Nana Tsetschladze, Rusudan Papiashvili, “Modern Program Packages (FLEx and Elan) of the Language Data Management and the Prospects of Their Usage”, *International Journal of Multilingual Education* 16, 2020, 32–42. [https://www.multilinguaeducation.org/storage/uploads/articles\\_contents/3.%20tandashvili.pdf](https://www.multilinguaeducation.org/storage/uploads/articles_contents/3.%20tandashvili.pdf).

## Batumi Linguoculturological Digital Archive: New Perspectives of the Documentation of the Adjara Region

Ramaz Khalvashi (Batumi)

The paper introduces the “Batumi Linguocultural Digital Archive”, an initiative that was implemented through financial support by the Shota Rustaveli National Science Foundation during the period of 2016 to 2019. The project was specifically designed to document the linguistic and cultural heritage of the Adjara region, Georgia. The archival holdings encompass an array of digital resources, including audio and video materials, totalling 100 hours of content formatted in MP3 and AVI.

Among these holdings, a subset of 30 hours consists of audio and video materials that have undergone transcription. Additionally, 10 hours of video materials have received comprehensive multimedia annotation, featuring transcriptions, glosses, and interlinearisation in the form of EAF-type resources. Furthermore, the archive includes 5 hours of video content with multimedia annotations that comprise English translations, also structured as EAF-type resources.

The development of the Linguistic and Cultural Digital Archive marked a pivotal moment in advancing digital documentation and archiving practices in Georgia. The initiative not only introduced innovative research methodologies but also leveraged digital resources, thereby significantly enhancing the efficiency of endeavours focussing on the research in the Humanities in Georgia. The primary achievement of the “Batumi Linguocultural Digital Archive” project lies in the creation of a modern, standardised framework for the documentation and archiving of cultural and linguistic materials in Georgia, a milestone that is poised to foster the development of interdisciplinary studies within the country.

Furthermore, it is worth highlighting that the resources acquired and catalogued under the project’s auspices have transcended national boundaries, serving as valuable assets with international applicability. The materials have been translated into English, rendering them accessible to a broader, non-Georgian speaking audience. This effort toward internationalisation underscores the project’s commitment to facilitating cross-cultural exchange and knowledge dissemination.